



**Politecnico  
di Torino**

**Politecnico di Torino**

Corso di Laurea Magistrale in Ingegneria Informatica

A.A. 2022/2023

Sessione di Laurea dicembre 2023

**Private Cloud per la ricerca:  
implementazione di una regione OpenStack  
basata sul modello federato GARR**

Relatore:  
Prof. Maurizio MORISIO

Candidata:  
Paola BELMONTE



# Indice

Indice .....	I
Elenco delle tabelle .....	IV
Elenco delle figure.....	V
Sommario .....	1
Capitolo 1. Introduzione.....	3
Capitolo 2. Il cloud computing, modelli e diffusione .....	7
2.1 <b>Concetti di base e modelli</b> .....	8
2.1.1    I modelli di servizio.....	8
2.1.2    I modelli di deployment .....	11
2.1.3    Virtualizzazione: VM e container .....	14
2.1.4    Interfacce di accesso ai servizi cloud .....	16
2.2 <b>Stato dell'arte e maturità del cloud</b> .....	17
2.2.1    Diffusione del cloud in Europa e in Italia.....	17
2.2.2    La maturità tecnologica del cloud computing .....	18
2.2.3    Casi studio e stato dell'arte nella ricerca .....	20
2.2.4    Punti di forza e sfide aperte.....	22
2.2.5    L'adozione nel settore pubblico e le politiche europee.....	27
2.2.6    Il cloud nelle università.....	31
2.2.7    Cloud repatriation.....	35
Capitolo 3. Il cloud nella Pubblica Amministrazione italiana.....	41
3.1 <b>Organizzazioni di riferimento</b> .....	42
3.2 <b>Piano Triennale per l'informatica 2022-2024</b> .....	42
3.3 <b>Censimento del patrimonio ICT della PA</b> .....	43
3.4 <b>Dal censimento alla Strategia Cloud Italia</b> .....	45
3.5 <b>Razionalizzazione dei data center</b> .....	47
3.6 <b>Il principio cloud first</b> .....	48
3.7 <b>Qualificazioni fornitori e servizi cloud</b> .....	49

3.8	Il Polo Strategico Nazionale.....	52
3.9	Formazione e cambio culturale .....	54
3.10	La spesa ICT nella PA.....	55
3.11	Cronologia norme e risultati .....	57
3.12	Ricerca e università .....	58
<b>Capitolo 4. Il modello federato GARR e la piattaforma OpenStack .....</b>		<b>60</b>
4.1	<b>Il modello Cloud GARR .....</b>	<b>61</b>
4.1.1	Il Consortium GARR.....	61
4.1.2	Il Community cloud nel contesto europeo delle NREN .....	62
4.1.3	Il progetto Cloud GARR.....	63
4.1.4	Architettura di riferimento .....	64
4.1.5	I servizi offerti.....	66
4.1.6	I vantaggi del Community cloud.....	68
4.2	<b>Approfondimento sulle soluzioni software.....</b>	<b>69</b>
4.2.1	MAAS.....	69
4.2.2	Juju.....	71
4.2.3	Ceph .....	77
4.3	<b>La piattaforma OpenStack .....</b>	<b>80</b>
4.3.1	Rilascio delle versioni.....	80
4.3.2	Stato dell'arte nella ricerca accademica .....	81
4.3.3	Panoramica dei casi d'uso .....	83
4.3.4	Architettura logica e microservizi .....	85
4.3.5	Ruoli e servizi nel cluster .....	87
4.3.6	Servizi del cloud controller.....	88
4.3.7	Configurazione di rete .....	89
4.3.8	Caratteristiche per la scalabilità .....	90
4.4	<b>L'architettura federata GARR.....</b>	<b>92</b>
4.4.1	Il modello multi-region.....	92
4.4.2	Concetti di Keystone.....	93
4.4.3	La delega amministrativa.....	94
4.4.4	Nuove linee di indirizzo.....	97
4.4.5	Federazione e autonomia.....	98
<b>Capitolo 5. Il cloud OpenStack Polito .....</b>		<b>99</b>
5.1	<b>Progettazione infrastruttura e servizi OpenStack .....</b>	<b>100</b>
5.1.1	Hardware .....	100
5.1.2	Pianificazione dei servizi OpenStack e ruoli server .....	101
5.1.3	Risorse disponibili e overcommit.....	103
5.1.4	Cluster di test .....	104
5.2	<b>Installazione cloud OpenStack Polito.....</b>	<b>105</b>

5.2.1	Installazione MAAS .....	105
5.2.2	Aggiunta server al catalogo MAAS.....	106
5.2.3	Installazione Juju .....	107
5.2.4	Deploy cloud OpenStack con Juju .....	109
5.2.5	Problemi riscontrati e risoluzione.....	113
5.2.6	Federazione GARR .....	114
<b>5.3</b>	<b>Utilizzo del cloud OpenStack Polito .....</b>	<b>115</b>
5.3.1	Profili disponibili per le VM .....	115
5.3.2	Modalità di utilizzo.....	116
5.3.3	Attività di evoluzione futura.....	117
<b>5.4</b>	<b>Bilancio del progetto.....</b>	<b>118</b>
<b>Capitolo 6. Analisi economica e confronto.....</b>		<b>120</b>
<b>6.1</b>	<b>Infrastruttura on-premise.....</b>	<b>121</b>
6.1.1	Modello costi on-premise.....	121
6.1.2	Costi hardware per server .....	122
6.1.3	Costi hardware per rack e accessori .....	125
6.1.4	Costi hardware per networking e accessori .....	127
6.1.5	Costi di licenza.....	129
6.1.6	Consumi energetici.....	130
6.1.7	Costi del personale.....	133
6.1.8	Riepilogo dei costi on-premise.....	134
<b>6.2</b>	<b>Infrastruttura su cloud pubblico .....</b>	<b>135</b>
6.2.1	Modello di tariffazione per VM Azure.....	135
6.2.2	Servizi a costi aggiuntivi .....	138
<b>6.3</b>	<b>Scenari di confronto .....</b>	<b>140</b>
6.3.1	Scenario 1 .....	140
6.3.2	Scenario 2 .....	142
6.3.3	Osservazioni finali.....	142
<b>Capitolo 7. Conclusioni e spunti futuri.....</b>		<b>144</b>
<b>Bibliografia .....</b>		<b>149</b>

# Elenco delle tabelle

Tabella 1. Elenco acronimi utilizzati .....	5
Tabella 2. Confronto tra i modelli di cloud deployment.....	13
Tabella 3. Cloud computing: punti di forza e sfide .....	24
Tabella 4. Adozione del cloud nei paesi europei: fattori e obiettivi strategici.....	29
Tabella 5. Data center della PA: risultato del censimento AgID.....	44
Tabella 6. Cloud computing nella PA italiana: cronologia norme e risultati.....	57
Tabella 7. OpenStack: ultime 5 versioni rilasciate.....	80
Tabella 8. OpenStack Polito: caratteristiche hardware dei server .....	102
Tabella 9. Ruoli dei nodi e distribuzione dei servizi OpenStack.....	102
Tabella 10. Elenco reti nella configurazione OpenStack.....	104
Tabella 11. OpenStack Polito: elenco dei flavor disponibili per le VM.....	115
Tabella 12. OpenStack Polito: progetti e risorse utilizzate.....	116
Tabella 13. OpenStack Polito: risorse utilizzate in un mese .....	117
Tabella 14. Infrastruttura on-premise: fasi dell'investimento .....	121
Tabella 15. Convenzione Consip: caratteristiche hardware server.....	123
Tabella 16. Acquisto server: stima delle risorse totali .....	123
Tabella 17. Costi on-premise: apparati di rete .....	128
Tabella 18. Licenze Microsoft.....	129
Tabella 19. Costi on-premise: consumi apparecchiature .....	131
Tabella 20. Analisi economica: costo annuale per consumi elettrici.....	133
Tabella 21. Costi on-premise: personale tecnico .....	133
Tabella 22. Costi on-premise: apparecchiature hardware .....	134
Tabella 23. Analisi economica: TCO on-premise su 1 e 5 anni di esercizio .....	134
Tabella 24. Cloud Azure: esempio tariffazione VM .....	136
Tabella 25. Analisi economica: ipotesi di carico on-premise e cloud .....	140
Tabella 26. Cloud Azure: scenario 1, preventivo costi VM.....	141
Tabella 27. Analisi economica: confronto costi on-premise e cloud .....	141
Tabella 28. Analisi economica: punto di pareggio Cloud e TCO on-premise.....	142

# Elenco delle figure

Figura 1. Cloud: modelli di servizio.....	9
Figura 2. Cloud: analogia “Pizza as a Service”.....	10
Figura 3. Confronto VM e container. ....	14
Figura 4. Diffusione del cloud nei paesi europei. ....	17
Figura 5. Adozione del cloud nei paesi europei per modelli di servizio.....	18
Figura 6. Gartner Hype Cycle 2023.....	19
Figura 7. Inversione della spesa per gli asset ICT entro il 2025. ....	20
Figura 8. Università: fattori determinanti per l’adesione al cloud. ....	33
Figura 9. Sondaggio “Cloud repatriation” di IDC. ....	35
Figura 10. Sondaggio “Cloud repatriation” di S&P Global. ....	36
Figura 11. Adozione della soluzione Multicloud, sondaggio IDC. ....	40
Figura 12. Censimento AgID: distribuzione data center della PA.....	44
Figura 13. Strategia Cloud Italia: sfide e fasi attuative.....	46
Figura 14. ACN: qualificazione dei servizi cloud .....	50
Figura 15. Strategia Cloud Italia: classificazione dati e servizi cloud. ....	51
Figura 16. PSN: dislocazione sul territorio italiano .....	53
Figura 17. La spesa ICT nella PA: spesa totale cloud.....	55
Figura 18. La spesa ICT nella PA: i servizi cloud utilizzati .....	56
Figura 19. La mappa della rete GARR .....	61
Figura 20. OpenStack, MAAS e Juju: schema generale.....	64
Figura 21. Cloud GARR: dati sull’utilizzo. ....	67
Figura 22. MAAS: interfaccia di gestione.....	69
Figura 23. MAAS: schema di funzionamento. ....	70
Figura 24. Juju: esempio di relazioni tra application unit. ....	71
Figura 25. Juju: esempio application unit bare metal e in container LXD.....	72
Figura 26. Juju controller: architettura ad alto livello. ....	75
Figura 27. OpenStack Report: utilizzo di Ceph come block storage.....	77
Figura 28. Ceph: opzioni di storage.....	77
Figura 29. Ceph: diagramma dell’archiviazione dati. ....	79
Figura 30. OpenStack: numero di sviluppatori di aziende e volontari.....	82
Figura 31. OpenStack: numero di commit per aziende e volontari.....	82
Figura 32. OpenStack al CERN: i numeri dell’infrastruttura.....	83

Figura 33. OpenStack: architettura logica e microservizi. ....	85
Figura 34. OpenStack: struttura interna di un modulo. ....	86
Figura 35. OpenStack: architettura multi nodo. ....	87
Figura 36. OpenStack: schema delle reti. ....	89
Figura 37. OpenStack: rete interna e rete floating. ....	90
Figura 38. OpenStack: schema di architettura multi-region. ....	91
Figura 39. Modello federato GARR: servizi globali e locali. ....	92
Figura 40. Modello federato GARR: schema di delega amministrativa. ....	95
Figura 41. OpenStack Polito: isola rack dedicata nel data center. ....	100
Figura 42. OpenStack Polito: ruoli e servizi sui nodi. ....	101
Figura 43. VM dell'ambiente di test OpenStack Polito. ....	104
Figura 44. Cloud Polito: MAAS. ....	106
Figura 45. GARR Cloud: login attraverso IDEM. ....	114
Figura 46. Convenzione Consip: configuratore server. ....	124
Figura 47. Costi on-premise: configurazione rack. ....	125
Figura 48. Convenzione Consip: configuratore server e accessori rack. ....	126
Figura 49. Modello di rete spine-leaf. ....	127
Figura 50. Convenzione Consip: tariffe energia elettrica lotto Piemonte. ....	130
Figura 51. Cloud Azure: catalogo delle VM. ....	135
Figura 52. Cloud Azure: calcolatore prezzi VM. ....	137
Figura 53. Cloud Azure: costi aggiuntivi da considerare. ....	138
Figura 54. Cloud Azure: costi per traffico di rete in uscita. ....	139





# Sommario

I dati della diffusione del cloud computing in Europa e in Italia mostrano che questa tecnologia è in una fase di piena maturità. La trasformazione innescata dai nuovi modelli ha raggiunto anche il settore pubblico che ha investito il cloud di un ruolo abilitante; non a caso molti paesi hanno inserito la politica cloud first nelle proprie linee di indirizzo. Enti di ricerca e università, spesso in veste di pionieri delle innovazioni, hanno adottato il nuovo paradigma cloud anche se con modalità differenti tra gli ambiti di didattica, ricerca e attività amministrative. L'analisi dei punti di forza e delle sfide dei modelli di deployment (Public, Private, Hybrid, Community cloud) rivela che non esiste un'unica soluzione per ogni esigenza. Inoltre il mercato dei servizi cloud è in crescita, ma meno del 15% dei provider è europeo: ciò porta a interrogarsi sulla dipendenza tecnologica dell'Europa dalle multinazionali statunitensi, fornitrici di servizi anche tramite i contratti di appalti pubblici.

In questo contesto si colloca il lavoro svolto per l'implementazione di un private cloud OpenStack, basato sul modello federato GARR, presso la direzione ISiAD del Politecnico di Torino. GARR, ente no profit che gestisce la rete nazionale a banda ultra larga dedicata agli istituti di istruzione e ricerca, ha sviluppato una soluzione che combina i modelli Private e Community cloud. Lo schema federato permette di usare servizi comuni come l'autenticazione centralizzata, mantenendo l'autonomia sulle proprie risorse. La collocazione delle infrastrutture di storage e calcolo in data center sotto il controllo della comunità garantisce la sovranità su dati e tecnologie, anche nel rispetto di vincoli legislativi come il GDPR. La base software totalmente open source riduce il rischio di vendor lock in; l'adesione alla federazione consente di creare e mantenere le competenze all'interno della comunità, favorendo scambi e collaborazione.

Fulcro dell'architettura è OpenStack, progetto cloud modulare ideato da NASA e Rackspace, basato su una vasta comunità di sviluppo. Per la virtualizzazione lato storage si usa Ceph, dotato di una struttura distribuita e molto robusta. Due gli strumenti di automazione nella gestione hardware e software del ciclo di vita dei servizi: MAAS (Metal-As-A-Service) per installare e gestire il parco hardware e Juju per orchestrare e distribuire i singoli servizi sui nodi del cluster.

L'infrastruttura cloud gestita da ISIAD conta 35 server in configurazione iperconvergente, per un totale di 30TB di RAM, 4500 core CPU, 1.2PB HDD e 85TB SSD di storage raw. L'acquisto di ulteriori 30 server raddoppierà le risorse a disposizione entro fine 2024. Il cloud è operativo, offre servizi IaaS e un virtual data center per l'esecuzione di macchine virtuali, usate da docenti e ricercatori per finalità di didattica e ricerca. L'adesione al modello federato GARR ha permesso di implementare un'architettura replicabile e progettata per lo stesso tipo di contesto, potendo inoltre contare sul supporto tecnico di specialisti cloud. Si espone un'analisi economica di sintesi; l'attività non può dirsi conclusa: molte sfide restano aperte. Sul piano organizzativo, nell'ottica di una maggiore diffusione del servizio, sarebbero opportune la definizione di policy di utilizzo e strategie per l'organizzazione del personale.

Dal punto di vista tecnico è cruciale ottimizzare le attività di gestione e di manutenzione evolutiva dell'infrastruttura, soprattutto in vista della crescente complessità hardware e software.

# Capitolo 1.

## Introduzione

Con questa tesi mi propongo di esplorare in dettaglio l'implementazione di un cluster OpenStack come regione del modello federato GARR presso il Politecnico di Torino.

L'idea nasce con l'intento di presentare in maniera sintetica l'impegnativo lavoro svolto. L'installazione di un sistema informativo non costituisce un elemento isolato bensì è parte integrante del contesto in cui si inserisce, di cui riflette le aspettative e i requisiti. Pertanto, prima di entrare nei dettagli tecnici, ho deciso di delineare lo scenario più ampio del *cloud computing*, a partire dai modelli di erogazione, allo stato dell'arte nella ricerca accademica e alle norme che regolano l'ambito della Pubblica Amministrazione italiana.

Negli ultimi anni, il paradigma del cloud computing ha rivoluzionato radicalmente il modo in cui le organizzazioni gestiscono le risorse informatiche. Attraverso questo lavoro il mio intento è di evidenziare le sfide e le opportunità che derivano dall'adozione di soluzioni cloud federate e comprendere il ruolo della piattaforma OpenStack in tale contesto. L'analisi si concentra sulla specifica configurazione presso il Politecnico di Torino, offrendo un quadro dettagliato dell'esperienza di implementazione, esaminando i benefici e le lezioni apprese durante il processo.

Il lavoro è articolato come segue. Nel capitolo 2, partendo dalle definizioni, viene presentata un'analisi dei modelli di distribuzione del cloud e della diffusione di questa tecnologia in Europa e in Italia, traendo i dati dalle statistiche pubblicate da Eurostat, Istat e dall'Osservatorio Cloud del Politecnico di Milano. Si prende poi in esame come la ricerca accademica ha affrontato alcuni tra i molteplici temi correlati al cloud, descrivendone i benefici e le principali criticità, passando all'adozione del cloud computing nel settore pubblico e nel contesto delle università, con una menzione alle politiche europee in materia. Infine viene analizzato il recente fenomeno chiamato *cloud repatriation* a dimostrazione che i modelli di adozione non sono statici e che le organizzazioni hanno compreso come modellare dinamicamente la tecnologia in base alle proprie necessità.

Il capitolo 3 è incentrato sull'adozione del cloud nella Pubblica Amministrazione italiana. Si analizzano le linee guida tracciate dal Piano Triennale per l'informatica

nella Pubblica Amministrazione per promuovere la trasformazione digitale dell'Italia e per pianificare il percorso di evoluzione tecnologica nelle PA. Viene descritto il principio *cloud first*, punto nodale all'interno della Strategia Cloud Italia e adottato anche in altri stati europei, e si esaminano i processi ad esso correlati. Tra questi la qualificazione dei fornitori e dei servizi cloud, il censimento degli asset informatici realizzato da AgID e i progetti di consolidamento e razionalizzazione delle strutture, comprese l'eventuale dismissione dei data center obsoleti e la migrazione dei servizi verso data center più affidabili oppure in cloud. Infine si descrive la situazione di università e enti di ricerca pubblici che, pur essendo a tutti gli effetti delle PA e rientrando nel campo di applicazione delle norme, sono stati oggetto di una deroga che consente di mantenere e far evolvere i propri data center on-premise.

Nel capitolo 4 viene dato ampio spazio al modello cloud federato GARR e alla piattaforma OpenStack. GARR, ente no profit che possiede e gestisce la rete nazionale dedicata alla comunità dell'istruzione e della ricerca, ha ideato un sistema federato che combina i modelli Private e Community cloud. Si approfondisce pertanto l'architettura di riferimento basata su OpenStack, la piattaforma open source di tipo modulare sviluppata da NASA e Rackspace Cloud. Inoltre si analizzano in dettaglio le soluzioni software a corredo quali gli strumenti di automazione Juju e MAAS e la virtualizzazione lato storage con Ceph. Infine si esaminano i componenti del cloud federato GARR che costituiscono il modello multi-region, approfondendo caratteristiche e vantaggi del *Community cloud* e della federazione.

Il capitolo 5 è dedicato all'implementazione del cloud OpenStack presso la direzione ISIAD del Politecnico di Torino. Vengono descritte l'infrastruttura hardware dedicata e le procedure per l'installazione e l'adesione al modello federato GARR. Si riportano inoltre i dettagli sull'utilizzo del cloud che, una volta operativo, offre servizi IaaS e un datacenter virtuale ad uso del personale docente e ricercatore.

Nel capitolo 6 viene esposta un'analisi economica di sintesi, confrontando i costi di un'infrastruttura *on-premise* con i costi per l'esecuzione delle macchine virtuali su un servizio di cloud pubblico.

Infine nel capitolo 7 si sintetizzano le conclusioni sul lavoro, descrivendo il bilancio del progetto e alcune considerazioni sulla convenienza economica. L'attività non può dirsi conclusa con il setup del cluster OpenStack, in quanto molte attività restano aperte e richiedono un'accurata pianificazione. In chiusura, si raccolgono gli sviluppi futuri e diversi spunti di miglioramento in vista della crescente complessità hardware e software e nell'ottica di una più ampia diffusione del servizio in ateneo.

Prima di cominciare, è doverosa una nota in merito alla terminologia impiegata. Ho utilizzato il termine *host* non nel senso generico di host connesso ad una rete TCP/IP ma con l'accezione, peraltro piuttosto diffusa, di virtualizzatore, vale a dire un sistema server in grado di virtualizzare le periferiche hardware e che ospita diverse *virtual machine* (VM). Inoltre, per semplicità di esposizione, spesso ho impiegato come fossero

sinonimi i due termini *host* e *hypervisor*, anche se denotati da sfumature concettuali differenti. Infine nella tabella seguente è riportato l'elenco completo degli acronimi utilizzati nel documento.

Tabella 1. Elenco acronimi utilizzati

<b>Sigla</b>	<b>Enunciato completo</b>
AMQP	Advanced Message Queuing Protocol
API	Application Programming Interface
AWS	Amazon Web Services
BTU	British Thermal Units
CLI	Command Line Interface
CPU	Central Processing Unit
CSP	Cloud Service Provider
DHCP	Dynamic Host Configuration Protocol
DNAT	Dynamic Network Address Translation
FQDN	Fully Qualified Domain Name
GCP	Google Cloud Platform
GRE	Generic Routing Encapsulation
GUI	Graphical User Interface
HA	High Availability
HCI	Hyper Converged Infrastructure
HDD	Hard Disk Drive
HTTP	HyperText Transfer Protocol
IaaS	Infrastructure As A Service
IoT	Internet of Things
IDEM	IDentity Management
IPMI	Intelligent Platform Management Interface
JSON	JavaScript Object Notation
KVM	Kernel-based Virtual Machine
KVM	Keyboard, Video and Mouse
LACP	Link Aggregation Control Protocol
LXD	Linux Container Daemon
NAT	Network Address Translation

<b>Sigla</b>	<b>Enunciato completo</b>
NREN	National Research and Education Network
NVME	Non-Volatile Memory Express
OIDC	OpenID Connect
OSCC	Open Source Cloud Computing
PaaS	Platform as a Service
PA	Pubblica Amministrazione
PAC	Pubblica Amministrazione Centrale
PAL	Pubblica Amministrazione Locale
PDU	Power Distribution Unit
PUE	Power Usage Effectiveness
PUN	Prezzo Unico Nazionale
PXE	Preboot Execution Environment
RAID	Redundant Array of Independent Disks
RAM	Random Access Memory
RBAC	Role-based access control
R&E	Research and Education
REST	REpresentational State Transfer
SaaS	Software as a Service
SAN	Storage Area Network
SDD	Solid State Drive
SLA	Service Level Agreement
SNAT	Source Network Address Translation
SOA	Service-oriented architecture
TCO	Total Cost of Ownership
TFLOPS	Tera Floating Point Operations Per Second
TFTP	Trivial File Transfer Protocol
ToR	Top-of-Rack
UE	Unione Europea
UPS	Uninterruptible Power Supply
VDI	Virtual Desktop Infrastructure
VLAN	Virtual Local Area Network
VM	Virtual Machine
VXLAN	Virtual eXtensible Local Area Network
XML	eXtensible Markup Language

## Capitolo 2.

# Il cloud computing, modelli e diffusione

Lo scorso 26 ottobre 2022 le principali testate italiane<sup>1</sup> hanno dedicato un titolo al venticinquesimo compleanno del cloud computing, convenzionalmente fatto coincidere con la prima definizione fornita nel 1997 dal prof. Ramnath Chellappa, illustrando ad una conferenza a Dallas il proprio lavoro “Intermediaries in Cloud-Computing: A New Computing Paradigm” [1]:

*Computing has evolved from a main-frame-based structure to a network-based architecture. While many terms have appeared to describe these new forms, the advent of electronic commerce has led to the emergence of ‘cloud computing’.*

È parere unanime che si tratti di una delle tecnologie più rivoluzionarie della seppur giovane storia dell’informatica, in grado di apportare trasformazioni profonde sia nell’ambito dell’IT sia in tutti i settori correlati.

La definizione universalmente accettata per il cloud computing è quella fornita dal NIST<sup>2</sup>, che in un breve rapporto del 2011 [2] propone lo schema composto da tre modelli di servizio (*service models*), quattro modelli di distribuzione (*deployment models*) e cinque caratteristiche essenziali. In base alla natura del servizio offerto agli utenti, emergono le tre categorie *as-a-Service* IaaS, PaaS, SaaS. In base all’implementazione e all’operatività si distinguono i quattro modelli *Public*, *Private*, *Community* e *Hybrid* cloud, caratterizzati dalla diversa collocazione fisica delle infrastrutture e da diversi perimetri di accesso dall’esterno. Infine, le cinque caratteristiche fondamentali che differenziano il modello cloud dall’approccio computing tradizionale si riassumono in:

---

<sup>1</sup> A titolo esemplificativo si veda: <https://www.infodata.ilsole24ore.com/2022/10/26/oggi-e-il-compleanno-del-cloud-computing-day-cinque-neri-sulla-piu-moderna-rivoluzione-informatica>

<sup>2</sup> National Institute of Standards and Technology: è un’agenzia del Dipartimento del Commercio USA oltre che uno dei più antichi laboratori di scienze e tecnologie. Definisce gli standard per vari settori scientifici e tecnologici, obbligatori negli USA e punto di riferimento globale.



utilizzo autonomo (*on-demand self-service*), accesso via rete, risorse di calcolo suddivise in pool e condivise, elasticità di crescita e misurabilità.

Dal punto di vista tecnologico l'avvento del cloud computing è una sorta di "risultato collettivo" basato sul progresso in diversi campi dell'informatica. Infatti gli sviluppi paralleli di molteplici tecnologie sono stati combinati tra loro e hanno contribuito a formarne le basi: hardware, reti e telecomunicazioni, architetture applicative (virtualizzazione e *service oriented architecture*), calcolo distribuito (*grid computing*), tecnologie web (*web services*) [3].

## 2.1 Concetti di base e modelli

In questa sezione si affrontano i modelli di servizio e di distribuzione del cloud con un riferimento ai principali concetti teorici che ne costituiscono i presupposti.

### 2.1.1 I modelli di servizio

In base alla natura del servizio offerto e al livello di astrazione, si distinguono i tre principali modelli [3, pp. 76-83], [4, pp. 13-18], [5].

- *Infrastructure as a Service* (IaaS). Fornisce funzionalità di data center virtuale, in particolare risorse informatiche virtualizzate come server, reti e storage. Gli utenti hanno il controllo completo sull'ambiente operativo, sui sistemi e i livelli software. Sono demandate al fornitore tutte le attività di gestione di data center e infrastrutture, quali ad esempio l'installazione hardware, i cablaggi, gli aggiornamenti firmware, i consumi energetici e degli impianti raffreddamento. IaaS è il pilastro su cui è costruita un'architettura di cloud computing. Amazon Web Services (AWS), Google Cloud Platform, Microsoft Azure.
- *Platform as a Service* (PaaS). Offre ambienti di hosting e framework di sviluppo per creare, testare, distribuire applicazioni web. Usa la stessa infrastruttura di base di IaaS, ma include anche sistemi operativi, middleware, strumenti di sviluppo, librerie runtime e sistemi di database.
- *Software as a Service* (SaaS). Rappresenta l'offerta di soluzioni software pronte per essere acquistate e utilizzate attraverso Internet, eseguite direttamente tramite browser web, senza doversi occupare dell'infrastruttura sottostante. Il fornitore gestisce l'infrastruttura, i sistemi operativi, il middleware e i dati necessari per fornire il programma, garantendo che il software sia disponibile quando e dove i clienti ne hanno bisogno. Spesso le applicazioni SaaS vengono, senza la necessità di download o installazioni.

Tra le aziende che operano come *Cloud Service Provider* (CSP) ne emergono alcune di grandi dimensioni che forniscono servizi su larga scala a livello enterprise. Tali organizzazioni, definite *hyperscaler*, sono caratterizzate da infrastrutture distribuite su

diverse regioni geografiche, connessioni di rete ad alta velocità e un'ampia gamma di servizi cloud. Rientrano in questa categoria Amazon Web Services, Google Cloud, Microsoft Azure, IBM Cloud e Alibaba Cloud.

Prendendo come base i servizi forniti dai principali CSP, alcuni esempi di IaaS sono Amazon EC2, lanciato nel 2006 e uno dei primi servizi di cloud computing, Google Compute Engine (GCE) e Microsoft Azure. Esempi di PaaS sono Microsoft Azure App Service e Google App Engine, mentre rientrano tra le casistiche di SaaS le soluzioni Microsoft 365 e Google Workspace (Gmail, Google Docs, Sheets, Drive, Calendar ecc..).

La Figura 1 presenta uno schema dei modelli di servizio in cui vengono confrontate le attività di gestione in carico al cliente e al fornitore.

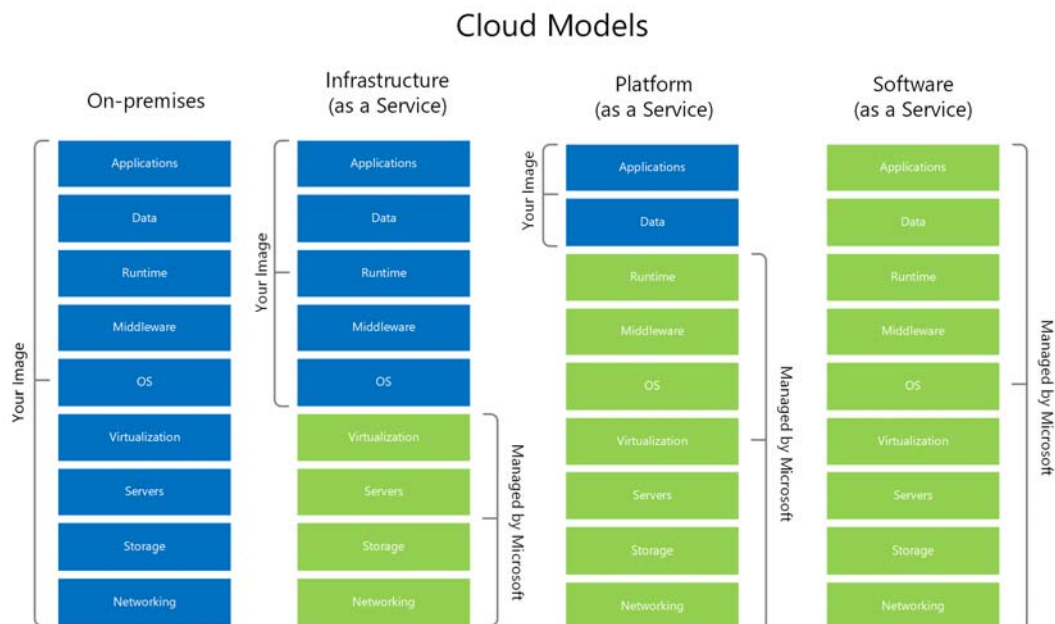


Figura 1. Cloud: modelli di servizio.

Fonte <https://learn.microsoft.com/it-it/azure>

Estendendo il modello di servizio oltre i tradizionali SaaS, PaaS, IaaS, l'approccio di erogazione di servizi in un formato basato su cloud ha portato alla creazione del termine *everything as a service* o XaaS [3, pp. 87-89], [6]. Con XaaS si intende uno schema in cui qualsiasi risorsa o funzionalità può essere fornita come un servizio attraverso la rete. Rientrano ad esempio in questa definizione:

- *Backup as a Service (BaaS)* per fornire servizi di backup e restore dei dati;
- *Database as a Service (DBaaS)* per fornire database come servizio, eliminando la necessità di gestire l'infrastruttura sottostante;
- *Desktop as a Service (DaaS)* per i servizi di gestione remota delle postazioni client aziendali, che includono manutenzione, applicazione di patch, intervento remoto;
- *Security as a Service (SECaaS)* per servizi di sicurezza, come firewall e antivirus.

Nell'ambito dei modelli di servizio, è diventata molto popolare per la sua originalità l'analogia "Pizza as a Service" che usa un insolito confronto con il cloud a partire dal concetto di pizza. Secondo Barron [7] se si volesse mangiare una pizza, ci sarebbero quattro scelte. Si potrebbe cucinare da zero, comprando tutti gli ingredienti necessari, preparando l'impasto e utilizzando il proprio forno per la cottura (*on-premises*). Si potrebbe acquistare al supermercato una pizza già pronta oppure surgelata; in questo caso si dovrebbero impiegare le proprie "infrastrutture" come il forno per la cottura e la cucina come luogo per mangiare (IaaS). Un'altra opzione è quella di ordinare una pizza da asporto con consegna a domicilio, quindi mangiarla a casa, occupandosi di incombenze quali apparecchiare la tavola e lavare le stoviglie (PaaS). Infine si potrebbe decidere di recarsi direttamente in una pizzeria (SaaS). L'obiettivo finale è comunque gustare una pizza, tuttavia in alcuni casi si è responsabili di svariate attività, mentre in altri casi il lavoro è demandato a persone specializzate.

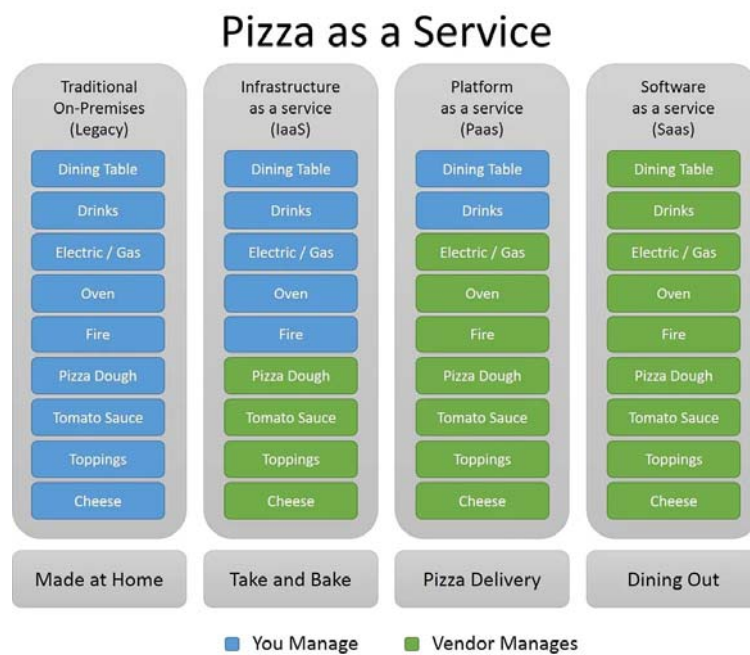


Figura 2. Cloud: analogia "Pizza as a Service"

Fonte [7]

### 2.1.2 I modelli di deployment

Nel cloud computing i modelli di deployment descrivono l'utilizzo, il perimetro di accesso e la collocazione fisica delle risorse rispetto all'ubicazione dell'utilizzatore del cloud [3, pp. 66-71.], [4, pp. 18-22], [5].

- *Public Cloud.* È il modello più diffuso, in cui un fornitore di servizi detto *Cloud Service Provider* (CSP) è il proprietario dell'intera infrastruttura hardware e software e si occupa in toto della sua gestione (manutenzione continua, upgrade, patch di sistema operativo, strato middleware e gestione del software applicativo). La creazione di ambienti virtuali logicamente separati consente la *multi-tenancy* ovvero l'utilizzo di risorse condivise da parte di più clienti (dall'inglese *tenant*, inquilino) che utilizzano Internet per effettuare l'accesso.
- *Private Cloud.* È chiamato anche modello interno o aziendale: le risorse sono ad uso esclusivo di un'azienda o un'organizzazione. Le infrastrutture hardware possono essere situate fisicamente nei data center locali dell'organizzazione (*on-premise*) oppure possono essere ospitate da un provider di servizi di terze parti (*off-premise*). Il compito di controllare i sistemi nonché effettuare la gestione ricade sull'organizzazione stessa, che può eventualmente affidare tali attività in outsourcing. Ad ogni modo le risorse rimangono dedicate esclusivamente agli utenti dell'organizzazione e sono tipicamente gestite in una rete privata.
- *Community Cloud.* È un modello che si colloca in una posizione intermedia: consente l'accesso alle risorse a un gruppo ristretto di organizzazioni o di utenti che appartengono a una comunità. Anche in questo caso le infrastrutture possono essere *on-premise*, ad esempio collocate presso la sede di uno o più membri della comunità, oppure *off-premise*, presso un fornitore esterno. L'obiettivo è poter beneficiare dei vantaggi del cloud pubblico come la multi-tenancy e la scalabilità con l'aggiunta di un livello di sicurezza, controllo e protezione dati.
- *Hybrid Cloud.* Questo modello vede la coesistenza di due o più modelli che rimangono distinti ma sono connessi tra loro con tecnologie proprietarie in modo che applicazioni e dati possano spostarsi tra le piattaforme senza problemi.

Nel modello Public, a cui si fa tipicamente riferimento quando si parla di cloud, a fronte dell'assenza di controllo dell'infrastruttura, si può contare su una scalabilità praticamente illimitata e su una elevata affidabilità in quanto sono disponibili risorse on demand su una rete di server bilanciati e in configurazione *fault tolerant*. Il sistema *pay per use* consente di pagare solo per i servizi usati senza necessità di investimenti per l'infrastruttura e per le attività di gestione e manutenzione, completamente a carico del provider. È adatto a molteplici realtà aziendali, ad esempio per società in cui l'infrastruttura informatica non rappresenta il core business, oppure che hanno carichi

di lavoro con volumi elevati o esigenze variabili nel tempo o ancora per chi si occupa di sviluppo software e progetti collaborativi. È anche indicato per piccole organizzazioni oppure per aziende in fase di avvio come le startup, in cui è possibile avere accesso ad una infrastruttura informatica tariffata a consumo o con contratto 1-3 anni senza dover anticipare per acquisti di hardware e software.

Nel Private cloud il livello hardware è costituito da un'infrastruttura fisica dedicata e non condivisa con altri clienti, quindi tale modello è adottato dalle organizzazioni che hanno la necessità di un maggiore controllo sulla sicurezza di server, rete e applicazioni. L'accesso alla rete interna è limitato ai membri autorizzati e questo assicura una maggiore protezione in caso si sia sottoposti a requisiti di sicurezza stringenti oppure si gestiscano dati critici, ad esempio sanitari o finanziari. L'azienda è responsabile del funzionamento, come avviene nelle architetture di computing tradizionali ed è necessario un livello elevato di competenze IT. Questo modello offre però una maggiore flessibilità di adattamento ai propri processi aziendali, garantisce conformità alle normative e un più ampio controllo dei costi in quanto gli esborsi sono noti in anticipo senza variazioni legate ai consumi o al traffico. Inoltre utilizzando un hardware separato le prestazioni dei carichi di lavoro sono maggiormente prevedibili e non influenzate dalla multi-tenancy.

Gli altri modelli nascono con l'intento di combinare i vantaggi del Public e del Private cloud e minimizzarne gli aspetti negativi. Nel Community cloud si ha l'ampia flessibilità di un progetto su misura che serva obiettivi comuni e specifici; l'eventuale condivisione delle infrastrutture e delle risorse rende questo schema più economico rispetto a molteplici singoli cloud privati, beneficiando dei vantaggi come ad esempio la multi-tenancy e la scalabilità con l'aggiunta di un livello di sicurezza, controllo e protezione dati.

Nel modello Hybrid in genere si parte da un cloud privato che viene successivamente esteso verso uno o più servizi pubblici esterni. In questo modo un'organizzazione può sia sfruttare i vantaggi del Public cloud sia esercitare maggiore controllo sull'ambiente, ad esempio differenziando la collocazione dei propri dati in base ai requisiti normativi oppure alle diverse necessità di protezione. Se un cloud privato è preferibile per carichi di lavoro con schemi d'uso prevedibili, i cloud ibridi sono considerati più versatili, perché possono ospitare ovunque qualsiasi carico di lavoro.

Si propone di seguito un riepilogo in forma tabellare, parzialmente tratto e rielaborato da [5], con le principali caratteristiche dei diversi modelli.

Tabella 2. Confronto tra i modelli di cloud deployment

	<b>Public</b>	<b>Private</b>	<b>Community</b>	<b>Hybrid</b>
<b>Ease of Setup</b>	Very easy, the CSP does most of the work	Very hard, internal IT expertise is needed	Easier because of community practices	Very hard due to interconnected systems
<b>Data control</b>	Low, the CSP provider has all control	Very high as you own the system	High (if members collaborate)	Very high (with the right setup)
<b>Reliability</b>	Prone to failures and outages	High (with the right team)	Depends on the community	High (with the right setup)
<b>Scalability</b>	High, most CSPs offer huge resources	High as there are no other tenants, but need to buy hw	Fixed capacity limits scalability	High (with the right setup)
<b>Security and privacy</b>	Very low, not a good fit for sensitive data	Very high, ideal for corporate data	High (if members collaborate on security policies)	Very high as you keep the data on a private cloud
<b>Setup flexibility</b>	Little to no flexibility, CSPs usually offer only predefined setups	Very flexible	Little flexibility, setups are usually predefined to an extent	Very flexible
<b>Cost</b>	Very inexpensive	Very expensive	Members share the costs	Cheaper than a private model, pricier than a public one
<b>Demand for in-house hardware</b>	No	In-house hardware is not a must but is preferable	No	In-house hardware is not a must but is preferable
<b>Ownership</b>	CSP	Single organization	Several organizations	Organization and CSP
<b>Performance</b>	Low to medium	Excellent	Very good	Good
<b>Location</b>	Off premise	Off or on premise	Off or on premise	Off or on premise
<b>Managed by</b>	CSP	Single organization	Several organizations or CSP	Organization and CSP
<b>User's control</b>	Limited control	Full control	High control but limited by community policies	Full control over private part and limited at public part
<b>Workload</b>	Normal workload with short-spikes in demand	Not suitable for handling large workload	Suitable for handling large workload	Highly dynamic or changeable
<b>Used By</b>	Anyone can access	Limited people can access	Depend upon number of cooperatives	Medium accessibility

### 2.1.3 Virtualizzazione: VM e container

Una delle tecnologie che ha contribuito fortemente alla genesi del cloud computing è senza dubbio la virtualizzazione che ha consentito di svincolarsi dal deployment tradizionale in cui ogni server fisico ospitava una o poche applicazioni, con un frequente sottoutilizzo delle risorse hardware.

Il concetto cardine è la possibilità di eseguire più di una *virtual machine* (VM) su una singola CPU fisica. Il disaccoppiamento delle applicazioni software dall'hardware è stato possibile grazie alla creazione di uno strato software sul sistema hardware in grado di simulare un intero ambiente fisico. Su tale livello di simulazione vengono create le VM per l'esecuzione dei servizi e dei carichi applicativi [3, p. 27]. Ogni VM è una macchina completa che, sopra l'hardware virtualizzato, esegue tutti i componenti necessari, incluso il proprio sistema operativo.

Il concetto di virtualizzazione ha subito ulteriori evoluzioni con l'introduzione dei *container* nei sistemi operativi Linux. I container nascono per fornire un'architettura più leggera per la gestione delle applicazioni e offrono nel loro piccolo degli ambienti *runtime* completi, autonomi e isolati. Pur disponendo di una segregazione di CPU, memoria, file system, i container condividono il kernel del sistema operativo host e isolano solo il sistema utente. Il disaccoppiamento dall'infrastruttura sottostante li rende portabili tra distribuzioni diverse e anche tra differenti cloud [8].

La Figura 3 mostra un confronto tra VM e container. Le VM, eseguendo il proprio kernel e il proprio spazio utente, forniscono un isolamento maggiore a scapito di tempi di avvio più lunghi e una maggiore richiesta di risorse computazionali e storage. I container condividono il kernel e le risorse del sistema operativo, limitando l'impatto sulle risorse e offrendo tempi di avvio più rapidi poiché non richiedono l'avvio di un sistema operativo completo.

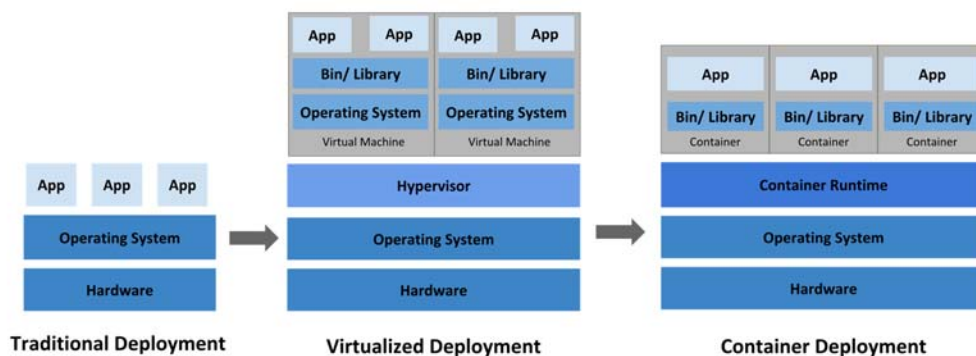


Figura 3. Confronto VM e container.

Fonte [8]

I container offrono un ambiente autonomo con librerie, file binari e applicazioni: questa peculiarità ha consentito una trasformazione del paradigma di sviluppo software

ponendo le basi dell'approccio a *microservizi*. Un'architettura a microservizi è composta da un insieme di servizi indipendenti e altamente specializzati che comunicano tramite API. Ogni singolo servizio può essere distribuito sulla rete e ridimensionato in maniera indipendente, garantendo maggiore flessibilità e scalabilità oltre a semplificare le modifiche e la manutenzione delle applicazioni rispetto alle soluzioni monolitiche. Le applicazioni progettate ad hoc per sfruttare l'elasticità e la scalabilità dell'infrastruttura cloud con l'uso di container, API e microservizi sono definite *cloud native*.

Tra gli strumenti open source, due in particolare hanno rivoluzionato l'ambito dello sviluppo software: Docker e Kubernetes<sup>3</sup>.

- *Docker* è un ambiente che permette di creare, distribuire e gestire container. Gli sviluppatori possono creare immagini di container con all'interno il codice applicativo, le dipendenze e tutte le risorse necessarie per eseguire l'applicazione.
- *Kubernetes* è una piattaforma di orchestrazione per la gestione di container su larga scala. Nato da un progetto di Google, è stato reso open source nel 2014. Offre funzionalità per coordinare il deployment, la scalabilità, il bilanciamento del carico di applicazioni containerizzate complesse [8].

Sulla base della diffusione delle tecnologie legate ai container, molti CSP hanno iniziato a proporre un nuovo modello di servizio cloud chiamato *Container as a Service* (CaaS), visto anche come un'estensione del PaaS [8]. CaaS offre un'infrastruttura pronta all'uso per le applicazioni basate su container, includendo anche servizi di orchestrazione, tra cui Kubernetes è il più popolare. L'idea di fondo è semplificare il ciclo di vita dallo sviluppo alla distribuzione e alla gestione dei container.

Volgendo lo sguardo ai servizi degli hyperscaler, Amazon AWS propone EKS (*Elastic Kubernetes Service*) e ECS (*Elastic Container Service*), Microsoft è presente con Azure Container Service e Google con Container Engine. Gli utenti possono accedere al servizio tramite chiamate API o attraverso un'interfaccia del portale web. Nel paragrafo successivo si coglie l'occasione per approfondire le principali interfacce utilizzate per l'interazione con i servizi cloud.

---

<sup>3</sup> Link ai rispettivi progetti <https://www.docker.com> e <https://kubernetes.io>.



### 2.1.4 Interfacce di accesso ai servizi cloud

Esistono diverse modalità con cui gli utenti oppure le applicazioni possono interagire e utilizzare le risorse offerte attraverso i servizi cloud. Si riportano di seguito le interfacce di accesso più comuni [3, pp. 51-52], [4, pp. 71-76].

- Interfacce grafiche (*Graphical User Interface*, GUI). Forniscono una rappresentazione visiva delle diverse funzionalità utilizzabili tramite un pannello dotato di elementi grafici quali icone, pulsanti, menu e immagini. Ne sono un esempio le pagine web di amministrazione dei servizi cloud: *AWS Management Console*, *Google Cloud Console*, *Azure Portal*.
- Interfacce da riga di comando (*Command Line Interface*, CLI). Consentono agli utenti di interagire con i servizi cloud attraverso comandi testuali inseriti in una console o in un terminale. Alcuni esempi sono: *AWS Command Line Interface* (AWS CLI), *Google Cloud SDK*, *Azure Command-Line Interface* (Azure CLI).
- Interfacce di programmazione applicazioni (*Application Programming Interface*, API). Sono un insieme di procedure che definiscono le operazioni e i formati dei dati per consentire l'integrazione tra diversi sistemi; in ambito cloud permettono agli sviluppatori di creare applicazioni che possono interagire con i servizi cloud come nel caso di *AWS API Gateway* e di *Google Cloud Endpoints*.
- Interfacce API RESTful. Sono API che seguono i principi dell'architettura REST<sup>4</sup> e le proprietà del protocollo HTTP, come l'uso di URIs per identificare le risorse, i metodi standard (GET, POST, PUT, DELETE), e la rappresentazione dei dati in formato JSON o XML. Molti servizi cloud offrono API RESTful come Facebook o Twitter per i contenuti social oppure Google Maps API.
- Interfacce specializzate. Sono interfacce progettate per casi d'uso specifici. Un servizio di storage cloud o di file sharing come Dropbox o Google Drive può offrire un'interfaccia di accesso basata su file per il caricamento e il download. Un servizio di database cloud può fornire un'interfaccia specifica per connessioni a fonti dati e per l'esecuzione di query.

Si evince che esistono molte somiglianze nei concetti generali e nelle funzionalità tra i vari CSP. Nonostante ciò, per i servizi di cloud computing non esistono degli standard ufficiali e ciascun fornitore può proporre le proprie implementazioni specifiche e offrire in alcuni casi funzionalità esclusive.

---

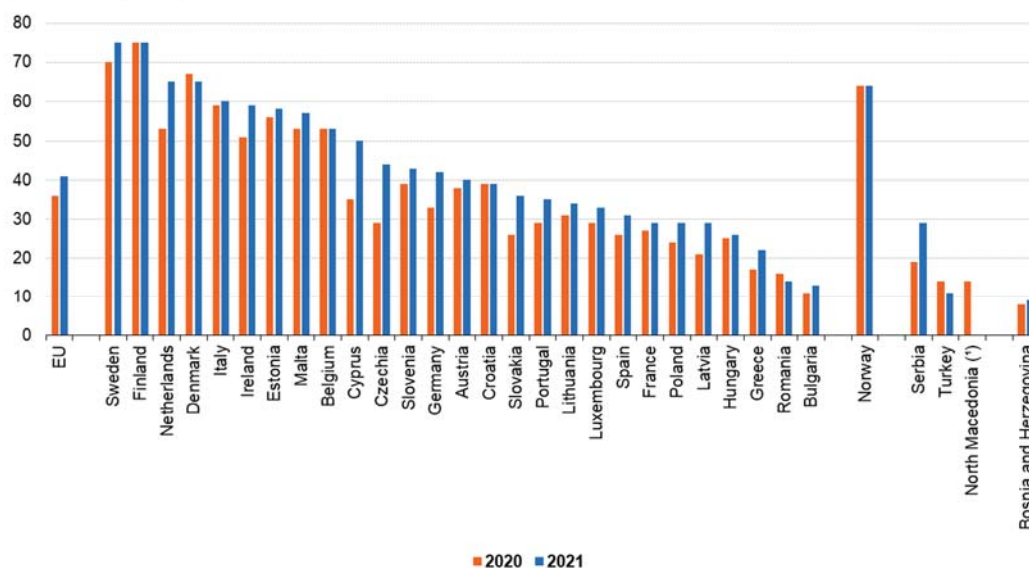
<sup>4</sup> Representational State Transfer, ideato da Roy Fielding nel 2000, è uno stile architetturale che definisce un insieme di principi e vincoli per il progetto di sistemi distribuiti.

## 2.2 Stato dell'arte e maturità del cloud

### 2.2.1 Diffusione del cloud in Europa e in Italia

Una fotografia della diffusione del cloud in Europa è fornita dalla raccolta di statistiche pubblicate da Eurostat [9]. Nel 2021 il 41% delle imprese UE ha utilizzato il cloud computing (5% in più rispetto al 2020), principalmente per la posta elettronica e l'archiviazione di file. Si osservano tuttavia differenze significative tra gli stati membri. Nei paesi del nord come Svezia, Finlandia, Paesi Bassi e Danimarca almeno il 65% delle imprese utilizza il cloud computing, mentre l'adozione avviene da parte di meno del 25% delle imprese in Grecia, Romania e Bulgaria (Figura 4).

**Use of cloud computing services, 2020 and 2021**  
(% of enterprises)



(\*) Data for 2021: not available yet.  
Note: Montenegro 2020 and 2021: data unreliable. Iceland: data not available  
Source: Eurostat (online data code: isoc\_cicce\_use)

eurostat

Figura 4. Diffusione del cloud nei paesi europei.

Fonte [9]

Guardando ai modelli di servizio (Figura 5), la maggior parte delle imprese UE che ha acquistato servizi cloud computing ha utilizzato almeno un servizio SaaS (94%) come applicazioni finanziarie o contabili, posta elettronica, software per ufficio. Un'ampia percentuale (74%) ha utilizzato anche almeno un servizio di tipo IaaS, per database aziendali, archiviazione di file oppure potenza di calcolo per eseguire specifiche applicazioni. Risulta meno utilizzato il modello PaaS, selezionato dal 21%, forse perché più di nicchia e maggiormente orientato a chi si occupa di sviluppo e test software.

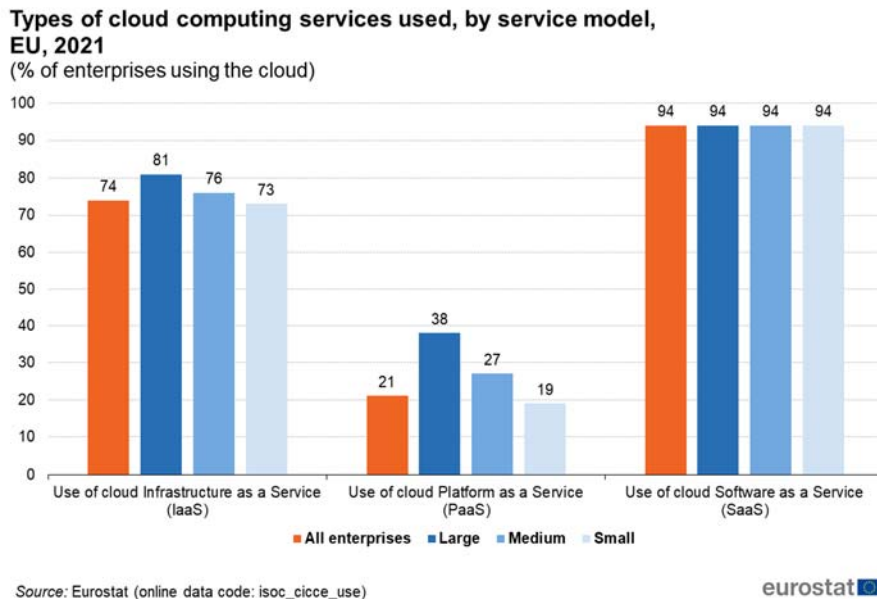


Figura 5. Adozione del cloud nei paesi europei per modelli di servizio.

Fonte [9]

Volgendo lo sguardo all'Italia, secondo il rapporto Istat 2021 [10], le imprese italiane che utilizzano servizi cloud sono passate in due anni dal 23% al 59% mentre quelle che adottano servizi cloud evoluti sono aumentate dall'11% al 32%, cifre su cui ha sicuramente inciso la crisi sanitaria innescata dal Covid-19. Dello stesso ordine di grandezza sono i dati [11] sulle PMI pubblicati dall'Osservatorio Cloud Transformation della School of Management del Politecnico di Milano. Nel 2022 il 52% nelle PMI ha utilizzato almeno un servizio cloud (7% in più rispetto al 2021). I modelli di servizio IaaS, PaaS e SaaS sono rispettivamente in crescita del 27%, 33% e 14%. Per ciò che concerne i modelli di deployment, è interessante che la dinamica di crescita più significativa sia stata quella relativa alla componente "Public e Hybrid cloud" (+22% rispetto al 2021), che è costituita dall'insieme di servizi forniti dai CSP e dall'interconnessione tra cloud pubblici e privati. Dal punto di vista economico, il mercato del cloud in Italia vale oltre 4,5 miliardi di euro, con una crescita del +18%, sebbene sia da evidenziare che l'aumento è composto da un +15% di crescita organica in continuità con gli anni precedenti e da un +3% dovuto ai rialzi dei prezzi.

### 2.2.2 La maturità tecnologica del cloud computing

La visione dell'informatica come servizio di tipo infrastrutturale è stata paragonata all'introduzione del servizio centralizzato dell'elettricità [3, p. 11]. Così come le aziende odierne senza il cloud devono implementare e gestire i propri sistemi informatici tradizionali, all'inizio del XX secolo le industrie costruivano e gestivano internamente i propri impianti di produzione di energia, prima dell'avvento della rete elettrica.

Per valutare la maturità del cloud computing, come proposto in [4, p. 5], interessante è l'analisi dell'Hype Cycle<sup>5</sup>. Si tratta di un grafico che descrive su una curva l'evoluzione e il ciclo di vita delle tecnologie emergenti, ponendo valore e aspettative sulle ordinate e il tempo sull'asse delle ascisse. Il percorso inizia con una fase di avvio tecnologico (*innovation trigger*) che genera interesse da parte delle imprese; alcuni casi di successo portano al picco caratterizzato da molte aspettative ed eccessivi entusiasmi (*inflated expectations*) a cui segue un periodo di disincanto (*trough of disillusionment*) in cui c'è un abbandono, crollano gli investimenti e la popolarità. Successivamente (*slope of enlightenment*) la tecnologia viene gradualmente ripresa e sviluppata, attirando nuovi investimenti. Se ne comprendono meglio applicabilità, rischi e benefici, fino al momento in cui essa acquisisce solidità e diffusione sul mercato (*plateau of productivity*).

Nel report 2023 della società di ricerca Gartner [12] il cloud computing è stato inserito nel punto di flesso della curva, dopo aver superato la fase delle aspettative e in prossimità della depressione della disillusione (Figura 6). Si tratta di una posizione a prima vista negativa ma che in realtà consente l'ascesa verso la fase di diffusione su larga scala e di maggiore stabilità. Considerati i dati descritti nel paragrafo 2.1.3 si può ritenere che il livello del cloud sia tale da potersi avviare verso le due fasi successive fino alla stabilità produttiva, in cui alla consapevolezza sulle potenzialità si affianca la realizzazione di progetti più maturi, sempre più inseriti nei processi aziendali.

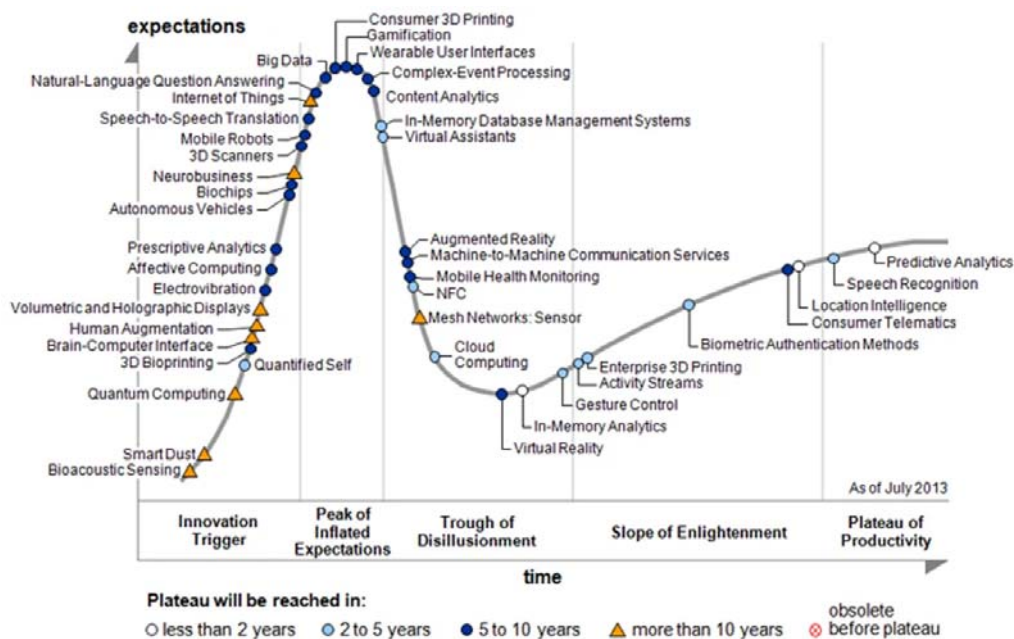
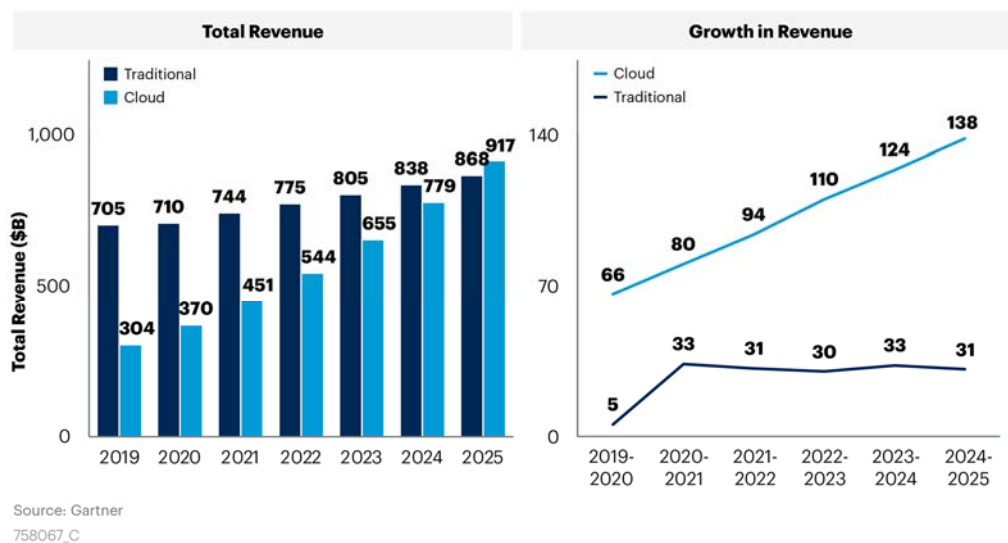


Figura 6. Gartner Hype Cycle 2023.

Fonte [12]

<sup>5</sup> Modello previsionale ideato da Gartner per rappresentare su una curva l'evoluzione delle tecnologie IT. Link: <https://www.gartner.com/en/research/methodologies/gartner-hype-cycle>

Un'altra recente indagine di Gartner dal titolo "Market Impact: Cloud Shift - 2022 Through 2025" [13] ha analizzato la transizione al cloud (*cloud shift*) in un sottoinsieme di aziende operanti nei settori dello sviluppo software, dei servizi legati ai processi aziendali e delle infrastrutture di sistema (Figura 7). Secondo le previsioni dei ricercatori, entro il 2025 in questo tipo di organizzazioni ci sarà un'inversione di tendenza della spesa per gli asset informatici: gli investimenti in cloud computing supereranno quelli per le installazioni on-premise.



Gartner.

Figura 7. Inversione della spesa per gli asset ICT entro il 2025.

Fonte [13]

### 2.2.3 Casi studio e stato dell'arte nella ricerca

Il tema del cloud computing tocca diversi settori disciplinari e la ricerca accademica ha affrontato i molteplici argomenti correlati: si tratta di aspetti socio-economici, architetture hardware e software, casi studio di migrazione, gestione delle applicazioni, fino a questioni più trasversali quali l'impatto energetico e le dinamiche di sicurezza.

Nella loro analisi, Khajeh-Hosseini et al. [14] (2010) studiano il caso di un'impresa, fornitrice di servizi ad aziende del settore petrolifero, che sposta il proprio sistema IT dal data center interno verso Amazon EC2. Sebbene non sia recente, l'indagine è interessante perché propone una parziale valutazione economica e un'ampia disamina dei potenziali vantaggi e dei rischi associati alla migrazione. I risultati mostrano che in 5 anni l'infrastruttura in esame costerebbe sul cloud EC2 oltre il 30% in meno, riducendo anche il numero di richieste di supporto relative alla piattaforma. L'analisi però va oltre il lato economico e valuta anche l'impatto complessivo sull'organizzazione aziendale, rilevando rischi significativi come ad esempio il peggioramento dell'assistenza

e della qualità del servizio ai clienti, a causa della dipendenza da una terza parte per i sistemi informatici, oppure il ridimensionamento dei dipartimenti IT e la perdita di competenze interne o ancora la necessità di avere staff motivato e con esperienza per far fronte alla migrazione e gestire eventuali problematiche post migrazione.

In merito alla transizione al cloud, già nel 2015 Gholami et al. [15] evidenziavano come il volume delle ricerche sulla migrazione fosse vasto ma piuttosto frammentato. Sottolineando la necessità di un modello di riferimento generico relativo ai processi di migrazione, hanno pertanto elaborato un quadro per la valutazione dei diversi approcci esistenti e una dettagliata selezione degli studi pubblicati fino ad allora.

Molti lavori accademici sono incentrati sul tema delle applicazioni con particolare riferimento alla modellazione dei processi di transizione. La progettazione di sistemi con un approccio basato sui servizi è alla base della ricerca di Scandurra et al. [16] (2015). Iniziando dal caso studio di migrazione di un sistema esistente a una soluzione cloud da parte di un'azienda italiana di commercio alimentare, gli autori propongono delle linee guida di analisi e progetto in cui adottano standard e strumenti di modellazione cloud, come ad esempio TOSCA (*Topology and Orchestration Specification for Cloud Applications*) e SCA (*Software Composition Analysis*). Data la complessità delle decisioni legate alla migrazione di risorse verso il cloud, Alkhalil et al. [17] (2017) propongono un modello di analisi basato sui principi *Diffusion Of Innovation* (DOI) e *Technology-Organization-Environment* (TOE), due basi teoriche che analizzano le modalità e i processi di adozione delle innovazioni tecnologiche all'interno delle organizzazioni.

Lo studio di Ahmad et al. [18] (2018) evidenzia l'importanza di un metodo consolidato per accompagnare la transizione al cloud e propone una procedura in cinque fasi: valutazione aziendale, analisi tecnica, strategia di migrazione, pianificazione ed esecuzione, monitoraggio e ottimizzazione.

La parte più complessa della migrazione delle applicazioni verso il cloud è l'individuazione del percorso più adatto. Le opzioni considerate sono le stesse che emergono nell'analisi di Linthicum [19] (2017), secondo quello che è chiamato il *modello delle 7R*. Alcune applicazioni possono essere mantenute senza modifiche (*retain*), alcune semplicemente spostate su diverse infrastrutture (*rehost*), altre richiedono modifiche al codice o all'architettura (*refactor*) oppure il passaggio a una diversa piattaforma (*replatform*). La migrazione al cloud può essere un'occasione per eliminare le applicazioni obsolete (*retire*) oppure sostituirle con servizi di tipo SaaS (*replace*) o infine consolidare applicazioni e servizi simili (*reuse*).

Da parte della comunità scientifica vi è anche particolare attenzione agli aspetti relativi ai consumi e all'impatto energetico dei data center. Già nel 2012 Lee e Zomaya [20] sottolineavano l'importanza di una strategia di allocazione delle risorse che, attraverso la virtualizzazione, tenesse conto dell'utilizzo delle risorse per una migliore efficienza energetica negli ambienti cloud. In generale a questo scopo vengono utilizzate

diverse tecniche tra cui la riduzione della frequenza dei processori, il consolidamento delle VM su un numero inferiore di host fisici, l'introduzione di stati di sospensione durante i periodi di inattività. Pur raggiungendo l'obiettivo della diminuzione del consumo energetico, questi meccanismi possono influire sulle prestazioni applicative. Nell'ottica di migliorare il compromesso tra consumo e prestazioni in [21] Rossi et al. (2017) analizzano le tecniche esistenti e presentano un prototipo di orchestrazione in grado di distribuire e combinare le diverse metodologie di risparmio energetico. Alla base del consolidamento nei data center vi è il concetto di *live migration*, il processo con cui si trasferisce una VM da un host fisico all'altro senza interruzione di servizio. Tale operazione richiede cicli aggiuntivi di CPU e ha una durata che dipende dalla dimensione della RAM della VM e dalla larghezza di banda di rete disponibile. Essa ha pertanto un extra costo energetico, tema che è stato analizzato nei lavori di Strunk [22] (2013) e di Rybina e Schill [23] (2016).

Un'altra indagine sulle diverse strategie per l'efficienza energetica nel cloud computing con la rassegna di studi in letteratura è quella di More e Ingle [24] (2017) mentre lo studio di Vafamehr e Khodayar [25] (2018) ha la particolarità di mettere in luce l'interdipendenza tra le due infrastrutture: la rete elettrica e il cloud computing. I fornitori da un lato propongono sul mercato i propri servizi cloud con le rispettive tariffe, ma dall'altro partecipano al mercato all'ingrosso dell'elettricità con offerte per l'acquisto dell'energia necessaria per le infrastrutture. Il prezzo dell'energia incide sulle tariffe dei servizi cloud e allo stesso tempo la domanda di servizi cloud influisce sull'energia che viene richiesta al mercato all'ingrosso dell'elettricità, che infine condiziona ulteriormente il prezzo dell'energia acquistata.

Ulteriori approfondimenti della letteratura sarebbero troppo ampi rispetto alle finalità del presente lavoro; la panoramica descritta ha la finalità di mostrare quanto sia esteso l'interesse per il cloud computing e la fitta interconnessione tra diversi aspetti non esclusivamente tecnologici.

#### **2.2.4 Punti di forza e sfide aperte**

Le diverse fonti in letteratura convergono nell'indicare i benefici del cloud computing; si vedano tra gli altri [3, pp. 42-45], [14], [26]. In particolare, prendendo principalmente in considerazione il modello Public cloud, viene posta l'enfasi sulla scalabilità e la flessibilità di crescita verso risorse computazionali e di storage virtualmente illimitate, con l'unico acquisto necessario per i sistemi client e la connettività Internet, in confronto all'approccio tradizionale in cui solo grandi aziende potevano permettersi di affrontare investimenti ingenti.

Ridotto l'investimento di capitale iniziale in infrastruttura, su chi sottoscrive un'offerta di cloud computing gravano esclusivamente i costi operativi relativi al

consumo del servizio, con lo slittamento da Capex verso Opex<sup>6</sup>. Inoltre il cloud non ha costi nei casi di inutilizzo e gli esborsi non sono fissi: ogni utilizzo viene contabilizzato e si sostengono solo i costi legati agli effettivi consumi. In caso di carichi elevati o picchi di attività, è possibile aumentare le risorse in uso ma anche effettuare uno *scale down* rapidamente quando le risorse non sono più necessarie.

Un altro vantaggio evidenziato da tutti è legato alla gestione. Con il modello stile outsourcing, gli utenti si liberano della responsabilità di gestione e si svincolano ad esempio da problematiche energetiche e di raffreddamento dei data center, nonché dalle attività di manutenzione continua su hardware e software, come aggiornamenti di sistema operativo e delle applicazioni.

Altri due aspetti di peso sono l'affidabilità e la resilienza: affidandosi a un cloud provider si può avere a disposizione un'infrastruttura operativa 24/7 e con service time che solitamente non scende sotto il 99,9%, dotata di sistemi di load balancing, di backup e ripristino in caso di fallimenti e con un supporto in termini di orario e di competenze superiore a quello che potrebbe avere uno staff IT interno. Tutto questo consente di ridurre l'impatto e la durata di interruzioni di servizio causate da circostanze avverse.

Per chi si occupa di progettazione e distribuzione del software, gli ambienti cloud offrono un provisioning delle risorse, VM o container, molto più rapido e in alcune situazioni anche automatizzato, rendendo più veloce il tempo di sviluppo e di test e di conseguenza il *time to market*. Inoltre nelle modalità PaaS e SaaS non è necessario acquisire alcuna licenza per utilizzare le applicazioni o gli ambienti di sviluppo, a differenza di quanto avviene nell'approccio tradizionale in cui le aziende devono dedicare un budget specifico per le licenze dei vari software in uso. Non da ultimo è da citare la possibilità di collaborazione più efficace grazie alla modalità di accesso e alle risorse e ai dati, disponibili ovunque, in qualsiasi momento e da qualsiasi client dotato di una connessione a Internet, senza vincoli di luogo o di specifici dispositivi. Sugli stessi temi si trovano concordi anche due studi che hanno declinato il tema dell'adozione del cloud da parte delle piccole e medie imprese (PMI), Salleh et al. [27] (2018) e più recentemente Jayeola et al. [28] (2022), con una revisione sistematica della letteratura.

La seguente tabella, elaborata sulla base di [3, pp. 42-49], riporta una sintesi dei vantaggi e delle sfide.

---

<sup>6</sup> Due tipi di spese aziendali, divise in base al modo in cui influiscono sulla contabilità. Opex (*Operating Expenses*) sono spese operative, rappresentano i costi di gestione ricorrenti (es. stipendi, affitti, spese per le utenze). Capex (*Capital Expenditures*) sono invece spese per investimenti a lungo termine effettuate per migliorare l'attività (es. acquisto di attrezzature e macchinari, software e materiale informatico, ristrutturazione di edifici).



Tabella 3. Cloud computing: punti di forza e sfide

<b>Punti di forza</b>	<b>Sfide</b>
Riduzione dei costi operativi	Minore prevedibilità costi, base consumo
Riduzione delle responsabilità di gestione	Minore controllo sull'infrastruttura
Modello di pagamento <i>pay per use</i>	Costi extra: larghezza di banda, storage
Illimitate risorse computazionali e storage	Rischio di <i>vendor lock-in</i>
Affidabilità infrastruttura	Conformità legislativa
Alta resilienza, disponibilità 24x7	Sicurezza dei dati
Accesso ubiquitario	Dipendenza dalla rete Internet
Rapida installazione e distribuzione appl.	Portabilità limitata tra diversi CSP
Software come servizio, senza spesa licenza	Problemi di interoperabilità applicativa
Minore impatto energetico	Distribuzione geografica e latenze

Come avviene per ogni tecnologia, il cloud computing non è adatto a tutti gli scenari aziendali e presenta alcune limitazioni che ne possono frenare l'adozione. Secondo Khajeh-Hosseini et al. [14] (2010) possono presentare una serie di rischi le questioni di sicurezza, legali e di privacy così come la posizione geografica del cloud, aspetti tecnici e organizzativi derivanti dal cambiamento che il cloud computing apporta alle imprese. Tale analisi trova continuità nel recente lavoro di Bashari et al. [29] (2017) che presenta un esame delle principali sfide, sia tecniche sia organizzative, legate all'adozione del cloud. I fattori più discussi riguardano, tra gli aspetti tecnici, la dipendenza dalla rete Internet e le proprietà di sicurezza delle informazioni (confidenzialità, integrità, disponibilità), mentre sul lato giuridico economico sono citati la selezione dei fornitori, la definizione degli SLA e i problemi di *lock-in*. Sono infine considerati gli aspetti organizzativi come l'importanza della personalizzazione delle soluzioni, la gestione del cambiamento e l'impatto sullo staff IT in termini di ruoli e competenze.

In generale, uno degli svantaggi più citati è l'inesistenza di meccanismi di dipendenza, in primo luogo dalla connessione Internet senza la quale non è possibile accedere a dati e risorse. In secondo luogo vi è la subordinazione a una terza parte sia per il controllo sull'infrastruttura sottostante sia in caso di problemi tecnici o malfunzionamenti sui sistemi del fornitore. Infine vi è il rischio di vincolarsi ad un fornitore (*lock-in*) rendendo complicato lo spostamento dei dati verso altri provider nel caso si desideri cambiare, ad esempio per inadempienza oppure per fallimento o cessata attività.

Non mancano le preoccupazioni relative a rischi per la sicurezza dei dati e minacce, soprattutto criticità legate alla custodia dei dati interni, del personale e di fornitori e clienti. Inoltre è da considerare la complessità di integrazione del cloud con i sistemi esistenti, che richiede personale IT specializzato oppure l'attivazione di percorsi di

formazione per acquisire competenze specifiche. Il meccanismo *pay per use* può condurre a spese impreviste derivanti dai consumi di storage o larghezza di banda non preventivati o comunque non prevedibili.

Tutte queste considerazioni trovano ampio consenso anche nei lavori sul mondo delle PMI precedentemente citati [27], [28] mentre un'indagine che si differenzia dalle altre è quella di Suhaimi et al. [30] (2018) che passa al vaglio i problemi e le sfide dell'adozione del cloud computing non solo dal punto di vista del cliente finale, come la maggior parte delle analisi prese in considerazione, ma anche da parte del fornitore, sulla base della ricerca in letteratura di vari paesi. Una delle principali preoccupazioni per i CSP è gestire in modo efficace le relazioni con i clienti affinché essi continuino a utilizzare i loro servizi, lavorando sia sugli aspetti tecnici sia sull'affidabilità e sulla relazione di fiducia. I punti di criticità legati agli aspetti tecnologici coinvolgono ad esempio la scelta delle architetture, i meccanismi di alta affidabilità, le strategie di migrazione delle applicazioni. Sul lato economico sono vincolanti l'avveduta gestione dei consumi e la capacità di negoziare i prezzi dell'energia all'ingrosso, mentre sul lato organizzativo incidono la disponibilità di personale affidabile e altamente specializzato nonché attente politiche interne in materia di sicurezza per costruire e mantenere alta la propria reputazione.

Elemento comune a tutti gli studi è il richiamo ai temi della sicurezza, aspetto che occupa un ruolo nodale nelle analisi e nei modelli di migrazione. In tale ambito, sono due le organizzazioni di riferimento che forniscono le linee di indirizzo. A livello europeo si colloca l'*European Union Agency for Cybersecurity* (ENISA)<sup>7</sup>, centro di competenza in materia di sicurezza delle reti e dell'informazione per l'UE, i suoi Stati membri, il settore privato e i cittadini europei. A livello globale opera invece la *Cloud Security Alliance* (CSA)<sup>8</sup>, un'organizzazione no-profit che ha come obiettivi generali la formazione, la sensibilizzazione e la promozione delle best practice di sicurezza sull'utilizzo del cloud computing.

Nel rapporto ENISA *Cloud Computing Security Risk Assessment* [31] (2009) un'analisi molto approfondita, realizzata con il supporto di esperti del mondo dell'industria, delle università e degli enti governativi, propone un modello con raccomandazioni pratiche per la valutazione dei rischi cloud, divisi in macro categorie.

- *Perdita di governance*: utilizzando le infrastrutture cloud, il cliente cede necessariamente il controllo al CSP su una serie di questioni che incidono sulla sicurezza, che non sempre sono contemplate negli SLA: si può quindi venire a creare una falla nelle difese di sicurezza.

---

<sup>7</sup> Link al sito ufficiale: <https://www.enisa.europa.eu>

<sup>8</sup> Link al sito ufficiale: <https://cloudsecurityalliance.org>

- *Lock-in*: in assenza di standard per i dati o per le interfacce di accesso ai servizi, è difficile che sia garantita la portabilità di dati, applicazioni e servizi, rendendo complicato migrare da un fornitore a un altro oppure riportare dati on-premise.
- *Mancaza di isolamento*: con la multi-tenancy in caso di potenziali punti deboli nei meccanismi di separazione delle risorse (storage, memoria, rete), un attacco ai sistemi di un cliente può compromettere anche l'operatività di altri clienti che condividono la stessa infrastruttura.
- *Rischi di conformità e protezione dei dati*: alcuni settori richiedono delle certificazioni specifiche o devono adeguarsi a standard di settore o requisiti normativi, come ad esempio l'ottemperanza del GDPR<sup>9</sup> in Europa. In alcuni casi potrebbe essere difficile per il cliente, in qualità di titolare del trattamento dei dati, controllare le pratiche di trattamento dei dati del fornitore cloud e quindi avere la certezza che i dati siano gestiti in modo lecito.
- *Compromissione dei sistemi di gestione web*: le interfacce di gestione dei sistemi cloud sono tipicamente delle applicazioni web accessibili tramite Internet e forniscono l'accesso a ampi blocchi di risorse. La superficie di attacco è maggiore, soprattutto se combinate con l'accesso remoto e le vulnerabilità del browser web.
- *Rischi sulla cancellazione dei dati*: quando un cliente decide di eliminare una risorsa cloud, la cancellazione potrebbe essere incompleta o non effettiva, perché il fornitore mantiene copie di backup oppure perché lo storage è una risorsa condivisa con altri clienti.
- *Azioni dolose del personale interno*: in assenza di politiche di gestione dei ruoli o di formazione adeguata, utenti dei CSP potrebbero accedere o mettere a rischio dei dati inavvertitamente o in maniera fraudolenta.

Se i rischi sopra menzionati concordano con le preoccupazioni raccolte negli studi citati nel paragrafo 2.2.4, il recente rapporto CSA *Top Threats to Cloud Computing Pandemic Eleven* [32] (2022) mostra un'interessante inversione di tendenza o perlomeno un'attenuazione dei timori. A fronte di un'indagine condotta tra professionisti della sicurezza ICT, rispetto ai rapporti sulle principali minacce degli anni precedenti, gli intervistati considerano meno preoccupati alcune questioni tradizionali (tecnologia condivisa, *Denial of Service*, perdita o violazione dei dati), tanto da essere addirittura escluse dal report. Entrano invece tra le criticità temi più complessi come la gestione della configurazione, le modalità di autenticazione e accesso, gli eventuali punti deboli del *control plane*, i meccanismi di interfaccia tra i vari livelli. Segno questo, secondo i ricercatori CSA, di una crescente fiducia e una maggiore comprensione del cloud da parte delle aziende nonché di una tecnologia più matura da parte dei fornitori.

---

<sup>9</sup> *General Data Protection Regulation*, regolamento dell'Unione europea UE 2016/679 in materia di trattamento dei dati personali e di privacy.

### 2.2.5 L'adozione nel settore pubblico e le politiche europee

La trasformazione innescata dai nuovi modelli tecnologici ha raggiunto anche il settore pubblico, in un processo di evoluzione delle applicazioni e delle infrastrutture nell'ottica del miglioramento dei servizi a cittadini e imprese. Per le sue caratteristiche, il cloud computing riveste un ruolo centrale e in diversi paesi le linee di indirizzo prevedono di optare per il cloud come approccio per nuovi investimenti nel campo ICT, politica spesso definita *cloud first*. Nonostante ciò, rispetto ad altri ambiti l'adozione è più lenta a causa dei vincoli aggiuntivi a cui è soggetta la pubblica amministrazione (PA) rispetto al settore privato. La transizione verso il cloud accomuna tutti gli Stati, sebbene siano differenti le modalità con cui è stata avviata, così come eterogenee sono le condizioni di partenza di ciascuno, dalla condizione delle infrastrutture tecnologiche esistenti al contesto legislativo e geopolitico.

Negli USA alla base dell'adozione del cloud tra gli enti governativi, c'è la roadmap sviluppata dal NIST che definisce in modo molto preciso i requisiti da soddisfare per la migrazione nonché gli standard e le linee guida per l'interoperabilità dei diversi sistemi, la sicurezza e la portabilità delle applicazioni. Uno studio della società di consulenza Deloitte [33] (2022) indica che la pandemia Covid-19 ha accelerato l'adozione del cloud da parte di enti governativi, agenzie federali e stati, soprattutto per le realtà che non avevano ancora affrontato la migrazione in modo considerevole, mentre altre organizzazioni si trovano già proiettate in avanti, cercando di capire come sfruttare al meglio le caratteristiche del cloud per trasformare i servizi pubblici.

Le diverse sfaccettature di questo tema sono molto presenti nella ricerca accademica. Un'ampia panoramica dei primi anni di adozione del cloud è proposta da Wyld [34] (2010), che analizza il rapporto tra enti pubblici e cloud spaziando dal governo federale degli Stati Uniti, esclusi gli usi militari, a Europa, Cina e Giappone. Quest'ultimo ha seguito un approccio comune ad altri stati, avviando un progetto di consolidamento delle infrastrutture IT governative verso un'infrastruttura cloud unica chiamata *Kasumigaseki Cloud*. Sallehudin et al. [35] (2018), citando gli esempi di USA e Regno Unito, focalizzano il lavoro sull'adozione del cloud negli enti governativi in Malesia. Un'analisi della situazione in Brasile, dove è attivo *G-Cloud* (Government Cloud) per il dominio del servizio pubblico, viene proposta da Branco et al. [36] (2019) che raccolgono una serie di raccomandazioni da parte dei manager che hanno già affrontato percorsi di migrazione. L'orizzonte geografico dell'Africa sub-sahariana e delle Nigeria è trattato da Dahiru e Abubakar [37] (2017) che analizzano nell'adozione del cloud sfide simili a quelle di altri contesti. Cidres et al. [38] (2020) includono, oltre l'Europa, anche Australia e USA proponendo uno strumento di valutazione chiamato *Cloud Calculator* a supporto delle scelte di migrazione nelle pubbliche amministrazioni. Ali Kh. et al. [39] (2018) analizzano con molto dettaglio i molteplici aspetti dell'adozione del cloud computing e presentano anche uno schema di Hybrid cloud per adattare la tecnologia al settore pubblico. In particolare il modello proposto si articola su tre livelli

di cloud, locale, regionale e nazionale, divisi su base organizzativa e geografica ma in relazione tra loro per la condivisione dei dati.

Nel panorama europeo, Jones et al. [40] (2019), partendo dall'esperienza di casi studio in tre enti pubblici sul territorio del Regno Unito, affrontano in maniera critica benefici e punti di attenzione del cloud, testando sul campo l'aderenza rispetto a quanto descritto in letteratura. Una fotografia dello stato dell'arte è offerta dal rapporto conclusivo del progetto *Cloud for Europe* di Draoli [41] (2014) e dall'indagine di Zaharia-Rădulescu e Radu [42] (2017) che delineano un confronto tra gli approcci al cloud in diversi stati europei. In generale il primo passo in tutti i paesi dell'Unione Europea è stato quello di allineare la propria legislazione nazionale con quella europea in materia di privacy e protezione dei dati personali.

Nei Paesi Bassi il governo possiede il suo cloud privato, distribuito su 4 data center, nato dal consolidamento di 66 data center tra il 2013 e il 2014. A livello applicativo, tale cloud ospita il *Rijks Application Store* (RAS) una sorta di archivio digitale centrale per smartphone o tablet da cui è possibile scaricare applicazioni per i servizi pubblici, in maniera simile a come si farebbe con Google Play Store oppure AppStore di Apple. Il Portogallo è intervenuto sulla situazione degli enti pubblici, dotati ciascuno di applicazioni e infrastrutture indipendenti e non comunicanti tra loro. Con l'obiettivo di ridurre i costi ICT e le ridondanze nella PA, ha lanciato l'iniziativa *GO-Cloud* per fornire un'infrastruttura condivisa e una serie di servizi progettati specificatamente per assicurare integrazione e uso di interfacce standard. In maniera analoga si muove l'Italia, come verrà approfondito successivamente nel capitolo 3.

Il Regno Unito utilizza da molto tempo la tecnologia per aumentare l'efficienza dei servizi pubblici; già nel 2013 è stata lanciata la politica *Government Cloud First* con lo slogan *Public cloud first; services, not servers* dando priorità all'adozione di soluzioni di cloud pubblico. In particolare l'iniziativa Government Cloud (G-Cloud) ha lo scopo di promuovere l'adozione del cloud computing a livello governativo, la semplificazione delle procedure di approvvigionamento e l'utilizzo di una piattaforma digitale per l'accesso ai servizi. A fare da base è la *Government Secure Intranet*, l'infrastruttura di rete privata che collega dipartimenti e agenzie governative. In questo c'è una similitudine con il contesto della Spagna, in cui tutti gli enti e le agenzie pubbliche sono connesse tramite la rete *Red SARA* (*Sistema de Aplicaciones y Redes para las Administraciones*) che costituisce anche il fondamento del cloud privato della PA spagnola.

La tabella tratta da [41] riporta l'analisi dei fattori motivanti e degli obiettivi strategici in un sottoinsieme di paesi europei oggetto dello studio.

Tabella 4. Adozione del cloud nei paesi europei: fattori e obiettivi strategici.

Country	Date of Strategy	Policy Drivers	Main Objectives
<b>Portugal</b>	2012	<ul style="list-style-type: none"> <li>- Cost savings</li> <li>- Economic impact for the private sector</li> <li>- Improvement of existing public services</li> </ul>	<ul style="list-style-type: none"> <li>- The introduction of cloud services in public administrations</li> <li>- Establish a framework agreement for the procurement of cloud computing services</li> <li>- Implementation of the GO-Cloud (Governmental Open Cloud), a platform with shared cloud services</li> </ul>
<b>Italy</b>	2013	<ul style="list-style-type: none"> <li>- Cost savings</li> </ul>	<ul style="list-style-type: none"> <li>- Rationalization of data centres</li> </ul>
<b>Netherlands</b>	2011	<ul style="list-style-type: none"> <li>- Improved and new modes of working for central government</li> <li>- Cost saving</li> </ul>	<ul style="list-style-type: none"> <li>- Closed cloud</li> <li>- Phased implementation</li> <li>- Cloud first policy</li> </ul>
<b>Spain</b>	2011	<ul style="list-style-type: none"> <li>- Cost savings</li> <li>- Economic impact for the private sector</li> <li>- Improvement of existing public services</li> <li>- Development of new public services</li> </ul>	<ul style="list-style-type: none"> <li>- The SARA network is a national network interconnecting national, regional and local administrations. It is evolving to provide cloud services, in a private cloud model. This network includes a datacentre dedicated to common services provided to administrations. This datacentre is currently growing over time, as the number of cloud services is growing.</li> </ul>
<b>Slovakia</b>	2013	<ul style="list-style-type: none"> <li>- Cost savings</li> <li>- More flexible and scalable information systems</li> <li>- Faster deployment of public services</li> <li>- Robustness of public administration information systems</li> </ul>	<ul style="list-style-type: none"> <li>- Creating a secure environment for citizens, businesses and public administration</li> <li>- Optimum use of information technologies in public administration through a shared services platform</li> <li>- Provide all type of cloud services</li> <li>- In order to make the use of cloud services easier for the consumers, they will be listed in a catalogue</li> <li>- Information systems to be set up as part of new projects will be implemented within the framework of the eGovernment cloud platform (the "cloud only" rule)</li> </ul>

Secondo un rapporto di Synergy Research Group<sup>10</sup> (2021) il mercato dei servizi cloud è in crescita, ma meno del 15% dei provider è europeo. Questo porta a interrogarsi sulla

<sup>10</sup> European cloud providers double in size but lose market share, 2021. Link: <https://www.srgresearch.com/articles/european-cloud-providers-double-in-size-but-lose-market-share>

dipendenza tecnologica dell'Europa dalle grandi multinazionali statunitensi, le cosiddette *Big Tech* (Amazon, Google e Microsoft), che sono fornitori di servizi cloud anche attraverso i contratti derivanti dagli appalti pubblici [43].

Su questo fronte vigila attivamente il Garante europeo per la protezione dei dati, *European Data Protection Supervisor* (EDPS), che nel maggio 2021 ha avviato due indagini<sup>11</sup> sugli appalti pubblici per la verifica della conformità al GDPR. La prima è relativa al trasferimento di dati personali nell'utilizzo dei servizi forniti da Microsoft e Amazon Web Services nell'ambito dei contratti Cloud II. L'altra indagine riguarda l'impiego di Microsoft Office 365 da parte della Commissione europea.

Se da un lato i dati personali sono oggetto di protezione in materia di privacy, dall'altro lato rappresentano al giorno d'oggi informazioni cruciali che vengono elaborate da algoritmi software per le applicazioni operanti in ogni settore della vita dei cittadini. Tali dati costituiscono un patrimonio tecnologico di valore e diventa fondamentale assicurare che, quando sono archiviati e trasportati attraverso i sistemi informatici, non siano resi disponibili a terzi mettendo a repentaglio la sovranità dello Stato. Questa riflessione è alla base del concetto di *sovranità digitale*.

Vista l'alta considerazione del cloud come tecnologia strategica, l'ENISA sta lavorando a un sistema europeo di certificazione della cybersicurezza per i servizi cloud. Inoltre l'UE ha in previsione una serie di norme per fornire un quadro unico europeo tramite un regolamento UE sul cloud computing e una guida sugli appalti pubblici dei servizi di trattamento dei dati. Le direttive della strategia digitale<sup>12</sup> della Commissione europea aprono verso due direzioni: da un lato verso la standardizzazione e l'interoperabilità dei servizi cloud e dall'altro proprio verso la sovranità digitale:

*The Commission is determined to make this Europe's "Digital Decade". Europe must now strengthen its digital sovereignty and set standards, rather than following those of others – with a clear focus on data, technology, and infrastructure.*

La sovranità digitale si fonda su due componenti interconnesse: la sovranità sui dati e quella sulle tecnologie, due assi lungo i quali si muove anche il programma europeo. La condivisione sicura dei dati e la definizione di regole di interscambio a livello globale sono alla base dell'architettura di *International Data Spaces Association* (IDSA)<sup>13</sup>.

---

<sup>11</sup> Rassegna stampa EDPS. Link: [https://edps.europa.eu/press-publications/press-news/press-releases/2021/edps-opens-two-investigations-following-schrems\\_en](https://edps.europa.eu/press-publications/press-news/press-releases/2021/edps-opens-two-investigations-following-schrems_en)

<sup>12</sup> UE strategy and policy. Link: [https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age\\_en](https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_en)

<sup>13</sup> Link al progetto: <https://internationaldataspaces.org/>

Nel contesto dell'indipendenza dal mercato cloud si inserisce il progetto di Gaia-X<sup>14</sup>, associazione fondata da Germania e Francia che coinvolge oltre 300 organizzazioni di diversi paesi. Il progetto ha l'obiettivo di costruire e gestire un'infrastruttura di cloud computing per fornire al settore pubblico una soluzione tecnologica indipendente dalle *Big Tech* statunitensi, nel rispetto degli standard europei. Utilizzando alla base il principio open source e una folta community di sviluppatori, l'idea è quella di creare un'infrastruttura dati federata, collegando diverse infrastrutture centralizzate e decentralizzate all'interno di un sistema omogeneo, mantenendo fin dalla progettazione il focus sui requisiti di sicurezza, affidabilità e trasparenza.

### 2.2.6 Il cloud nelle università

Le università e le istituzioni accademiche svolgono spesso il ruolo di pionieri delle innovazioni e sono in continua evoluzione, cercando di rinnovarsi tecnologicamente e di mantenere sempre alto il livello della didattica e della ricerca.

Come evidenzia Fuller [44] (2023), il contesto delle università ha alcune peculiarità che lo rendono unico. L'amministrazione centrale può avere prospettive diverse rispetto alle strutture decentralizzate come dipartimenti o collegi, che spesso hanno anche l'autorità di agire in modo autonomo rispetto al reparto IT. Ogni dipartimento può avere obiettivi discordi e gestire i propri budget con discrezionalità, creando sul versante tecnologico un insieme di sistemi disparati. I centri di ricerca hanno la necessità di infrastrutture e risorse informatiche ad alte prestazioni, siano esse presso le sedi locali oppure nel cloud. Infine, rispetto a tante organizzazioni del settore privato, molte università esistono da decenni e hanno quindi un'eredità di decisioni e architetture IT in cui i sistemi legacy devono coesistere con nuove applicazioni, garantendo innovazione senza interrompere i servizi esistenti.

I processi interni di un ateneo possono essere raggruppati in base alle quattro attività principali: didattica, ricerca, trasferimento tecnologico e attività amministrative a sostegno delle precedenti. Ciascuna presenta esigenze diverse e utilizza le tecnologie più adatte alle proprie finalità. La didattica richiede strumenti di e-learning ma anche laboratori, attrezzature e licenze software. Il mondo della ricerca è sempre più caratterizzato dalla necessità di potenza computazionale per l'esecuzione di esperimenti scientifici e simulazioni, mentre la gestione dei processi tecnici e amministrativi da parte dello staff si confronta con il rinnovo tecnologico e le nuove esigenze, come quelle innescate dalla pandemia Covid-19, dalla gestione remota dei sistemi allo *smart working*. Come sottolineato da Mircea e Andreescu [45] (2011), non esiste sicuramente un unico modello di cloud applicabile, al contrario per ciascuna funzione potrebbe essere indicata una scelta differente. La maggior parte delle organizzazioni utilizza modelli ibridi, mantenendo gli elementi chiave all'interno dell'infrastruttura privata e in un

---

<sup>14</sup> Link al progetto: <https://gaia-x.eu/>



Public cloud esterno i componenti ritenuti meno sensibili. Proprio in merito alla protezione dei dati, è importante cercare un punto di equilibrio tra due contesti: da un lato mantenere i dati sensibili all'interno del proprio data center ed esternalizzare gli altri con il rischio di aumentare la latenza per alcune applicazioni oppure accettare e gestire potenziali rischi per la sicurezza esternalizzando tutti i dati per ottenere alte prestazioni e scalabilità.

Tra le pubblicazioni accademiche sono presenti molteplici lavori che affrontano l'adozione del cloud nelle università. Sono quasi tutti accomunati da una parte introduttiva, spesso ridondante, dedicata a definizioni, modelli, benefici e criticità. Molti svolgono una revisione sistematica della letteratura scientifica esistente e utilizzano uno schema composto da indagini con questionari o interviste presso le strutture coinvolte. Alcuni autori si concentrano su quali modelli di cloud siano i più indicati [45] o propongono dei modelli astratti di adozione o transizione [46], [47]. Altri lavori discutono nel dettaglio le soluzioni tecniche, ponendo il focus sull'infrastruttura di rete necessaria, come ad esempio per il Private cloud [48], [49]. Dal punto di vista dei prodotti utilizzati, le università citate negli studi hanno optato per l'utilizzo di soluzioni commerciali come Amazon o Google [50], VMware [49] o software open source come OpenStack [51].

Il lavoro di Sultan [50] (2010) è meno recente rispetto ad altri ma presenta diverse casistiche tra le università in vari paesi, descrivendo anche alcuni esempi delle prime adozioni del cloud. A Berkeley la *University of California*, grazie a una donazione di Amazon Web Services, nel 2009 spostò uno dei suoi corsi di sviluppo software dal data center locale al cloud. Il *Medical College of Wisconsin* portò avanti un avanzato programma di proteomica<sup>15</sup> utilizzando la potenza computazionale di server Google Cloud in grado di analizzare il vasto flusso di dati generati dalla strumentazione impiegata. Sultan riporta inoltre il caso studio della *University of Westminster* nel Regno Unito che nell'anno accademico 2008-2009 attuò la migrazione del proprio sistema di posta sulla piattaforma Google, sfruttando la possibilità di utilizzare caselle email gratuite e calendario condiviso, con un'ampia capacità di spazio disco per ciascun utente.

Tra l'altro, un passaggio simile è quello affrontato anche dall'Università degli Studi di Torino che dal 2014 ha scelto il servizio di Google Apps for Education sia per la posta degli studenti sia per il personale dell'ateneo<sup>16</sup>.

Nella sua indagine Ali M. B. [52] (2019) cerca di esaminare da molteplici prospettive tutti i fattori che influenzano l'adozione del cloud, soffermandosi in particolare su quelli

---

<sup>15</sup> Branca della biologia molecolare che studia il proteoma, l'insieme di tutte le proteine espresse dal genoma di una cellula in un particolare momento, per comprendere i meccanismi alla base dell'insorgenza delle malattie (fonte: Dizionario di Medicina Treccani).

<sup>16</sup> Link: <https://www.unito.it/servizi/servizi-line/webmail/webmail-del-personale>

tecnologici, organizzativi, ambientali e personali. Interessante notare che tra i fattori organizzativi, i più influenti sono il supporto da parte dei livelli manageriali, la dimensione dell'ateneo e le competenze IT mentre tra i fattori ambientali spiccano il sostegno e le politiche del governo. Inoltre l'implementazione del cloud è strettamente legata alla situazione tecnologica dei vari paesi e ai requisiti di base, primi fra tutti le forniture stabili di energia elettrica o l'esistenza di reti di telecomunicazioni e connessioni Internet affidabili e veloci.

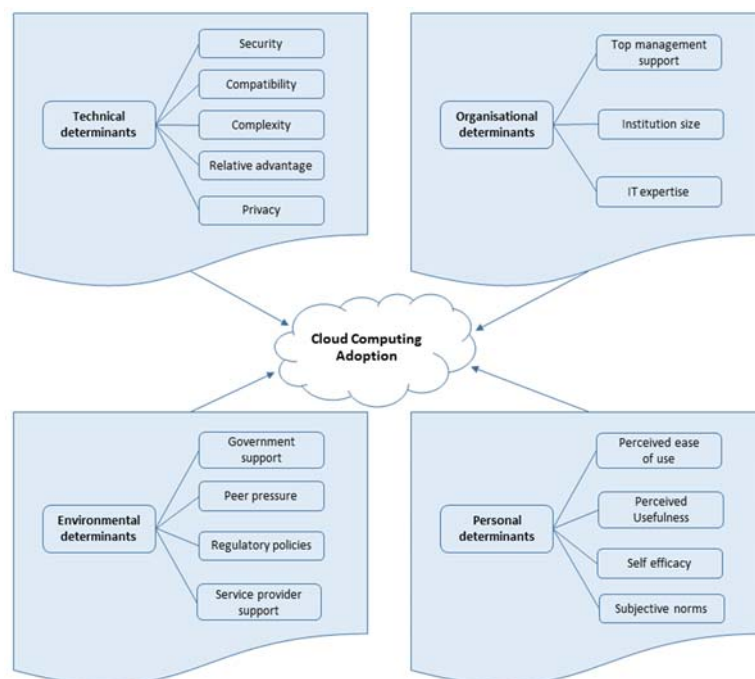


Figura 8. Università: fattori determinanti per l'adesione al cloud.

Fonte [52]

Il lavoro di Aydin [46] (2021) prende il via da un'indagine in Turchia, in cui oltre il 43% delle università, pur usando la virtualizzazione, dichiara di non disporre di un'infrastruttura cloud. L'autore delinea un'ipotetica architettura per l'IT applicando il modello Hybrid cloud, ma a differenza di [45] che vede l'abbinamento di un cloud pubblico e uno privato, qui si affiancano un *Private University Cloud* e un *Community University Cloud* tra i quali è configurata la sincronizzazione di dati. Il primo, di proprietà e gestito in autonomia da ciascuna università, permette di ospitare le attività e le applicazioni critiche nella rete interna e dietro il firewall. Il secondo è formato da servizi e applicazioni fornite dai vari atenei che partecipano ed è indicato per la condivisione, con la comunità accademica e altri enti affiliati, delle informazioni caratterizzate da requisiti di sicurezza meno stringenti.

Un altro modello astratto, battezzato *Campus Cloud Computing*, è quello proposto per le università in Iraq da Jasim et al. [47] (2022), in cui si ricalca la divisione standard

dei livelli di servizio del cloud computing riformulati in Campus-IaaS, Campus-PaaS e Campus-SaaS. Nel primo vengono forniti *as a service* le infrastrutture di ateneo quali server, storage, dispositivi di rete. Il secondo offre la base per lo sviluppo di proprie applicazioni o servizi; infine l'ultimo livello distribuisce agli utenti le applicazioni web based del sistema universitario come posta elettronica, strumenti e-learning, procedure amministrative e di immatricolazione, gestione di aule e laboratori.

Due scenari simili incentrati sulla realizzazione di un Private cloud, sono esaminati da Alzoubaidi [48] (2016) e da Kyriakou et al. [49] (2023). Alzoubaidi ambienta la sua ricerca presso la *Al-Balqa Applied University* (BAU), un'università multi-campus con 13 sedi in Giordania, caratterizzata da un modello IT decentrato, con un data center presso il campus principale nella città di Salt e 12 data center regionali in campus remoti, con evidente duplicazione delle risorse sia hardware e software sia di personale per la gestione. Viene proposto uno schema operativo con le fasi salienti per consolidare i servizi e le infrastrutture IT in un unico cloud privato ridonato, concentrando maggiormente l'analisi sugli aspetti di networking e meno sulle soluzioni di virtualizzazioni scelte. In modo analogo Kyriakou et al. propongono il caso della rete della *University of the Aegean*, un ateneo dalla spiccata peculiarità multi-campus, in quanto dislocato su sei diverse isole nel mar Egeo. Ne descrivono il passaggio da una topologia di rete organizzata per campus, in cui ogni isola aveva il proprio insieme di server, con funzioni spesso ridondanti, ad un'infrastruttura di tipo Private cloud, con un data center centralizzato e uno di backup. La realizzazione vede l'impiego di soluzioni di virtualizzazione di tipo proprietario (VMware), lasciando comunque la porta aperta all'eventuale espansione verso un'infrastruttura IT ibrida.

L'adozione di soluzioni open source per la realizzazione di cloud privati è al centro della pubblicazione di Marken e Jain [51] (2020) che, dopo un breve confronto delle piattaforme disponibili, discutono l'implementazione di OpenStack in un ambiente di test presso l'ateneo indiano *Punjab Agricultural University* a Ludhiana.

Dall'analisi dei documenti in letteratura, si può affermare che anche nel settore universitario emergono gli stessi benefici attesi e timori analizzati nel paragrafo 2.2.4 per l'adozione del cloud in generale. Si aggiungono, citati da alcuni studi, la tutela della proprietà intellettuale e la protezione dei risultati della ricerca, sia durante il trasporto in rete sia in fase di archiviazione. In conclusione emerge la riflessione che l'abilitazione al cloud non deve essere vista come uno spostamento totale dei propri asset verso nuove soluzioni, ma molto spesso la soluzione più vantaggiosa è legata all'ibridazione, soprattutto in funzione di particolari esigenze infrastrutturali e organizzative.

### 2.2.7 Cloud repatriation

Il livello di maturità raggiunto dal cloud computing consente alle organizzazioni di comprendere meglio come modellare la tecnologia in base alle proprie necessità. In questo processo di adattamento si inserisce il recente fenomeno chiamato “rimpatrio dal cloud” oppure in termini anglosassoni *cloud repatriation*. Si tratta dello spostamento di dati e applicazioni dal cloud pubblico verso altre soluzioni, principalmente ad opera di aziende di medio-grosse dimensioni del settore tecnologico. Come riporta Jewargi [53] (2023), molti staff IT hanno riscontrato che la migrazione e la compatibilità applicativa non erano così semplici da gestire, dopo gli entusiasmi iniziali e le brillanti pubblicità dei fornitori. Si sono resi conto che era inadeguato utilizzare il cloud pubblico come si usa il proprio data center interno aziendale, oltretutto rilevando l'aumento oltre le aspettative dei costi per la spesa cloud.

Hanno affrontato il tema due aziende note a livello mondiale: *International Data Corporation* (IDC), specializzata in ricerche e analisi di mercato per il settore ICT e *S&P Global*, società di servizi di informazione finanziaria. Nel 2018 IDC ha condotto il sondaggio *Cloud and AI Adoption* [54] rivolto agli utenti finali delle strutture IT, con un campione di 400 aziende in USA, Regno Unito, Singapore, India, Australia e Nuova Zelanda. L'80% degli intervistati segnalava di aver spostato carichi di lavoro dai cloud pubblici e che prevedevano di dislocare il 50% delle loro applicazioni nei successivi due anni. Tuttavia secondo gli analisti questi risultati non significano che il consumo nel cloud pubblico scenderà del 50% né che vi possa essere un rallentamento della crescita nella sua adozione. Sicuramente molte aziende hanno effettuato una transizione significativa verso lo sviluppo di applicazioni modulari, da eseguire nei cloud pubblici ma più facilmente portabili o estendibili anche in ambienti cloud privati o dedicati.



Figura 9. Sondaggio “Cloud repatriation” di IDC.

Fonte [54]

I fattori principali dietro a questa scelta, secondo l'analisi, sono le preoccupazioni relative alla sicurezza (19%), alle prestazioni (14%), ai costi (12%), nonché necessità di maggiore controllo (12%) e di riduzione dello *shadow IT*<sup>17</sup>.

Numeri leggermente diversi ma con la stessa tendenza emergono dall'indagine di S&P Global [55] con il sondaggio *Voice of the Enterprise: Datacenters 2021*: su oltre 600 intervistati, il 48% ha indicato di aver trasferito, nell'anno precedente, un carico di lavoro o un'applicazione da un public cloud ad un'altra sede. Anche in questo caso, come commentano gli analisti, ciò non significa che metà dei carichi di lavoro sono stati spostati, ma è un indicatore dell'esperienza comune di diverse aziende nel ripensare la collocazione dei propri asset. In particolare quasi l'86% ha scelto un proprio datacenter e il 14% una struttura di *colocation*, opzione però citata anche da un 28% che afferma di aver scelto un'infrastruttura gestita da terzi, suggerendo che nel sondaggio ci sia una sovrapposizione seppur lieve di risposte.



Figura 10. Sondaggio “Cloud repatriation” di S&P Global.

Fonte [55]

Anche se con percentuali differenti, ci sono alcune analogie tra i due studi relativamente alle motivazioni della scelta. Secondo S&P Global prevalgono lo *shadow IT*, con l'uso non autorizzato del cloud pubblico (25%), la sicurezza (23%), la gestione del ciclo di vita delle applicazioni (22%), le norme o requisiti di governance (18%) e la localizzazione o sovranità dei dati (16%).

Lo scenario si è arricchito con due popolari casi di società del settore IT che hanno annunciato scelte di *cloud repatriation*. Il più citato è quello dell'azienda statunitense

<sup>17</sup> Utilizzo di dispositivi o soluzioni informatiche all'interno di un'organizzazione senza l'esplicita approvazione del reparto IT.

Dropbox, leader dei servizi di archiviazione file online che era nata sul cloud usando AWS e che dal 2015 ha iniziato a spostare i dati dei suoi 500 milioni di utenti su un'infrastruttura di proprietà<sup>18</sup>. In realtà Dropbox ha costruito una vera e propria rete con sistemi progettati ad hoc dagli ingegneri interni e con un sistema software realizzato su misura per le esigenze peculiari di archiviazione e condivisione di file e per ottimizzare la gestione dello storage. Secondo i report finanziari, dal 2015 al 2016, la società ha risparmiato 40 milioni di dollari sui costi operativi dell'infrastruttura e ulteriori 35 milioni nel 2017 [56]. Nei luoghi in cui non ci sono le condizioni per creare propri data center o il mercato è meno prevedibile, come ad esempio in Europa, Giappone e in generale al di fuori dell'America, l'azienda continua comunque a utilizzare il cloud di Amazon e gestisce le connessioni di rete attraverso dei POP (*Point of Presence*) locali.

Un altro caso è quello sollevato nel 2022 da David Heinemeier Hansson, creatore di *Ruby on Rails*, framework open source per lo sviluppo di applicazioni web, e fondatore di progetti come Basecamp e HEY, rispettivamente uno strumento di project management web-based e un sistema alternativo di e-mail. Sul suo blog ha raccontato in diversi post<sup>19</sup> come dopo anni di utilizzo del cloud di Amazon e Google per le sue aziende, abbia deciso di abbandonare questa soluzione. In alcuni suoi testi a volte ironici e provocatori, oltre a criticare i grandi profitti dei grandi hyperscaler, egli pone anche questioni di principio:

*But this isn't just about cost. It's also about what kind of internet we want to operate in the future. It strikes me as downright tragic that this decentralized wonder of the world is now largely operating on computers owned by a handful of mega corporations. If one of the primary AWS regions go down, seemingly half the internet is offline along with it. This is not what DARPA designed!*<sup>18</sup>

Dal punto di vista tecnico, secondo Hansson il cloud è perfetto se le applicazioni sono semplici e generano poco traffico oppure quando il carico è molto irregolare, con picchi e oscillazioni non prevedibili. Secondo le sue stime, con l'abbandono del cloud pubblico il risparmio per le sue aziende ammonterebbe a 1,5 milioni di dollari all'anno. Per portare i servizi on-premise e tenerli aggiornati, sono stati acquistati nuovi server Dell per un totale di 4.000 vCPU, 7.680 GB di RAM e 384 TB di storage NVM. L'hardware, ammortizzato in cinque anni, è costato circa mezzo milione di dollari, meno di un terzo del risparmio annuale previsto. La soluzione scelta tuttavia non è stata quella di creare

---

<sup>18</sup> Dropbox.Tech. Link: <https://dropbox.tech/infrastructure/magic-pocket-infrastructure>

<sup>19</sup> "Why we're leaving the cloud", <https://world.hey.com/dhh/why-we-re-leaving-the-cloud-654b47e0>. "We have left the cloud", <https://world.hey.com/dhh/we-have-left-the-cloud-251760fb>

e gestire un proprio data center, come nel caso di Dropbox, ma collaborare con Deft<sup>20</sup>, un'azienda specializzata in questo tipo di servizi.

Un altro spunto di interesse è fornito dall'azienda *Andreessen Horowitz* [56]. Nel 2021 un team di analisti finanziari ha studiato 50 delle principali società di software quotate in borsa e ha valutato che l'ingente spesa per il cloud influisce negativamente sull'efficienza a breve termine di una società: se le aziende riescono a garantire ampi risparmi riducendo tali uscite, questo non solo aumenta i margini di profitto ma consente loro di acquisire un maggiore valore sui mercati azionari. In particolare hanno stimato che la capitalizzazione di mercato aumenta in media di circa 24-25 volte il risparmio netto sui costi derivanti dal rimpatrio del cloud.

Lo scenario appena delineato sembra in conflitto con l'indagine di Gartner [13] di cui al paragrafo 2.2.2, con la previsione che entro il 2025 le spese per il cloud computing saranno superiori a quelle per le installazioni on-premise. In realtà l'analisi di Gartner comprende un campione più ampio: non solo società che operano nel settore del software ma anche di servizi legati ai processi aziendali e alle infrastrutture di sistema. Gli esempi e gli studi riportati sul *cloud repatriation* si collocano principalmente all'interno di organizzazioni di ampie dimensioni e che lavorano nel settore tecnologico, quindi in grado di affrontare progetti complessi e di acquisire elevate competenze interne. In particolare Dropbox, le società di Hansson e in genere le aziende il cui core business è la tecnologia hanno maggiori necessità di prestazioni e controllo, quindi la ricollocazione dei servizi è una mossa strategica. Al contrario, per le aziende non tecnologiche, il cloud pubblico rimane più che valido in termini di costi e opportunità.

Non sarebbe comunque appropriato considerare il fenomeno del *cloud repatriation* con accezione negativa. Guardando a Dropbox, HEY e Basecamp, il cloud pubblico ha consentito a piccole startup di utilizzare, senza investimenti di capitale, risorse computazionali e di storage per poter avviare le proprie attività. Questi esempi, piuttosto, possono aiutare le aziende aumentando la consapevolezza delle implicazioni sui costi a lungo termine del cloud e diffondendo nuovi approcci alla gestione delle infrastrutture e delle operazioni IT. Gli specialisti hanno ormai compreso che non tutti i carichi di lavoro di dati e applicazioni sono adatti al cloud pubblico e sta diventando sempre più importante valutare dove e come eseguire i servizi. La maggior parte delle aziende non cerca più una risposta unica e onnicomprensiva alle proprie esigenze IT quanto piuttosto una struttura che possa soddisfare i requisiti di costo, prestazioni, e controllo dei diversi carichi di lavoro, concetto conosciuto come *best execution venue* [53], [54]. Sono candidati ideali allo spostamento i servizi caratterizzati da modelli d'uso noti con consumi di risorse stabili e prevedibili, o che risentono in caso di latenze eccessive, lasciando invece sul cloud pubblico le applicazioni contraddistinte da picchi di utilizzo e traffico imprevedibile che richiedono scalabilità. Nonostante siano spesso

---

<sup>20</sup> Società che fornisce servizi IT gestiti per le grandi aziende. Link: <https://deft.com/>.

usati come sinonimi, non è detto che *cloud repatriation* indichi obbligatoriamente l'opzione *back-to-on-premise*. La ricollocazione dei carichi di lavoro non significa adottare le stesse configurazioni on-premise precedenti alla transizione cloud, anzi si può convergere verso modelli alternativi: oltretutto la migrazione è facilitata nel caso siano stati usati approcci di tipo *cloud native* e modulari. Non tutte le società inoltre hanno le possibilità e l'interesse a investire per creare un data center on-premise.

Un'alternativa piuttosto comune come destinazione per lo spostamento dei carichi di lavoro è la cosiddetta *colocation* o *housing*, soluzione che prevede l'affitto di un'infrastruttura completa in un ambiente privato o condiviso presso una sala operativa. Sono molti infatti i fornitori che hanno iniziato a proporre il noleggio dell'hardware IT secondo il modello *as-a-Service* con sistemi preconfigurati, dotati di capacità di espansione, forniti in armadi rack e fatturati in base al consumo [57].

Questa opzione è particolarmente apprezzata dalle aziende perché consente di ridurre i costi di capitale (*Capex*) e trasformare l'hardware fisico in un costo operativo (*Opex*)<sup>21</sup> con l'eventualità di negoziare maggiormente i contratti e personalizzare le soluzioni di fornitura. L'indagine di S&P Global [55] conclude con la previsione che a tendere il cloud pubblico coesisterà con le soluzioni private *on-premise/colocation*, come indicato dal 57% delle organizzazioni che puntano agli ambienti ibridi per il loro IT. Portando un paragone alla situazione delle PMI italiane, è interessante notare che nello studio dell'Osservatorio Cloud Transformation del Politecnico di Milano, citato nel paragrafo 2.1.3, la dinamica di crescita più significativa sia stata quella relativa alla componente *Public e Hybrid Cloud* (+22% rispetto al 2021), voce in cui si sommano i servizi forniti dai CSP e l'interconnessione tra cloud pubblici e privati.

Come è naturale, lo sviluppo dei prodotti e i servizi sul mercato segue l'evoluzione della domanda. Da un lato i grandi CSP cosiddetti hyperscaler hanno avviato iniziative *cloud-to-ground* (AWS Outposts, Microsoft Azure Stack, Google Anthos) per fornire servizi privati e infrastrutture dedicate collegate al cloud ma distribuite in data center locali. Dall'altro lato gli storici fornitori di hardware hanno stanno cercando il proprio spazio nel mercato orientandosi verso modelli di *Hardware-as-a-Service*: rientrano in questo scenario HPE con l'offerta Greenlake<sup>22</sup> e Dell con Apex<sup>23</sup>.

---

<sup>21</sup> Sono due tipi di spese aziendali, divise in base al modo in cui influiscono sulla contabilità. *Opex* (*Operating Expenses*) sono spese operative, rappresentano i costi di gestione ricorrenti (es. stipendi, affitti, spese per le utenze). *Capex* (*Capital Expenditures*) sono invece spese per investimenti a lungo termine effettuate per migliorare l'attività (es. acquisto di attrezzature e macchinari, software e materiale informatico, ristrutturazione di edifici).

<sup>22</sup> Link: <https://www.hpe.com/it/it/greenlake.html>

<sup>23</sup> Link: <https://www.dell.com/it-it/dt/apex>



Anche se allo stato attuale coinvolge principalmente grosse aziende che lavorano o che investono molto nel settore tecnologico, il fenomeno del *cloud repatriation* è comunque da ritenersi significativo. Non si assisterà al declino dei fornitori di servizi cloud pubblici, ma sicuramente quando le aziende raggiungeranno una maggiore esperienza nell'adozione della tecnologia, saranno in grado di decidere la migliore dislocazione per eseguire servizi e applicazioni, se pubblica/off-premise oppure privata/on-premise o ancora privata/off-premise (Figura 11).

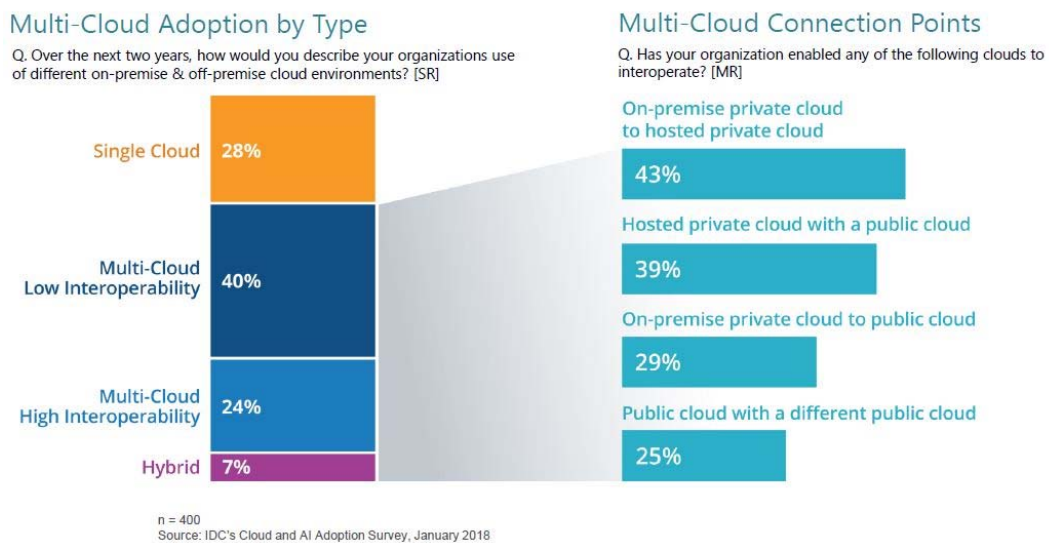


Figura 11. Adozione della soluzione Multicloud, sondaggio IDC.

Fonte [54]

Gli analisti di S&P Global [52] ritengono però che sarebbe preferibile utilizzare un termine diverso da *cloud repatriation*, in quanto il concetto di rimpatrio può veicolare l'idea di un ritorno indietro dopo scelte non ottimali o approcci inefficaci al cloud [52], mentre spesso si tratta di una decisione strategica e consapevole. Suggestiscono il termine *workload mobility*, una definizione più indicata per descrivere la natura dinamica dell'IT e la tendenza alla mobilità delle applicazioni.

Come avviene in molteplici settori, la chiave del successo risiede nell'integrazione tra diverse soluzioni: in questo caso mantenendo le caratteristiche più apprezzate del modello operativo cloud, l'agilità e l'erogazione efficiente dei servizi, la direzione è sicuramente verso un'infrastruttura IT ibrida o in un ambiente ibrido/multi-cloud.

# Capitolo 3.

## Il cloud nella Pubblica Amministrazione italiana

L'adozione del cloud da parte delle PA si colloca all'interno dello scenario composto dal Piano Triennale per l'informatica nella PA 2022-2024 (PT), dalla Strategia Nazionale di Cybersicurezza 2022-2026 e dal Piano Nazionale di Ripresa e Resilienza (PNRR). Inoltre come riporta Montagnani [58] (2022)

*la digitalizzazione delle pubbliche amministrazioni si inserisce in un disegno di più ampio respiro di matrice europea, parte integrante della strategia UE per attuare i 17 obiettivi di sviluppo sostenibile previsti dall'Agenda 2030<sup>24</sup> e per realizzare una rete di punti di incontro tra i diversi stakeholder (imprese, terzo settore, cittadini, PA) che daranno vita alle future città e comunità digitali.*

In tutti gli aggiornamenti del Piano Triennale che si sono susseguiti negli anni, il cloud è stato sempre indicato come una tecnologia abilitante; al capitolo 4, obiettivo 4.1 [59, p. 41] sono rimarcate le linee di azione per la transizione: “*migliorare la qualità e la sicurezza dei servizi digitali erogati dalle amministrazioni attuando la strategia Cloud Italia e migrando verso infrastrutture e servizi cloud qualificati*”.

Sono ingenti anche gli investimenti destinati: il PNRR, all'interno della Missione 1 “*Digitalizzazione, innovazione, competitività, cultura e turismo*”, riserva il 21% delle risorse alla transizione digitale dell'Italia, inserendo a budget 1,9 miliardi di euro per la voce cloud [59, p. 39].

---

<sup>24</sup> Programma d'azione sottoscritto nel 2015 dai 193 Paesi membri ONU, formato da 17 obiettivi per lo sviluppo sostenibile da raggiungere entro il 2030. Link: <https://unric.org/it/agenda-2030>.

### 3.1 Organizzazioni di riferimento

Si indicano di seguito le principali organizzazioni coinvolte nei processi di adozione del cloud nelle PA, riportando le descrizioni dai rispettivi siti web istituzionali.

- *Agenzia per l'Italia Digitale (AgID)*<sup>25</sup>. Istituita con la legge n. 134 del 7 agosto 2012, è l'agenzia tecnica della Presidenza del Consiglio che ha il compito di garantire la realizzazione degli obiettivi dell'Agenda digitale italiana e contribuire alla diffusione dell'utilizzo delle tecnologie dell'informazione e della comunicazione, favorendo l'innovazione e la crescita economica.
- *Dipartimento per la Trasformazione Digitale (DTD)*<sup>26</sup>. È la struttura di supporto alla Presidenza del Consiglio dei Ministri per la promozione ed il coordinamento delle azioni del Governo finalizzate alla definizione di una strategia unitaria in materia di trasformazione digitale e di modernizzazione del Paese attraverso le tecnologie digitali.
- *Agenzia per la cybersicurezza nazionale (ACN)*<sup>27</sup>. Istituita con il D.L. n. 82 del 14 giugno 2021, è l'autorità a tutela degli interessi nazionali nel campo della cybersicurezza. L'Agenzia ha il compito di tutelare la sicurezza e la resilienza nello spazio cibernetico. Si occupa di prevenire e di mitigare il maggior numero di attacchi cibernetici e favorire il raggiungimento dell'autonomia tecnologica.

### 3.2 Piano Triennale per l'informatica 2022-2024

Il Piano Triennale per l'informatica nella Pubblica Amministrazione (PT) [59], introdotto per il periodo 2017-2019, è uno strumento volto a promuovere la trasformazione digitale dell'Italia. Rappresenta un documento di supporto e di orientamento per la PA italiane che devono pianificare il percorso di evoluzione tecnologica. L'edizione 2020-2022 e il relativo aggiornamento 2021-2023 erano incentrati sulle linee guida di implementazione e monitoraggio dei risultati. L'attuale aggiornamento 2022-2024 offre un quadro strutturato dei vari ambiti di intervento, in ottica di razionalizzazione e standardizzazione, e si caratterizza per una presenza dominante del PNRR. Il piano si fonda sui seguenti principi guida:

- *digital & mobile first* (digitale e mobile come prima opzione): i nuovi sistemi devono essere realizzate primariamente digitali;
- *digital identity only* (accesso esclusivo mediante identità digitale): i servizi delle PA devono essere accessibili in via esclusiva con sistemi di identità;

---

<sup>25</sup> Sito ufficiale: <https://www.agid.gov.it/>

<sup>26</sup> Sito ufficiale: <https://innovazione.gov.it/>

<sup>27</sup> Sito ufficiale: <https://www.acn.gov.it/>

- *cloud first* (cloud come prima opzione): nei nuovi progetti deve essere adottato primariamente il modello cloud, cercando di prevenire il rischio lock-in;
- *servizi inclusivi e accessibili*: i servizi delle PA devono andare incontro alle esigenze dei diversi cittadini e dei territori;
- *dati pubblici un bene comune*: in quanto bene fondamentale per lo sviluppo del Paese, il patrimonio dei dati della PA deve essere valorizzato e reso disponibile in forma aperta a persone e imprese;
- *interoperabile by design*: i servizi delle PA devono essere progettati in modalità integrata esponendo le opportune API per la comunicazione con altri servizi;
- *sicurezza e privacy by design*: già nella fase di progettazione dei servizi delle PA deve essere rispettata la protezione dei dati personali secondo il GDPR;
- *user-centric, data driven e agile*: le PA devono sviluppare i servizi digitali basandosi sull'esperienza utente e sul continuo monitoraggio di dati di utilizzo, prevedendo modalità agili di miglioramento;
- *once only*: le PA devono evitare di chiedere a cittadini e imprese informazioni già fornite in precedenza;
- *transfrontaliero by design* (concepito come transfrontaliero): le PA devono rendere disponibili a livello transfrontaliero i servizi pubblici digitali rilevanti;
- *codice aperto*: per le applicazioni si deve preferire l'uso di software con codice aperto; nello sviluppo software deve essere reso disponibile il codice sorgente.

La funzione di coordinamento delle amministrazioni nel percorso di attuazione del PT è stata affidata all'agenzia AgID. Con il progetto “*Razionalizzazione infrastruttura ICT e migrazione al cloud*”<sup>28</sup> avviato a ottobre 2017, AgID si è posta due macro obiettivi: razionalizzare i data center della PA sul territorio e incrementare l'uso del cloud come soluzione per accelerare il miglioramento e l'organizzazione efficace dei servizi pubblici.

### 3.3 Censimento del patrimonio ICT della PA

Tra le prime azioni del PT realizzate da AgID vi è l'organizzazione del censimento degli asset informatici della PA. I data center sono stati divisi in 3 categorie [60]:

1. *Poli strategici nazionali (PSN)*: strutture di eccellenza tecnica e organizzativa
2. *Gruppo A*: strutture di qualità con carenze considerate minori, che possono continuare ad operare ma non essere oggetto di investimenti per l'evoluzione
3. *Gruppo B*: strutture che non garantiscono requisiti minimi di sicurezza e affidabilità, da consolidare verso i PSN o verso il cloud.

---

<sup>28</sup> Link al progetto: <https://www.agid.gov.it/it/agenzia/progetti-pon-governance/razionalizzazione-ict-migrazione-cloud>

Alla prima e seconda fase del censimento, nel corso del 2018, hanno partecipato 778 amministrazioni e sono stati registrati 927 data center. Alla terza fase (luglio-settembre 2019) hanno partecipato 212 nuovi enti, per un totale di 990 partecipanti e 1.252 data center complessivi. Dei 1.252 data center censiti, 35 sono risultati candidabili all'utilizzo da parte del polo strategico nazionale, 27 sono stati classificati nel gruppo A e i rimanenti 1.190 nel gruppo B [61].

Tabella 5. Data center della PA: risultato del censimento AgID

Tipologia	Descrizione	Censiti	Futuro
Candidabili	Strutture con requisiti di eccellenza tecnica e organizzativa	35	Candidabili all'utilizzo da parte di un PSN
Gruppo A	Strutture di qualità con carenze considerate minori	27	Possono continuare ad operare ma non essere oggetto di investimenti per l'evoluzione
Gruppo B	Strutture che non garantiscono requisiti minimi di sicurezza e affidabilità	1.190	Da dismettere, consolidare verso i PSN, altri DC o verso il cloud

L'82% delle PA censite ha dichiarato di possedere un data center di proprietà. Il 28% ha avuto avvio prima del 2000, il 56% tra il 2000 e il 2014, solo il 16% è costituito dai più recenti, realizzati a partire dal 2015. L'ultimo intervento di rinnovo è stato effettuato dopo il 2015 per il 64% delle strutture, prima del 2015 per il restante 36%.

### Il censimento del patrimonio ICT - Report

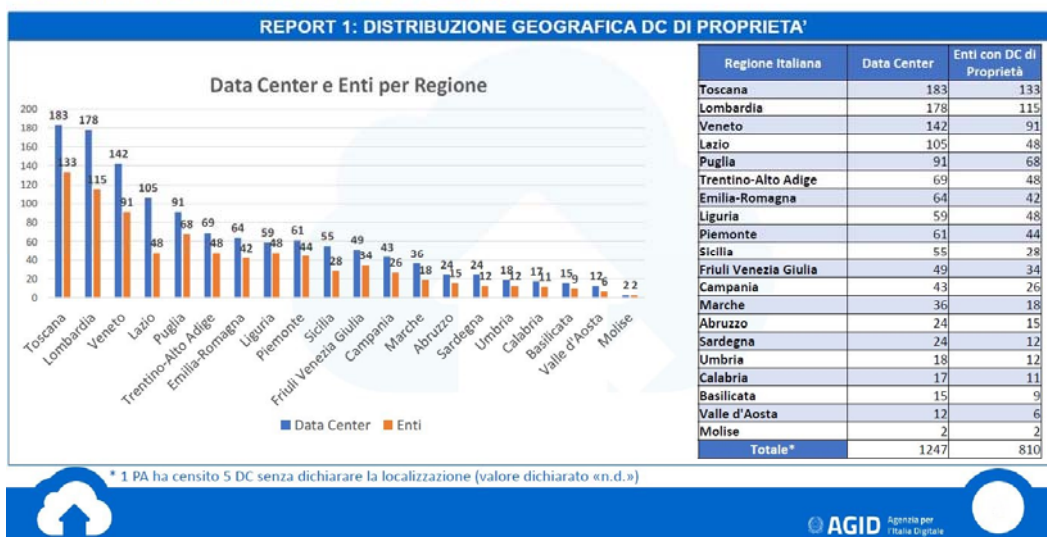


Figura 12. Censimento AgID: distribuzione data center della PA

Fonte [61]

Considerando la partecipazione di 990 enti sul totale di circa 13.000 PA italiane<sup>29</sup>, il censimento non individua la totalità delle risorse ICT a disposizione e rileva meno dell'8% delle situazioni sul territorio. Inoltre, la mancata partecipazione al censimento, se non opportunamente motivata, ha comportato la classificazione automatica della PA nel Gruppo B delle infrastrutture carenti.

### 3.4 Dal censimento alla Strategia Cloud Italia

Il censimento ha consentito di fotografare, seppur parzialmente, la situazione delle infrastrutture ICT ad uso delle PA sul territorio. Successivamente il progetto “*Razionalizzazione infrastruttura ICT e migrazione al cloud*” è proseguito su due componenti parallele. Da un lato si è sviluppato il piano di ottimizzazione delle infrastrutture digitali, che comprende attività quali il consolidamento e la razionalizzazione delle strutture, l'eventuale dismissione dei data center obsoleti e la migrazione dei servizi verso data center più affidabili oppure in cloud. Contemporaneamente è emersa l'adozione del *Modello del Cloud della PA*, un insieme di procedure per acquistare e gestire in cloud i servizi pubblici, utilizzando fornitori qualificati da AgID che rispettino i criteri di affidabilità e sicurezza.

La Strategia Cloud Italia si inserisce in questo contesto per condurre la PA verso la migrazione al cloud e per implementare il criterio *cloud first* descritto sia nel PNRR sia nel PT [59, pp. 6-8], con l'obiettivo che tutti i servizi pubblici siano sviluppati in modalità nativa sulla base dei modelli cloud. Il documento di indirizzo è stato pubblicato nel settembre 2021 dal DTD e ACN, indicando le linee guida per scegliere le soluzioni cloud più adeguate per i servizi pubblici e la tutela dei dati. La strategia, la cui realizzazione è prevista entro il 2025, si articola in tre fasi [62, pp. 9-13]:

1. *Classificazione dei dati e dei servizi*: definizione di un processo di classificazione dei dati per indirizzare la migrazione dei dati e servizi della PA sul cloud;
2. *Qualificazione dei servizi cloud*: realizzazione di un processo sistematico di scrutinio e qualificazione dei servizi cloud utilizzabili dalla PA;
3. *Polo strategico nazionale*: creazione di un'infrastruttura nazionale per i servizi cloud, la cui gestione e controllo siano autonomi da soggetti extra UE.

Il filo conduttore di queste iniziative è dotare la PA di una regolamentazione chiara e rendere uniforme l'adozione del cloud tra i diversi enti, oltre che sfruttare l'economia di scala per riuscire a ridurre i costi di gestione mantenendo alta la qualità dei servizi.

In linea con le considerazioni già descritte nel capitolo 2, il documento di strategia approfondisce le sfide poste dal cloud computing [62, pp. 7-8]. Riveste un ruolo di primo piano l'autonomia tecnologica. La constatazione che grandi aziende extra UE detengono la maggior parte dei servizi e delle piattaforme digitali, ma soprattutto le infrastrutture

---

<sup>29</sup> Dati Istat, giugno 2023. Link: <https://www.istat.it/it/censimenti/istituzioni-pubbliche>

che ne consentono il funzionamento, pone le PA degli stati europei in una posizione di “debolezza contrattuale”. Per promuovere l’adozione massiva di tecnologie IT abilitanti per lo sviluppo del Paese, si ritiene fondamentale esercitare un diretto controllo sui dati e sui servizi, evitando che organizzazioni di paesi terzi accedano a dati sensibili o comunque strategici per le istituzioni. Con questo obiettivo viene quindi delineata una procedura di classificazione per identificare le tipologie di dati che possono essere gestite in un cloud pubblico da un fornitore extra UE e quelle che necessitano di CSP aderenti a specifici requisiti di sicurezza. Un’altra classificazione riguarda i fornitori di cloud pubblico, per i quali era già stata predisposta una procedura di qualificazione AgID.

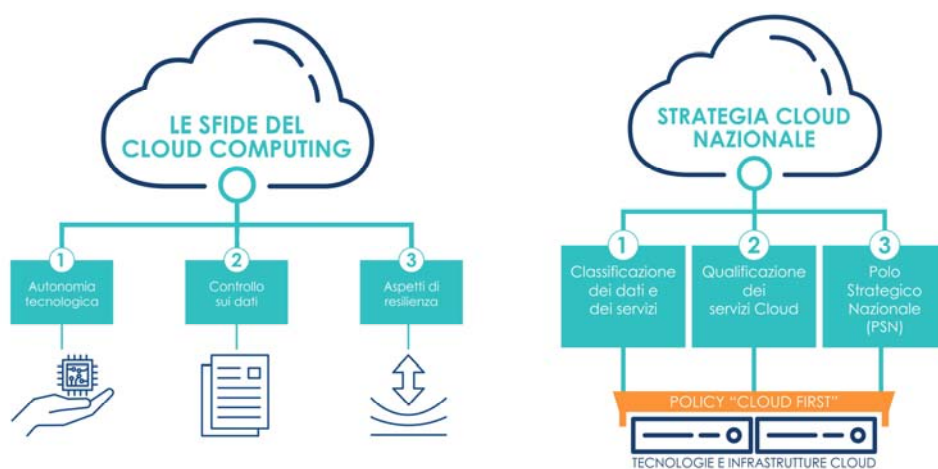


Figura 13. Strategia Cloud Italia: sfide e fasi attuative.

Fonte [62]

Un altro aspetto su cui la Strategia Cloud si sofferma con dettaglio è la resilienza, indicando la necessità di mettere in atto controlli di sicurezza nonché procedure organizzative e tecniche per assicurare la ridondanza, la continuità di servizio e il *disaster recovery* in siti geograficamente distribuiti sul territorio nazionale. Infine un’altra linea di indirizzo coinvolge la standardizzazione e l’interoperabilità di servizi e applicazioni, ribadendo la partecipazione dell’Italia all’obiettivo di sviluppare requisiti comuni per un’infrastruttura dati europea e al progetto Gaia-X (si veda capitolo 2).

### 3.5 Razionalizzazione dei data center

Dopo la chiusura del censimento del patrimonio ICT nel 2019, alle PA non è più consentito deliberare spese per i data center, se non estremamente necessarie. Come riporta la circolare 1/2019 di AgID [60]:

*in materia di spesa le PA non possono effettuare spese o investimenti in materia di data center, ma previa approvazione di AgID possono procedere agli adeguamenti dei propri data center esclusivamente al fine di evitare problemi di interruzione di pubblico servizio, anticipare processi di dismissione per migrare al cloud della PA, consolidare i propri servizi sui data center di altre PA per ottenere economie di spesa.*

In termini di asset tecnologici, viene effettuata una distinzione tra le infrastrutture che gestiscono servizi considerati strategici, legati a funzioni vitali del Paese (ad esempio telecomunicazioni, energia, mobilità), e tutte le altre infrastrutture gestite da PA centrali e locali, usate per erogare i servizi pubblici ordinari. Nel primo caso l'obiettivo è spostare le infrastrutture sotto la gestione diretta di un PSN; nel secondo caso esse devono essere razionalizzate. Ciò implica la dismissione dei data center più obsoleti con la relativa migrazione dei servizi verso data center più affidabili oppure verso servizi cloud qualificati nell'ottica di un percorso, condiviso come indicato di seguito [59, pp. 37-38,43].

- Le PA proprietarie di data center nel Gruppo A:
  - continuano nella gestione e manutenzione, con comunicazione ad AgID delle eventuali spese necessarie
  - avviano piani di adeguamento per i livelli minimi di sicurezza e affidabilità dei data center previsti da AgID
  - possono proporre al PSN di utilizzare le proprie risorse per alcuni servizi e sono coinvolte per supportare PA con data center nel gruppo B.
- Le PA proprietarie di data center nel Gruppo B devono inviare ad AgID i piani di migrazione:
  - verso i servizi cloud qualificati o verso PSN, in caso di PAC (PA Centrali)
  - verso i servizi cloud qualificati o verso data center di gruppo A, in caso di PAL (PA Locali).

Da febbraio 2023 è operativa la piattaforma online *PA digitale 2026* da utilizzare per la trasmissione dei piani di migrazione redatti dalle PA e su cui è pubblicato il modello da utilizzare, predisposto da DTD e ACN.



Gli strumenti messi a disposizione delle PA per realizzare il processo di migrazione sono descritti nel PT [59, p. 39] e sono i seguenti.

- Dal punto di vista economico, il PNRR ha previsto finanziamenti per un totale di 1,9 miliardi di euro per l'obiettivo *cloud first*
  - *Investimento 1.1: Infrastrutture digitali* per PA Centrali, ASL e Aziende Ospedaliere
  - *Investimento 1.2: Abilitazione e facilitazione migrazione al cloud* per Comuni, Scuole, ASL e Aziende Ospedaliere.
- Dal punto di vista procedurale e tecnico, è stato realizzato il *Manuale di abilitazione al cloud*<sup>30</sup>, una raccolta di spunti e buone pratiche per fornire una metodologia di lavoro utile per pianificare e attuare la migrazione. Tale risorsa, parte del cosiddetto *Cloud Enablement Kit*, è frutto della collaborazione con alcune PA che hanno avuto il ruolo di apripista nell'ambito del Programma nazionale di abilitazione al cloud attraverso sperimentazioni e progetti pilota.
- Dal punto di vista degli approvvigionamenti, sono disponibili gli strumenti Consip<sup>31</sup> per la negoziazione telematica come il Mercato Elettronico (MEPA) e il Sistema Dinamico di Acquisizione della PA (SDAPA), oltre che le Gare strategiche ICT di Consip (es. Accordo Quadro Public Cloud). Queste opzioni consentono alle PA di limitare i tempi sia per gli acquisti di servizi public cloud IaaS e PaaS sia di servizi professionali in caso occorra acquisire sul mercato le competenze necessarie per le procedure di migrazione.

### 3.6 Il principio cloud first

Il *cloud first*, adottato anche in altri stati europei, tra cui Paesi Bassi [41, p. 30] e Regno Unito [38, p. 134], è una linea di indirizzo che specifica di optare per il cloud come approccio predefinito per investimenti e progetti ICT di nuova ideazione. All'interno della Strategia Cloud Italia l'obiettivo è che, a tendere, tutti i servizi erogati dalle PA siano basati su applicazioni *cloud-native*, sviluppate cioè nativamente sulla base dei paradigmi cloud. Questo non significa solo adottare uno dei modelli SaaS, PaaS, IaaS bensì anche costruire architetture software modulari e scalabili, basate su microservizi, API e container (si rimanda al paragrafo 2.1.3).

---

<sup>30</sup> Disponibile al link: <https://docs.italia.it/italia/manuale-di-abilitazione-al-cloud/manuale-di-abilitazione-al-cloud-docs/it/bozza/index.html>

<sup>31</sup> Concessionaria Servizi Informativi Pubblici: è la centrale degli acquisti della PA, società interamente partecipata dal Ministero dell'Economia e delle Finanze.

Un esempio di applicazione di *cloud first* è quello indicato nella prima versione del Piano Triennale 2017-2019<sup>32</sup>. In fase di sviluppo di una nuova applicazione, una PA può valutare in via prioritaria se è possibile utilizzare i servizi SaaS, verificando se nel *Cloud Marketplace* sono già presenti componenti adatti alle proprie esigenze, anche sulla base delle regole sul riuso del software nella PA. In secondo luogo possono essere prese in considerazione soluzioni PaaS e infine di tipo IaaS.

Secondo i documenti strategici, la realizzazione di queste macro-azioni permetterà di armonizzare e regolamentare l'adozione del cloud nella PA, nonché applicare economie di scala per favorire una riduzione dei costi di gestione offrendo servizi digitali più affidabili e resilienti.

### 3.7 Qualificazioni fornitori e servizi cloud

Per dotare le PA di informazioni certe e facilmente confrontabili sui servizi cloud, nell'aprile 2018 sono stati predisposti un iter di qualificazione e un catalogo dei servizi chiamato *Cloud Marketplace*, contenente il registro pubblico dei CSP e la lista dei servizi qualificati di tipo SaaS, PaaS e IaaS. Tali iniziative sono state avviate da AgID con due circolari<sup>33</sup>: la circolare n. 2/2018 *Criteri per la qualificazione dei Cloud Service Provider per la PA* che riporta i requisiti e la relativa procedura di qualificazione per i CSP e la circolare n. 3/2018 *Criteri per la qualificazione di servizi SaaS per il Cloud della PA* che delinea i requisiti e la relativa procedura di qualificazione dei servizi SaaS erogabili sul cloud della PA. Dal 1 aprile 2019 le PA possono acquisire esclusivamente servizi cloud qualificati.

Dall'analisi della procedura emerge tuttavia che i fornitori inviano una modulistica in cui descrivono il proprio servizio e sostanzialmente autocertificano di aderire ai requisiti richiesti, ma non sono effettuate verifiche puntuali da parte di AgID.

Lo scorso 19 gennaio 2023 è avvenuto il passaggio di competenze da AgID a ACN sulla qualificazione dei servizi cloud. Per non interrompere la continuità dei servizi certificati già in uso dalle PA è stato istituito un regime transitorio, inizialmente previsto fino al 31 luglio 2023 e recentemente prorogato al 31 gennaio 2024<sup>34</sup>. Fino a tale data rimangono valide le qualificazioni già in possesso dai fornitori, mentre a partire dal 1 febbraio 2024 le istanze dovranno essere nuovamente inoltrate con la procedura messa a disposizione da ACN. Il *Cloud Marketplace* è stato trasferito sul sito web di ACN al link: <https://catalogocloud.acn.gov.it>.

---

<sup>32</sup> Cloud Enablement e principio Cloud First, link: <https://docs.italia.it/italia/piano-triennale-ict/cloud-docs/it/stabile/cloud-enablement.html>

<sup>33</sup> Cloud della PA, <https://cloud-italia.readthedocs.io/projects/cloud-italia-circolari/it/latest>

<sup>34</sup> Qualificazione dei Servizi Cloud ACN, link <https://www.acn.gov.it/strategia/strategia-cloud-italia/qualificazione-cloud>

La novità più rilevante di questo cambio di competenza è che le richieste di qualificazione, in precedenza basate sull'autocertificazione, saranno soggette a una verifica da parte di ACN che ha facoltà di effettuare accertamenti tecnici, anche periodici, e sopralluoghi. I fornitori che non rispettano i nuovi requisiti possono incorrere nella sospensione o nella revoca della qualifica [63].

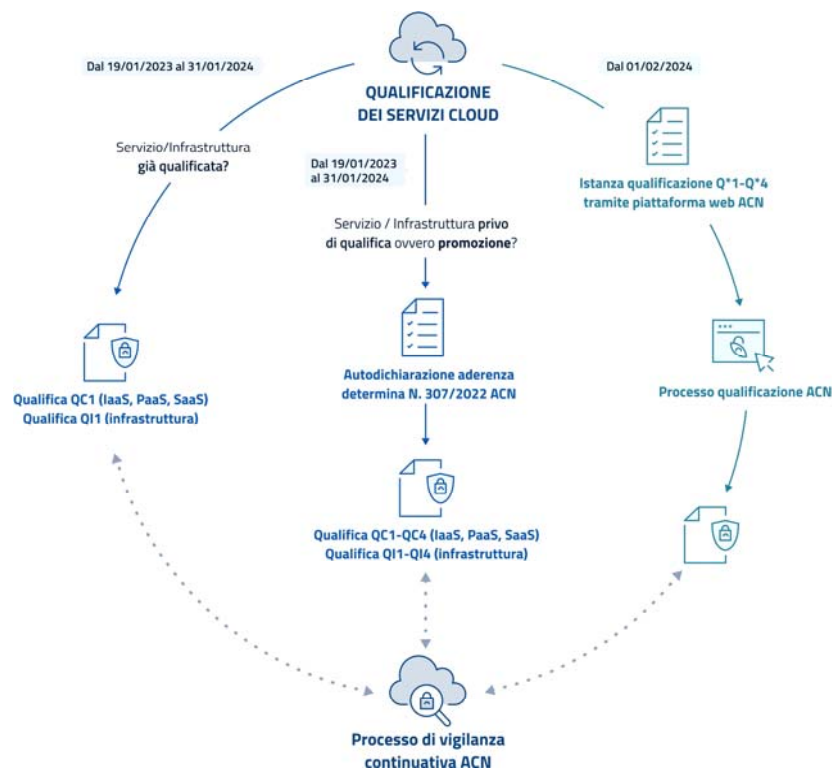


Figura 14. ACN: qualificazione dei servizi cloud

Fonte [63]

La Strategia Cloud Italia impone inoltre la classificazione dei dati e dei servizi da attuarsi ad opera di ACN con la collaborazione delle singole PA [62, pp. 13-15]. Sono state individuate tre categorie di dati, usando come criterio il danno che provocherebbe al Paese una loro eventuale compromissione:

- *strategici*: la compromissione può avere un impatto sulla sicurezza nazionale;
- *critici*: la compromissione può pregiudicare funzioni rilevanti per la società, la salute, la sicurezza e il benessere economico e sociale del Paese;
- *ordinari*: la compromissione non provoca un'interruzione di servizi dello Stato o un pregiudizio per il benessere economico e sociale del Paese.

Nel documento strategico trovano indicazione anche i parametri utilizzati per la qualificazione dei servizi cloud [62, p. 15]. Uno di questi è la gestione operativa, legata agli standard tecnico-organizzativi utilizzati e alle misure di controllo sui dati. A seguire vi sono i requisiti di sicurezza applicati, come ad esempio controlli di sicurezza adottati

e la modalità di gestione delle chiavi di cifratura. Infine sono indicate le condizioni contrattuali applicate al servizio, come SLA e rendicontazione.

Le due suddette classificazioni convergono infine nella catalogazione dei servizi cloud.

- *Cloud Pubblico non qualificato (extra UE/UE)*: servizi che non rispettano i criteri tecnico-organizzativi e normativi;
- *Cloud Pubblico qualificato (UE)*: servizi che rispondono ai requisiti di sicurezza tecnico-organizzativi e compatibili con le normative (ad esempio GDPR e NIS);
- *Cloud pubblico Criptato (IT)*: servizi che offrono maggior livello di controllo delle infrastrutture, come la verifica on-premise dei meccanismi di sicurezza;
- *Cloud privato e ibrido*: servizi che consentono la localizzazione dei dati in Italia e un maggior isolamento dalle regioni pubbliche dei principali CSP, distinte in:
  - *Cloud privato/ibrido “su licenza” (IT)*: soluzioni basate su tecnologia hyperscaler licenziata da uno o più CSP;
  - *Cloud Privato Qualificato (IT)*: soluzioni basate su tecnologie commerciali qualificate mediante procedure di scrutinio e certificazione tecnologica.

A ciascuna classe di dati corrisponde una diversa gestione. Le opzioni *Cloud Pubblico Qualificato* e *Pubblico Criptato* sono idonee per ospitare dati e servizi ordinari. Le strutture di *Cloud Pubblico Criptato*, *Privato/Ibrido “su licenza”* e *Privato Qualificato* possono essere usate per contenuti critici mentre le uniche opzioni qualificate per dati e servizi strategici sono *Cloud Privato/Ibrido “su licenza”* e *Privato Qualificato*.

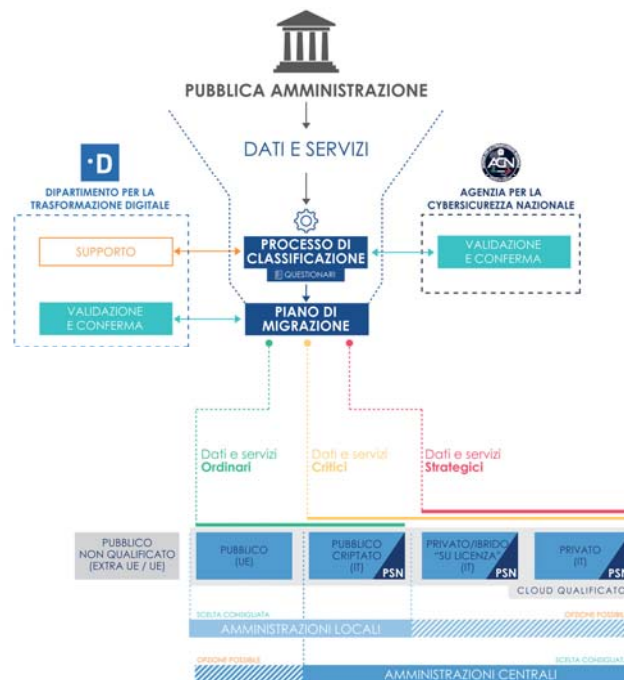


Figura 15. Strategia Cloud Italia: classificazione dati e servizi cloud.

Fonte [62]

### 3.8 Il Polo Strategico Nazionale

Un tassello fondamentale della Strategia Cloud Italia è la creazione di una nuova infrastruttura informatica sul territorio nazionale a servizio della PA, il cosiddetto *Polo Strategico Nazionale* (PSN) [62, p. 18]. Obiettivo del PSN è fornire alla PA tecnologie e infrastrutture cloud con elevati standard di affidabilità, resilienza e indipendenza.

Inizialmente, nella fase del censimento AgID, i data center con requisiti di eccellenza tecnica e organizzativa erano candidati ad essere utilizzati dal PSN, ma non era chiaro con quali forme di aggregazione. Le intenzioni parevano quelle di realizzare una federazione di data center esistenti, mantenendo operative le strutture classificate nel gruppo A e dismettendo quelle del gruppo B. La strategia è rimasta stabile con il susseguirsi dei diversi piani triennali finché nel 2020 si è verificato un cambio di rotta e le 62 strutture non in gruppo B (35 candidabili e 27 gruppo A) sono state ridefinite come data center di tipo A. A tali strutture è stato quindi riconosciuto il ruolo di supporto alle PAL mentre la PAC è stata orientata verso l'uso del PSN, creato ex-novo e non da una federazione di data center preesistenti.

La realizzazione del PSN è stata inserita nella Strategia Cloud Italia presentata a settembre 2021. Come riporta un approfondimento del DTD [64], il progetto ha visto il completamento in tempi relativamente brevi, circa 15 mesi, probabilmente anche grazie alle tempistiche stringenti imposte dal PNRR.

- A dicembre 2021 il DTD, tra le tre proposte ricevute, ha selezionato il progetto presentato dalla cordata TIM, Leonardo, Sogei, Cassa Depositi e Prestiti per la realizzazione e la gestione del PSN sul modello di partenariato pubblico-privato.
- La procedura è stata affidata a Difesa Servizi S.p.A., una società in-house del Ministero della Difesa, e nel gennaio 2022 è stato pubblicato il bando di gara europea con base d'asta pari a 4,4 miliardi di euro.
- A conclusione del procedimento, a luglio 2022 l'aggiudicazione è andata al gruppo che aveva presentato il progetto e che ha prevalso sull'altra cordata Fastweb-Aruba esercitando il diritto di prelazione.
- Il PSN è attivo da dicembre 2022 e viene gestito dalla società chiamata proprio *Polo Strategico Nazionale S.p.A.* e partecipata da TIM, Leonardo, Cassa Depositi e Prestiti e Sogei.

L'infrastruttura tecnologica è ospitata all'interno di 4 data center distribuiti sul territorio e distanti tra loro centinaia di chilometri. La configurazione prevede una doppia regione nord e sud, ciascuna formata da una coppia di data center in configurazione di *business continuity*, distanti tra loro alcune decine di chilometri. Le sedi prescelte sono Acilia e Pomezia nel Lazio, Rozzano e Santo Stefano Ticino in Lombardia. I due data center della stessa regione sono configurati in alta affidabilità e la distribuzione dei carichi di lavoro è effettuata in modalità trasparente.

In Europa un esempio simile è quello dei Paesi Bassi in cui il governo possiede un suo cloud privato, distribuito su 4 data center, nato però non ex-novo ma dal consolidamento di 66 data center tra il 2013 e il 2014 [41]. In riferimento al PNRR, la realizzazione del PSN ha permesso di completare la prima milestone della Missione 1, componente 1, investimento 1.1 Infrastrutture digitali [65, p. 92].

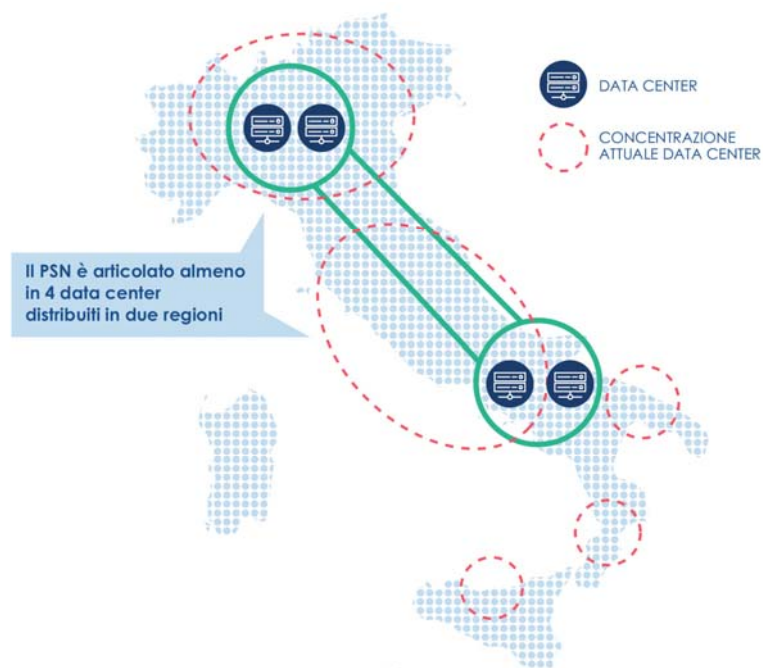


Figura 16. PSN: dislocazione sul territorio italiano

Fonte [62]

Il PSN può essere visto a tutti gli effetti come un nuovo fornitore per la PA. Seguendo la classificazione descritta al paragrafo 3.7, il PSN è progettato per offrire tutti i servizi cloud privato/ibrido (Cloud Privato/Ibrido “su licenza” e Cloud Privato Qualificato) nonché i servizi di Cloud Pubblico Criptato, consentendo l’uso di strumenti di cifratura on-premise. Per le PA il processo di adesione e l’attivazione del contratto sono su base volontaria. Ad oggi nell’ambito dell’iniziativa 1.2 del PNRR *Abilitazione e facilitazione migrazione al cloud*, sono stati pubblicati tre avvisi per la migrazione sul PSN di dati critici o strategici [66].

- *febbraio - maggio 2023*: destinatarie le PA Centrali, hanno aderito più di 40 amministrazioni centrali grazie a 157 milioni di euro del PNRR;
- *marzo - luglio 2023*: destinatarie le Aziende Sanitarie Locali (ASL) e Aziende Ospedaliere, hanno aderito più di 190 strutture per un investimento complessiva di 300 milioni di euro;
- *luglio - ottobre 2023*: destinatarie le PA Centrali, nuovo avviso da 280 milioni di euro, ancora in corso al momento della stesura del presente lavoro.

Il processo di adesione prevede che la PA invii un piano dei fabbisogni, a cui il PSN risponde con un'offerta con la descrizione dei servizi, un piano di migrazione di massima e la quotazione economica basata sul listino di gara. La PA può richiedere eventuali modifiche, il PSN provvede agli aggiornamenti e una volta che il progetto è approvato dalla PA, le due parti procedono alla stipula del contratto.

La milestone iniziale si prefigge la migrazione delle prime 100 amministrazioni entro settembre 2024 [66]. L'obiettivo della misura nel PNRR è di portare, entro il 2026, il 75% delle PA in Italia ad utilizzare servizi cloud [64].

### 3.9 Formazione e cambio culturale

Da quanto emerso finora è evidente come il processo di adozione del cloud nella PA sia sottoposto a maggiori vincoli rispetto al settore privato: è necessario creare una legislazione, definire degli standard e fornire linee guida sia tecnologiche sia logistiche. In aggiunta, la migrazione al cloud comporta anche un cambio organizzativo e culturale perché induce una reingegnerizzazione dei processi e una evoluzione delle logiche di lavoro in tutti i settori, non solo l'ambito IT. Criticità comuni sono le resistenze interne al cambiamento, la carenza di organico e la necessità di acquisire ruoli diversi e competenze professionali supplementari.

Il tema della formazione del personale interno è parte fondamentale del processo di trasformazione e proprio il PT ribadisce con chiarezza che *“gli obiettivi potranno essere raggiunti solo attraverso azioni di sensibilizzazione e di formazione che coinvolgano necessariamente i dipendenti della Pubblica Amministrazione”*. [59]

Si inserisce in questo ambito l'iniziativa *Syllabus per la formazione digitale*<sup>35</sup>, promossa a gennaio 2022 dal Ministero per la Pubblica Amministrazione e finanziata dal PNRR (Missione 1, componente 1, investimento 2.3 Competenze e capacità amministrativa [65, p. 97]). Al progetto hanno aderito oltre 2.000 enti<sup>36</sup> erogando ai propri dipendenti i percorsi formativi gratuiti, in modalità e-learning, strutturati ad hoc sulle competenze digitali.

Una ulteriore iniziativa intrapresa da AgID è rappresentata dall'iter di selezione di istituzioni universitarie e enti di ricerca con esperienza in ambito cloud computing per la formazione dei dipendenti pubblici. Lo scorso maggio 2023 è stato pubblicato l'elenco<sup>37</sup> delle organizzazioni accreditate: oltre al Politecnico di Torino - Dipartimento

---

<sup>35</sup> La piattaforma Syllabus, link: <https://www.funzionepubblica.gov.it/formazione/syllabus>

<sup>36</sup> Articolo al link: <https://www.funzionepubblica.gov.it/articolo/ministro/03-03-2022/oltre-2mila-amministrazioni-aderiscono-al-progetto-syllabus-la>

<sup>37</sup> AgID, elenco istituzioni per la formazione dei dipendenti pubblici sul cloud computing, link: <https://www.agid.gov.it/it/agenzia/stampa-e-comunicazione/notizie/2023/05/16/pubblicato-lelenco-istituzioni-formazione-dipendenti-pubblici-sul-cloud-computing>

di Automatica ed Informatica, sono stati qualificati il Politecnico di Milano e l'Istituto TeCIP (Istituto di Telecomunicazioni, Informatica e Fotonica).

A settembre 2023 sono partiti i corsi di formazione<sup>38</sup> rivolti ai dipendenti pubblici, realizzati da AgID in collaborazione con il Politecnico di Milano. Anche in questo caso i percorsi formativi sono a partecipazione gratuita e fruibili a distanza, sia in diretta streaming sia *on-demand*. L'obiettivo è quello di coinvolgere nell'implementazione del cloud tutte le figure professionali interessate direttamente e indirettamente, realizzando un percorso di evoluzione quanto più ampio e condiviso.

### 3.10 La spesa ICT nella PA

La spesa ICT nelle PA viene periodicamente monitorata da AgID. L'ultimo rapporto sulla rilevazione 2022 ha coinvolto un campione di 77 amministrazioni centrali (PAC), regionali e locali (PAL) [67]. Dall'analisi dei dati emerge che nel 2022 la PA italiana ha speso oltre 7 miliardi di euro in ICT e nel complesso le organizzazioni proseguono nel percorso di innovazione digitale. Per il periodo 2022-2024 vi è una tendenza della spesa in crescita (+5,2%), incremento in cui convergono sia gli investimenti residuali in risposta alla pandemia Covid-19, sia in larga parte l'utilizzo dei fondi del PNRR.

I servizi cloud sono pervasivi all'interno del campione: ne fanno uso 73 enti su 77 e quelli che non lo utilizzano hanno già piani per l'adozione futura. In particolare i tassi di diffusione sono i seguenti:

- PAL (28 Enti): 100%
- Regioni e Province autonome (20 Enti): 95%
- PAC (25 enti): 89%

Nel 2022 le spese sostenute in ambito cloud ammontano a 146 milioni di euro, in crescita rispetto al 2021 (+29%). Esse sono legate in particolare all'uso di servizi IaaS (41%), seguiti da servizi SaaS (37%) e infine PaaS (22%), come riportato in Figura 17.



Figura 17. La spesa ICT nella PA: spesa totale cloud

Fonte [67]

<sup>38</sup> Cloud PA, link: <https://www.agid.gov.it/it/infrastrutture/cloud-pa/formazione>



Le istituzioni rilevate usano in prevalenza servizi di tipo IaaS, in particolare servizi di archiviazione e/o backup, web hosting per i siti web e hosting per i server. In seconda battuta vi sono i servizi SaaS, impiegati principalmente per le applicazioni ritenute meno strategiche, come ad esempio gestione mense, gestione del verde e delle potature. Sono diffusi in maniera minore i servizi di tipo PaaS, probabilmente in quanto le attività di sviluppo software non sono molto presenti nelle PA, ad eccezione delle PAC e delle Regioni e Province autonome che hanno contratti con società in-house per l'affidamento di tali attività in outsourcing (Figura 18).

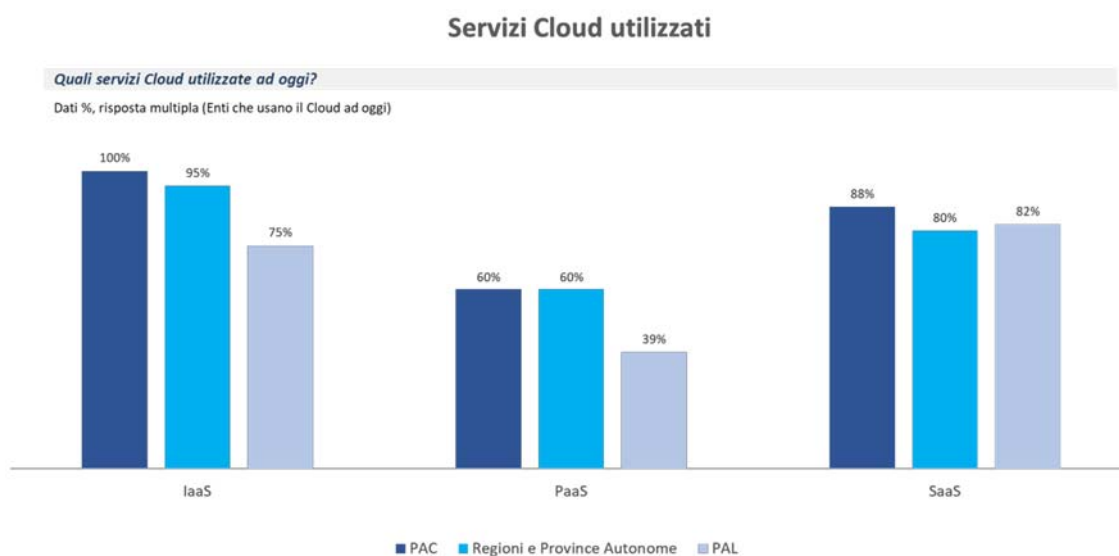


Figura 18. La spesa ICT nella PA: i servizi cloud utilizzati

Fonte [67]

### 3.11 Cronologia norme e risultati

In conclusione si riepiloga nella Tabella 6 la cronologia con i diversi passaggi normativi, gli eventi significativi e i risultati della PA italiana nell'adozione del cloud.

Tabella 6. Cloud computing nella PA italiana: cronologia norme e risultati

Data/periodo	Descrizione eventi
agosto 2012	Legge 134/2012: istituzione dell'Agenzia per l'Italia Digitale (AgID)
maggio 2017	Pubblicazione primo Piano Triennale per l'informatica nella PA 2017-2019
aprile 2018	Avvio qualificazioni e realizzazione registro pubblico dei CSP Realizzazione marketplace della PA, la piattaforma con il catalogo dei servizi cloud qualificati da AgID
giugno 2018	Pubblicazione portale <i>cloud.italia.it</i> per la strategia cloud nella PA
2018-2019	AgID: censimento patrimonio ICT della PA
gennaio 2020	Pubblicazione nuova piattaforma Marketplace dei servizi qualificati
luglio 2020	Approvazione del Piano Triennale per l'informatica nella PA 2020-2022
settembre 2020	Legge 120/2020: nuove regole e strumenti per i servizi PA Obbligo migrazione verso infrastrutture ad alta affidabilità o cloud
giugno 2021	DL 82/2021: istituzione dell'Agenzia per la cybersicurezza nazionale (ACN)
luglio 2021	Approvazione del Piano Nazionale di Ripresa e Resilienza (PNRR)
settembre 2021	Pubblicazione Strategia Cloud Italia
dicembre 2021	Selezione del progetto PSN e avvio gara europea per la realizzazione
aprile 2022	ACN: Classificazione dati e servizi per avviare la migrazione cloud PA
maggio 2022	ACN: Pubblicazione Strategia Nazionale di Cybersicurezza 2022-2026
giugno 2022	Conclusione gara per il PSN
dicembre 2022	Attivazione PSN ridonato su 4 sedi, 2 in Lazio e 2 in Lombardia
dicembre 2022	Approvazione del Piano Triennale per l'informatica nella PA 2022-2024
gennaio 2023	Trasferimento competenza da AgID a ACN Regime transitorio in attesa delle nuove procedure di qualificazione per fornitori e servizi
gennaio 2023	PA: trasmissione piani di migrazione sulla piattaforma PA digitale 2026
maggio 2023	Pubblicazione elenco di istituzioni per la formazione dei dipendenti pubblici sul cloud computing: presente anche il Politecnico di Torino
maggio 2023	Adesione di oltre 40 PA centrali al progetto per migrare dati e servizi in cloud, a partire da quelli critici e strategici, verso il PSN
luglio 2023	ACN: proroga regime transitorio, a gennaio 2024 termine per qualificazione dei servizi e razionalizzazione dei livelli minimi per strutture e servizi cloud
gennaio 2024	PA: scadenza presentazione relazione di conformità e adozione dei requisiti, possibilità di estensione a ottobre 2024 per progetti complessi già avviati

### 3.12 Ricerca e università

All'interno delle organizzazioni R&E<sup>39</sup>, le università e gli enti di ricerca pubblici sono a tutti gli effetti delle PA e rientrano pertanto nel campo di applicazione delle regole fin qui descritte. È già stata introdotta al paragrafo 3.5 la circolare AgID 1/2019 [60] che delinea il processo di razionalizzazione delle infrastrutture e impone blocchi agli investimenti per i data center. Nel testo di tale circolare è prevista un'eccezione per le attività di ricerca:

*sono esclusi dalla richiesta di approvazione gli adeguamenti che prevedono acquisti nei seguenti ambiti: progetti di ricerca a titolarità di istituzioni universitarie e/o enti di ricerca; sistemi a supporto della diagnostica clinica.*

All'interno di un ateneo, pertanto, tre delle quattro attività principali (si faccia riferimento al paragrafo 2.2.6) rimangono soggette al divieto degli investimenti per i CED e all'obbligo di spostare i servizi in data center affidabili o in cloud: la didattica, il trasferimento tecnologico e tutte le attività amministrative a sostegno delle altre funzioni. Solo alla ricerca viene garantita maggiore flessibilità; se sicuramente questo ne tutela l'indipendenza, la disparità di trattamento provoca una asimmetria che si riflette nei piani di evoluzione e nelle relative scelte tecnologiche.

Probabilmente l'impianto degli interventi di razionalizzazione AgID è stato concepito avendo come destinatari altri tipi di PA, più standardizzate e con meno peculiarità rispetto al mondo accademico e della ricerca, che quindi appare rappresentato solo in parte. Dopo l'emissione della circolare AgID 1/2019, un'iniziativa congiunta della *Conferenza dei rettori delle università italiane* (CRUI) e del *Convegno dei Direttori generali delle Amministrazioni Universitarie* (CODAU) ha portato all'attenzione di AgID le diverse perplessità sull'attuazione delle norme, da parte del mondo accademico e del gruppo di coordinamento dei *Responsabili per la Transizione al Digitale* (RTD) degli atenei [68], [69].

Il punto di vista esposto dai gruppi di studio è che in generale, pur comprendendo le motivazioni, la maggior parte degli atenei si troverebbe davanti a onerosi interventi di adeguamento per soddisfare dei requisiti eccessivamente stringenti. Ne vengono elencati alcuni, ad esempio che i data center siano gestiti con turni operativi 24/7/365, abbiano indice di disponibilità almeno pari al 99,6% e che siano presenti certificazioni ISO 27001 o procedure per la sicurezza IT come ISO 27002. La richiesta è quindi quella di lasciare agli atenei una maggiore flessibilità, eventualmente valutando gli scenari caso per caso, con la possibilità di mantenere on-premise servizi infrastrutturali già disponibili e acquistare applicazioni non incluse nei servizi SaaS dal marketplace AgID.

---

<sup>39</sup> Research and Education

Alcuni aspetti critici sono legati alla specificità di molti servizi propri del contesto universitario, utilizzati sia in ambito didattico sia di ricerca, talvolta con soluzioni personalizzate implementate on-premise e di difficile (o non conveniente) migrazione verso il cloud. Tra questi vi sono la gestione PC di aule e di laboratori, i sistemi di e-learning e di supporto alla didattica, la gestione delle licenze di software complessi e legati ad attrezzature scientifiche. Altri servizi infrastrutturali di base (sistemi di controllo e gestione impianti, rete e fonia, monitoraggio, autenticazione, controllo accessi), anche se non specifici del quadro universitario, hanno tuttavia un grado di integrazione molto elevato. Un'altra peculiarità è sul fronte del software utilizzato: talvolta sono necessari acquisti, anche in tempi stretti, di applicazioni di nicchia non elencate nel marketplace di AgID, prodotte da aziende estere e spesso installabili solo su sistemi *on-premise*, dotati di chiavi hardware di licenza o che necessitano di server locali di licenza.

Come è descritto da un articolo a firma GARR<sup>40</sup> nel luglio 2019 è stata siglata un'intesa per la collaborazione tra CRUI e AgID, concordando che in materia di data center saranno gli atenei a proporre il proprio piano di migrazione, in accordo con AgID, e tenendo conto delle peculiarità del settore. Per ciò che concerne più in generale la transizione al cloud, Attardi [70] (2017) in un articolo su *Agenda Digitale* espone come gli enti di ricerca quali ENEA, INFN e CNR siano propensi all'adozione marginale di modelli Public cloud, ad esempio per far fronte a picchi della domanda di calcolo, ma tendano verso soluzioni alternative per una maggiore sicurezza dei dati e dei risultati scientifici. Le università hanno posizioni più varie e molte, accettando i rischi di lock-in, utilizzano le offerte di cloud pubblico per i servizi legati alla didattica se non per quelli della ricerca. Negli anni si è fatta strada l'idea delle organizzazioni R&E di fare sistema e affrontare il cloud condividendo risorse di infrastruttura e di competenze.

Nel capitolo che segue si analizzerà una soluzione federata come approccio al cloud computing per la ricerca. Si tratta di una infrastruttura progettata e messa in opera da GARR, il consorzio che gestisce la rete nazionale a banda ultralarga dedicata alla comunità dell'istruzione e della ricerca e i servizi innovativi per le attività correlate.

---

<sup>40</sup> Articolo al link: <https://www.garrnews.it/rubriche-interne-21/nuvola-di-ricerca-istruzione-21/717-data-centre-d-eccezione>

## Capitolo 4.

# Il modello federato GARR e la piattaforma OpenStack

Le università e gli enti di ricerca pubblici, in deroga alle disposizioni AgID, hanno la facoltà di mantenere e far evolvere i propri data center on-premise. Tuttavia il cloud ha da tempo attirato l'interesse del mondo scientifico sia perché non tutte le organizzazioni R&E hanno fondi e competenze per gestire dei propri data center sia perché alcune caratteristiche del modello cloud sono particolarmente adatte a svolgere attività computazionali che richiedono un numero di risorse storage e di elaborazione sempre crescente. Scrive Attardi [70]:

*Si tratta quindi non tanto di decidere se utilizzare il cloud o meno, ma piuttosto di scegliere la strategia migliore per adottarlo e come prepararsi a questa transizione. Dal punto di vista della comunità scientifica e accademica, si tratta di capire come cavalcare l'onda del cloud senza restarne travolti. Esiste un modo per passare al cloud da protagonisti anziché ridursi a semplici clienti delle grandi cloud pubbliche? È possibile restare padroni dei propri dati senza fare rinunce in termini di efficienza?*

L'ampia comunità R&E è formata da membri multidisciplinari e distribuiti a livello globale, ma esigenze e preoccupazioni sono largamente condivise, prime tra tutte l'indipendenza della ricerca e la sovranità digitale. Molte reti NREN<sup>41</sup>, le reti europee di ricerca tra cui GARR, si stanno da tempo muovendo per valutare soluzioni cloud che possano armonizzare le necessità crescenti mantenendo alti standard tecnologici.

---

<sup>41</sup> *National research and education networks*: forniscono connettività di rete di alta qualità e servizi correlati alle università e agli istituti di ricerca del proprio paese, collegando campus e istituzioni tra loro e al resto di Internet (fonte: <https://about.geant.org/nrens/>).

## 4.1 Il modello Cloud GARR

Questa sezione analizza il modello cloud GARR, descritto sul sito <https://cloud.garr.it>, con riferimenti al contesto europeo, alle scelte architettoniche e dell'offerta dei servizi.

### 4.1.1 Il Consortium GARR

Il Consortium GARR (Gruppo per l'Armonizzazione delle Reti per la Ricerca) è un ente senza fini di lucro fondato e finanziato dai principali centri di ricerca e dalle università italiane, rappresentate dalla Fondazione CRUI. GARR possiede e gestisce la rete nazionale a banda ultra larga riservata alla comunità dell'istruzione e della ricerca: conta 4,5 milioni di docenti, ricercatori e studenti serviti con 20.000km di fibra ottica con dorsale recentemente migrata a 20 Tbps. La rete GARR è direttamente connessa a GÉANT<sup>42</sup>, a sua volta formata dalla collaborazione di tutte le NREN.

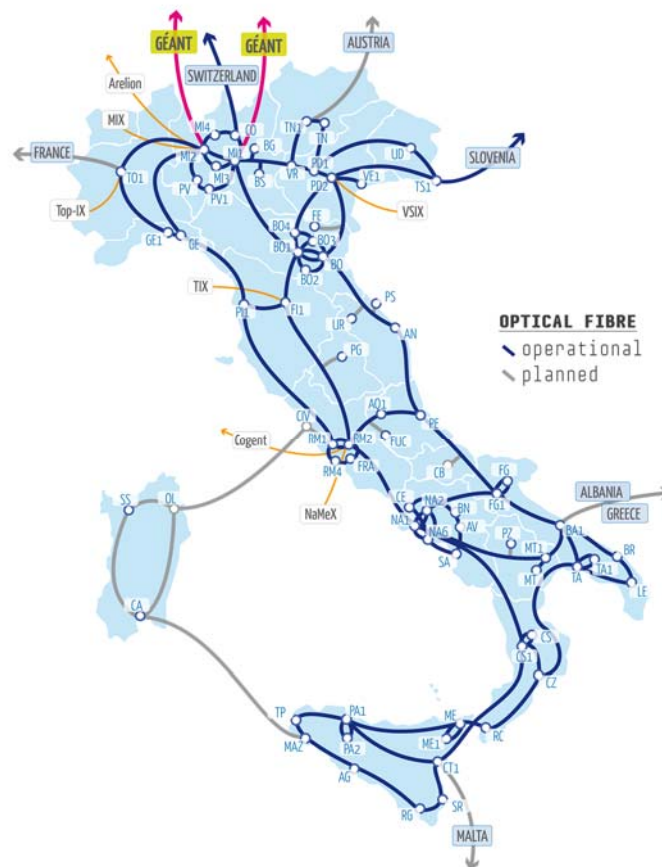


Figura 19. La mappa della rete GARR

Fonte <https://www.garr.it/it/infrastrutture/rete-nazionale>

<sup>42</sup> Sito web istituzionale: <https://geant.org/>

#### 4.1.2 Il Community cloud nel contesto europeo delle NREN

Le direttive della strategia digitale della Commissione europea, come già introdotto nel capitolo 2, riconoscono il valore del cloud quale tecnologia abilitante e sono focalizzate sui temi della sovranità digitale e sulla standardizzazione e l'interoperabilità dei servizi.

Frutto di questo contesto è la costituzione all'interno di GÉANT del gruppo di lavoro SIG-CISS<sup>43</sup> (*Cloudy Interoperable Software Stacks*) che dal 2017 riunisce le persone che nel proprio paese si occupano di progettare o gestire cloud di ricerca e sono accomunate dalla volontà di condividere conoscenze, strategie, buone pratiche nel ciclo di vita delle infrastrutture cloud. Si inseriscono in questo scenario le attività relative alla creazione e alla gestione di Community cloud portate avanti da diverse NREN come servizio per i propri utenti [61]. Oltre all'infrastruttura GARR descritta in seguito è interessante una breve panoramica per comprendere lo stato dell'arte.

- *SURF, Paesi Bassi*: finanziato dal governo olandese, dai membri di SURF e da varie collaborazioni, fornisce servizi cloud basati su OpenStack, Ceph, Swift, Owncloud e Kubernetes, utilizzabili dal portale Research Cloud.
- *RENATER, Francia*: fornisce l'interconnessione di infrastrutture informatiche e servizi distribuiti (PaaS e SaaS) per la comunità R&E accessibili dalla federazione di identità a livello nazionale e da eduGain a livello internazionale.
- *GRNET, Grecia*: è il principale abilitatore di infrastrutture/servizi per Open Science in Grecia e dal 2010 gestisce una grande infrastruttura IaaS OpenStack distribuita in grandi data center (oltre 1800 server e 7000 VM).
- *CESNET, Repubblica Ceca*: dal 2013 propone un cloud IaaS e PaaS, inizialmente gestito con OpenNebula e più recentemente con OpenStack, integrato con alcune federazioni di identità locali e internazionali.
- *CSC/Funet, Finlandia*: finanziato dal Ministero della Pubblica Istruzione e da varie collaborazioni, fornisce servizi cloud utilizzando OpenStack e OpenShift, oltre a fare da broker di servizi cloud commerciali verso le università finlandesi.
- *KIFU, Ungheria*: ha iniziato a proporre servizi cloud nel 2011 con la prima generazione, su una piattaforma Open Nebula. La seconda generazione di Community Cloud (2015) e la terza (2021) sono basate su OpenStack. I servizi offerti sono IaaS e SaaS containerizzato su VM.
- *PSNC, Polonia*: offre un ampio catalogo di servizi: oltre a IaaS, PaaS, data center e storage distribuito, anche servizi di gestione dei dati come piattaforme di Data-Science aaS, basi di dati scalabili aaS, BigData aaS.

La diffusione delle iniziative di Community Cloud a livello internazionale manifesta la forte intenzione dei membri della comunità R&E di non restare semplici spettatori o utilizzatori della tecnologia ma di voler costruire un ruolo attivo nei processi di

---

<sup>43</sup> Pagina wiki del gruppo, link: <https://wiki.geant.org/display/CISS/>

abilitazione al cloud [60]. GARR, direttamente connesso a GÉANT, appartiene alle reti NREN e apporta un contributo al gruppo SIG-CISS con la presenza costante di uno dei suoi specialisti cloud.

#### 4.1.3 Il progetto Cloud GARR

Nel capitolo 2 si sono trattati i modelli di deployment e le loro caratteristiche. Tra i due scenari opposti di Public cloud e Private cloud si collocano l'approccio ibrido e il Community cloud. In particolare quest'ultimo consente l'accesso alle risorse a un gruppo ristretto di organizzazioni o utenti che appartengono a una comunità, sulla base di infrastrutture on-premise, ad esempio collocate presso la sede di uno o più membri, oppure off-premise, presso fornitori esterni.

Secondo Attardi et al. [60], [62] nell'ambito R&E il Community cloud può risultare vincente: nella sua proposta introduce un approccio federato come alternativa a un approccio centrale, con l'idea di costruire dal basso un'infrastruttura che riesca a scalare adeguatamente, formata dalla combinazione di altre strutture di grandi/medie dimensioni sul territorio.

Il progetto GARR Cloud afferisce al Dipartimento Calcolo e Storage Distribuito (CSD), nato nel 2014 con due principali obiettivi. Da un lato c'è l'intenzione di proporsi come provider di risorse per la comunità R&E, in particolare per la cosiddetta *long tail of science*, ossia i gruppi di ricerca di piccole dimensioni che non hanno sufficienti risorse per creare e gestire in autonomia un'infrastruttura di calcolo. Dall'altro lato c'è il proposito ambizioso di diventare un aggregatore di risorse, proporre un'architettura di riferimento e creare una piattaforma cloud a cui altre organizzazioni possano unirsi in un modello federato [63]. Sono quattro i concetti cardine su cui si fonda la realizzazione di un Community Cloud nel progetto di GARR [62], [63]:

4. l'adozione di piattaforme software robuste e mature, in grado di garantire prestazioni comparabili ai servizi di cloud pubblici;
5. l'impiego di soluzioni open source, sviluppate attraverso una collaborazione aperta, per evitare vincoli commerciali e di licenza;
6. l'uso di strumenti di automazione e monitoraggio per minimizzare i costi di gestione del sistema complesso, lato hardware e lato software;
7. l'enfasi sulla collaborazione e la condivisione delle risorse infrastrutturali.

Secondo Al-Baroodi et al. [6] fattori significativi nel vaglio di una soluzione *Open Source Cloud Computing* (OSCC) sono la robustezza, la resilienza in caso di malfunzionamenti, l'interoperabilità e la flessibilità d'uso in diversi domini applicativi, la varietà di hypervisor supportati e infine la possibilità di creare VM indipendenti e compatibili con altre piattaforme. Nel paragrafo successivo si descrivono le scelte architetturali di GARR.



#### 4.1.4 Architettura di riferimento

Dal punto di vista delle risorse fisiche, dopo l'aggiornamento hardware del 2022 il cloud GARR conta ad oggi un totale di 7.600 core CPU, 60TB di RAM, 13,5PB di storage di cui il 15% SSD, schede GPU per complessivi 333 TFLOPS. Ai quattro data center presenti nel progetto originario del 2014 (Catania, Cosenza, Napoli, Palermo) se ne è aggiunto un quinto presso l'Università di Torino [64]. Per l'implementazione software, GARR ha scelto la piattaforma IaaS OpenStack, uno dei prodotti open source più stabili, e in particolare il progetto *Canonical OpenStack Charms* [63], [64], [65].

Questa sezione fornisce una breve panoramica dell'architettura di riferimento, che viene approfondita nei paragrafi successivi (4.2 e 4.3).

OpenStack è una soluzione cloud formata da diversi moduli che collaborano per gestire le risorse data center in una sorta di astrazione per le risorse computazionali (RAM e CPU), di archiviazione (dischi e storage) e di rete (NIC, switch e router). L'obiettivo è la replica software di un'infrastruttura IT con un livello di tipo IaaS.

Una piattaforma di cloud computing è formata, lato hardware, da un numero elevato di server e storage e, lato software, dall'interconnessione di molteplici componenti: per la gestione efficace di tale complessità è necessario utilizzare strumenti di automazione nell'intero ciclo di vita dei servizi. A tale scopo, il progetto OpenStack Charms propone MAAS (Metal-As-A-Service) e Juju<sup>44</sup>, il primo per installare e gestire il parco hardware e il secondo per l'orchestrazione e la distribuzione dei singoli servizi sui nodi del cluster.

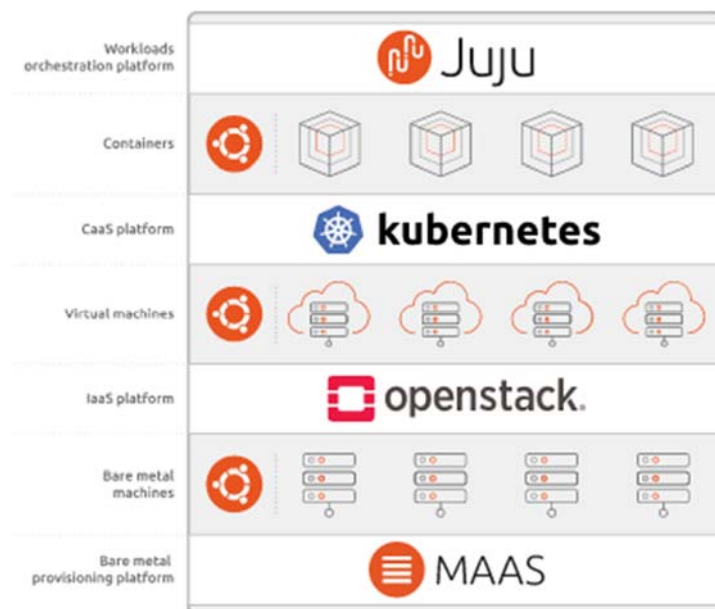


Figura 20. OpenStack, MAAS e Juju: schema generale

Fonte <https://ubuntu.com/openstack>

<sup>44</sup> Progetti open source curati da Canonical, link <https://maas.io/> e <https://jaas.ai/>

### *Gestione hardware fisico*

MAAS consente di gestire l'infrastruttura fisica con un approccio più standardizzato rispetto all'impiego di singoli script personalizzati per l'installazione. In particolare permette di installare in batteria i server fisici, anche geograficamente distribuiti, mantenendo una configurazione uniforme grazie alla possibilità di definire dei profili (ad esempio nodi controller, compute, network).

MAAS presenta un'architettura modulare che lo rende particolarmente adatto agli ambienti distribuiti e multi-site, come il cloud GARR [63]. Un *region controller* ospita il catalogo dei server e la console di gestione web, mentre in ogni data center si può collocare un numero variabile di *rack controller* per la gestione dell'infrastruttura locale.

Quando si istanziano i server fisici, MAAS coordina le attività di *commissioning*: effettua il rilevamento dell'hardware, fornisce le informazioni di rete e le configurazioni necessarie, seleziona un'immagine del sistema operativo dal proprio catalogo e provvede all'installazione. Infine MAAS si occupa anche del *decommissioning* ovvero di rilasciare i server quando non sono più necessari e renderli disponibili per nuovi carichi di lavoro.

### *Distribuzione applicazioni*

Juju è uno strumento per la modellazione di applicazioni e un orchestratore di risorse cloud. Può essere usato per distribuire, configurare, gestire e scalare applicazioni cloud su diversi tipi di piattaforme: su un catalogo di server fisici (in collaborazione con MAAS), su container LXD o su soluzioni cloud IaaS commerciali o private (tra cui OpenStack, Google Cloud Platform GCP, Amazon Web Services AWS). I principali punti di forza sono:

- focus sulle applicazioni e le relazioni tra loro;
- possibilità di schemi ripetibili per le fasi del ciclo di vita delle applicazioni;
- focus sui modelli di distribuzione delle applicazioni, gestione e scalabilità.

Juju utilizza il concetto dei modelli per rappresentare l'architettura dei servizi e le relative interdipendenze. Per l'installazione delle applicazioni vengono utilizzati file di script specifici chiamati *charm*, messi a disposizione attraverso il repository CharmHub<sup>45</sup> e sviluppati sia da Canonical sia dalla comunità open source. Le funzionalità di Juju consentono la scalabilità delle applicazioni e la gestione dell'intero ciclo di vita: installazione, configurazione, aggiornamento e manutenzione. L'aggiunta o la rimozione dinamica di unità può essere usata per gestire eventuali picchi di carico, oppure ripristinare configurazioni in HA e servizi in caso di problemi hardware.

Nel progetto OpenStack Charms i singoli componenti OpenStack possono essere distribuiti via Juju sia direttamente sulle macchine fisiche sia in container LXD. Come

---

<sup>45</sup> Link al repository ufficiale: <https://charmhub.io/>

nel caso del cloud GARR, combinando MAAS come fornitore di server bare-metal e Juju come orchestratore per i servizi (su host fisici, container LXD e VM), è possibile generare sia servizi infrastrutturali OpenStack su larga scala sia applicazioni cloud native [62], [63], [64].

#### *Gestione storage*

La virtualizzazione e la gestione delle risorse storage sono basate su Ceph<sup>46</sup>, progetto open source caratterizzato da un'architettura distribuita. Le caratteristiche che hanno influito sulla scelta di GARR sono la versatilità, con diverse possibilità di provisioning dello storage, la scalabilità, la struttura dotata di componenti ridondati e l'assenza di *single point of failure* e infine il supporto di una comunità di sviluppatori piuttosto ampia e reattiva [63], [66].

Ceph può essere gestito utilizzando MAAS e Juju, ma il primo cluster GARR è stato installato con Ceph-Ansible, essendo stato creato prima della selezione e definizione degli strumenti di automazione cloud. Tali strumenti sono stati poi impiegati per il deploy dei cluster Ceph successivi e per quelli di test. Nell'infrastruttura cloud GARR, i servizi Ceph sono eseguiti in container, tranne quelli bare-metal sui server su cui sono fisicamente presenti i dischi. Attualmente Ceph è utilizzato per erogare due servizi:

- un servizio di *block storage*, per fornire i volumi utilizzati dalle VM OpenStack;
- un servizio di *object storage*, che consente agli utenti di caricare oggetti, che vengono replicati sul Cloud storage GARR.

#### **4.1.5 I servizi offerti**

Gli utenti finali del cloud GARR sono sia gli enti sia le figure professionali come gli specialisti IT e i ricercatori. Sono tre le categorie di servizi offerti [64]:

- 1) *IaaS tradizionale*: erogazione sia di VM in diversi tagli e versioni sia di virtual data center, costituiti da un insieme di risorse che dagli amministratori possono raggruppare in progetti e assegnare agli utenti;
- 2) *PaaS e DaaS*: include una piattaforma container basata su Kubernetes e un livello di servizio aggiuntivo *Deployment as a Service* (DaaS) che combina le funzionalità di IaaS e PaaS, usato per distribuire applicazioni specifiche;
- 3) *SaaS*: applicazioni cloud fruibili da un catalogo in modalità *self-provisioning*. Tra queste vi sono *Workplace*, una suite collaborativa open source basata su Onlyoffice per la gestione di documenti e la piattaforma *GARR.tv*, il servizio di streaming di contenuti audio/video prodotti dalla comunità [63].

---

<sup>46</sup> Link al sito del progetto: <https://ceph.io>

Attraverso la piattaforma DaaS è possibile effettuare il deploy di applicazioni cloud, disponibili nel catalogo, con l'ausilio di Juju. Al momento sono disponibili oltre 200 applicazioni in diversi domini (e-learning, CMS, database, analisi Big Data, sviluppo web) che possono essere distribuite attraverso una interfaccia di gestione web [62].

La Figura 21 mostra alcuni dati sull'utilizzo del cloud GARR. È interessante notare che GARR è allo stesso tempo sia amministratore sia utilizzatore del proprio sistema cloud, in quanto tutti gli strumenti di automazione utilizzati per la distribuzione e la gestione dei servizi si appoggiano sul cloud stesso.

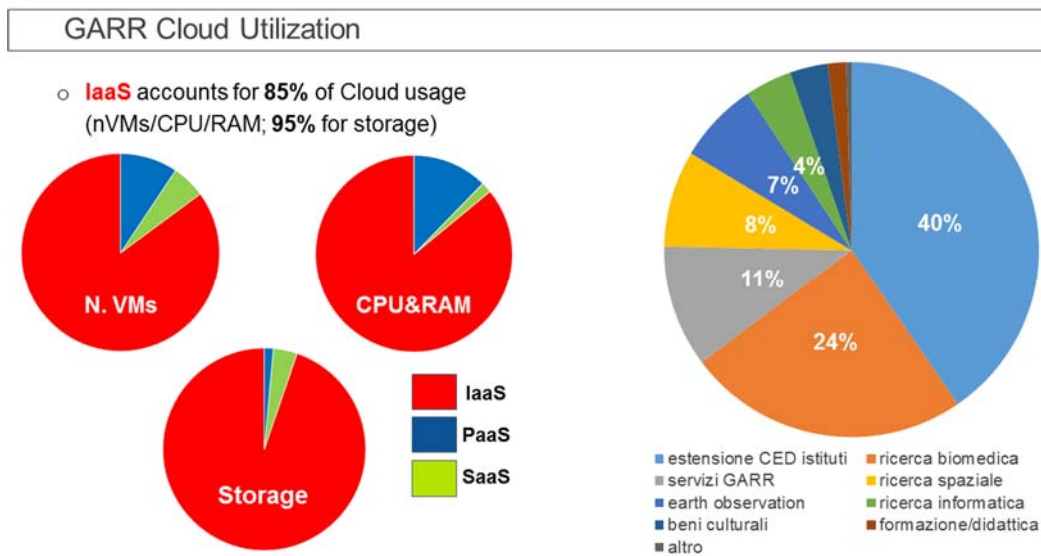


Figura 21. Cloud GARR: dati sull'utilizzo.

Fonte [64]

Come emerge dal grafico in Figura 21, il servizio più richiesto è quello IaaS, che utilizza l'85% delle risorse fisiche CPU e RAM e il 95% dello storage. Quest'ultimo (grafico a torta di destra) per il 40% è impiegato per l'estensione del cloud GARR nei data center degli enti associati (servizi come server DNS secondari, siti istituzionali, backup e disaster recovery). Per la restante quota, lo storage è usato per le attività di ricerca, tra cui spicca il 24% per il settore biomedico. Rilevante citare in questo contesto che il Ministero della Salute è stata la prima organizzazione a stipulare una convenzione con il GARR per la ricerca, in particolare per la fornitura di risorse agli Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS) [64].

#### 4.1.6 I vantaggi del Community cloud

Nel loro recente lavoro, Ardizzone et al. [63] (2022) descrivono le motivazioni e i vantaggi del cloud federato GARR. Molti temi sono condivisi anche dal documento di visione strategica [61] del gruppo di lavoro SIG-CISS di GÉANT. Il primo focus è sui dati: le infrastrutture di storage e calcolo sono situate nel paese e in data center gestiti direttamente; l'intero ciclo di vita dei dati è sotto il controllo della comunità che mantiene la sovranità sia dei dati sia del software e può garantire il rispetto dei vincoli legislativi quali GDPR e NIS.

Dal punto di vista del networking, i servizi sfruttano le reti ad alta velocità fornite dalle NREN. Inoltre, rimanendo sulla rete R&E si limitano gli accordi di peering<sup>47</sup> ad alta capacità con fornitori commerciali. Le offerte di cloud commerciali sono solitamente precostruite e molto rigide, eventuali personalizzazioni incrementano enormemente i costi; al contrario nel Community Cloud si ha la flessibilità di un progetto su misura che serve obiettivi comuni e specifici. Inoltre adottando l'open source gli strumenti software sono liberi da licenze o vincoli commerciali e si possono plasmare in base alle esigenze della comunità, riducendo il rischio di *vendor lock-in*.

Infine il Community cloud consente di creare e mantenere le competenze all'interno della comunità, favorendo lo scambio e la collaborazione a partire dalla progettazione, all'implementazione, ottimizzazione e integrazione dell'intero ambiente cloud. Come sostiene Attardi [60], se per i grandi fornitori di servizi cloud e gli hyperscaler, in perenne concorrenza tra loro, la condivisione non è prevista dal modello di business, per le comunità di ricerca essa invece può rappresentare una interessante ricchezza.

---

<sup>47</sup> Un accordo tra due o più reti per realizzare un'interconnessione fisica e scambiare informazioni sulle reti raggiungibili attraverso i protocolli di routing.

## 4.2 Approfondimento sulle soluzioni software

In questa sezione vengono analizzate le soluzioni software open source che, oltre a OpenStack, corredano l'architettura GARR cloud. Si approfondiscono i due strumenti di automazione Juju e MAAS, entrambi di progetti di Canonical, e la virtualizzazione lato storage con Ceph. Una sezione a parte sarà dedicata a OpenStack (si veda 0).

### 4.2.1 MAAS

MAAS è una piattaforma progettata per automatizzare il processo di provisioning di server hardware all'interno di infrastrutture data center di ampie dimensioni, come quelle destinate al cloud. È molto adatto agli ambienti distribuiti e multi-site grazie all'architettura modulare, basata su due ruoli, configurabili in HA [67].

- Il *region controller* è il componente centrale su cui sono eseguiti i servizi centralizzati di DNS, NTP e proxy; coordina tutti i *rack controller*, sincronizza i dati e mantiene il database locale Postgres. Questo gli consente di avere una visione globale dell'intera infrastruttura; ospita il catalogo dei server e fornisce un'interfaccia web per la gestione amministrativa.
- I *rack controller* sono distribuiti in più luoghi o data center e gestiscono le risorse hardware all'interno del loro ambito. Forniscono servizi locali DHCP, IPMI, PXE, TFTP, si occupano di rilevare i nodi hardware, effettuare il provisioning dei sistemi operativi, avviare e configurare i server fisici. Comunicano su base periodica con il region controller per sincronizzare e condividere le informazioni sullo stato dell'infrastruttura.

The screenshot displays the MAAS web interface for a machine named 'valued-akita.maas'. The interface is organized into several sections:

- Overview:** Shows the machine is 'Deployed' with the OS '18.04 LTS "Bionic Beaver"'. It lists hardware specifications: CPU (4 cores, 1.8 GHz, Intel(R) Xeon(R) CPU E5-2403 0), Memory (15.6 GiB), and Storage (2000.4 GB over 2 disks).
- Machine summary:** A navigation bar with tabs for Network, Storage, Commissioning, Tests, Logs, Events, and Configuration.
- System Information:** A table listing system details such as Vendor (HP), Product (ProLiant DL360e Gen8), Version, Serial, Mainboard, and Firmware.
- NUMA Node:** Details for Node 0, including CPU cores (4), Memory (15.6201171875 GiB), Storage (2.00 TB over 2 disks), and Network (4 interfaces).
- Network:** A table listing network interfaces (eno1-eno4) with their MAC addresses, link speeds, fabric types, DHCP status, and SRIOV support.

Figura 22. MAAS: interfaccia di gestione.

Fonte [67]

L'automazione di MAAS consente di installare i server fisici, anche geograficamente distribuiti, mantenendo una configurazione uniforme. Sono supportati anche backend commerciali come AWS, Microsoft Azure, Google Cloud Platform.

L'amministratore può definire dei profili (ad esempio nodi *controller*, *compute*, *network*) per creare delle categorie in cui collocare ogni server. Quando si istanziano i server fisici, MAAS coordina le attività di commissioning tramite i protocolli PXE, TFTP e DHCP. In particolare:

- effettua il rilevamento dell'hardware: CPU, RAM, layout dei dischi, periferiche;
- fornisce le informazioni di rete e le configurazioni necessarie;
- utilizza un'immagine del sistema operativo dal proprio catalogo;
- provvede all'installazione del sistema operativo e alla configurazione di base.

Infine MAAS si occupa anche del decommissioning ovvero di rilasciare i server quando non sono più necessari e renderli disponibili per nuovi carichi di lavoro. La figura seguente mostra lo schema sintetico di funzionamento.

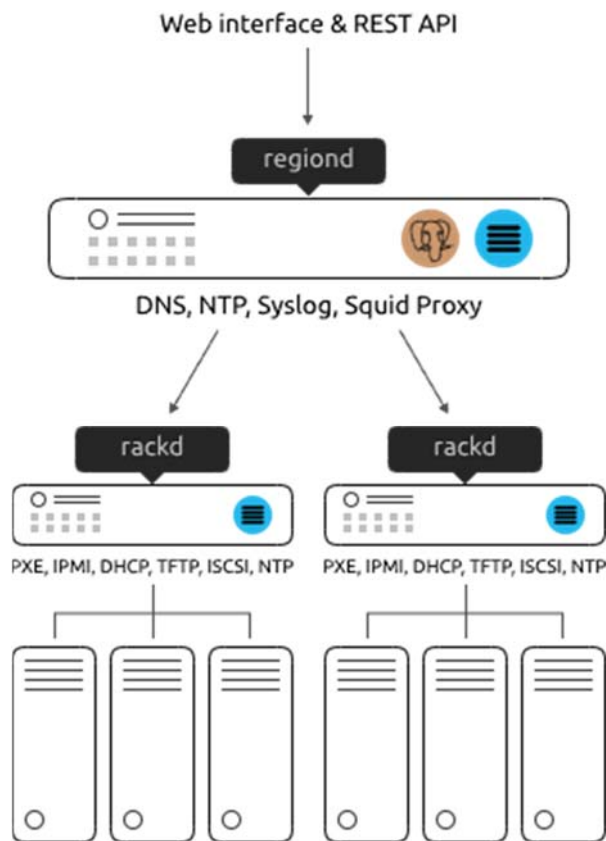


Figura 23. MAAS: schema di funzionamento.

Fonte [67]

## 4.2.2 Juju

Juju è uno strumento sofisticato per la distribuzione e la gestione di applicazioni complesse in un ambiente cloud, in grado di orchestrare risorse su diversi tipi di piattaforme. Il punto centrale per la gestione dell'ambiente cloud è il *controller* Juju: ne viene istanziato uno per ciascun cloud che si desidera gestire (ad esempio MAAS, OpenStack, GCP, AWS) e consente di configurare i diversi servizi, detti anche carichi di lavoro. Il componente client è il software CLI (interfaccia a riga di comando) usato per inviare i comandi e gestire uno o più controller Juju [68].

Juju installa applicazioni, le cui istanze prendono il nome di *unit*, e per la loro distribuzione (*deploy*) si utilizzano caratteristici file di script chiamati *charm*, messi a disposizione attraverso il repository CharmHub<sup>48</sup> e sviluppati sia da Canonical sia dalla comunità open source. Ogni charm è specifico per un'applicazione e descrive come distribuirla e configurarla.

In generale le applicazioni *cloud native* sono raramente istanziate in modo isolato, anzi sono composte da diversi microservizi. Si pensi ad esempio che anche una semplice applicazione web richiede almeno un database e un server http collegati tra loro. I charm implementano l'intelligenza necessaria per connettere le diverse applicazioni; tali connessioni, dette *relazioni* (o *integrazioni* nella nuova versione di Juju), sono formate collegando opportunamente gli *endpoint* forniti dalle applicazioni stesse [68].

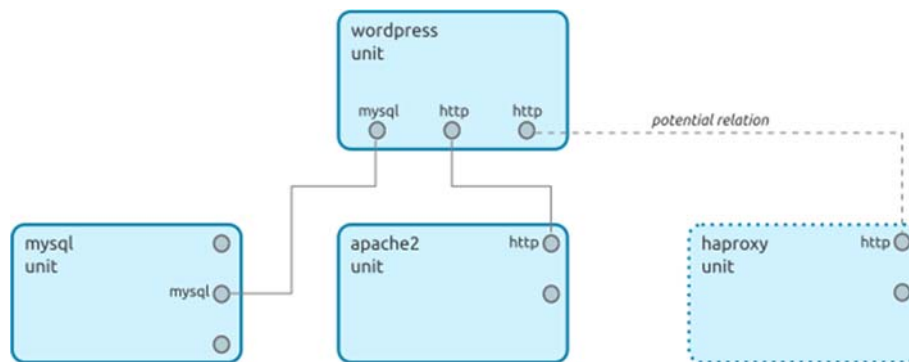


Figura 24. Juju: esempio di relazioni tra application unit.

Fonte [68]

Un sistema complesso, formato da diversi componenti interconnessi, viene detto *Juju model*. Per la sua realizzazione si raccolgono i riferimenti ai charm di tutti i servizi occorrenti in un unico file YAML chiamato *bundle*. Esso costituisce una sorta di ricetta<sup>49</sup> e contiene tutte le configurazioni delle applicazioni e le relazioni necessarie. Si può dire che se un charm rappresenta una singola applicazione o servizio, un bundle rappresenta un intero modello Juju [63]. Nell'ambito della terminologia, è importante

<sup>48</sup> Link al repository ufficiale: <https://charmhub.io/>

<sup>49</sup> In ambito cloud si usa proprio il termine inglese *recipe*.



evidenziare una distinzione tra i due concetti di *machine* e di *unit*, come riportato in Figura 25:

- *machine*: rappresenta un'istanza di risorsa hardware, che sia macchina fisica o virtuale, su cui è possibile eseguire servizi e applicazioni;
- *unit*: rappresenta un'istanza di un'applicazione, creata da un charm specifico, in esecuzione all'interno di una macchina.

È possibile distribuire una singola unità per macchina oppure, per ottimizzare le risorse, distribuire sulla stessa macchina più unità della stessa o di applicazioni diverse.

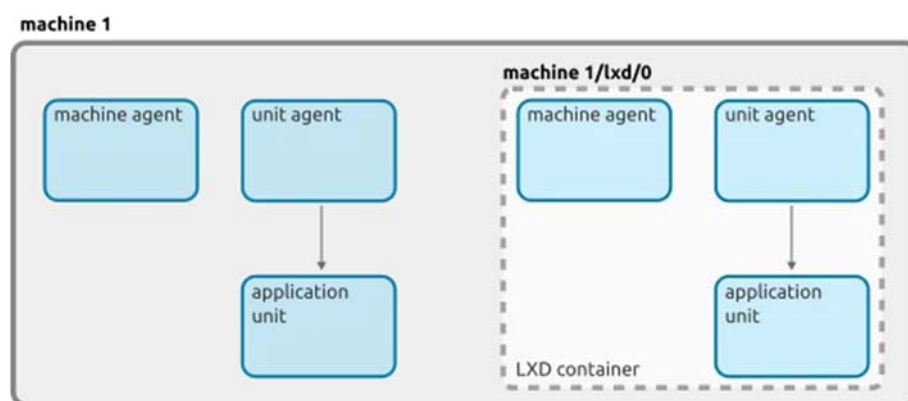


Figura 25. Juju: esempio application unit bare metal e in container LXD.

Fonte [68]

Il modello Juju viene generato, configurato e modificato attraverso l'interazione del Juju controller con la rete formata dagli *agent*, i componenti software eseguiti su ogni machine e su ogni unit:

- *Machine agent*: lavorano a livello più ampio e si occupano della distribuzione, configurazione e monitoraggio sulla macchina virtuale o fisica sottostante;
- *Unit agent*: lavorano a livello della singola unità applicativa e gestiscono lo stato della singola istanza.

Entrambe le tipologie di agent comunicano con il controller Juju inviando il proprio stato e ricevendo eventuali assegnazioni di attività da effettuare per aggiornare lo stato del modello.

#### *Esempio di un bundle Juju*

Si riporta di seguito una breve descrizione delle parti principali che compongono un bundle YAML per la distribuzione con Juju, tratta dalla documentazione ufficiale<sup>50</sup>.

<sup>50</sup> Link: <https://discourse.charmhub.io/t/bundle/1158>

- *series*

Imposta la versione di Ubuntu da utilizzare per tutte le macchine e i container (può essere specificato anche a livello di singola macchina o applicazione):

```
series: bionic   indica Ubuntu 18.04 Bionic Beaver
series: focal   indica Ubuntu 20.04 Focal Fossa
series: jammy   indica Ubuntu 22.04 Jammy Jellyfish
```

- *variables*

Definizione facoltativa delle variabili come ancore. I valori sono preceduti da un carattere “&”. Quando la variabile è usata nel bundle (in genere nelle opzioni di un’applicazione) è preceduta da un carattere \*. Ad esempio:

```
variables:
  data-port: &data-port 8080
  expected-osd-count: &expected-osd-count 5
  expected-mon-count: &expected-mon-count 3
```

- *machines*

Indica gli ID delle macchine disponibili su cui installare le applicazioni con l’opzione “to”. Ad esempio:

```
machines:
  '0':
  '1':
  '2':
```

- *constraints*

Definisce i vincoli per l’applicazione, che saranno applicati anche in caso di aggiunta unit di una stessa applicazione). Alcuni esempi:

```
constraints: root-disk=8G
constraints: cores=4 mem=4G root-disk=16G
constraints: arch=amd64 tags=controller-node
```

- *applications*

In questa sezione vengono inserite le caratteristiche delle applicazioni e di tutte le unit, usando una struttura nidificata di parametri.

- *charm:*

Indica quale charm è usato per il deploy dell’applicazione.

- *bindings:*

Consente di connettere gli endpoint di un’applicazione agli spazi di rete, nel caso in cui le unit abbiano più schede di rete e si utilizzino subnet diverse per separare i servizi.

- *num\_units*:  
Indica il numero di unità da distribuire
- *to*:  
Indica la destinazione delle unità distribuite. Ad esempio:
  - to: 0            Installazione bare-metal sulla macchina 0
  - to: "lxd:2"    Installazione in container LXD nella macchina 2
- *options*:  
Permette di definire la configurazione di un'applicazione. Le opzioni sono diverse per ciascuna applicazione e si trovano nella documentazione del relativo (file metadata.yaml).

Esempio:

```
applications:
  nova-compute:
    charm: cs:nova-compute-301
    num_units: 3
    to:
      - "0"
      - "1"
      - "2"
    options:
      action-managed-upgrade: true
      debug: false
      enable-live-migration: true
      enable-resize: true
    bindings:
      cloud-compute: openstack-mng-space
      ceph: openstack-storage-space
```

- *relations*

Indica le relazioni da aggiungere tra le applicazioni. Ogni relazione si sviluppa su due righe, in cui ciascuna riga rappresenta un lato di una relazione: la prima inizia con due trattini e la seconda inizia con un trattino singolo. La relazione è nel formato # '<application>:<endpoint>'. Ad esempio:

```
relations:
- - openstack-dashboard:ha
  - horizon-hacluster:ha
- - neutron-api:amqp
  - rabbitmq-server:amqp
- - nova-compute:image-service
  - glance:image-service
```

### *Approccio dichiarativo*

Dalla struttura appena esposta si può notare che il bundle Juju utilizza una modellazione dichiarativa per effettuare il deploy dei diversi servizi che costituiscono il

modello. In generale, l'approccio dichiarativo consente di descrivere le parti che compongono il sistema e le loro relazioni di fornitore/consumatore senza descrivere il modo procedurale per ottenere il risultato.

Come esposto nel lavoro di Attardi et al. [62], un modello di distribuzione dei servizi imperativo prevede di specificare come eseguire il deployment, fornendo un piano esplicito. Si pensi alla tipica distribuzione del software effettuata attraverso script procedurali, in cui si elencano tutti i passaggi da eseguire, senza indicare motivi e obiettivi. La manutenzione di tali script ha lo svantaggio di richiedere una conoscenza generale del funzionamento di tutte le parti del sistema e presenta il rischio di perdere vincoli importanti tra i diversi componenti, se essi non vengono opportunamente documentati dal programmatore. Al contrario, un modello di distribuzione dei servizi dichiarativo specifica un obiettivo e fornisce una descrizione di alto livello dei requisiti e dei vincoli, oltre che la composizione logica dei servizi. Una sorta di motore di distribuzione ha il compito di interpretare la descrizione e di eseguire il piano delle azioni per ottenere la configurazione desiderata.

Nel caso specifico, dato un bundle, Juju opera in diversi passaggi trasformando via via lo stato del sistema fino a convergere quando raggiunge la configurazione che soddisfa tutti i vincoli. Questo meccanismo è basato sulle interazioni esterne e interne in cui è coinvolto il Juju controller [68].

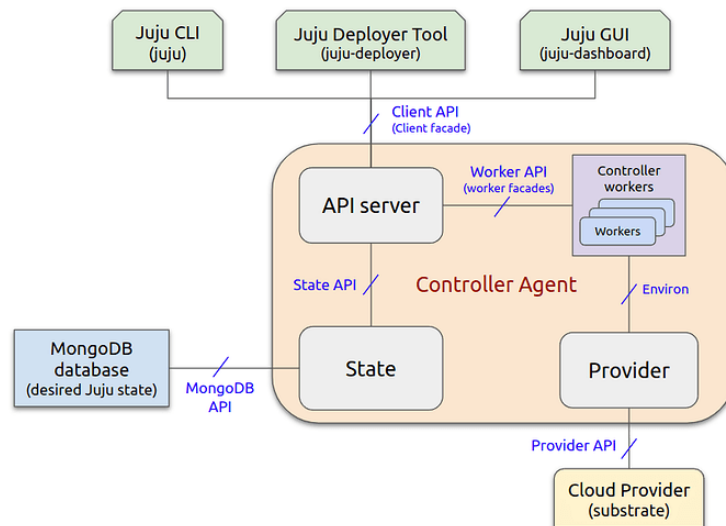


Figura 26. Juju controller: architettura ad alto livello.

Fonte [68]

In riferimento allo schema della Figura 26, le interazioni del Juju controller verso l'esterno sono sostanzialmente tre:

- un database locale MongoDB, su cui il Juju controller mantiene una rappresentazione dello stato desiderato dei modelli e di tutte le applicazioni all'interno di ciascun modello (servizi installati, configurazioni e relazioni);
- una connessione verso il cloud provider prescelto su cui organizzare le risorse di elaborazione, rete e storage;
- un'interfaccia su cui ricevere comandi dal Juju client, via CLI, via interfaccia grafica o con gli strumenti di sviluppo.

Verso l'interno Juju controller gestisce una serie di processi *worker* che comprendono tutti i vari agent a livello unit e a livello machine. Tali agent interagiscono direttamente con il cloud per eseguire operazioni per applicazioni e unità, secondo il seguente flusso.

8. Il Juju controller mantiene una rappresentazione dello stato desiderato dei modelli e di tutte le applicazioni all'interno di ciascun modello (servizi installati, configurazioni e relazioni), in un database interno MongoDB.
9. Gli agent a livello unit e livello machine inviano periodicamente al controller informazioni sul loro stato.
10. Il Juju controller verifica lo stato del modello; se è richiesta una variazione, assegna un compito specifico agli agent preposti. Alcuni esempi sono: installare o aggiornare il software, modificare un parametro di configurazione, scalare un'applicazione aumentando le unit, aggiungere relazioni.
11. Gli agent inviano dettagli sull'esito del proprio compito (success/failure) al Juju controller.
12. In base al riscontro ricevuto, il Juju controller aggiorna lo stato del modello oppure prende le eventuali contromisure nel caso in cui alcune attività siano fallite, ad esempio ritentare il processo.
13. Il Juju controller effettua un monitoraggio dei modelli a intervalli regolari, quindi il processo si ripete.

L'utilizzo di Juju e dell'approccio dichiarativo porta con sé ulteriori vantaggi [62]. Da un lato la componibilità dei servizi consente agli utenti di condividere i vari modelli, riutilizzando oggetti sviluppati da altri e costruendo eventualmente solo le parti mancanti, necessarie alle proprie esigenze. D'altro lato si ottiene un sistema agile perché le modifiche alla configurazione possono essere apportate rapidamente, lasciando al motore di distribuzione il compito di pianificare i passaggi necessari per l'attuazione. Tutti i meccanismi operano astraendosi completamente dall'infrastruttura cloud sottostante perché Juju si occupa di indirizzare le azioni di distribuzione utilizzando le opportune chiamate API dei diversi provider. Infine, è interessante notare che nell'architettura GARR Cloud l'approccio dichiarativo e i bundle Juju sono utilizzati per effettuare l'installazione sia della piattaforma cloud stessa sia delle applicazioni che gli utenti possono distribuire in modalità self-service [62].

### 4.2.3 Ceph

Ceph è un progetto open source caratterizzato da un'architettura distribuita, robusta e molto versatile. Come indicato dai report ufficiali *OpenStack Survey 2022*, la maggior parte delle installazioni di cloud OpenStack utilizza Ceph per lo storage a blocchi.

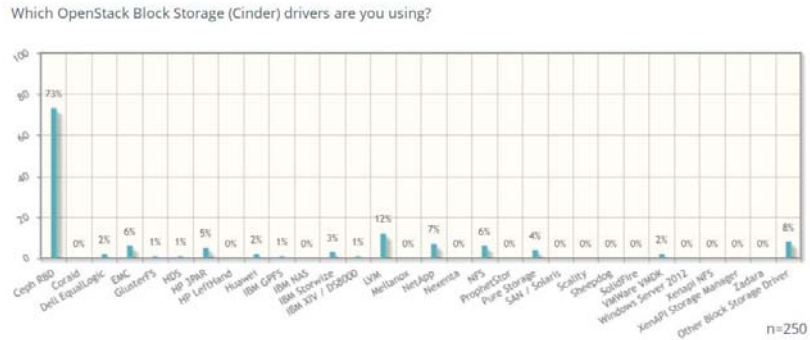


Figura 27. OpenStack Report: utilizzo di Ceph come block storage

Fonte <https://www.openstack.org/analytics>

Una delle caratteristiche molto interessanti di Ceph è la possibilità di rendere i dati disponibili via rete in diversi modi, come illustrato in Figura 28 [66]:

- come archivio oggetti (*object storage*), compatibile con i protocolli Amazon S3 e Swift, utilizzando il gateway RADOS;
- come volumi di storage a blocchi (*block storage*) tramite il driver RBD;
- come *file system* compatibile POSIX tramite CephFS.

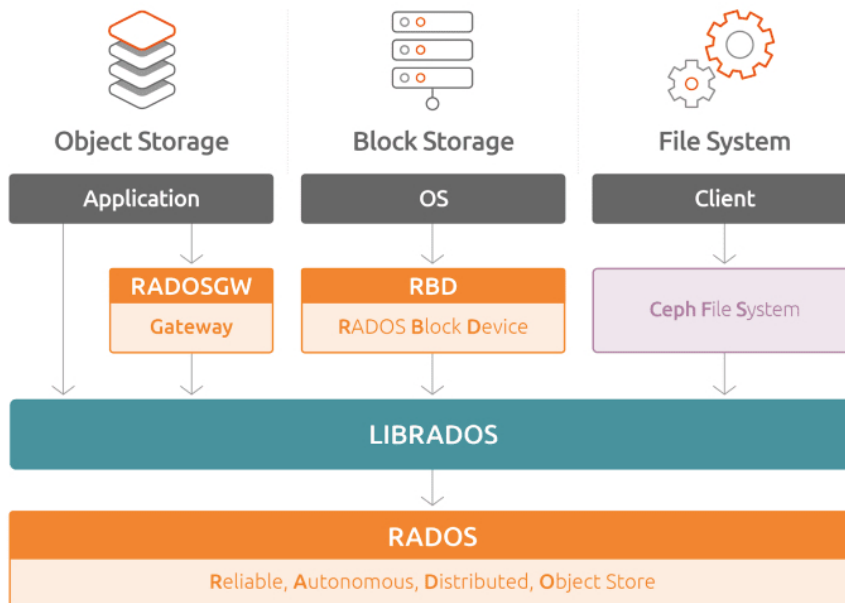


Figura 28. Ceph: opzioni di storage

Fonte <https://ubuntu.com/blog/ceph-storage-on-ubuntu-an-overview>

Nell'architettura Ceph, si distinguono tre componenti di base:

- i nodi *Monitor* (ceph-mon) eseguono la logica applicativa operando in un cluster con HA, mantengono la configurazione e la mappa dello stato del cluster e gestiscono l'autenticazione tra i client e i ceph-osd;
- i nodi *Object Storage Daemon* (ceph-osd) operano più a basso livello e archiviano i dati sui dischi fisici. Esiste un ceph-osd e il relativo processo in esecuzione per ciascun disco fisico esistente;
- i nodi *Manager* registrano le metriche del cluster e forniscono l'interfaccia verso sistemi esterni di monitoraggio.

Elemento fondamentale per i meccanismi Ceph è l'algoritmo CRUSH [69] (*Controlled Replication Under Scalable Hashing*) che calcola come e dove archiviare (o recuperare) i dati nel cluster con l'obiettivo di distribuirli in modo uniforme su tutti i dispositivi di storage. La mappa CRUSH, formata dall'albero degli osd e dalle regole CRUSH, descrive l'intera struttura fisica del cluster ed è mantenuta dai nodi ceph-mon. L'amministratore può personalizzare le regole CRUSH che definiscono le politiche di posizionamento dei dati, ad esempio indicando quale classe di dispositivo utilizzare, o quanti blocchi scrivere in caso di *erasure code*. Ciascun nodo del cluster, usando la mappa, è in grado di calcolare dove è stato archiviato un elemento di dati, quindi non è necessario mantenere una sorta di directory centrale. Inoltre i client che devono connettersi alle risorse storage ricevono la mappa dai nodi ceph-mon e possono poi interagire direttamente con i nodi ceph-osd.

I dischi fisici vengono catalogati in *device class*. Accanto alle classi di default *hdd*, *ssd*, *nvme* si possono anche definire delle etichette personalizzate per le proprie esigenze. L'archiviazione dei dati nel cluster Ceph avviene all'interno di *pool* distinti, ovvero dei contenitori logici con proprietà e caratteristiche specifiche. Ne esistono di due tipi, a seconda dell'approccio usato per garantire la ridondanza dei dati [66], [69].

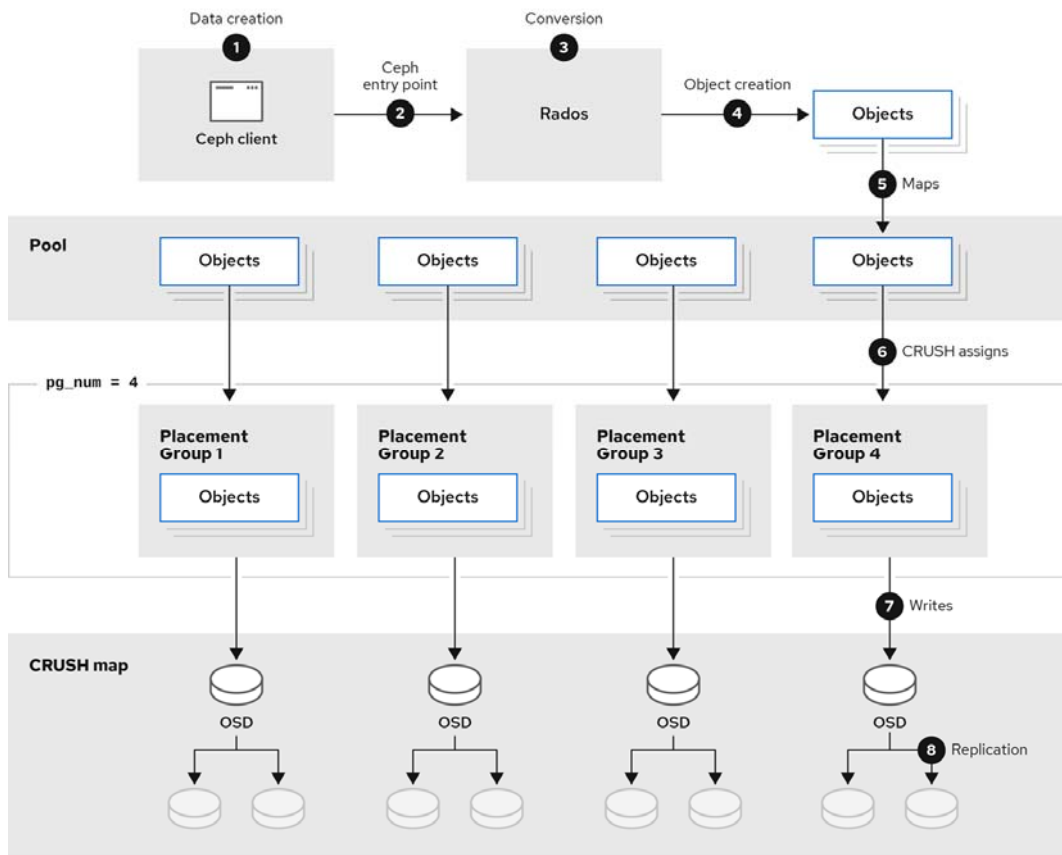
- Pool con *replica*: mantengono copie ridondanti dei dati su più nodi, garantendo alta affidabilità ma richiedendo più spazio di archiviazione. Il livello minimo consigliato è 3; sebbene il 2 sia configurabile non fornisce adeguata protezione ma espone a rischi di perdita dati.
- Pool con *erasure code*: i dati sono suddivisi in due tipi di blocchi (*chunk*), blocchi di dati e di parità. Dai chunk dati, con opportuni algoritmi, vengono calcolate porzioni ridondanti dette *parity chunks*, utilizzate per ricostruire i dati nel caso di guasti.

La replica garantisce una elevata affidabilità ma è meno efficiente in termini di spazio di archiviazione. L'opzione erasure coding richiede meno spazio rispetto al numero di repliche, in quanto sono necessarie solo alcune porzioni ridondanti. Tuttavia il calcolo della parità può richiedere elaborazioni aggiuntive e influire sulle prestazioni in fase di

scrittura. In generale, la scelta dipende dal compromesso che si intende raggiungere tra affidabilità, uso dello spazio e prestazioni generali del sistema.

All'interno di un pool i dati vengono divisi in *placement group* e successivamente abbinati ai dischi fisici osd. L'uso dei placement group contribuisce al bilanciamento del carico all'interno del cluster e consente a Ceph di scalare orizzontalmente. In caso di aggiunta o rimozione di dispositivi storage, Ceph può ridistribuire automaticamente i placement group in modo che il carico sia uniforme tra i dispositivi disponibili.

La strategia di archiviazione di Ceph è completamente trasparente all'utente, che avrà come unica interfaccia il pool [69]. La figura che segue mostra il flusso completo dell'archiviazione dei dati.



©2018 Ceph, 0621

Figura 29. Ceph: diagramma dell'archiviazione dati.

Fonte [69]



### 4.3 La piattaforma OpenStack

OpenStack<sup>51</sup> è una piattaforma open source di tipo modulare sviluppata per ambienti cloud privati e pubblici. Rilasciato sotto licenza Apache, nasce nel 2010 come progetto di NASA e Rackspace Cloud (una grande società di web hosting IT), ma ad oggi apportano il loro contributo oltre 500 aziende, tra cui AMD, Canonical, Cisco Systems, Dell, Hewlett-Packard, IBM, Intel, Red Hat, VMware, Oracle.

I componenti OpenStack sono microservizi interdipendenti, a ciascuno dei quali corrisponde un diverso progetto open source, che interagiscono tra loro in un modello di architettura *service oriented*. L'obiettivo è creare una sorta di astrazione per le risorse computazionali (RAM e CPU), di archiviazione (dischi e storage) e di rete (NIC, switch load balancer e router) per replicare in software un'intera infrastruttura IT e fornire un livello IaaS per l'esecuzione di macchine virtuali (VM). Le VM sono separate in diversi progetti, concetto usato da OpenStack per realizzare la cosiddetta *multi-tenancy*, la possibilità di creare sulla stessa infrastruttura hardware ambienti virtuali logicamente separati, utilizzati da più clienti (dall'inglese *tenant*, inquilino).

#### 4.3.1 Rilascio delle versioni

OpenStack ha un ciclo di rilascio di 6 mesi<sup>52</sup>. Ogni ciclo include la pianificazione delle nuove funzionalità da sviluppare e i miglioramenti da apportare, lo sviluppo, la fase di testing e il rilascio di finale.

Tutti i team dei progetti ufficiali lavorano sotto la supervisione del comitato tecnico di OpenStack e concorrono alla realizzazione dei vari risultati finali. In base a una convenzione, le versioni di OpenStack sono denominate in ordine alfabetico e cambiano ad ogni rilascio. Dopo *Zed*, da marzo 2023 si è ripartiti dalla *A* con *Antelope*.

Tabella 7. OpenStack: ultime 5 versioni rilasciate.

Series	Status	Initial Release Date	Next Phase
2024.1 Caracal	Development	2024-04-03 estimated	Maintained estimated 2023-04-03
2023.2 Bobcat	Maintained	2023-10-04	Extended Maintenance estimated 2025-04-04
2023.1 Antelope	Maintained	2023-03-22	Extended Maintenance estimated 2024-09-22
Zed	Maintained	2022-10-05	Extended Maintenance estimated 2024-04-05
Yoga	Maintained	2022-03-30	Extended Maintenance estimated 2023-11-02

<sup>51</sup> Link al sito del progetto: <https://www.openstack.org/>

<sup>52</sup> OpenStack Releases, link: <https://releases.openstack.org/>

#### 4.3.2 Stato dell'arte nella ricerca accademica

OpenStack è nato nel 2010 e da oltre un decennio molti ricercatori si sono interessati a questa piattaforma perché, a differenza di altre soluzioni cloud open source, ha alle spalle l'adesione di grossi nomi del panorama IT e una solida comunità di sviluppo.

Nel lavoro di Al-Baroodi et al. [6] (2013) vengono esaminate diverse piattaforme *Open Source Cloud Computing* (OSCC): OpenStack, Abicloud, Eucalyptus, Nimbus e OpenNebula. I risultati vedono prevalere OpenStack che, data la velocità di sviluppo, viene considerato una valida opzione a lungo termine e versatile per l'impiego in diversi settori. Similarmente, una comparazione tra le diverse soluzioni OpenStack, Eucalyptus e OpenNebula è alla base della pubblicazione di Manacero et al. [70] (2013) secondo cui OpenStack presenta maggiori vantaggi sui concorrenti anche in virtù della presenza di una buona documentazione on-line e una community molto attiva, pronta alla gestione e alla correzione di bug. Secondo Raj et al. [71] (2013) il confronto in termini di facilità di implementazione e orchestrazione dei servizi pone OpenStack in vantaggio rispetto ad altri strumenti. Giocano inoltre a favore il supporto e il finanziamento da parte di molte grandi aziende e l'ampio bacino di sviluppo che questo produce. Sulla stessa linea si colloca lo studio comparativo di Kumar et al. [72] (2014) in cui si ritrovano a confronto le soluzioni Eucalyptus, OpenStack, CloudStack e OpenNebula.

Le stesse piattaforme vengono testate da Shahzadi et al. [73] (2017) con particolare riguardo alle prestazioni e all'architettura. Sposta invece il focus dai test al deploy in produzione il lavoro dei ricercatori del CERN, Bell et al. [74] (2015) che analizzano come scalare il proprio private cloud, operativo dal 2013, che al tempo contava 5.500 nodi e oltre 12.000 VM in esecuzione. Più recente lo studio di Ciptaningtyas et al. [75] (2023) sulla realizzazione di un private cloud OpenStack presso l'ateneo indonesiano ITS come base per l'uso didattico della piattaforma cloud KYPO Cyber Range.

Riconoscendo il ruolo cruciale degli strumenti di automazione nella gestione di cloud di tipo IaaS, è interessante la pubblicazione di Coullon et al. [76] (2017) che modellano il processo di distribuzione di OpenStack e lo usano come base per un confronto dettagliato tra 5 software di automazione: Kolla, Enos, Juju, Kubernetes e TripleO.

Ricca di spunti è l'analisi di Lima et al. [77] (2019) sullo stato dell'arte dei moduli del progetto OpenStack e le principali attività in tema di sviluppo. Applicando metodi statistici ed esaminando i repository del codice sorgente, si raccolgono dettagli sulla collaborazione della comunità e le influenze delle aziende IT che supportano OpenStack Foundation. Red Hat è stata il primo contributore, seguita da Rackspace, HP, Mirantis e IBM. Dell si è concentrata sulla fornitura di hardware e servizi di consulenza nella progettazione dei componenti; Canonical ha acquisito rilievo grazie a Ubuntu, diventato il sistema operativo principale nei setup di OpenStack; non sorprende trovare VMware tra i membri più esperti sui temi della virtualizzazione. Sullo stesso tema, la presenza delle aziende nello sviluppo dell'open source, hanno lavorato anche Zhang et al. [78] (2019) con uno studio esplorativo su OpenStack. Da un dataset creato

dall'analisi dei commit su Git, dal lancio di OpenStack (21/07/2010) al 18/01/2017, gli autori quantificano i contributi dei programmatori volontari e quelli delle aziende tramite gli sviluppatori affiliati su 14 versioni complete. Dai risultati emerge un maggiore apporto delle aziende rispetto ai volontari (Figura 30) con una distribuzione assai sbilanciata: circa il 20% delle aziende ha fornito l'80% degli sviluppatori e il 10% delle aziende ha contribuito per l'80% dei commit (Figura 31).

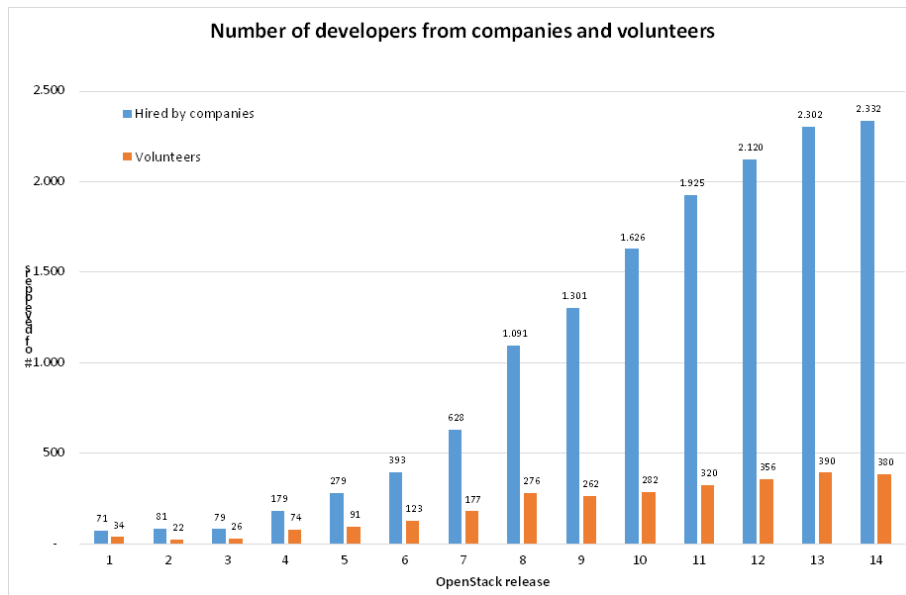


Figura 30. OpenStack: numero di sviluppatori di aziende e volontari.

Fonte [78]

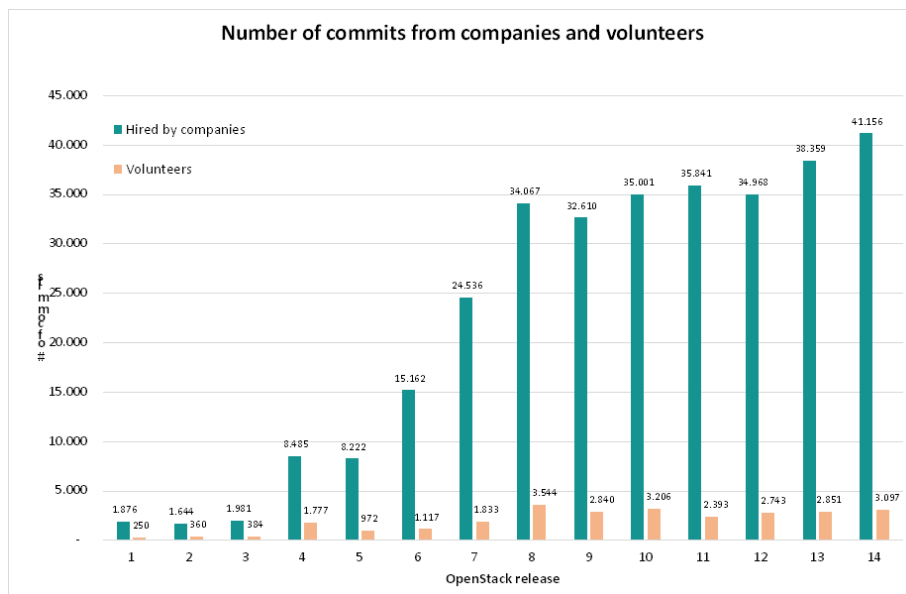


Figura 31. OpenStack: numero di commit per aziende e volontari.

Fonte [78]

### 4.3.3 Panoramica dei casi d'uso

Il forte supporto dell'industria e della comunità di sviluppo ha contribuito alla diffusione di OpenStack in numerosi contesti di produzione, come si evince dall'elenco dei casi d'uso sul sito ufficiale<sup>53</sup>.

Hanno adottato questa piattaforma diversi istituti bancari come gli spagnoli Banco Santander e BBVA (Banco Bilbao Vizcaya Argentaria), la cinese PSBC (Postal Savings Bank of China), SBAB in Svezia e infine una delle banche più grandi al mondo, Commonwealth Bank of Australia. Sono molte le implementazioni da parte di aziende del settore telecomunicazioni, tra cui Deutsche Telekom, China Mobile, Saudi Telecom STC, colosso dell'Arabia Saudita e il gruppo statunitense T-Mobile. Canonical ha inoltre diffuso sul suo blog<sup>54</sup> che *Charmed OpenStack* è stato scelto per i propri cloud privati da Telefonica Brazil e da Nayatel, uno dei principali fornitori Telco in Pakistan.

In Europa una delle installazioni più rilevanti è quella del CERN, l'organizzazione europea per la ricerca nucleare con il più grande laboratorio al mondo di fisica delle particelle, situato vicino a Ginevra al confine tra Francia e Svizzera. CERN utilizza OpenStack per il suo private cloud dal 2013 [74]; gli specialisti IT sono molto attivi nella comunità, condividendo le proprie esperienze sul blog<sup>55</sup> e nei numerosi interventi durante i Summit OpenStack nel mondo. L'infrastruttura operativa al momento conta oltre 6.500 nodi fisici e quasi 35.000 VM [79], come emerge dalla Figura 32.

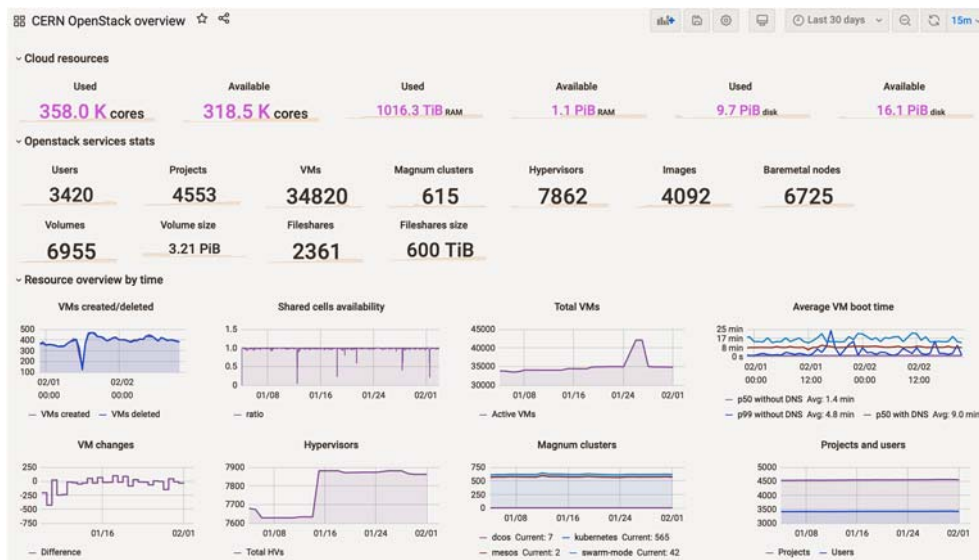


Figura 32. OpenStack al CERN: i numeri dell'infrastruttura.

Fonte [79]

<sup>53</sup> Link: <https://www.openstack.org/use-cases>

<sup>54</sup> Link: <https://ubuntu.com/blog/openstack-is-dead>

<sup>55</sup> CERN TechBlog, link: <https://techblog.web.cern.ch/techblog>

Un esempio di piattaforma cloud federata per la ricerca, molto simile al modello GARR, è il progetto bwCloud<sup>56</sup> sviluppato in Germania da diverse università locali dello stato del Baden-Württemberg per formare il cosiddetto *IaaS for Science and Education* [62]. L'architettura di bwCloud, basata su OpenStack, offre oltre 100 nodi di calcolo, 2.700TB di storage raw e 1.800 vCPU con i servizi dislocati nei quattro siti di Mannheim, Freiburg, Karlsruhe e Ulm.

Nello scenario italiano, oltre alle infrastrutture gestite da GARR, si possono citare i cloud privati dell'Istituto Nazionale di Fisica Nucleare (INFN)<sup>57</sup> e dell'Istituto Nazionale di Geofisica e Vulcanologia (INGV). In particolare quest'ultimo, con il supporto degli specialisti del Dipartimento Calcolo e Storage Distribuito di GARR, ha realizzato nel 2020 il progetto di una sala operativa virtuale basata su un sistema cloud OpenStack multi-region, con i nodi distribuiti nei data center delle sezioni di Roma, Napoli e Catania [80].

Tra le aziende di servizi commerciali, due casi d'uso coinvolgono Fastweb e Aruba Cloud. Fastweb ha lanciato nel 2016 un'offerta cloud per le aziende italiane basata sulla piattaforma OpenStack di Red Hat<sup>58</sup>, mentre Aruba Cloud propone OpenStack nelle sue configurazioni cluster, accanto alle altre soluzioni proprietarie come VMware e Microsoft Hyper-V<sup>59</sup>.

Infine, conclude questa panoramica il sistema cloud open source Nivola<sup>60</sup> sviluppato interamente sulla tecnologia OpenStack dal CSI Piemonte (Consorzio per il Sistema Informativo), un consorzio di enti pubblici che opera come fornitore e partner tecnologico delle PA sul territorio piemontese. CSI Piemonte è un cloud provider certificato da AGID per offrire infrastrutture e servizi per la PA; con Nivola propone sia servizi di tipo IaaS con server virtuali, configurabili combinando virtual CPU, RAM e spazio storage, sia servizi aggiuntivi quali DBaaS (*DataBase as a Service*), STaaS (*Storage as a Service*) e BaaS (*Backup as a Service*).

---

<sup>56</sup> Progetto bwCloud, link: <https://www.bw-cloud.org/>

<sup>57</sup> INFN Cloud Documentation, link: <https://guides.cloud.infn.it/docs/admins-guides>

<sup>58</sup> Fastweb, link: <https://www.fastweb.it/corporate/media/comunicati-stampa/fastweb-lancia-la-prima-offerta-cloud-per-le-aziende-italiane-basata-su-piattaforma-red-hat-openstack/>

<sup>59</sup> Aruba hypervisor: differenze tra i sistemi offerti, link: <https://kb.cloud.it/computing/cloud-server/creare-e-configurare-un-cloud-server/hypervisor-differenze-tra-i-sistemi-offerti.aspx>

<sup>60</sup> CSI e progetto Nivola, link: <https://www.nivolapiemonte.it>

#### 4.3.4 Architettura logica e microservizi

Il cuore di OpenStack è un insieme di moduli interconnessi, ognuno basato su un diverso progetto open source, che interagiscono tra loro in un modello *service oriented* per fornire i servizi di un data center virtuale: orchestrazione, gestione VM, rete, storage a oggetti e a blocchi. Di seguito i principali microservizi [81, pp. 12-23] [82, pp. 3-23]:

- *Nova*: gestisce il servizio *Compute* per il ciclo di vita delle VM. Può usare diversi hypervisor, spesso la scelta ricade su KVM, integrato nel kernel Linux;
- *Neutron*: è il servizio che si occupa del *Networking* fornendo l'infrastruttura di rete per creare reti virtuali, assegnare sottoreti e gestire indirizzi IP;
- *Keystone*: è il componente responsabile dell'*Identity Service*, servizio che svolge l'autenticazione degli utenti e l'assegnazione dei diritti sulle risorse. Fornisce inoltre un catalogo di *endpoint* per tutti i servizi OpenStack;
- *Cinder* è il servizio che fornisce il *Block Storage*, la memoria permanente sotto forma di memoria a blocchi, come ad esempio i dischi rigidi per le VM;
- *Glance*: esegue l'*Image Service*: mette a disposizione un catalogo di immagini di sistemi operativi o dischi da cui possono essere generate le VM;
- *Swift*: è il servizio che offre l'*Object Storage* cioè la possibilità di utilizzare l'archiviazione a oggetti per dati non strutturati;
- *Horizon* è il componente che fornisce la *Dashboard*, l'interfaccia grafica web based usata per la gestione delle operazioni sul sistema OpenStack.

---

### OpenStack revealed

---

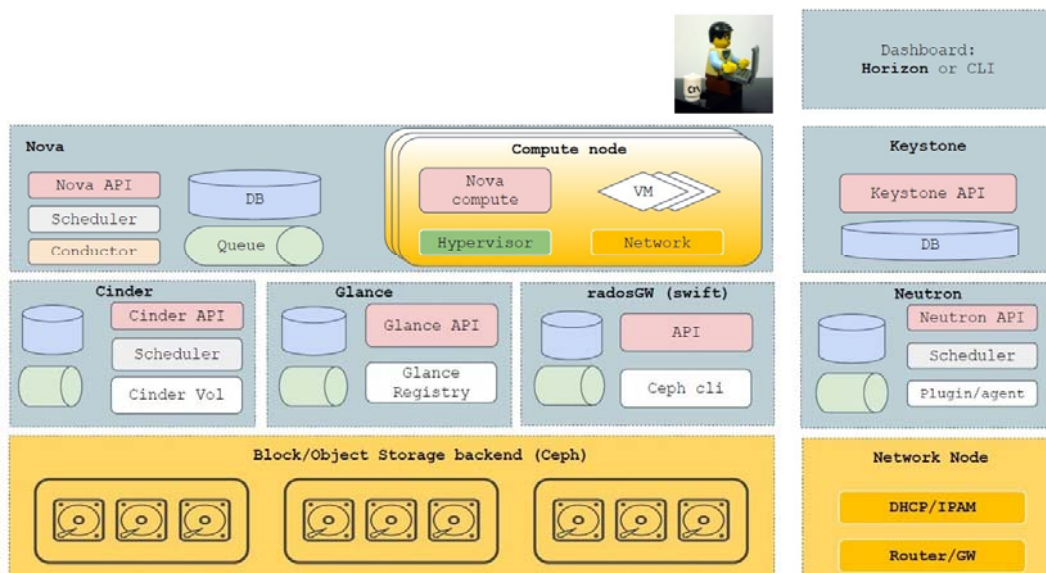


Figura 33. OpenStack: architettura logica e microservizi.

Fonte [64]

Tutti i componenti realizzano in software i vari moduli che sono normalmente presenti in un'infrastruttura hardware. Neutron fornisce il collegamento in rete, Nova i server hypervisor su cui far girare le VM, basate su immagini dei sistemi operativi gestite da Glance, i servizi per l'archiviazione storage sono offerti da Cinder e Swift, infine Keystone è responsabile della parte di autenticazione e autorizzazione degli utenti. Seppur svolgendo compiti diversi, tutti i servizi condividono il modello di architettura [64]. [81, pp. 12,23]; i moduli sono sviluppati in Python e ciascuno presenta:

- un'interfaccia esterna di *frontend* basata su Apache, chiamata *endpoint*, che fornisce le REST API usate per comunicare con gli altri servizi e con gli utenti;
- un database, basato su MySQL, per tenere traccia dello stato, delle risorse assegnate e delle operazioni eseguite;
- una coda di messaggi, basata su RabbitMQ, per gestire le comunicazioni e sincronizzare le azioni tra i vari componenti, necessaria per la distribuzione del carico di lavoro in un ambiente multiutente;
- una serie di librerie, driver e componenti *plug-in* per accedere ai servizi di *backend* o all'hardware sottostante.

L'immagine seguente mostra uno schema dell'architettura di un modulo; prendendo come esempio Cinder, il backend è rappresentato da Ceph che fa da interfaccia verso lo storage sottostante.

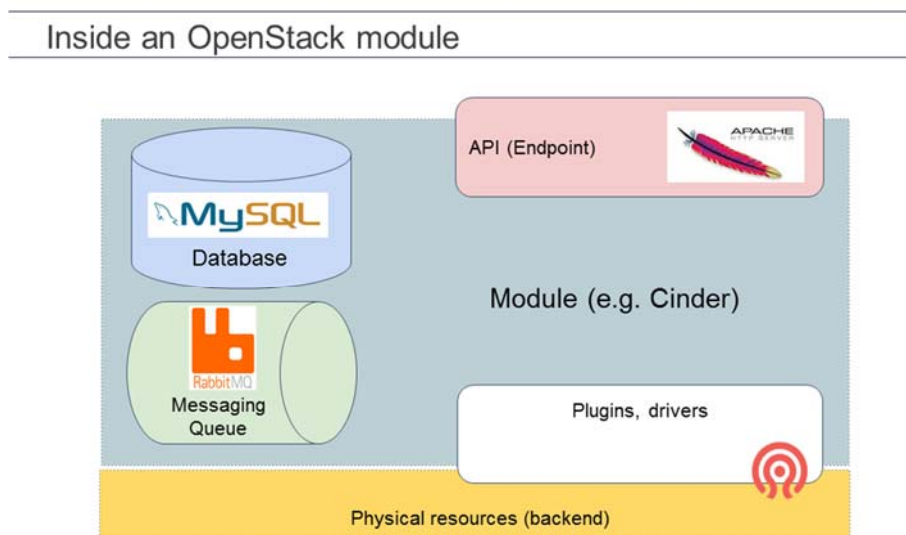


Figura 34. OpenStack: struttura interna di un modulo.

Fonte [64]

#### 4.3.5 Ruoli e servizi nel cluster

L'architettura di OpenStack nasce per essere scalabile orizzontalmente [82, pp. 31-38], [83] e quindi molti servizi sono distribuiti su più nodi, soprattutto nelle installazioni di larga scala e che necessitano di alta affidabilità.

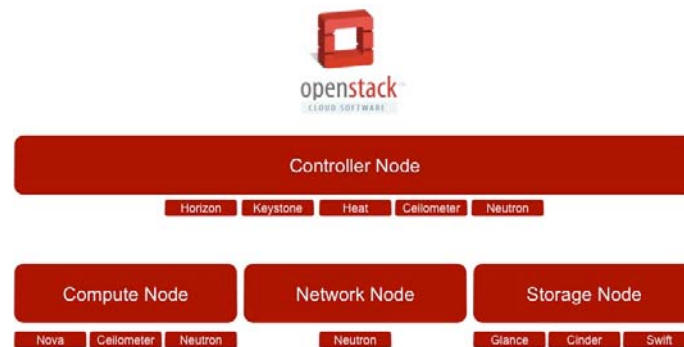


Figura 35. OpenStack: architettura multi nodo.

Fonte [83]

All'interno del cluster è possibile identificare quattro macro ruoli principali.

- *Nodo controller.* Rappresenta il *control plane* dell'ambiente OpenStack: esegue i servizi di natura centralizzata tra cui autenticazione, database, code di messaggi, scheduling, endpoint API.
- *Nodo compute.* Esegue un hypervisor tra quelli supportati (es. KVM, ESX, Microsoft Hyper-V), gestisce il servizio Nova per l'intero ciclo di vita delle VM e il servizio agent Open vSwitch<sup>61</sup> per il networking delle VM.
- *Nodo network.* Esegue i servizi Neutron e gli agent per i servizi di rete necessari: DHCP, Open vSwitch, metadata, Layer3 e NAT. Gestisce tutte le reti e il routing dei progetti. Neutron opera sopra Open vSwitch e crea tre bridge di rete: *br-int*, *br-tun* e *br-ex*. Il bridge *br-int* collega tutte le VM, *br-tun* connette le VM alla NIC fisica dell'hypervisor. Il bridge *br-ex* collega le VM alle reti esterne utilizzando i floating IP [83].
- *Nodo storage.* è responsabile della gestione dello storage e può ospitare servizi come Cinder (block storage), Swift (object storage) o di virtualizzazione dello storage come Ceph. Il ruolo è presente nell'architettura quando lo storage è separato su server specifici, oppure può coesistere con altri ruoli in caso venga utilizzata una configurazione HCI, detta anche iperconvergente<sup>62</sup>.

---

<sup>61</sup> Progetto open source che offre un'implementazione software di uno switch virtuale di classe enterprise. Link: <https://www.openvswitch.org>.

<sup>62</sup> L'infrastruttura iperconvergente o HCI (*Hyper Converged Infrastructure*) combina in un unico sistema tutti gli elementi del data center tradizionale: storage, elaborazione, networking.



#### 4.3.6 Servizi del cloud controller

Per garantire resilienza e tolleranza ai guasti, alcune funzioni di gestione sono centralizzate su più *control plane*, anche definiti *cloud controller* [81, pp. 72-88] che lavorano in configurazione HA.

Il cloud controller gestisce l'autenticazione e l'invio di notifiche a tutti i sistemi tramite una coda di messaggi, esegue il servizio di identità, alcune porzioni dei servizi Compute e Network, alcuni agent di rete, nonché servizi di supporto al cloud. L'elenco che segue descrive maggiori dettagli.

- *Database*: il controller mantiene un'istanza di database per ciascun servizio con le informazioni su tutti gli oggetti (es. utenti, VM, reti, volumi).
- *Servizi di Message Queue*: il controller esegue i servizi RabbitMQ-server e opera come broker per la gestione delle code di messaggi. Lo scambio di notifiche tra i vari componenti, basato sul protocollo AMQP (*Advanced Message Queuing Protocol*) è essenziale per coordinare le operazioni sul cloud.
- *Servizi di Conductor*: svolgono il ruolo di proxy verso il database, coordinando l'accesso e inoltrando le richieste dei vari servizi. Nella configurazione di default OpenStack, ad esempio tutti i compute node non comunicano direttamente con la coda dei messaggi e il server del database ma usano un servizio sul controller chiamato *nova-conductor*.
- *Autenticazione e autorizzazione*: il controller opera come identity provider eseguendo il componente centralizzato Keystone per l'intera gestione delle identità e degli accessi. Una volta effettuata l'autenticazione dell'utente, viene generato un token, valido per un periodo di tempo limitato, che fornisce all'utente l'autorizzazione per accedere ai servizi OpenStack.
- *Servizi di gestione delle immagini*: archivia e fornisce le immagini dei sistemi operativi per l'utilizzo nel cloud.
- *Servizi di scheduling*: il controller coordina le richieste di allocazione di risorse e le indirizza ai servizi appropriati. Il componente principale per la gestione del servizio di scheduling è Nova, che permette di pianificare e distribuire le VM in modo efficiente tra i nodi fisici disponibili.
- *Dashboard utente*: fornisce una console web based che consente agli utenti di utilizzare e gestire i servizi cloud OpenStack.
- *Endpoint API*: il servizio di identity provider mantiene il catalogo dei servizi e il controller consente l'accesso alle API REST di ciascuno di essi.

Il cloud controller gestisce un numero elevato di servizi diversi e deve essere in grado di reggere la quantità di traffico. Per un opportuno dimensionamento, nelle architetture di riferimento è di solito consigliato di avere per ciascun server almeno due interfacce da 10 Gbps per tutte le connessioni di rete tra il server e gli switch di rete [65].

### 4.3.7 Configurazione di rete

L'architettura di riferimento richiede in genere la disponibilità di almeno due reti fisiche separate: una rete per IPMI<sup>63</sup>, usata per il controllo dei server fisici, e una rete per il traffico dati [65], [81]. Sulla rete dati vengono convogliate diverse tipologie di traffico, che è opportuno isolare a livello 2 con l'impiego di vLAN.

- *Admin* network: rete usata per l'accesso amministrativo ai servizi.
- *Management* (o *internal*) network: rete privata su cui avviene la comunicazione tra i principali servizi OpenStack.
- *API* (o *public*) network: rete frontend pubblica per i servizi del cloud controller (ad es. accesso REST API verso i servizi, traffico per il database, dashboard).
- *Data* network: rete privata gestita dal servizio Neutron, veicola il traffico tra le VM e tra le VM e i servizi OpenStack. Isola il traffico dei diversi progetti e lo separa dal percorso usato per la comunicazione tra i componenti OpenStack.
- *External* network: rete esterna gestita dal servizio Neutron. Usa IP instradabili, fornisce connettività dall'esterno verso le VM e dalle VM verso Internet.
- *Storage data* network: rete privata utilizzata per il traffico tra gli Storage Node e i Compute node per accedere allo storage a blocchi o oggetti.
- *Storage cluster* network: rete privata utilizzata per replicare i dati storage tra i Ceph Storage Node.

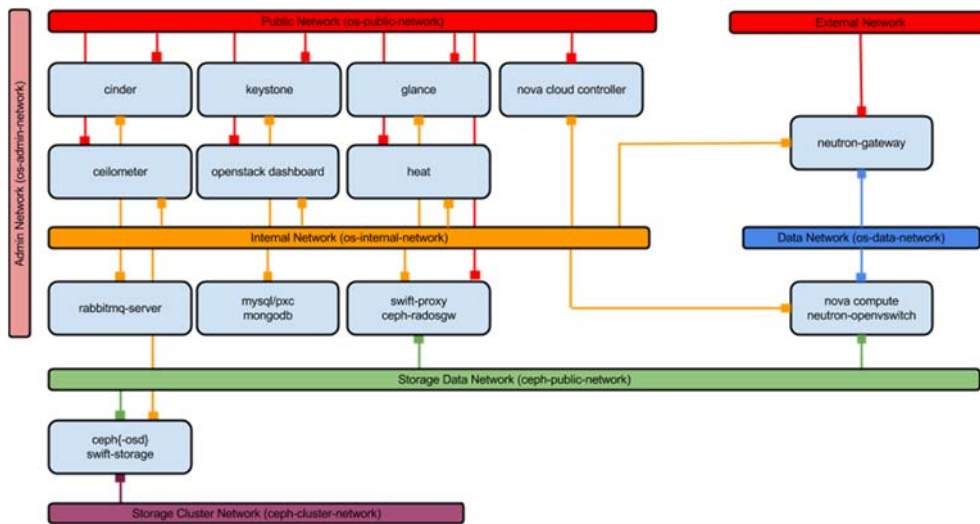


Figura 36. OpenStack: schema delle reti.

Fonte [81]

<sup>63</sup> Intelligent Platform Management Interface è un'interfaccia di gestione per monitorare lo stato fisico e la configurazione hardware di un server. Sui server di fascia enterprise c'è una NIC. L'architettura di riferimento richiede in genere la disponibilità di almeno due reti fisiche dedicate sulla scheda madre (ILO di HPE, iDRAC di Dell).

I nodi di rete Neutron vengono distribuiti su server fisici dedicati (network node), su cui gli agent L3 sono configurati in HA per garantire la ridondanza della rete virtuale. La separazione del traffico tra i vari progetti si ottiene con una configurazione di VLAN distinte e, all'interno dei network node fisici, spazi *netSPACE* e virtual router separati a cui fanno capo i gateway delle reti private dei progetti. Infine il traffico tra gli host fisici viene mantenuto diviso utilizzando meccanismi di incapsulamento o reti virtuali sovrapposte con i protocolli di tunnelling/overlay come VXLAN o GRE<sup>64</sup> [65].

Nella configurazione OpenStack le VM sono in esecuzione in una sottorete IP interna; ciascuna è dotata di un indirizzo IP privato per la comunicazione con le VM dello stesso progetto. Le VM sono raggiungibili attraverso il NAT di Neutron, se la sessione inizia dall'interno, oppure possono essere contattate direttamente dal mondo esterno solo se dispongono di un indirizzo instradabile chiamato *floating IP*. La figura che segue, tratta dal cluster di test, mostra la divisione tra rete interna (default) e rete floating.

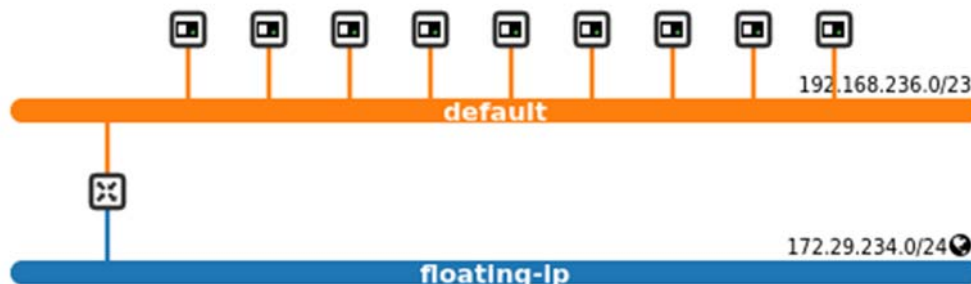


Figura 37. OpenStack: rete interna e rete floating

Fonte: ambiente di test OpenStack

#### 4.3.8 Caratteristiche per la scalabilità

OpenStack, progettato puntando alla robustezza e alla scalabilità, può scalare fino a migliaia di nodi, come dimostrano alcuni test riportati dalla documentazione ufficiale [84] e l'implementazione presso il CERN [74], [79]. L'estensione su diversi data center e su altre zone geografiche deve comunque garantire sia la ridondanza dei componenti sia una bassa latenza nei servizi endpoint API. Si utilizzano a questo scopo i meccanismi offerti dal servizio Nova quali *availability zone*, *host aggregate*, *region* e *cell* [81, pp. 109-113] che sono schematizzati in Figura 38.

<sup>64</sup> Due protocolli di incapsulamento. GRE incapsula i pacchetti a livello 2 con un nuovo header. VXLAN trasporta frame Ethernet all'interno di pacchetti UDP/IP, quindi una VXLAN può anche essere vista come una rete di livello 2 (overlay) su una rete di livello 3 (la rete IP underlay).

- *Availability zone (AZ)* o zona di disponibilità. È un raggruppamento di nodi compute sulla base dei domini di errore, concetto simile esistente a livello hardware (*failure domain*). Ad esempio, una AZ può comprendere tutti i nodi di calcolo situati in un determinato armadio rack, oppure alimentati dalla stessa PDU, o ancora collegati allo stesso switch *Top-of-Rack* (ToR). Un compute node può appartenere a una unica AZ e il dominio di errore è costituito dalla dipendenza comune da un'unica risorsa infrastrutturale.
- *Host Aggregate* o gruppo di host. È un insieme di compute node accomunati da caratteristiche particolari o funzionalità specializzate, come ad esempio una classe di processore o schede di rete più performanti, una specifica classe di storage disco oppure un sistema operativo hypervisor diverso.
- *Cell* o cella. Consente di estendere il cloud separando i nodi di calcolo in gruppi ma mantenendo un singolo endpoint API. Nella configurazione OpenStack di default, tutti i compute node devono comunicare con la coda dei messaggi e il database centralizzato tramite il servizio *nova-conductor* sul controller (si veda il paragrafo 4.3.5). La crescita del cloud può generare un carico eccessivo sulle code dei messaggi e sul database e conseguenti colli di bottiglia. Le celle Nova permettono di dividere tale carico: ogni cella esegue il proprio database, una coda di messaggi e un set completo di servizi compute Nova, condividendo solo il servizio endpoint API con le altre celle.
- *Region* o regione. È un raggruppamento di risorse, non necessariamente su base geografica, che ha un'installazione Nova completa, con i propri compute node e un endpoint API. Tutte le regioni Nova del cloud OpenStack condividono lo stesso servizio Keystone per l'autenticazione e il catalogo dei servizi API.

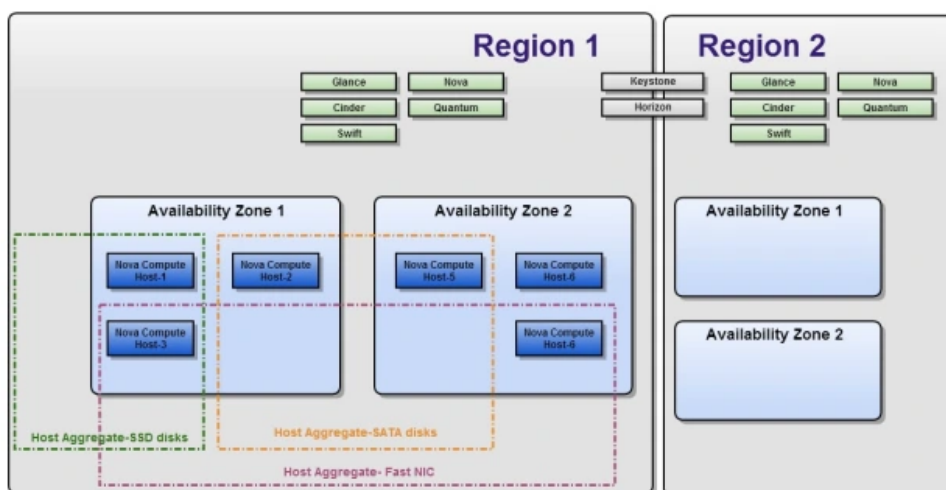


Figura 38. OpenStack: schema di architettura multi-region.

Fonte [65]

## 4.4 L'architettura federata GARR

Dopo aver descritto l'architettura del modello cloud GARR e analizzato le principali caratteristiche di OpenStack, si affronta con maggiore dettaglio la struttura federata.

### 4.4.1 Il modello multi-region

OpenStack non dispone di strutture native per la creazione di federazioni tra diversi provider di servizi cloud, quindi la soluzione ideata da GARR si fonda sull'ipotesi di federare sistemi cloud compatibili tra loro e implementati in modo coordinato. In particolare viene definita un'architettura OpenStack multi-region (si veda la Figura 19) modellando il concetto di regione Nova per raggruppare le risorse che ciascun ente decide di impegnare nella federazione [62], [65].

GARR ha identificato una classe di moduli per offrire servizi comuni, da configurare in alta affidabilità geografica: il servizio di identità Keystone, il repository di immagini Glance e lo storage a oggetti. Tali servizi, definiti globali, sono distribuiti su tre data center per garantire resilienza in caso di un guasto completo di un sito. Tutti gli altri moduli OpenStack (ad es. compute, network e block storage) appartengono a installazioni indipendenti, collocate in diversi data center, ciascuna a formare una regione nel cloud. Il concetto di regione viene anche usato come base per la federazione cloud: ogni regione federata ha la sua installazione OpenStack indipendente, ad esempio gestita in autonomia da enti federati, su cui sono in esecuzione i servizi di endpoint API, compute e network separati.

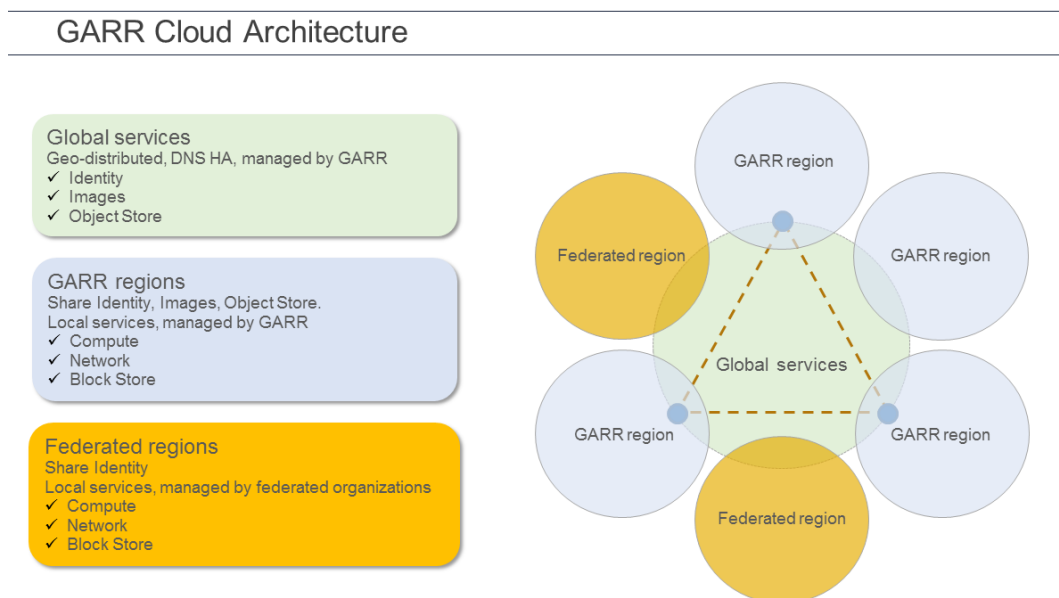


Figura 39. Modello federato GARR: servizi globali e locali.

Fonte [64]

Come descritto nel paragrafo 4.3.8, in Nova le regioni condividono lo stesso servizio Keystone e il catalogo dei servizi API.

Nel caso del modello GARR, le regioni oltre al servizio di identità condividono la dashboard Horizon che fornisce l'interfaccia web di gestione; opzionalmente possono anche condividere il repository di immagini e l'object store [63], [64]. In questo modo le organizzazioni partner possono utilizzare servizi comuni come l'autenticazione centralizzata, mantenendo al contempo l'autonomia sulle proprie risorse; inoltre il servizio di autenticazione a livello globale permette agli utenti di essere abilitati a utilizzare le risorse dell'intera federazione, accedendo con un unico set di credenziali. Il sistema di autenticazione consente l'utilizzo delle credenziali fornite dai provider di identità della federazione IDEM<sup>65</sup> e dai provider OpenID Connect (OIDC)<sup>66</sup> [63].

Riprendendo le definizioni dei modelli di deployment descritte nel capitolo 2, emerge che tale schema federato combina i modelli Private e Community cloud. Come nel Private cloud le risorse, collocate in una rete privata, sono ad uso esclusivo di un'organizzazione, su cui ricadono la gestione e il compito di controllare i sistemi. L'autenticazione centralizzata, il catalogo degli endpoint e la console di gestione sono tuttavia risorse condivise, caratteristica propria del Community cloud. Inoltre, in virtù del servizio di autenticazione comune, l'accesso alle risorse private di un ente può essere concesso anche agli utenti appartenenti ad altre organizzazioni federate, che comunque appartengono tutte alla più ampia comunità delle università e degli istituti di ricerca.

#### 4.4.2 Concetti di Keystone

Si ritiene utile riportare brevemente alcuni concetti utilizzati da Keystone che sono alla base dei meccanismi utilizzati dalla federazione [81, pp. 76-80].

- *Users* o utenti: rappresentano gli account che accedono alla piattaforma OpenStack. Ogni utente ha un nome e una password oppure usa altri metodi di autenticazione per accedere alle risorse del cloud; può avere ruoli specifici associati all'interno dei progetti.
- *Projects* o progetti: sono raggruppamenti di risorse (ad esempio VM, immagini, reti, volumi) che fungono da spazi di lavoro logici all'interno di OpenStack.
- *Domains* o domini: definiscono i confini amministrativi per la gestione delle entità. Un dominio può rappresentare una divisione logica tra diverse parti di

---

<sup>65</sup> IDentity Management per l'accesso federato: è la Federazione italiana di Infrastrutture di Autenticazione e Autorizzazione della comunità dell'istruzione e della ricerca, gestita da GARR. Consente di accedere ai servizi federati usando le proprie credenziali istituzionali (es. accesso a riviste elettroniche, servizi Wi-Fi).

<sup>66</sup> OpenID Connect è lo standard di autenticazione più diffuso per le applicazioni web e mobile, basato sul protocollo OAuth 2.0.

un'organizzazione oppure uno spazio di proprietà di un ente o di un operatore; ad alto livello è visto come un contenitore di progetti, utenti, gruppi.

- *Roles* o ruoli: definiscono le autorizzazioni e i privilegi all'interno di un progetto, quindi quali operazioni possono essere svolte e quali invece non sono consentite.
- *Authentication token* o token di autenticazione. Quando un utente viene autenticato da Keystone, gli viene generato un token con una determinata validità temporale. Tale token viene utilizzato quando l'utente cerca di accedere alle risorse per verificarne l'identità e autorizzare l'accesso.

I progetti non possiedono utenti, ma utenti o i gruppi di utenti hanno accesso a uno o più progetti utilizzando il concetto di assegnazioni di ruolo (RBAC, *role-based access control*). Il ruolo definisce il tipo di accesso, ad esempio *administrator*, *member*, *guest* e determina quali operazioni sono permesse. Al di là dei ruoli predefiniti, è possibile creare ruoli personalizzati in caso di esigenze specifiche.

Ogni servizio OpenStack definisce le policy di accesso per le proprie risorse in un file di policy. Per risorsa si può intendere qualsiasi oggetto collegato all'operatività di un servizio: ad esempio un accesso API, la possibilità di avviare una VM, di collegare un volume o una porta di rete [81, p. 232]. Le regole individuali che definiscono quali utenti e ruoli possono accedere a quali risorse sono specificate in formato JSON; ogni servizio dispone di un proprio file *policy.json* che viene analizzato da un motore di policy eseguito dal servizio stesso.

#### 4.4.3 La delega amministrativa

OpenStack opera con due domini predefiniti: il *service domain* che contiene gli account utilizzati dai servizi e lo *users domain* che contiene gli account per gli utenti. Ogni componente detiene un set di credenziali per un account di servizio: lo username è noto, Keystone genera una password, che salva in hash nel proprio database. Per il loro funzionamento tutti i servizi, sia quelli di altri cloud sia i servizi GARR, validano il proprio token verso Keystone.

Per realizzare la federazione, GARR ha definito una procedura che consiste nell'esportazione dal Keystone locale delle credenziali e degli endpoint API della regione e la successiva importazione nel Keystone centralizzato. Per garantire la sicurezza delle credenziali il meccanismo è cifrato; vengono solo trasferiti gli hash, mentre le credenziali rimangono eventualmente in chiaro solo nei file di configurazioni dei singoli servizi locali alla regione [63]. In ogni installazione di OpenStack i servizi dei vari progetti hanno sempre la stessa denominazione (es. nova, neutron, glance). Per evitare sovrapposizioni e riuscire a distinguere i servizi delle diverse regioni, quando si installa una nuova regione ogni servizio viene rinominato specificando un opportuno prefisso da assegnare agli utenti e ai progetti di servizio, attraverso i parametri di configurazione *service-admin-prefix* e *service-tenant* del charm Keystone. Prendendo l'esempio della

regione Polito, con il comando `juju config keystone service-admin-prefix=polito_` i nomi utente dei servizi saranno `polito_nova`, `polito_neutron` e così via. Con il comando `juju config keystone service-tenant=services_polito keystone` viene creato un progetto `services_polito` a cui gli utenti saranno associati).

Per ogni regione viene creato un dominio Keystone separato, attraverso il quale l'amministratore locale può gestire le risorse. Le attività sono delegate seguendo una modalità gerarchica [63] come riportato in Figura 40. L'amministratore del cloud può:

- creare, aggiornare, eliminare, elencare regioni e amministratori di regioni;
- creare, aggiornare, eliminare, elencare domini e utenti;
- impostare le quote per i domini;
- concedere, revocare, elencare ruoli nei domini.

L'amministratore del cloud delega per ogni regione un amministratore decentrato a svolgere le seguenti attività, limitatamente alla propria regione:

- aggiungere, rimuovere, attivare, disattivare hypervisor (compute node);
- aggiungere, rimuovere reti esterne;
- concedere, revocare, elencare ruoli agli amministratori di progetto;
- impostare le quote per i progetti principali (virtual datacenter);
- creare, aggiornare, eliminare, elencare progetti;
- concedere, revocare, elencare ruoli agli utenti e su progetti del proprio dominio.

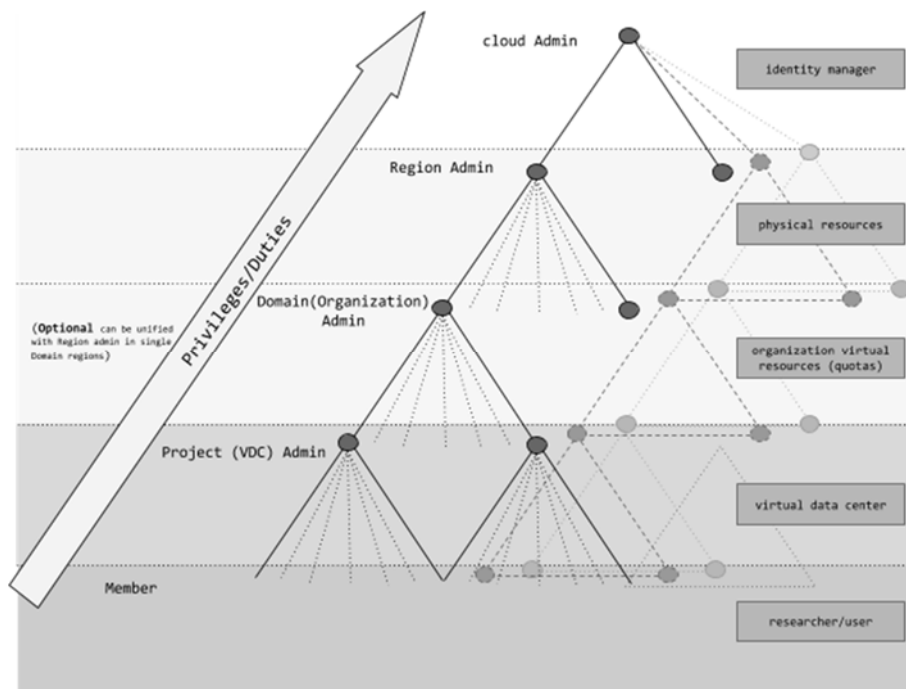


Figura 40. Modello federato GARR: schema di delega amministrativa.

Fonte [63]



Ogni regione federata ha pertanto il proprio dominio Keystone, all'interno del quale possono essere organizzati vari progetti, anch'essi in gerarchia. L'amministratore del cloud definisce le quote per i domini, specificando la quantità massima di una determinata risorsa che ciascun dominio ha a disposizione (ad es. numero VM, vCPU, RAM, volumi). L'amministratore del dominio può distribuire tale quantità tra tutti i progetti del proprio dominio, sfruttando il meccanismo chiamato quote nidificate (*nested quotas*) che consente di assegnare impostare una quantità di quote del progetto genitore ai relativi sotto progetti [63], [65].

Un esempio di quote è fornito dall'output del comando *openstack quota show* mostrato di seguito (il valore -1 rappresenta una quota illimitata).

```
# openstack quota show --fit-width
```

Field	Value
backup-gigabytes	1000
backups	10
cores	1600
fixed-ips	-1
floating-ips	80
gigabytes	81920
gigabytes_SSD	-1
groups	10
injected-file-size	10240
injected-files	5
injected-path-size	255
instances	150
key-pairs	100
networks	10
per-volume-gigabytes	-1
ports	250
ram	5242880
rbac_policies	10
routers	2
secgroup-rules	1250
secgroups	125
server-group-members	10
server-groups	10
snapshots	50
snapshots_SSD	-1
snapshots__DEFAULT__	-1
subnet_pools	-1
subnets	10
volumes	125
volumes_SSD	-1
volumes__DEFAULT__	-1

#### 4.4.4 Nuove linee di indirizzo

Durante la stesura del presente lavoro, dall'8 al 10 novembre 2023 si è svolto a Roma il convegno *Net Makers*, l'annuale workshop della comunità GARR<sup>67</sup>. Nell'ambito infrastrutture e servizi, oltre alla presentazione di esperienze e casi d'uso, sono state introdotte soluzioni di accounting e monitoraggio per le infrastrutture cloud ma soprattutto sono state delineate le nuove linee di indirizzo con particolare attenzione all'evoluzione delle tecnologie cloud offerte.

Nella presentazione *Evoluzione del modello cloud GARR*, il vicedirettore Carboni ha ribadito che rimane tra gli obiettivi mantenere ed evolvere il sistema cloud, ma che tuttavia è necessaria una revisione del modello tecnico e operativo. In particolare ha esposto i principali obiettivi per il consolidamento dell'infrastruttura OpenStack [85]:

- razionalizzazione del numero di regioni, introducendo una soglia minima di massa critica delle risorse per la creazione di una regione;
- razionalizzazione della federazione (per un approccio più strutturato in termini di protocolli) sull'esperienza di IDEM e di altre federazioni;
- messa in produzione del sistema di monitoring, account e reportistica;
- revisione della proposta tecnico/economica, che definisca il ruolo primario dei soci GARR come principali beneficiari dell'infrastruttura cloud;
- definizione di un modello economico da applicare agli utenti della comunità GARR che tenga conto dei costi GARR oltre che dell'infrastruttura;
- evoluzione dell'infrastruttura IaaS al fine di supportare soluzioni Cloud Native;
- attività di sviluppo *green field* per consentire la messa in funzione congiunta dell'infrastruttura mista di calcolo, storage a supporto dei diversi servizi di tipo IaaS, PaaS, SaaS.

È opinione di Carboni che molti soci di GARR possano contare su risorse in grado di coprire la propria domanda interna e di altri soggetti afferenti, quindi per GARR è fattibile snellire i processi amministrativi e tecnici, avendo come interlocutore diretto le organizzazioni strutturate piuttosto che i singoli membri della comunità R&E. Identifica inoltre l'approccio cloud native come focus su cui investire, sviluppando la transizione verso soluzioni in container e potenziando la componente Kubernetes nell'ottica di creare un modello agnostico rispetto al livello sottostante delle tecnologie di virtualizzazione. Molte organizzazioni hanno già aderito a soluzioni di cloud pubblico quindi elementi quali semplicità, riuso, condivisione e standardizzazione dei container potrebbero agevolare la costituzione di infrastrutture di tipo ibrido o multi cloud. In linea a quanto analizzato nel capitolo 2 (si vedano 2.2 e 2.8) è su questi modelli che GARR vede incentrato il futuro del cloud.

---

<sup>67</sup> GARR, concluso 'Net Makers', link: <https://www.garr.it/it/news-e-eventi/2279-concluso-net-makers-il-workshop-annuale-della-comunita-garr>

#### 4.4.5 Federazione e autonomia

L'attuale organizzazione della struttura federata e l'autenticazione centralizzata aprono alla possibilità di collaborazione anche in termini di scambio di risorse. Agli utenti registrati, a tutti gli effetti utenti globali del cloud, possono essere assegnate risorse in tutte le regioni. Questo consente agli enti di mettere a fattore comune eventuali risorse private, ad esempio accordandosi per ospitare o trasferire VM in un'altra regione nel caso di esigenze temporanee.

La prima regione federata è stata quella del Politecnico di Torino. Si sono aggiunte successivamente due regioni: l'Istituto Nazionale di Geofisica e Vulcanologia (INGV) e la più recente dell'Istituto Zooprofilattico Sperimentale (IZS) dell'Abruzzo e Molise.

Oltre a ciò, anche senza aderire alla federazione, il modello di deploy progettato da GARR può essere impiegato come base da seguire per la creazione di cluster OpenStack indipendenti. Ha scelto questa opzione il Dipartimento di Informatica dell'Università di Torino, implementando un'infrastruttura OpenStack e Kubernetes creata con Juju nell'ambito dell'iniziativa HPC4AI<sup>68</sup>, un laboratorio per *High-Performance Computing e Artificial Intelligence*.

Anche INGV, oltre alla regione integrata nel Cloud GARR, ha installato e gestisce un cluster indipendente sulla base dello stesso modello di deploy. In particolare si tratta di un cloud OpenStack per l'esecuzione dei servizi critici di una sala operativa virtuale per il monitoraggio sismico. Nel 2020, con il supporto del Dipartimento CSD di GARR è stato progettato un sistema multi-region, con opportuni meccanismi di replica e nodi distribuiti nei data center di Roma, Napoli e Catania, con l'obiettivo di garantire la ridondanza geografica sul territorio e preservare l'allineamento dello stato e delle configurazioni anche in caso di eventi naturali avversi [80].

Infine, nel capitolo che segue viene dato spazio all'implementazione della regione OpenStack in esecuzione presso il Politecnico di Torino, analizzando le caratteristiche dell'infrastruttura e i processi di installazione.

---

<sup>68</sup> Progetto HPC4AI, link: <https://hpc4ai.unito.it/#systems>

# Capitolo 5.

## Il cloud OpenStack Polito

Tutte le università pubbliche italiane, a seguito del censimento effettuato da AgID, hanno visto i propri data center classificati nel gruppo B. Solo due atenei rappresentano un'eccezione, l'Università di Pisa<sup>69</sup> e l'Università di Genova<sup>70</sup> che dispongono di strutture dotate delle misure organizzative e tecniche tali da afferire al gruppo A (per maggiori dettagli si rimanda al capitolo 3).

Il Politecnico di Torino, in qualità di pubblica amministrazione (PA), si trova pertanto nella condizione di dover redigere e presentare il piano di migrazione al cloud per le infrastrutture dei servizi amministrativi, mentre può pianificare scenari evolutivi nei propri data center on-premise per le attività correlate alla ricerca.

L'obiettivo di ampliare l'offerta dei servizi di ateneo ai dipartimenti e centralizzare le risorse hardware destinate alla ricerca ha portato la Divisione Area IT del Politecnico di Torino a valutare alcune soluzioni di private cloud. Nell'ambito dei sistemi erano già in funzione cluster per la fornitura di VM, sia utilizzando strumenti open source come oVirt sia con soluzioni proprietarie come VMware vSphere. Il modello GARR è stato scelto perché offre un'architettura replicabile e progettata per la stessa tipologia di contesto, priva inoltre di costi di licenza in quanto basata su software open source.

È stato quindi costituito un gruppo di tecnici, selezionati tra il personale dell'Area IT e il personale afferente ai dipartimenti di Informatica e di Elettronica, per avviare la fase di studio del private cloud OpenStack con la collaborazione degli specialisti del dipartimento CDS di GARR. Dopo un'implementazione pilota, il cluster OpenStack è diventato operativo e la gestione è stata presa in carico da Area IT.

Da novembre 2022, a seguito di una riorganizzazione delle strutture amministrative dell'ateneo, la Divisione Area IT è diventata Direzione ISIAD - Infrastrutture Servizi Informatici e Amministrazione Digitale.

---

<sup>69</sup> L'infrastruttura del data center UniPi, link <https://start.unipi.it/computing/infrastructure/>

<sup>70</sup> L'infrastruttura IT di UniGe, link: <https://openscience.unige.it/genOAwk2021/materiali>

## 5.1 Progettazione infrastruttura e servizi OpenStack

Le attività per la creazione di una regione OpenStack federata con GARR sono state suddivise nelle seguenti fasi operative:

14. installazione e configurazione degli strumenti di automazione MAAS e Juju su server Polito ad hoc;
15. registrazione delle risorse fisiche (nodi server, reti, VLAN, storage) nel catalogo MAAS;
16. deploy via Juju dei componenti OpenStack sul gruppo di server destinati a costituire la regione Polito;
17. trasferimento delle credenziali di sistema e importazione degli endpoint API Polito (procedura con GARR), aggiunta di endpoint Keystone centralizzato.

### 5.1.1 Hardware

Dal punto di vista hardware, in uno dei due data center presso la sede centrale del Politecnico di Torino è stata allestita un'isola dedicata con 6 armadi rack. L'isola è alimentata da due circuiti separati, uno connesso ad un sistema di alimentazione ininterrotta (UPS) e l'altro connesso un gruppo elettrogeno a gasolio, in grado di garantire una fornitura di energia stabile ed evitare interruzioni improvvise.

Ciascun armadio rack è dotato di due alimentazioni verso i due circuiti separati e anche i server sono provvisti di alimentazioni ridondate e bilanciate. Dal punto di vista network, ogni server è collegato a due switch diversi in modo da garantire la continuità delle comunicazioni e prevenire eventuali interruzioni di rete causate da guasti agli apparati di rete.



Figura 41. OpenStack Polito: isola rack dedicata nel data center.

### 5.1.2 Pianificazione dei servizi OpenStack e ruoli server

Come descritto nel capitolo 4, i servizi OpenStack sono distribuiti su diversi nodi e la logica di controllo del cluster è basata su container in HA, ognuno dei quali viene replicato sui tre controller node. È stata scelta un'installazione iperconvergente con la coesistenza dei ruoli storage e compute sullo stesso hardware [83].

Per il progetto iniziale è stato impiegato un primo blocco formato da 11 server per i servizi OpenStack; gli strumenti di automazione e gestione del cloud MAAS e Juju, sono stati installati su 2 server dedicati. Nel tempo si sono aggiunti un'infrastruttura di test e ulteriori server, fino a raggiungere l'attuale numero di 33, le cui caratteristiche hardware sono riportate nella Tabella 8. Di seguito le tre tipologie di nodo:

- 3 *controller node* che ospitano i servizi centralizzati in HA (servizi di scheduling Nova, database, mySQL-Innodb, coda di messaggi *RabbitMQ*, endpoint API, Keystone, Nova Controller, immagini Glance, dashboard Horizon) oltre ai servizi per lo storage Ceph: il control plane *Ceph-mon*;
- 2 *network node* su cui sono in esecuzione gli agent per i servizi di rete Neutron: DHCP, metadata, Layer2, Layer3 e NAT;
- 28 *compute node* che hanno il ruolo di KVM hypervisor (Nova compute) per l'esecuzione delle VM; ospitano inoltre i servizi di gestione dischi *Ceph-osd*.

Per schematizzare le tipologie di ruolo e la dislocazione dei diversi servizi OpenStack nell'installazione cloud Polito, è stato realizzato il seguente diagramma.

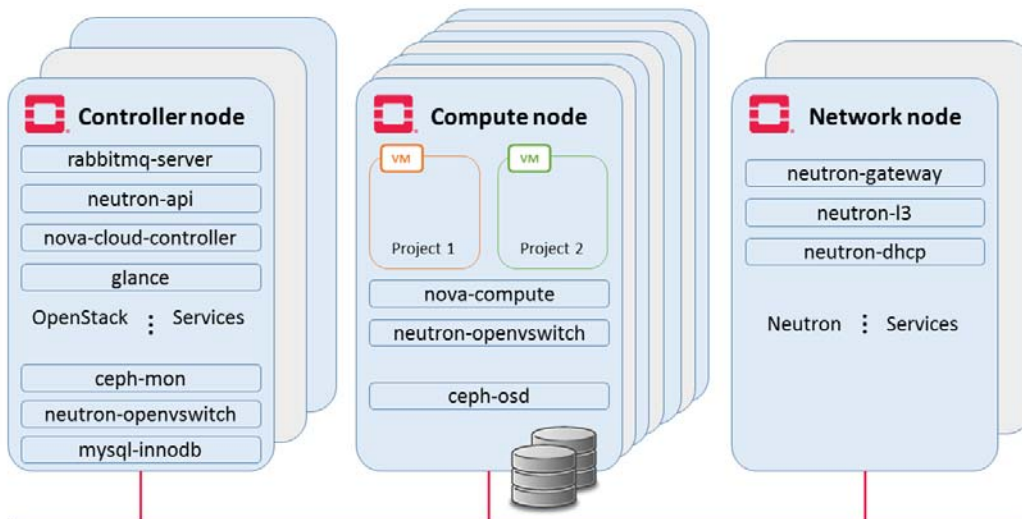


Figura 42. OpenStack Polito: ruoli e servizi sui nodi.

Tabella 8. OpenStack Polito: caratteristiche hardware dei server

Caratteristiche	Descrizione
Marca e modello	11x HPE Gen10 ProLiant DL580 e 22x HPE ProLiant DL560
CPU	4x Intel Xeon Gold 6140M (2.3GHz, 18Core/36Thread, 140W)
Memoria RAM installata	1 TB (256GB per ogni CPU)
RAID controller	PERC H740P 8Gb
Dischi interni installati	2x SSD da 1.9TB e 22x HDD SAS da 1.8TB Endurance DWPD

Il modello GARR ha messo a disposizione un'architettura replicabile e progettata per la stessa tipologia di contesto. Tuttavia nell'implementazione al Politecnico di Torino per alcuni componenti sono state effettuate scelte architetturali diverse per armonizzare le esigenze e la disponibilità di hardware.

Dal punto di vista storage, il dipartimento CSD di GARR gestisce un cluster Ceph su hardware ad hoc mantenendo i ruoli compute e storage separati. Per il cloud Polito è stata preferita una configurazione iperconvergente, che utilizza server dotati di ampie risorse CPU, RAM e un numero considerevole di dischi per eseguire sugli stessi nodi sia i servizi compute sia quelli storage di Ceph. Mentre GARR utilizza il servizio Nova in container, per il cloud Polito si è valutato di installare tale servizio bare metal, mantenendo invece in container i servizi Ceph-osd, per ottimizzare il ruolo ibrido compute e storage node. La Tabella 9 riassume il deploy dei servizi OpenStack sui diversi tipi di nodo all'interno del cloud Polito.

Tabella 9. Ruoli dei nodi e distribuzione dei servizi OpenStack

#Server	Tipologia	Descrizione e servizi in esecuzione
3	Controller node	Servizi centralizzati, eseguiti in HA: database <i>mySQL-InnoDB</i> , servizi di scheduling <i>Nova</i> , endpoint API per l'accesso REST ad ogni servizio, message queue <i>RabbitMQ</i> , dashboard web based <i>Horizon</i> , servizi di gestione delle immagini <i>Glance</i> , servizi di storage <i>Ceph-monitor</i>
2	Network node	Servizi di rete <i>Neutron</i> necessari al funzionamento del cloud: virtual switch, routing Layer3, DHCP, NAT, metadata agent
28	Compute node	Servizi <i>Nova</i> , hypervisor KVM per la gestione del ciclo di vita delle VM, servizi <i>Neutron</i> locali per le VM, servizi di storage distribuito <i>Ceph-osd</i>

### 5.1.3 Risorse disponibili e overcommit

Si riepilogano in questo paragrafo le risorse complessive, considerando che al momento l'infrastruttura cloud risulta composta da 35 server:

- RAM: 30TB
- core CPU: oltre 4.500
- storage: 1,2PB HDD e 85TB SSD

L'acquisto di ulteriori 30 server, in attesa di consegna, porterà a raddoppiare le risorse a disposizione entro la fine del 2024.

Le indicazioni quantitative riportate sono da considerarsi lorde. Una parte delle risorse, soprattutto di CPU e RAM, è utilizzata per l'esecuzione dei servizi stessi di OpenStack, Ceph e per gli hypervisor. Ad esempio Ceph ha un overhead di circa 4GB per ogni demone osd in esecuzione (uno per ogni disco sul sistema) e ogni nodo compute e storage usa circa 80-90GB di RAM per i servizi che esegue.

Anche la componente storage raw su Ceph subisce una decurtazione in base alla configurazione di ridondanza applicata; non è definibile a priori perché essa dipende dalla combinazione delle policy applicate sui diversi pool (replica N oppure erasure coding, si veda il paragrafo 4.2.3). In una semplificazione approssimativa, impiegando solo pool Ceph con replica 3, lo spazio effettivamente utilizzabile dai servizi e dalle VM si ridurrebbe a un terzo.

OpenStack permette di gestire le risorse CPU e RAM dei compute node in modalità *overcommit*, è consentito cioè assegnare alle VM una quantità di risorse superiore rispetto alle risorse fisiche effettivamente presenti. L'overcommit si basa sulla ricollocazione dinamica di risorse inutilizzate e in stato inattivo; ha il vantaggio di aumentare il numero di potenziali VM in esecuzione ma può anche avere un impatto negativo sulle prestazioni sia delle VM stesse sia degli hypervisor [81, pp. 118-120].

Il servizio Nova presenta dei valori di default per l'overcommit (*allocation ratio*), recentemente modificati dal team di sviluppo e diventati effettivi a partire dalla versione OpenStack Yoga<sup>71</sup>. Il rapporto di allocazione CPU è passato da 16:1 a 4:1; lo scheduler può quindi allocare fino a 4 core virtuali per ciascun core fisico. Ad esempio, nel caso di un server del cloud Polito dotato di 4 processori Intel Xeon Gold 6140M con 18 core e hyperthreading (si veda la Tabella 8) il totale dei core pari a 144 sale a 576 con il valore di allocation ratio 4:1.

Il rapporto di allocazione RAM è invece passato da 1.5:1 a 1:1; con il nuovo valore lo scheduler non può quindi superare la quantità di RAM fisica, mentre in precedenza su un server dotato di 1TB di RAM sarebbe stato possibile allocare VM fino a 1,5TB di RAM. La modifica più conservativa di tale valore è dovuto al fatto che l'overcommit

---

<sup>71</sup> Change default overcommit values, link: <https://blueprints.launchpad.net/nova/+spec/nova-change-default-overcommit-values>



della RAM è più rischioso rispetto all'overcommit di CPU. La carenza di memoria può determinare un drastico calo delle prestazioni dei carichi di lavoro, inoltre se non gestito correttamente può influire anche sui servizi degli hypervisor e causare gravi interruzioni o periodi di inattività.

#### 5.1.4 Cluster di test

Oltre all'infrastruttura di produzione è stato realizzato anche un cluster di test per avere un banco di prova su cui effettuare configurazioni e aggiornamenti, prima del deploy effettivo in produzione. L'ambiente di test è virtualizzato su un server HPE ProLiant DL560 Gen10 con

- CPU: 4 processori Intel Xeon Gold 5220 2.20GHz con 18 core
- RAM 1 TB, 256GB per ciascun processore
- Storage: 2 dischi 280GB in RAID1 per disco di boot, 1 Disco SSD 1TB come datastore per le VM, 21 dischi SAS 1,8TB: datastore a disposizione da 37,8TB.

Sul server è installato VMware ESXi 7.0.3 e sono state create le VM che riproducono l'ambiente di produzione (MAAS controller, Juju controller, 3 server con Nova compute e 2 server Neutron), con un'ulteriore VM che crea una struttura di rete isolata con il software open source OPNSense.

Juju-controller	Powered On	✓ Normal	35.08 GB	35.08 GB	197 MHz	2.04 GB
MAAS-test01-Monitor	Powered On	✓ Normal	103.08 GB	103.08 GB	307 MHz	3.05 GB
OPNSense-Upgrade	Powered On	✓ Normal	56.19 GB	37.72 GB	65 MHz	8.01 GB
Test-OS-01	Powered On	✓ Normal	396.1 GB	189.52 GB	5.86 GHz	36.25 GB
Test-OS-02	Powered On	✓ Normal	396.1 GB	193.59 GB	3.89 GHz	36.24 GB
Test-OS-03	Powered On	✓ Normal	396.1 GB	194.57 GB	4.15 GHz	36.25 GB
Test-OS-04	Powered On	✓ Normal	294.09 GB	82.7 GB	548 MHz	21.78 GB
Test-OS-05	Powered On	✓ Normal	294.09 GB	88.81 GB	943 MHz	24.12 GB
Ubuntu-Desktop	Powered On	✓ Normal	24.08 GB	24.08 GB	21 MHz	8.07 GB

Figura 43. VM dell'ambiente di test OpenStack Polito

La tabella che segue elenca le reti utilizzate nella configurazione OpenStack. È da aggiungere la rete interna *192.168.236.0/23*, usata dalle VM in esecuzione, che esiste esclusivamente all'interno di OpenStack e viene gestita da Neutron.

Tabella 10. Elenco reti nella configurazione OpenStack

Nome	Sottorete	Descrizione
ceph-priv-test	10.0.1.0/24	Rete interna per i nodi Ceph
ceph-pub-test	10.0.2.0/24	Rete comunicazione nodi Ceph e altri servizi
endpoint-test	172.29.168.0/24	IP per servizi raggiungibili dall'esterno
floating-instance-test	172.29.234.0/24	IP per VM raggiungibili dall'esterno
os-mng-test	172.29.157.0/24	IP per management
pxe-test	172.29.169.0/24	IP per il deploy dei nodi OpenStack via MAAS

## 5.2 Installazione cloud OpenStack Polito

Per avere la ridondanza degli strumenti di automazione e gestione del cloud, vengono utilizzati due server di management (su hardware HPE DL560 Gen8Intel e un HPE DL360 Gen6 con CPU Intel Xeon E5-4610 v2 2.30GHz, 256GB):

- sistema operativo Ubuntu 20.04.6 LTS;
- MAAS installato in un container LXC;
- una VM basata su KVM per Juju controller.

### 5.2.1 Installazione MAAS

L'installazione di MAAS su Ubuntu può essere effettuata utilizzando i pacchetti APT oppure via Snap<sup>72</sup>. È stata scelta la prima opzione e di seguito si riportano i principali macro passaggi della procedura di configurazione.

- installazione pacchetti e dipendenze (es. *maas-cli*, *maas-dhcp*, *maas-proxy*)
- creazione account amministrativo
- configurazione proxy, fonte delle immagini dei sistemi operativi
- importazione delle immagini di boot
- definizione delle reti (fabric, vlan, subnet e space)
- configurazione servizio DHCP.

Conclusa l'installazione, il comando *maas status* permette di verificare i servizi MAAS in esecuzione. Di seguito l'esempio tratto dal cluster di test.

```
$ sudo maas status
[sudo] password for maasadmin:
bind9          RUNNING      pid 5025, uptime 3 days, 20:33:37
dhcpd          RUNNING      pid 5259, uptime 3 days, 20:33:08
dhcpd6        STOPPED      Not started
http           RUNNING      pid 5255, uptime 3 days, 20:33:08
ntp            RUNNING      pid 5004, uptime 3 days, 20:33:39
proxy          RUNNING      pid 5018, uptime 3 days, 20:33:38
rackd          RUNNING      pid 4894, uptime 3 days, 20:33:51
regiond        RUNNING      pid 4896, uptime 3 days, 20:33:51
syslog         RUNNING      pid 5003, uptime 3 days, 20:33:39
```

Per aggiungere al catalogo MAAS i server destinati ai servizi OpenStack è stato configurato su ciascuno di essi il boot da rete via PXE<sup>73</sup>.

---

<sup>72</sup> APT (Advanced Package Tool) è il sistema di gestione dei tradizionali pacchetti deb in Ubuntu. La versione 16.04 LTS ha introdotto il supporto ai pacchetti in formato snap, di cui è possibile consultare lo store sul sito *snappycraft.io* (fonte: <https://wiki.ubuntu-it.org/>).

<sup>73</sup> PXE (*Preboot Execution Environment*) è una specifica che usa protocolli client-server per consentire l'avvio dei computer su una rete da un server centralizzato.

## 5.2.2 Aggiunta server al catalogo MAAS

I server da installare destinati a OpenStack hanno un'interfaccia connessa alla stessa sottorete che MAAS può raggiungere e gestire, pertanto al riavvio essi effettuano il boot PXE tramite MAAS. Dopo che MAAS ha effettuato il primo discovery dei nodi, i server sono etichettati come *new* e si procede con le ulteriori impostazioni:

- modifica impostazioni dell'interfaccia BMC<sup>74</sup> di tipo IPMI (es. tipo di boot, indirizzo IP, credenziali per la gestione remota);
- aggiunta di tag per le tipologie di server da installare, es. *controller-node*, *compute-node*, *network-node*, *juju-controller*;
- creazione delle interfacce sulle VLAN e configurazione dei bond per le schede
- avvio della procedura di commissioning.

Per la gestione completa dei server, MAAS dev'essere in grado di spegnere e riaccendere ciascun nodo durante la fase di rilevamento; a questo scopo vengono utilizzati i parametri dell'interfaccia IPMI (chiamata *ILO* nei server HPE, *iDRAC* nei server Dell). I tag invece vengono utilizzati per identificare i nodi con hardware specifico o che devono ricoprire un determinato ruolo; Juju può sfruttare questa funzionalità per collocare i servizi durante il deploy delle applicazioni. Per ogni server, la fase di commissioning avvia la macchina e carica una versione temporanea del sistema operativo Ubuntu in RAM. MAAS usa tale sistema operativo per effettuare una scansione della configurazione hardware: CPU, RAM, layout dischi, dispositivi PCI e USB, GPU. Al termine della procedura, se non vi sono stati errori, i server sono etichettati come *ready* [67]. Con i passaggi successivi, quando i nodi saranno installati con i servizi OpenStack, lo stato sarà *deployed* (Figura 44).

<input type="checkbox"/> FQDN   MAC IP	POWER	STATUS	OWNER TAGS	POOL NOTE	ZONE SPACES	FABRIC VLAN	CORES ARCH	RAM	DISKS	STORAGE
<input type="checkbox"/> Deployed 7 machines										
<input type="checkbox"/> Test-OS-01.maas 172.29.169.131 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest controller-node	default	default 5 spaces	fabric-0 Default V...	4 amd64	6 GiB	6	193.3 GB
<input type="checkbox"/> Test-OS-02.maas 172.29.169.132 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest controller-node	default	default 5 spaces	fabric-0 Default V...	4 amd64	6 GiB	6	193.3 GB
<input type="checkbox"/> Test-OS-03.maas 172.29.169.133 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest controller-node	default	default 5 spaces	fabric-0 Default V...	4 amd64	6 GiB	6	193.3 GB
<input type="checkbox"/> Test-OS-04.maas 172.29.169.134 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest network-node	default	default 5 spaces	fabric-0 Default V...	4 amd64	6 GiB	6	193.3 GB
<input type="checkbox"/> Test-OS-05.maas 172.29.169.135 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest network-node	default	default 5 spaces	fabric-0 Default V...	4 amd64	6 GiB	6	193.3 GB
<input type="checkbox"/> Test-OS-06.maas 172.29.169.25 (PXE) (+4)	On Vmware	Ubuntu 20.04 LTS	maas:stest compute-nod...	default	default 5 spaces	fabric-0 Default V...	4 amd64	12 GiB	6	214.7 GB
<input type="checkbox"/> juju-controller.maas 172.29.169.253 (PXE)	On Vmware	ubuntu/bionic	maas:stest virtual	default	default pxe-test	fabric-0 Default V...	2 amd64	2 GiB	1	35.4 GB

Figura 44. Cloud Polito: MAAS

<sup>74</sup> *Baseboard Management Controller* (BMC): un microcontroller sulla scheda madre dei server che implementa la logica dell'interfaccia IPMI e l'interazione tra software di gestione e hardware.

### 5.2.3 Installazione Juju

Dopo aver predisposto MAAS da utilizzare come base per Juju, sono stati configurati i due componenti: *Juju controller* e *Juju client*.

L'installazione del Juju client, il software per l'esecuzione dei comandi di gestione, è stata effettuata nel container del MAAS controller con i successivi passaggi:

- creazione del *cloud definition file* in cui inserire le opportune informazioni per consentire a Juju la gestione di MAAS come cloud;
- aggiunta del cloud con il comando *juju add-cloud*;
- creazione del file contenente le credenziali MAAS usate da Juju per autenticarsi (recuperando la chiave API dalla console di gestione MAAS);
- aggiunta del cloud con il comando *juju add-credential*.

Per il cloud Polito sono stati creati i file *maas-cloud.yaml* e *juju-credential.yaml*. Si riportano di seguito il contenuto dei file e i comandi utilizzati nell'ambiente di test.

```
$ more maas-cloud.yaml
clouds:
  MAAS-TEST:
    type: maas
    auth-types: [oauth1]
    endpoint: http://172.29.169.2:5240/MAAS/

$ more juju-credential.yaml
credentials:
  MAAS-TEST:
    maastest:
      auth-type: oauth1
      maas-oauth: *****

juju add-cloud MAAS-TEST maas-cloud.yaml
This operation can be applied to both a copy on this client and to the one
on a controller. No current controller was detected and there are no
registered controllers on this client: either bootstrap one or register
one.
Cloud "MAAS-TEST" successfully added to your local client. You will need to
add a credential for this cloud (`juju add-credential MAAS-TEST`) before you
can use it to bootstrap a controller (`juju bootstrap MAAS-TEST`) or to
create
a model (`juju add-model <your model name> MAAS-TEST`).

juju add-credential MAAS-TEST -f juju-credential.yaml
Credential "maastest" added locally for cloud "MAAS-TEST".
```

Per il *Juju controller*, il nodo di gestione dell'ambiente cloud, è stata utilizzata una VM sul server di management ed è stato istanziato un nuovo controller con il comando *juju bootstrap*. Questa procedura genera un modello Juju di base, ma successivamente verrà creato il modello per tutti i componenti OpenStack (comando *juju add-model*). Si riportano di seguito i comandi utilizzati nell'ambiente di test.

```

juju bootstrap MAAS-TEST ctrl_polito_test --constraints mem=2G --to juju-
controller.maas
Creating Juju controller "ctrl_polito_test" on MAAS-TEST/default
Looking for packaged Juju agent version 2.8.10 for amd64
Launching controller instance(s) on MAAS-TEST/default...
- amrhxy (arch=amd64 mem=2G cores=2)
Installing Juju agent on bootstrap instance
Fetching Juju Dashboard 0.6.2
Waiting for address
Attempting to connect to 172.29.169.253:22
Connected to 172.29.169.253
Running machine configuration script...
Bootstrap agent now started
Contacting Juju controller at 172.29.169.253 to verify accessibility...

Bootstrap complete, controller "ctrl_polito_test" is now available
Controller machines are in the "controller" model
  Initial model "default" added

```

Creato il controller Juju, è possibile verificare la presenza del cloud MAAS usando il comando *juju list-clouds*. È anche possibile utilizzare il comando *juju clouds --all* che elenca sia i cloud configurati (con Credentials 1) sia tutti tutti i tipi di cloud supportati da Juju anche se non configurati (con Credentials 0).

```

$ juju list-clouds
Only clouds with registered credentials are shown.
There are more clouds, use --all to see them.

```

```

Clouds available on the controller:
Cloud      Regions  Default  Type
MAAS-TEST  1         default  maas

```

```

Clouds available on the client:
Cloud      Regions  Default  Type      Cred  Source  Description
MAAS-TEST  1         default  maas      1     local   Metal As A Service
Localhost  1         localhost lxd       0     built-in LXN Hypervisor

```

```

$ juju clouds -all
Clouds available on the controller:
Cloud      Regions  Default  Type
MAAS-TEST  1         default  maas

Clouds available on the client:
Cloud      Regions  Default  Type      Cred  Source  Description
MAAS-TEST  1         default  maas      1     local   Metal As A Service
aws        21        us-east-1  ec2       0     public  Amazon Web Serv
aws-china  2         cn-north-1  ec2       0     public  Amazon China
aws-gov    2         us-gov-west1  ec2       0     public  Amazon (US Govern)
azure      40        centralus  azure     0     public  Microsoft Azure
azure-china 4         chinaeast  azure     0     public  MS Azure China
cloudsigma 12        dub        cloudsigma 0     public
google     21        us-east1   gce       0     public  Google Cloud Plat
localhost  1         localhost  lxd       0     built-in LXN Hypervisor
oracle     4         us-phoenix1  oci       0     public  Oracle Cloud Serv
rackspace  6         dfw        rackspace 0     public

```

## 5.2.4 Deploy cloud OpenStack con Juju

GARR fornisce un bundle YAML di riferimento che ricalca il modello con tutti i servizi interdipendenti di OpenStack. La versione installata è Wallaby.

Prima di effettuare il deploy sui server Polito è stato necessario personalizzare le impostazioni quali ad esempio nomenclatura, numero fisico dei nodi, nomi delle reti, oltre che apportare una serie di modifiche per adeguare l'installazione alle scelte architetturali diverse, descritte nel paragrafo 0 (Nova bare metal anziché in container, Ceph-osd sui nodi compute).

Il file di bundle è piuttosto lungo e articolato in quanto contiene le impostazioni relative a tutti i servizi da installare. Per offrire una panoramica della sintassi e della struttura, si riporta di seguito un breve estratto del bundle dell'ambiente di test.

```
applications:
  ceph-mon:
    charm: cs:ceph-mon
    channel: stable
    revision: 62
    series: focal
    num_units: 3
    to:
      - lxd:0
      - lxd:1
      - lxd:2
    options:
      config-flags: mon osd allow
      primary affinity = true
      nagios_context: nagios-polito
      nagios_servicegroups: ceph,store
      source: cloud:focal-wallaby
      annotations:
        gui-x: "1200"
        gui-y: "400"
      constraints: arch=amd64
      bindings:
        "": os-mng-test
        admin: os-mng-test
        bootstrap-source: os-mng-test
        client: os-mng-test
        cluster: ceph-priv-test
        dashboard: os-mng-test
        mds: os-mng-test
        mon: os-mng-test
        nrpe-ext-master: os-mng-test
        osd: os-mng-test
        prometheus: os-mng-test
        public: os-mng-test
        radosgw: os-mng-test
        rbd-mirror: os-mng-test
  ceph-osd:
    charm: cs:ceph-osd
    channel: stable
    revision: 316
    series: focal
    num_units: 5
    to:
      - "0"
      - "1"
      - "2"
      - "3"
      - "4"
    options:
      nagios_context: nagios-polito
      nagios_servicegroup: ceph,store
      osd-devices: /dev/sdb /dev/sdc
      /dev/sdd /dev/sde /dev/sdf
      osd-format: xfs
      source: cloud:focal-wallaby
      use-syslog: true
      annotations:
        gui-x: "1400"
        gui-y: "400"
      constraints: arch=amd64
      storage:
        bluestore-db: loop,1024M
        bluestore-wal: loop,1024M
        osd-devices: loop,1024M
        osd-journals: loop,1024M
      bindings:
        "": os-mng-test
        cluster: ceph-priv-test
        mon: os-mng-test
        nrpe-ext-master: os-mng-test
        public: os-mng-test
        secrets-storage: os-mng-test
  ceph-radosgw:
    charm: cs:ceph-radosgw
    channel: stable
    revision: 300
    series: focal
    num_units: 3
    to:
      - lxd:1
      - lxd:2
      - lxd:0
```

```

options:
  ceph-osd-replication-count: 3
  nagios_context: nagios-polito
  nagios_servicegroups:
ceph,object,storage
  region: polito
  restrict-ceph-pools: true
  rgw-buckets-pool-weight: 5
  rgw-lightweight-pool-pg-num: 8
  source: cloud:focal-wallaby
  use-syslog: true
  vip: 172.29.157.103
annotations:
  gui-x: "1000"
  gui-y: "0"
  constraints: arch=amd64

bindings:
  "": os-mng-test
  admin: os-mng-test
  certificates: os-mng-test
  cluster: os-mng-test
  gateway: os-mng-test
  ha: os-mng-test
  identity-service: os-mng-test
  internal: os-mng-test
  master: os-mng-test
  mon: os-mng-test
  nrpe-ext-master: os-mng-test
  object-store: os-mng-test
  public: os-mng-test
  radosgw-user: os-mng-test
  slave: os-mng-test

```

In questo contesto, le operazioni effettuate sono state la creazione del modello Juju (comando *juju add-model*) e il test del bundle per la verifica di eventuali errori di sintassi (*deploy* con opzione *dry-run*) e infine l'installazione vera e propria secondo quanto riportato nel bundle (comando *juju deploy*).

```

juju add-model openstack-test
juju deploy ./bundle/2023.06.20.TEST.bundle.txt --debug --dry-run
juju deploy ./bundle/2023.06.20.TEST.bundle.txt --debug

```

La procedura di *deploy* ha avuto una durata alquanto lunga. È possibile monitorare l'andamento dell'installazione con il comando *juju status* e osservare lo stato dei diversi componenti mentre viene eseguito il piano delle azioni per ottenere la configurazione desiderata. Come descritto nel capitolo 4, dato un bundle, Juju opera in diversi passaggi trasformando via via lo stato del sistema fino a convergere quando raggiunge la configurazione che soddisfa tutti i vincoli.

Al termine, è possibile verificare nella console MAAS che i diversi server sono stati installati con il relativo tag.

Il risultato dell'installazione si può osservare con il comando *juju status --relations* che mostra lo stato completo di servizi e relazioni. In particolare vi sono 4 parti:

- 1) modello Juju e applicazioni installate;
- 2) elenco di tutte le unit applicative e il loro stato;
- 3) elenco di tutte le macchine (fisiche e container) e relativo stato;
- 4) elenco di tutte le relazioni tra le applicazioni.

La prima parte del comando *juju status* permette di vedere tutte le applicazioni presenti, la relativa versione, il charm da cui sono installate e lo stato.

```

$ juju status --relations
Model          Controller      Cloud/Region    Version  SLA          Timestamp
openstack-test ctrl_polito_test MAAS-TEST/default 2.9.31  unsupported  16:45:59Z

```

App	Version	Status	Scale	Charm	Channel	Rev	Message
Barbican	12.0.0	active	3	barbican	stable	46	Unit is ready
barbican-hacluster		active	3	hacluster	stable	81	Ready clustered
barbican-mysql-router	8.0.35	active	3	mysql-router	stable	26	Unit is ready
barbican-vault	3.7.1	active	3	barbic-vault	stable	27	Unit is ready
ceph-mon	16.2.9	active	3	ceph-mon	stable	183	Unit is ready
ceph-osd	16.2.9	active	5	ceph-osd	stable	564	Unit is ready
ceph-radosgw	16.2.9	active	3	ceph-radosgw	stable	300	Unit is ready
ceph-radosgw-hacluster		active	3	hacluster	stable	81	Ready clustered
cinder	18.2.0	active	3	cinder	stable	650	Unit is ready
cinder-ceph	18.2.0	active	3	cinder-ceph	stable	527	Unit is ready
cinder-ceph-ssd	18.2.0	active	3	cinder-ceph	stable	527	Unit is ready
cinder-hacluster		active	3	hacluster	stable	81	Ready clustered
cinder-mysql-router	8.0.34	active	3	mysql-router	stable	26	Unit is ready
controller-hacluster		active	3	hacluster	stable	81	Ready clustered
controller-host	20.04	active	3	ubuntu	stable	24	
easysrsa	3.0.1	active	1	easysrsa	stable	441	CA connected
etcd	3.3.19	active	3	etcd	stable	691	Healthy, 3peers
glance	22.1.1	active	3	glance	stable	575	Unit is ready
glance-hacluster		active	3	hacluster	stable	81	Ready clustered
glance-mysql-router	8.0.35	active	3	mysql-router	stable	26	Unit is ready
keystone	19.0.0	active	3	keystone	stable	654	App Ready
keystone-hacluster		active	3	hacluster	stable	81	Ready clustered
keystone-mysql-router	8.0.35	active	3	mysql-router	stable	26	Unit is ready
memcached		active	3	memcached	stable	36	Unit is ready
mysql-innodb	8.0.34	active	3	mysql-innodb	stable	56	Unit is ready
nagios-server		active	1	nagios	stable	56	ready
network-host	20.04	active	2	ubuntu	stable	24	
neutron-api	18.3.0	active	3	neutron-api	stable	561	Unit is ready
neutron-api-hacluster and clustered		active	3	hacluster	stable	81	Unit is ready
neutron-api-mysql-router	8.0.35	active	3	mysql-router	stable	26	Unit is ready
neutron-gw	18.3.0	active	2	neutron-gw	stable	512	Unit is ready
neutron-openvswitch	18.3.0	active	7	neutron-ovsw	stable	524	Unit is ready
nova-cloud-controller	23.2.0	active	3	nova-control	stable	690	Unit is ready
nova-compute	23.2.0	active	3	nova-compute	stable	699	Unit is ready
nova-mysql-router	8.0.34	active	3	mysql-router	stable	26	Unit is ready
ntp	3.5	active	5	ntp	stable	47	chrony: Ready
octavia	8.0.1	active	3	octavia	stable	39	Unit is ready
octavia-hacluster		active	3	hacluster	stable	83	Ready clustered
octavia-mysql-router	8.0.35	active	3	mysql-router	stable	26	Unit is ready
os-dashboard	19.2.0	active	1	os-dashboard	stable	598	Unit is ready
placement	5.0.1	active	3	placement	stable	80	Unit is ready
placement-hacluster		active	3	hacluster	stable	83	Ready clustered
placement-mysql-router	8.0.34	active	3	mysql-router	stable	26	Unit is ready
rabbitmq-srv	3.8.2	active	3	rabbitmq-srv	stable	182	Ready clustered
vault	1.5.9	active	3	vault	stable	54	Unit is ready
vault-hacluster		active	3	hacluster	stable	120	Ready clustered
vault-mysql-router	8.0.34	active	3	mysql-router	stable	26	Unit is ready

La seconda parte del comando consente di vedere tutte le istanze delle applicazioni, eventuali istanze subordinate e dove sono installate. Per sintetizzare l'output si presenta solo l'esempio di due applicazioni ceph-mon (installata in container) e nova (installata bare metal).



L'asterisco rappresenta per Juju la *unit leader* cioè è l'unità che costituisce la fonte autorevole per lo stato e la configurazione di un'applicazione [68].

Unit	Workload	Agent	Machine	Public address	Ports	Message
ceph-mon/0*	active	idle	0/lxd/1	10.0.1.142		Ready clustered
nrpe-cephmon/1*	active	idle		10.0.1.142	icmp,5666/tcp	Ready
ceph-mon/1	active	idle	1/lxd/1	10.0.1.141		Ready clustered
nrpe-cephmon/0	active	idle		10.0.1.141	icmp,5666/tcp	Ready
ceph-mon/2	active	idle	2/lxd/1	10.0.1.143		Ready clustered
nrpe-cephmon/2	active	idle		10.0.1.143	icmp,5666/tcp	Ready
nova-compute/0	active	idle	0	10.0.1.131		Unit is ready
neutron-openvswitch/0	active	idle		10.0.1.131		Unit is ready
nrpe-novacompute/0	active	idle		10.0.1.131		Ready
nova-compute/1	active	idle	1	10.0.1.133		Unit is ready
neutron-openvswitch/2	active	idle		10.0.1.133		Unit is ready
nrpe-novacompute/2	active	idle		10.0.1.133		Ready
nova-compute/2*	active	idle	2	10.0.1.132		Unit is ready
neutron-openvswitch/1*	active	idle		10.0.1.132		Unit is ready
nrpe-novacompute/1*	active	idle		10.0.1.132		Ready

La terza parte del comando visualizza l'elenco di tutte le macchine (fisiche e container), il loro stato e indirizzo IP.

Machine	State	Address	Inst id	Series	AZ	Message
0	started	10.0.1.131	Test-OS-01	focal	default	Deployed
0/lxd/0	started	172.29.157.181	juju-4f8502-0-lxd-0	focal	default	Container started
0/lxd/1	started	10.0.1.142	juju-4f8502-0-lxd-1	focal	default	Container started
0/lxd/2	started	172.29.157.157	juju-4f8502-0-lxd-2	focal	default	Container started
[...]						
1	started	10.0.1.133	Test-OS-03	focal	default	Deployed
1/lxd/0	started	172.29.157.179	juju-4f8502-1-lxd-0	focal	default	Container started
1/lxd/1	started	10.0.1.141	juju-4f8502-1-lxd-1	focal	default	Container started
[...]						
2	started	10.0.1.132	Test-OS-02	focal	default	Deployed
2/lxd/0	started	172.29.157.187	juju-4f8502-2-lxd-0	focal	default	Container started
2/lxd/1	started	10.0.1.143	juju-4f8502-2-lxd-1	focal	default	Container started
[...]						
3	started	10.0.1.134	Test-OS-04	focal	default	Deployed
4	started	10.0.1.135	Test-OS-05	focal	default	Deployed
6	started	10.0.1.136	Test-OS-06	focal	default	Deployed

Infine, la quarta e ultima parte del comando visualizza l'elenco di tutte le relazioni esistenti tra le applicazioni del modello Juju. L'intreccio è complesso, per dovere di sintesi si riportano a titolo esemplificativo solo le relazioni per l'applicazione *ceph-mon* nel cluster di test.

Relation	provider	Requirer	Interface	Type
ceph-mon:client		cinder-ceph-ssd:ceph	ceph-client	regular
ceph-mon:client		cinder-ceph:ceph	ceph-client	regular
ceph-mon:client		glance:ceph	ceph-client	regular
ceph-mon:client		nova-compute:ceph	ceph-client	regular
ceph-mon:mon		ceph-mon:mon	ceph	peer
ceph-mon:nrpe-ext-master		nrpe-cephmon:nrpe-ext-master	nrpe-ext-master	subordinate
ceph-mon:osd		ceph-osd:mon	ceph-osd	regular
ceph-mon:radosgw		ceph-radosgw:mon	ceph-radosgw	regular

### 5.2.5 Problemi riscontrati e risoluzione

Le procedure non si sono concluse con successo al primo tentativo, come avviene spesso quando si affrontano nuove installazioni, ma è stato necessario risolvere diversi problemi riscontrati durante i test. Di seguito se ne descrivono sinteticamente alcuni.

- I server di management avevano quattro schede configurate in bond<sup>75</sup>, a cui lato switch corrispondeva un aggregato LACP<sup>76</sup> su cui transitava il traffico delle diverse VLAN. In fase di provisioning dei server, MAAS non era in grado di raggiungere i server sulla rete PXE (il bond non può funzionare sui server finché il sistema operativo non è caricato). È stato necessario rimuovere una scheda di rete dal bond e dal canale LACP e dedicare tale scheda di rete al traffico PXE.
- I primi deploy con Juju non andavano a buon fine per problemi di rete, in particolare per l'impossibilità dei servizi in container di contattare i repository via web per le installazioni e gli aggiornamenti. È stato necessario rivedere con i colleghi dell'ufficio preposto la configurazione delle reti destinate ai servizi OpenStack, con l'obiettivo di definire quali sottoreti inserire dietro il NAT e su quali invece configurare l'accesso a Internet tramite il proxy di ateneo.
- Per i server con il ruolo compute/storage node su cui era in esecuzione *ceph-osd* è stato necessario disabilitare nel BIOS il controller delle schede SD per evitare che il nome *sda* risultasse già usato durante l'installazione e poter avere una denominazione uniforme dei dischi su tutti i nodi.
- In merito ai diversi componenti OpenStack, sono state affrontate problematiche di vario genere, dovute alla complessità delle configurazioni e delle relazioni dei diversi servizi interconnessi. risolte applicando delle fix o modificando le configurazioni. Tra questi vi sono stati errori sulle code dei messaggi gestite dal cluster *rabbitmq-server*, problemi di sincronizzazione delle unità di applicazioni in HA, problemi di comunicazione tra il servizio *nova-compute* e alcuni sotto servizi sui nodi di rete *neutron*.

---

<sup>75</sup> Il bonding o teaming è una configurazione in cui più schede di rete vengono aggregate in modo da formare un'unica interfaccia virtuale.

<sup>76</sup> Il Link Aggregation Control Protocol (LACP) è uno standard IEEE (802.3ad) che consente di raggruppare più porte fisiche di uno switch per ottenere un singolo canale logico.

## 5.2.6 Federazione GARR

Per inserire il cluster OpenStack Polito come regione federata, è stata seguita la procedura definita da GARR e descritta nel paragrafo 4.4.3.

Dal lato Polito sono stati esportati dal Keystone locale i dati delle credenziali e degli endpoint API della regione. Per garantire la sicurezza sono stati trasferiti solo gli hash e non le credenziali in chiaro. Di seguito le due query SQL utilizzate per estrarre i dati dal database Keystone locale.

```
/* Selezione credenziali locali */
SELECT local_user.name AS username,
password.password_hash AS pass_hash,
project.name AS project,
(SELECT name from project where id=local_user.domain_id) AS user_domain,
(SELECT name from project AS p where p.id=project.domain_id) AS project_domain
FROM local_user, user, password, project
WHERE (local_user.user_id=user.id
AND local_user.id=password.local_user_id
AND user.default_project_id=project.id)
AND ((local_user.domain_id=(SELECT id from project where name='default') )OR
local_user.domain_id=(SELECT id from project where name='service_domain'));

/* Selezione degli endpoint API locali */
SELECT endpoint.interface, endpoint.url, endpoint.region_id, service.type,
service.extra
FROM endpoint, service
WHERE service.id=endpoint.service_id
AND service.type <> 'identity';
```

Dal lato GARR, i tecnici del Dipartimento CDS hanno successivamente effettuato le operazioni di importazione nel Keystone centralizzato e le configurazioni necessarie per le deleghe amministrative. In virtù della federazione, gli utenti del Politecnico di Torino possono accedere alla dashboard OpenStack con le proprie credenziali di ateneo e visualizzare le risorse come progetti e VM a cui sono stati abilitati.

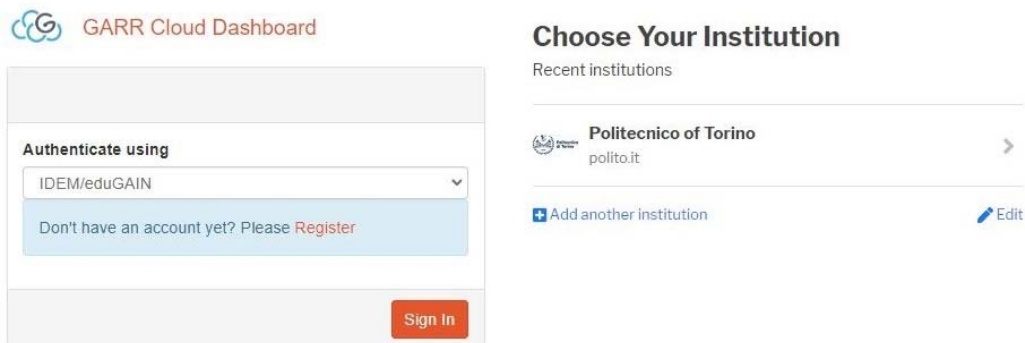


Figura 45. GARR Cloud: login attraverso IDEM

### 5.3 Utilizzo del cloud OpenStack Polito

Il cloud OpenStack Polito è operativo e offre servizi di tipo IaaS: sia datacenter virtuali (progetti) sia singole VM, usate dal personale docente e ricercatore per le finalità di didattica e di ricerca.

#### 5.3.1 Profili disponibili per le VM

Per la creazione delle VM si utilizzano combinazioni preimpostate, che in OpenStack sono chiamate *flavor*, con un determinato numero di vCPU, dimensione RAM e disco. In generale, i profili bilanciati forniscono un rapporto vCPU/RAM adatto ai carichi di lavoro più comuni. Altri profili possono essere più sbilanciati verso le vCPU (profili di calcolo) oppure verso la RAM (profili di memoria): si tratta di combinazioni indicate per carichi di lavoro con richieste intensive di CPU oppure per carichi di lavoro con uso intensivo di memoria. La Tabella 11 riporta l'elenco dei flavor definiti nel cloud Polito e le relative combinazioni di risorse.

Tabella 11. OpenStack Polito: elenco dei flavor disponibili per le VM.

Nome flavor	RAM [GB]	Disk [GB]	vCPUs
e1.xlarge-s2	64	20	64
e3.xlarge-s2	192	20	64
e4.s2	64	10	64
l1.jumbo	16	320	6
m1.jumbo	16	320	6
m1.large	4	80	2
m1.medium	2	40	2
m1.small	1	20	1
m1.tiny	1	10	1
m1.xlarge	8	160	4
m1.xxlarge	32	256	32
p1.2xlarge	32	64	8
p1.4xlarge	64	128	16
p1.4xlarge-s2	64	128	16
p1.8xlarge	128	256	32
p1.xlarge	16	32	4
p2.32xlarge	512	80	64

### 5.3.2 Modalità di utilizzo

Attualmente il cloud OpenStack Polito ospita in totale 177 VM e 10 progetti. Oltre a quelli richiesti da alcuni dipartimenti o gruppi di ricerca, è stato creato un progetto comune per ospitare le singole VM richieste da docenti e ricercatori. Inoltre un progetto contiene le VM di servizio e sono presenti alcuni progetti di test.

La Tabella 12 mostra la situazione attuale relativa al numero dei progetti, le VM e le relative risorse assegnate. I nomi dei progetti sono stati sostituiti con etichette generiche poiché i dati sono tratti dall'ambiente in produzione.

Tabella 12. OpenStack Polito: progetti e risorse utilizzate

Progetto	Tot VM	Tot vCPU	Tot RAM (GB)	Tot Disco (GB)
Dipartim1	4	96	384	768
Dipartim2	1	8	32	64
Dipartim3	7	196	1.028	1.184
Dipartim4	8	24	46	920
Dipartim5	36	928	3.680	8.512
Dipartim6	7	64	256	512
Dipartim-Test	2	7	17	660
Polito-Test	16	104	276	1.332
Polito-VM	81	1.122	4.116	13.316
Services	15	30	30	600
Totale complessivo	177	2.579	9.865	27.868

Dai dati riportati in tabella, considerato il totale di 177 VM, si può calcolare una dimensione media di 16 vCPU, 64GB di RAM e 160 GB di disco per ciascuna VM.

Dal punto di vista dei sistemi operativi, circa l'85% delle VM usa distribuzioni Linux (Ubuntu, Debian, CentOS, Alma Linux, Rocky), quindi utilizzabili senza licenza. Il restante 15% utilizza Windows: in questi casi si tratta di utenti che in generale hanno più dimestichezza con tale sistema operativo rispetto a Linux oppure che necessitano di utilizzare software commerciali disponibili solo su tale piattaforma.

In un sistema cloud è possibile monitorare le risorse di calcolo utilizzate da un'applicazione o un servizio durante un determinato periodo di tempo valutando i parametri *RAM MB hours*, *CPU hours* e *Disk GB Hours*. Queste metriche, spesso utilizzate per la fatturazione dei servizi nei cloud commerciali, rappresentano il prodotto tra la durata dell'utilizzo e la quantità di risorsa impiegata: memoria, potenza di elaborazione della CPU o spazio di archiviazione disco.

Per offrire un'idea di quanto siano usate le risorse nel cloud Polito si verifica con il comando *openstack usage list* e si riportano nella Tabella 13 i dati aggregati.

L'utilizzo è per 30 giorni, nell'intervallo tra il 16/10/2023 e il 15/11/2023.

Tabella 13. OpenStack Polito: risorse utilizzate in un mese

Progetto	Tot VM	Tot RAM TB-Hours	Tot CPU Hours	Tot Disk TB-Hours
Dipartim1	4	203	51.980	406
Dipartim2	1	21	5.376	42
Dipartim3	7	675	131.712	777
Dipartim4	8	30	16.128	604
Dipartim5	36	2.415	623.616	5.586
Dipartim6	7	168	43.008	336
Dipartim-Test	2	11	4.704	433
Polito-Test	16	181	69.888	874
Polito-VM	81	2.701	753.955	8.738
Services	15	20	20.160	394
Totale complessivo	177	6.425	1.720.527	18.190

### 5.3.3 Attività di evoluzione futura

L'attività descritta nel presente lavoro non può dirsi conclusa con il setup e la messa in produzione del cluster OpenStack Polito.

Il sistema è in continua evoluzione e sono molteplici le attività da pianificare dal punto di vista tecnologico. Sicuramente tra le più impegnative vi è l'aggiornamento dell'infrastruttura, attualmente alla versione Wallaby, ad una release OpenStack più recente, di concerto con GARR e le altre regioni federate. Inoltre, in virtù di quanto emergerà dalle nuove linee di indirizzo per il cloud GARR, descritte nel paragrafo 4.4.4, potrebbe essere necessario approfondire eventuali scenari legati alle policy e alle deleghe amministrative, oltre a modifiche sui servizi di autenticazione Keystone.

Limitatamente al cluster Polito, l'imminente consegna dei nuovi 30 server richiederà il coordinamento delle fasi di test, installazione e configurazione, cercando di calibrare gli strumenti di automazione MAAS e Juju per ottimizzare tutte le attività di gestione. Lo stesso Juju dovrà essere oggetto di upgrade dall'attuale versione 2.9 alla 3.1 per poter beneficiare di alcune nuove funzionalità.

Per ottenere maggiore resilienza e affidabilità, sono in programma la reinstallazione su altro hardware dei 3 nodi controller e 2 nodi network, oltre che la messa in funzione di un sistema di monitoraggio basato su Prometheus e Grafana. Infine, un progetto più complesso vede l'espansione dell'infrastruttura, al momento collocata in un unico locale, verso un secondo data center. Questo richiederà la progettazione di una soluzione multi-site che possa assicurare i meccanismi di HA per tutti i servizi OpenStack e per la replica storage Ceph.

## 5.4 Bilancio del progetto

La realizzazione del cloud privato OpenStack all'interno del Politecnico di Torino è stato ed è tutt'ora un progetto sicuramente stimolante e ambizioso.

L'adesione al modello cloud GARR ha permesso di implementare un'architettura replicabile e progettata per lo stesso contesto, potendo contare sul supporto tecnico di specialisti cloud, sia sugli aspetti progettuali sia implementativi.

Dal punto di vista organizzativo questo progetto ha consentito la collaborazione e lo scambio anche tra il personale dell'amministrazione centrale (ex Area IT, ora Direzione ISIAD) e dei due dipartimenti (Dipartimento di Automatica e Informatica DAUIN e Dipartimento di Elettronica e Telecomunicazioni DET), portando ad una crescita di competenze tecniche per tutti i partecipanti.

Per l'ateneo è innegabile la ricaduta positiva grazie all'ampliamento dei servizi per la comunità della ricerca, a cui è offerta la possibilità di disporre di ampie risorse computazionali talvolta difficilmente acquisibili e gestibili a livello di singolo gruppo di ricerca o di dipartimento. L'organizzazione della struttura federata e l'autenticazione centralizzata aprono inoltre alla opportunità di collaborazione anche in termini di scambio di risorse, consentendo agli atenei o a gruppi di enti di ricerca di mettere a fattore comune eventuali risorse private, come ad esempio trasferire VM in un'altra regione federata per fronteggiare necessità temporanee.

Per ciò che concerne l'organizzazione interna, con il tempo si è ridotto notevolmente il numero di tecnici che si dedicano al cluster OpenStack, attualmente in gestione alla direzione ISIAD, *Servizio Cybersecurity & IT for Research*. Oltre ad essere esiguo rispetto al carico di lavoro, al momento il personale non svolge tale attività per la totalità del proprio tempo e affersce inoltre a due Servizi diversi.

Convogliare una parte delle esigenze tecnologiche della ricerca dai dipartimenti verso l'infrastruttura OpenStack e centralizzarne la gestione tecnica, può creare un sovraccarico amministrativo nella direzione ISIAD e ritardare gli sviluppi futuri.

Oltre alle attività tecniche da pianificare descritte nel paragrafo 5.3.3, vi sono infatti anche altre prospettive da valutare. Sul piano organizzativo, nell'ottica di una maggiore diffusione del servizio, sarebbe opportuno perfezionare le strategie di comunicazione in ateneo e definire delle policy per regolare l'utilizzo del cloud. Per i gruppi di ricerca che lavorano su progetti, visto anche l'aumento delle attività legate al PNRR, sarebbe molto utile implementare un sistema di contabilizzazione che permetta di inserire i costi delle VM impiegate nelle rendicontazioni, in modo da realizzare una compartecipazione alle spese di gestione dell'infrastruttura.

Dal punto di vista tecnico, OpenStack è un sistema estremamente complesso e articolato: nel cluster Polito la logica di controllo in esecuzione sui nodi controller e network è distribuita su oltre 100 unit Juju, containerizzate su 5 server fisici. Questo da un lato offre robustezza e resilienza, ma dall'altro complica le fasi di debug; in caso

di eventuali problemi l'individuazione e la risoluzione possono richiedere tempo. I singoli componenti appartengono inoltre a progetti open source diversi, su cui si innesta ulteriormente il progetto *Canonical OpenStack Charms*. Nonostante la comunità di sviluppatori sia molto ampia, non sempre i problemi legati a bug o malfunzionamenti sono risolvibili in modo rapido. In caso di criticità, i tecnici del Dipartimento CDS di GARR sono sempre disponibili a fornire assistenza. Sarebbe tuttavia molto utile acquisire un contratto di supporto Canonical in modo da avere un canale diretto per l'assistenza in caso di bug e per problematiche che richiedono competenze trasversali.

Lavorare sull'infrastruttura OpenStack richiede competenze avanzate sui sistemi operativi Linux e Windows, conoscenze dell'architettura cloud e dei differenti servizi, competenze di networking, storage distribuito, database in HA, scripting, sicurezza, strumenti di automazione, oltre a soft skill come la capacità di analizzare e risolvere i problemi in modo efficiente.

Da un lato, in vista della crescente complessità hardware e software, si rivela fondamentale ottimizzare le attività di gestione puntando su eventuali nuovi strumenti che forniscano maggiore automazione e capacità di monitoraggio. Dall'altro lato è necessario poter contare su un team dedicato e adeguatamente preparato, rafforzando le competenze interne sia attraverso iniziative di formazione sia con strategie per l'organizzazione e il reclutamento di personale.



## Capitolo 6.

# Analisi economica e confronto

Dopo aver presentato nei capitoli precedenti il modello federato GARR e il cloud privato Polito, in questo capitolo si espone un'analisi economica di sintesi. In particolare si mettono a confronto i costi per un'infrastruttura on-premise e i costi per l'esecuzione delle VM in un cloud pubblico.

La valutazione dei costi relativi all'hardware esistente del cluster Polito richiederebbe l'esame di documenti amministrativi passati e una procedura laboriosa, in quanto gli acquisti sono stati dilazionati in tempi diversi. Si propone pertanto un modello di alto livello simulando l'implementazione di una nuova infrastruttura on-premise che possa offrire risorse paragonabili a quelle dell'attuale cluster OpenStack. La valutazione è ugualmente dettagliata e ha il vantaggio di essere allineata ai prezzi attuali. Si procede quindi con una stima del TCO<sup>77</sup> di questo scenario e lo si pone in confronto all'alternativa di eseguire su un cloud pubblico un carico di VM equiparabile a quello che consente la capacità on-premise.

Per ciò che concerne le quotazioni economiche, in qualità di PA, il Politecnico di Torino deve rispettare procedure amministrative specifiche per l'approvvigionamento di beni e servizi. Oltre a bandire gare d'appalto, esiste la possibilità di acquistare attraverso Consip (Concessionaria Servizi Informativi Pubblici), società interamente partecipata dal Ministero dell'Economia e delle Finanze che opera come centrale degli acquisti della PA. Consip organizza procedure di gara per le forniture e i servizi più comuni richiesti dalle PA italiane e si occupa di stipulare gli accordi quadro con fornitori; successivamente le PA possono acquistare direttamente tramite tali accordi, semplificando il processo di acquisizione.

Per l'analisi economica, ove possibile, si è pertanto fatto riferimento ai prezzi indicati nelle convenzioni Consip disponibili al momento della redazione del presente lavoro.

---

<sup>77</sup> *Total cost of ownership* (TCO): indica il costo totale nel ciclo di vita delle apparecchiature, considerando tutti i costi sostenuti per l'acquisto, l'installazione, l'esecuzione e la manutenzione.

## 6.1 Infrastruttura on-premise

In questa sezione si analizza la situazione on-premise, partendo dal presupposto che esista presso la sede un data center operativo e attrezzato con gruppi di continuità e impianto di raffrescamento.

La simulazione di acquisto si pone l'obiettivo di offrire risorse dell'ordine di grandezza di 30TB RAM, 4.500 vCPU, 1PB di disco raw, paragonabili a quelle dell'attuale cluster Polito, senza considerare il raddoppio programmato dell'infrastruttura. La Tabella 14 riporta le diverse fasi dell'investimento e i relativi costi.

Tabella 14. Infrastruttura on-premise: fasi dell'investimento

Fase	Descrizione Costi
Selezione e acquisto	Costi per la progettazione. Costi per apparecchiatura hardware: server, apparati di rete, armadi rack
Sviluppo	Costi per installazione e messa in opera dell'hardware Costi per configurazione e deploy software
Operativa	Costi di esercizio: sottoscrizioni per licenze software, personale, consumi elettrici
Dismissione	Costi di smaltimento apparecchiature (normalmente inclusi nei nuovi contratti di fornitura)

### 6.1.1 Modello costi on-premise

Il modello si compone di 3 nodi controller, 2 nodi network e 32 nodi con la funzione compute e storage e 2 nodi per i software di gestione MAAS e Juju. I costi per la messa in opera di una nuova infrastruttura on-premise sono strutturati come segue.

- *Hardware.* Vengono considerati i costi per l'acquisto dei server necessari, degli apparati di rete per la connessione, compresi i cablaggi e gli armadi rack per ospitare le apparecchiature nel data center.
- *Licenze e software.* OpenStack è una soluzione open source, così come gli strumenti di automazione MAAS e Juju. Si valutano quindi solo i costi di licenza per l'esecuzione di VM con sistema operativo Windows.
- *Manutenzione e assistenza.* Rientrano in questa sezione i costi associati all'assistenza tecnica di server e apparati di rete, oltre alle eventuali estensioni di garanzia dei prodotti.
- *Servizi accessori.* Sono compresi in questa categoria i costi per i servizi di trasporto, posa in opera, installazione, configurazione e avvio operativo dei sistemi, asporto degli imballaggi, ritiro e dismissione dell'esistente. Tali costi possono essere inclusi nei contratti di fornitura hardware.

- *Consumi energetici.* Vengono considerati i costi legati al consumo energetico dell'hardware sopra descritto e per il raffreddamento nel data center.
- *Personale.* Si propone una stima i costi relativi al personale IT necessario per la gestione e la manutenzione dell'infrastruttura OpenStack.

Per semplificare la valutazione, non si calcolano in questa sede i costi di progettazione, in quanto si utilizza un modello di servizi esistente, e i costi per lo smaltimento delle apparecchiature, normalmente inclusi nei nuovi contratti di fornitura oppure oggetto di procedure amministrative specifiche.

### 6.1.2 Costi hardware per server

Tra le gare Consip già aggiudicate viene presa in esame la *Convenzione Tecnologie Server edizione 4 Lotto 6*, il cui avviso di aggiudicazione è stato pubblicato in data 20 gennaio 2023. Dal sito del fornitore aggiudicatario *Italware S.r.l.* si ricava la guida alla convenzione corredata dalla tabella predisposta per la personalizzazione della configurazione e il calcolo del costo della fornitura. L'accordo ha una durata di 12 mesi, prorogabile di ulteriori 6 mesi. I singoli contratti stipulati dalle PA hanno una durata di 3 anni a partire dalla data di accettazione delle forniture e possono essere estesi fino a 5 anni acquistando il servizio opzionale di estensione della garanzia [86].

La suddetta convenzione Consip propone la fornitura di server *Dell Poweredge R940, quad-processor con CPU Intel Xeon Gold 6252N 24 core*, in configurazione con 1 TB RAM, dotati di controller con funzionalità RAID hardware, alimentazioni ridondate hot swap e possibilità di ospitare fino a 24 dischi HDD, oppure 12 HDD+12 NVME<sup>78</sup>. Le caratteristiche tecniche complete sono dettagliate in Tabella 15. I server selezionati hanno caratteristiche migliori rispetto ai server attuali e sono equiparabili a quelli acquistati per il raddoppio dell'infrastruttura OpenStack.

Per i 32 nodi compute e storage, nella personalizzazione della configurazione si selezionano 14 dischi HDD, oltre ai 2 SSD già compresi come dischi di sistema nel modello in convenzione. Si aggiungono 2 schede di rete dual port 10GbE e si include l'opzione di estensione della garanzia a coprire in totale 5 anni. La Figura 46 riporta un esempio di configuratore per la scelta dei componenti dei server.

La stessa configurazione viene usata come base per i nodi controller e network, con la differenza di soli 2 dischi aggiuntivi anziché 14 e, per i due server di management, solo 512GB di RAM invece di 1TB.

---

<sup>78</sup> NVME (*Non-Volatile Memory Express*) è un'interfaccia che consente di collegare dispositivi di archiviazione di massa usando lo standard PCI Express. Utilizza livelli di parallelismo che abbinati ai dischi SSD forniscono prestazioni di alto livello.

Tabella 15. Convenzione Consip: caratteristiche hardware server.

Fonte [86]

Caratteristiche	Descrizione
Marca e modello	Dell PowerEdge R940. Codice Prodotto conf. standard: TS4L6-SRV
Benchmark prestazionali	CPU2017 Integer Rates (I.R.), valore “base result”: 597 CPU2017 Floating Point Rates (F.P.R.), valore “base result”: 522
Form Factor	3RU
Quantità socket e chipset	4 socket. Chipset Intel C621A
CPU installate	2x Intel Xeon Gold 6252N (2.3GHz, 24Core/48Thread, 150W)
Memoria RAM installata	16 X 32GB RDIMM, 3200MT/s. Espandibilità 1,5 TB
Espandibilità I/O	8 x PCIe Gen3 slots
RAID controller	PERC H740P 8Gb
Dischi interni installati	2 X SSD da 960 GB con Endurance DWPDP >=0.9
Espandibilità dischi interni	24X 2.5 SAS/SATA o 12X 2.5 SAS/SATA + 12 X 2.5 NVME
Alimentatori	2 X Dual, Hot-plug, Redundant Power Supply (1+1), 1600W
Sistema di gestione	iDRAC9 Enterprise
Sistemi operativi supportati	Canonical Ubuntu Server LTS, VMware ESXi, Microsoft Windows Server LTSC with Hyper-V, Citrix Hypervisor, Oracle Linux Red Hat Enterprise, Linux SUSE Enterprise Server

L’acquisto di 32 nodi compute e storage offre un totale di 32TB di RAM, 6.144 core (con hyperthreading), circa 1PB di spazio disco HDD raw, come indicato in Tabella 16.

Tabella 16. Acquisto server: stima delle risorse totali

Num Server	Num CPU	Num Core	RAM [TB]	Num HDD	HDD [TB]	Tot Core hyperthread	Tot RAM [TB]	Tot HDD [TB]
32	4	24	1	14	2,4	6.144	32	1.075,20
al netto di overhead 10%						5.530	28,8	

Sui nodi compute e storage, come descritto nel capitolo 5, parte delle risorse è usata per l’esecuzione dei servizi di OpenStack, Ceph e per gli hypervisor (si veda il paragrafo 5.1.3). Si stima pertanto un overhead del 10% per CPU e RAM, che decurta le risorse disponibili per le VM. Anche la componente storage raw subisce una riduzione, variabile in base alla configurazione di ridondanza applicata. Con un calcolo approssimativo, usando pool Ceph con replica 3, lo spazio realmente utilizzabile si ridurrebbe a un terzo.

La dimensione media delle VM nel cluster Polito (riferimento 5.3.2) è di 16 vCPU e 64GB di RAM. Considerando solo il valore vCPU, le risorse permetterebbero l’esecuzione di 345 VM. Tale valore è puramente teorico perché è buona prassi che gli

hypervisor non operino mai al massimo delle loro capacità ma mantengano risorse libere per garantire eventuali migrazioni di VM in caso di guasti o manutenzioni.


Convenzione Tecnologie Server 4 - Lotto 6		Guida alla Convenzione	Totale Configurazione	Totale Ordine Iva Esclusa
DELL POWEREDGE R940		Quantità Server	1	14.704,79 €
				
<b>Base</b>	PowerEdge R940 Server (General Purpose)	TS4L6-SRV	1	5.754,51 €
<b>Chassis Configuration</b>	Chassis with Up to 12X 2.5 SAS/SATA + 12X 2.5 for 4CPU Config			0,00 €
<b>Processor</b>	2 x Intel Xeon Gold 6252N, 2.3G, 24C/48T, 10.4GT/s, 35.75M Cache, Turbo, HT (150W) DDR4-2933			
<b>Additional Processor</b>	2 x Intel Xeon Gold 6252N, 2.3G, 24C/48T, 10.4GT/s, 35.75M Cache, Turbo, HT (150W) DDR4-2933	TS4L6-CPU	1	1.908,90 €
<b>Memory Capacity (GB)</b>	1024	TS4L6-RAM32	16	4.218,24 €
		TS4L6-RAM64	0	
<b>Operating System STANDARD</b>	No Operating System		0	0,00 €
	Windows Server 2022 Standard Edition, Add License,2CORE	TS4L6-WINSRV2C		0,00 €
<b>Disk Controller (PCI SLOT 1)</b>	PERC H740P RAID Controller, 8Gb NV Cache, Full Height			0,00 €
Hard Drivers Bays n° 24 - 2 bay occupate dai dischi 960GB SSD di		<b>Usati</b>	<b>Disponibili</b>	
		16	8	
<b>2.4TB 10k SAS ISE 12Gbps</b>		TS4L6-HDD2TB	14	1.930,18 €
<b>960GB SSD SAS ISE RI 12Gbps</b>		TS4L6-SSD-RI800GB	2	0,00 €
<b>Power Supply</b>	Dual, Hot-plug Power Supply (1+1), 1400W	TS4L6-PSU1400W	1	
<b>Embedded Systems Management</b>	iDRAC9,Enterprise			
<b>PCI Slot 1 Processor 1 x8 PCIe for HL/FH</b>	Riservato per PERC			
<b>PCI Slot 2 Processor 1 x16 PCIe for HL/FH</b>	Broadcom 57412 Dual Port 10Gb, SFP+, PCIe, Full Height + Optics	TS4L6-LAN10	1	149,99 €
<b>PCI Slot 3 Processor 1 x16 PCIe for HL/FH</b>	Broadcom 57412 Dual Port 10Gb, SFP+, PCIe, Full Height + Optics	TS4L6-LAN10	1	149,99 €
<b>SFP+ Transceivers 10 Gbps</b>	SFP+ SR Optic for all SFP+ ports	TS4L6-OT10GB	4	70,20 €
<b>Garanzia</b>	Estensione della manutenzione in garanzia per ulteriori 24 mesi	TS4L6-OpzEst24	1	522,78 €

Figura 46. Convenzione Consip: configuratore server.

Fonte [86]

### 6.1.3 Costi hardware per rack e accessori

Tra gli elementi opzionali della *Convenzione Tecnologie Server edizione 4 Lotto 6* sono elencati anche gli armadi rack e gli accessori necessari per le attrezzature in un datacenter, come i moduli per *Power Distribution Unit* (PDU), gli switch KVM<sup>79</sup> e le unità con monitor e cavi [86].

Considerando che i server Dell PowerEdge R940 selezionati hanno un ingombro di 3U<sup>80</sup> ciascuno, si ipotizzano 4 armadi rack da 42U che consentono di collocare 10 server per rack, lasciando lo spazio necessario agli apparati di rete. Si inseriscono altrettanti switch KVM e unità monitor per fornire la console grafica locale ad ogni server. Per la ridondanza e la distribuzione bilanciata del carico, si calcolano sei PDU per rack.

Nella Figura 47 è riportato un esempio di rack realizzato con lo strumento di configurazione Dell *Enterprise Infrastructure Planning Tool* [87].



Figura 47. Costi on-premise: configurazione rack

Fonte [87]

<sup>79</sup> Console e switch KVM (*Keyboard, Video and Mouse*) costituiscono un dispositivo hardware che consente agli utenti di controllare più server da un unico blocco tastiera, monitor e mouse.

<sup>80</sup> Un'unità rack U (*rack unit*) è l'unità di misura che indica l'altezza dei componenti installati in un armadio rack da 19" o da 23". Un armadio rack da 19" contiene 42U.

Non è necessario conteggiare i costi relativi ai servizi di trasporto, posa in opera, installazione, configurazione e avvio operativo dei sistemi, asporto dell'imballaggio in quanto sono inclusi nella convenzione Consip [86].

La Figura 48 riporta l'esempio del configuratore Excel con la scelta completa per un armadio rack con i relativi accessori.

Convenzione Tecnologie Server 4 - Lotto 6		Totale altre opzioni	
Opzioni		4.863,72 €	
<b>OpzRack DELL</b> Rack 42RU DellEmc	TS4L6-RACK-D	0	- €
<b>OpzPDU Mono</b> PDU 220 mono output 20-C13 + 4-C19 IEC 320, input C20	TS4L6-PDU-M	6	3.240,00 €
<b>OpzPDU Trifase</b> PDU 220 Trifase output 6-C19 e 36-C13, input Spina 3P+T 400V 16A	TS4L6-PDU-T	0	- €
<b>OpzKVM DELL</b> Dell KVM + Switch 8 porte + cavi	TS4L6-KVM-D	0	- €
<b>Opz CMA</b> Cable Management Arm for PowerEdge Systems	TS4L6-CMA	0	- €
<b>Opz JBOD (NO DISK)</b> Enclosure di espansione JBOD 12 slot 3,5" + HBA (H840)	TS4L6-JBOD	0	- €
<b>Opz 12TB 7,2 RPM - per JBOD</b> 12TB 7.2K RPM NLSAS ISE 12Gbps 512e 3.5in Hot-plug Hard	TS4L6-12TBNLSAS	0	- €
<b>Opz 16TB 7,2 RPM - per JBOD</b> 16TB Hard Drive SAS ISE 12Gbps 7.2K 512e 3.5in Hot-Plug	TS4L6-16TBSAS	0	- €
<b>Opz 20TB 7,2 RPM - per JBOD</b> 20TB HDD SAS ISE 12Gbps 7.2K 512e 3.5in Hot-Plug	TS4L6-20TBSAS	0	- €
<b>OpzRack</b> Armadio Rack 42U	TS4L6-RACK	1	579,63 €
<b>OpzGUI</b> Unità KVM - monitor 17", cavo combi PS/2&USB da 1,8mt, tastiera italiana, touch pad	TS4L6-GUI	1	491,62 €
<b>OpzKVM</b> KVM Switch IP per GUI 16P	TS4L6-KVM	1	552,47 €

Figura 48. Convenzione Consip: configuratore server e accessori rack.

Fonte [86]

#### 6.1.4 Costi hardware per networking e accessori

Per la parte di networking, si propone un modello di tipo *spine-leaf*, l'architettura ormai più diffusa per i data center, composta da due livelli di switch (Figura 49).

Al livello leaf operano gli switch di accesso L2 che forniscono la connettività tra gli endpoint nel data center e aggregano il traffico dai server. Il livello spine L3 esegue il routing e fornisce l'interconnessione ad alta velocità tra gli apparati leaf, collegati in una topologia full-mesh [87]. I vantaggi di questo modello sono la scalabilità e la riduzione del numero di hop e della latenza della rete.

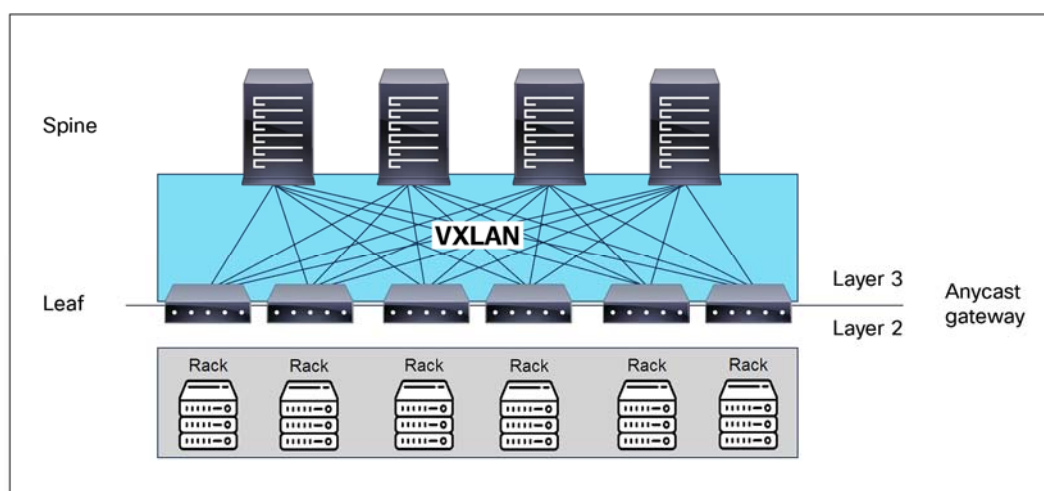


Figura 49. Modello di rete spine-leaf.

Fonte [87]

Si ipotizza uno schema simile a quanto proposto negli esempi sulla documentazione OpenStack [84]. I 4 switch leaf ToR<sup>81</sup> sono collegati a 2 apparati spine con uplink 40GbE, che a loro volta sono connessi al resto della rete e ai router di frontiera. I server nei rack sono collegati agli switch leaf ToR con link port-channel da 40GbE (4x10GbE con aggregazione LACP). Tutti i collegamenti utilizzano cavi in fibra multimodale. In aggiunta si devono considerare le connessioni in rame, sia per la rete PXE sia per IPMI, in quanto ogni server ha la necessità di collegare:

- 4 porte 10GbE per un totale di 128 porte fibra 10GbE
- 1 porta 1GbE per la rete PXE, per un totale di 39 porte rame 1GbE
- 1 porta 1GbE per la scheda IPMI, per un totale di 39 porte rame 1GbE.

Per la scelta degli apparati si fa riferimento alla Convenzione Reti Locali 8, aperta da Consip lo scorso 15 novembre 2023, in cui rientra la fornitura di apparati di rete attivi e di materiali per la realizzazione di cablaggi strutturati. L'accordo ha una durata

<sup>81</sup> Tor (*Top of Rack*): switch di rete posizionato sulla parte superiore dell'armadio rack, a cui sono collegati direttamente i server dello stesso rack.



contrattuale di 18 mesi ed è prorogabile fino ad ulteriori 6 mesi. I singoli contratti stipulati dalle PA hanno una durata di 12 mesi, così come gli eventuali servizi di assistenza e manutenzione acquistati come opzione. Con la fornitura sono incluse l'installazione e la configurazione degli apparati attivi [88].

Dalle tabelle della guida alla convenzione si possono estrarre le caratteristiche tecniche degli apparati dei fornitori Cisco, Aruba, Extreme Networks, Juniper e Huawei, divisi per fascia di prodotto. Per la parte in rame 1GbE, sono sufficienti gli apparati di tipo 2, mentre per la sezione leaf è necessario selezionare apparati di livello 8, aggiungendo gli opportuni connettori transceiver da 10GbE e 40GbE per il numero necessario di porte. L'analisi comparativa sui listini, non riportata in questa sede per ragioni di sintesi, mostra che la soluzione Juniper risulta la più conveniente a livello economico, a parità di caratteristiche tecniche. Inoltre il fornitore include una garanzia tipo *Enhanced Limited Lifetime* che offre la sostituzione hardware per tutta la vita degli apparati [88]. Si selezionano pertanto:

- 4 switch *Juniper EX2300-24P-VC-C*, con 24x1GbE + 4x1/10GbE SFP/SFP+
- 4 switch *Juniper EX4650-48Y-AFO-C*, con 48x10/25GbE SFP28/SFP+/SFP+ e uplink 8x40/100GbE QSFP28/QSFP+

Si inserisce anche il servizio di assistenza e manutenzione con il profilo HP (supporto attivo h24, 7/7), quotato pari al 3,00% dell'importo di fornitura degli apparati attivi.

La convenzione Consip non offre apparati di fascia alta con porte di accesso 40GbE che possano essere collocati al livello spine. Per la quotazione si fa pertanto riferimento ad un'offerta economica pervenuta a ISIAD nell'ambito di un recente progetto di rete.

Il prodotto è lo switch Cisco Catalyst C9500X-28C8D dotato di 28 porte 40/100GbE QSFP28 da utilizzare come downlink verso gli switch leaf e 8 porte 40/100/200/400GbE di uplink. Naturalmente l'importo dell'offerta non usufruisce di sconti e agevolazioni, come i prezzi dei listini Consip, ma risulta più vicino al prezzo di mercato.

I costi complessivi per il networking sono riportati nella Tabella 17.

Tabella 17. Costi on-premise: apparati di rete

Apparato	Descrizione	Quantità	Prezzo	Totale
Switch Juniper EX2300-24P	Switch 1GbE	4	654,09 €	2.616,36 €
Switch Juniper EX4650-48Y e transceiver	Switch 10GbE (leaf)	4	6.452,97 €	25.811,88 €
Switch C9500X-28C8D e transceiver	Switch 40GbE (spine)	2	38.000,00 €	76.000,00 €
Cablaggi e servizi	Cavi rame, fibre e posa	1	1.256,50 €	1.256,50 €
<b>Totale IVA esclusa</b>				<b>105.684,74 €</b>

---

---

**Totale IVA inclusa****128.935,38 €**

---

---

### 6.1.5 Costi di licenza

Prendendo come esempio il cluster Polito, circa l'85% delle VM usa sistemi operativi basati su distribuzioni Linux, mentre per il restante 15% che utilizza Microsoft Windows è necessario considerare un'opportuna licenza. Per le VM di test sarebbe possibile ricorrere alla licenza temporanea della durata di 180 giorni, ma per i sistemi destinati alla ricerca è preferibile utilizzare le licenze definitive.

Grazie all'accordo Microsoft-CRUI<sup>82</sup> le università possono accedere a listini scontati rispetto alle aziende commerciali. È stato contattato l'ufficio della Direzione ISIAD che si occupa delle convenzioni software per avere una quotazione economica.

Le licenze sono vendute in forma di sottoscrizioni annuali. La licenza di tipo Windows Server Datacenter è fornita in blocchi da 2 core, con un minimo per singolo server pari a 8 licenze (16 core). Ogni licenza consente di eseguire un numero illimitato di VM con sistema operativo Windows Server Datacenter su un hypervisor, anche non Microsoft. Si ipotizza di acquisire una sottoscrizione annuale considerando almeno 6 server, con le caratteristiche descritte nel paragrafo 6.1.2. In questo modo sono coperti da licenza anche gli scenari di migrazione di VM da un host all'altro in caso di manutenzione.

Tabella 18. Licenze Microsoft

Num Server	Num CPU	Num Core	Totale Core	Pacchetto 2 core	Costo unitario server	Costo Totale
6	4	24	96	35,00 €	1.680 €	10.080 €
<b>Totale IVA inclusa</b>						<b>12.298</b>

---

<sup>82</sup> Link alla convenzione: <https://ict.crui.it/microsoft/>

### 6.1.6 Consumi energetici

Lo scorso 16 ottobre 2023 Consip ha aggiudicato il nuovo contratto per la fornitura di energia elettrica alle PA per un quantitativo disponibile di 18TWh e un valore di oltre 3 miliardi di euro. La convenzione *Energia elettrica 20* è strutturata su più lotti divisi su base geografica; il lotto 1 per Valle d'Aosta e Piemonte è stato aggiudicato al fornitore: AGSM AIM Energia S.p.A. I singoli contratti hanno una durata di 12 mesi e si basano su una struttura tariffaria variabile (Figura 50) che prevede l'applicazione di parametri chiamati spread, offerti in sede di gara, alle quotazioni mensili del PUN<sup>83</sup> fissate dal Gestore dei Mercati Energetici S.p.A. [89].

Le fasce orarie sono suddivise in F1 (lun-ven dalle ore 8.00 alle 19.00), F2 (lun-ven dalle 7.00 alle 8.00 e dalle 19.00 alle 23.00, sabato dalle 7.00 alle 23.00, escluse festività nazionali), F3 (ore notturne, lun-sab dalle 23.00 alle 7.00, domenica e giorni festivi).

**acquistinretepa**

**Energia elettrica 20**  
 Convenzione per la fornitura di energia elettrica a prezzo variabile e servizi connessi per tutte le Pubbliche Amministrazioni

**Configuratore corrispettivi a Prezzo Variabile per la fornitura di Energia Elettrica edizione 20**

**Lotto 1: Valle d'Aosta, Piemonte**

Sconto PA Virtuose S	0,10 Euro/MWh
Opzione Verde	3,00 Euro/MWh

**Forniture a Prezzo variabile**

Fascia oraria	ΨVAR	dic-22	gen-23	feb-23	mar-23	apr-23	mag-23	giu-23	lug-23	ago-23	set-23	ott-23	nov-23
F0	9,18	304,09	183,67	170,25	145,56	144,15	114,91	114,52	121,27	121,07	124,88	143,44	
F1	9,18	369,91	205,42	183,51	148,96	144,73	119,17	117,38	124,09	119,44	127,73	153,74	
F2	9,18	319,14	193,42	182,07	161,13	161,23	128,99	127,14	131,50	137,30	137,26	157,81	
F3	9,18	254,12	164,28	153,40	133,84	135,58	104,29	105,79	113,52	113,66	115,17	128,26	

Lotto 1: Valle d'Aosta, Piemonte - Prezzi variabili al netto delle perdite di rete (Euro/MWh)

Figura 50. Convenzione Consip: tariffe energia elettrica lotto Piemonte

Fonte [89]

Nella pianificazione delle risorse e nell'infrastruttura complessiva è necessario valutare il consumo massimo possibile e le situazioni di carico elevato, mentre per i consumi in esercizio si effettua una stima in condizioni normali di utilizzo. A questo proposito si usano due informazioni riportate nelle specifiche tecniche delle apparecchiature [87].

- Potenza in ingresso (*input power*): rappresenta la quantità di potenza elettrica assorbita dal dispositivo quando è in funzione. Questo valore riflette il consumo di energia in condizioni normali di operatività.
- Potenza potenziale massima (*maximum potential power*): rappresenta la potenza che può essere richiesta in determinate circostanze, ad esempio durante picchi di utilizzo o situazioni di carico massimo.

<sup>83</sup> *Prezzo Unico Nazionale*: il prezzo all'ingrosso dell'energia elettrica, acquistata sul mercato della Borsa Elettrica Italiana (Italian Power Exchange, IPEX).

La potenza in ingresso è utile per comprendere il consumo energetico medio di un apparato durante il suo funzionamento normale. La potenza massima invece è rilevante in fase di progettazione per garantire che la capacità elettrica sia sufficiente anche durante picchi di utilizzo. Generalmente per le apparecchiature IT la richiesta di energia più copiosa avviene in fase di boot, quando server e apparati si avviano e non sono ancora attivi i controlli sull'efficienza. Una volta in esecuzione, essi si assestano sui consumi medi in base alle attività e alle condizioni in cui operano.

I server *Dell Poweredge R940* selezionati (paragrafo 6.1.2) sono dotati di due alimentatori ridondati. Solitamente in queste configurazioni un alimentatore funziona come primario e l'altro come alimentatore secondario o di backup. In condizioni normali, solo uno dei due alimentatori è attivo e fornisce la potenza necessaria. Se uno degli alimentatori dovesse fallire, l'altro può subentrare per garantire la continuità del servizio. Quindi, per il calcolo del consumo energetico di esercizio viene considerato solo un alimentatore attivo, poiché l'altro alimentatore in standby non contribuisce al consumo attivo finché non è richiesto. La stessa considerazione vale anche per gli switch di rete con alimentazioni ridondanti.

Le schede tecniche fornite dai produttori segnalano le informazioni sui consumi in maniera non uniforme. Ad esempio per lo switch *Juniper EX2300-24P-VC-C* è indicato *Max System Power Consumption 80W*. Per il modello di fascia più alta usato per la sezione leaf *Juniper EX4650-48Y-AFO-C* con due alimentatori ridondanti da 650W, le specifiche indicano più dettagli: *max load 450W; typical load 260W; idle load 160W*.

Per i server *Dell Poweredge R940* tali informazioni si reperiscono utilizzando uno strumento chiamato *Enterprise Infrastructure Planning Tool*, un'applicazione web che Dell mette a disposizione per la configurazione di infrastrutture data center [87].

I ragionamenti sopra esposti convergono nella Tabella 19 che riporta il calcolo dei consumi stimando un carico, in condizioni normali, pari al 60% della potenza massima.

Tabella 19. Costi on-premise: consumi apparecchiature

Apparato	Potenza [W] 1 alimentatore	Potenza max [W]	Stima potenza di esercizio 60% [W]	Numero apparati	Potenza totale di esercizio [kW]
Server	1.400	1.300	780	39	30,42
Switch 1GbE	100	80	48	4	0,19
Switch leaf	650	450	270	4	1,08
Switch spine	950	950	570	2	1,14
KVM switch	20	20	12	4	0,05
KVM console	35	35	21	4	0,08
				<b>Somma</b>	<b>32,96</b>
Raffreddamento					<b>10,99</b>
				<b>Totale</b>	<b>43,95</b>
				<b>Corrispondenza in MWh</b>	<b>0,044</b>

L'installazione di nuovi armadi rack con diverse apparecchiature introduce un carico termico aggiuntivo che può influire sull'impianto di raffreddamento del data center. Per una stima di massima si può considerare che il carico termico sia uguale al consumo energetico supponendo che tutta l'energia consumata dai dispositivi venga convertita in calore [90]. Si tratta di un'approssimazione che non tiene conto dell'efficienza energetica degli apparati, che permetterebbe di ridurre il carico termico generato.

Per una stima dei consumi di raffreddamento è necessario conoscere il coefficiente di prestazione COP<sup>84</sup>; sono state pertanto richieste alcune informazioni all'ufficio manutenzioni del *Settore Building Automation Sistemi di Climatizzazione* di ateneo per un'indicazione di massima. Per l'impianto di raffreddamento si decide di utilizzare un coefficiente pari a 3 che indica la capacità del sistema di rimuovere 3 volte più calore di quanto non consumi in energia elettrica. È importante notare che il COP può variare in base al tipo e all'efficienza del sistema di condizionamento e anche alle condizioni ambientali, assumendo valori più alti in inverno e inferiori nella stagione estiva. Il valore indicato rappresenta una stima, basata sull'esperienza, in grado di compensare tali variazioni. Nella Tabella 19 è stato pertanto calcolato il costo stimato per il consumo elettrico dovuto alle maggiori esigenze di raffreddamento.

Un'osservazione più generale riguardo all'installazione di nuove apparecchiature in un data center esistente è legata alla necessità di un confronto preventivo con il personale competente addetto alla progettazione degli impianti dei locali tecnici. Se l'impianto di raffreddamento esistente è dimensionato per operare efficacemente a carico parziale, in previsione delle nuove attrezzature, l'aggiunta dei rack non dovrebbe comportare aumenti significativi dei consumi energetici e il sistema potrebbe avvicinarsi alla sua capacità massima in modo ottimale. Se invece l'impianto di raffreddamento esistente è già vicino al suo massimo, l'aggiunta del nuovo carico termico potrebbe superare la capacità di raffreddamento disponibile. Per mantenere le temperature desiderate ed evitare il surriscaldamento delle attrezzature, l'impianto di raffreddamento si troverebbe a operare in una condizione subottimale che potrebbe potenzialmente mettere a rischio l'intero locale data center. In questo caso sarebbe opportuna un'eventuale riprogettazione prima della collocazione dei nuovi rack.

È importante anche considerare che la crescita del carico termico può comportare un aumento del *Power Usage Effectiveness* (PUE), l'indicatore dell'efficienza energetica di un data center. Il PUE è calcolato come il rapporto tra l'energia totale consumata dal data center e l'energia assorbita dalle apparecchiature IT. Consente pertanto di valutare quanto dell'energia consumata è effettivamente impiegata per le attrezzature informatiche, rispetto al consumo complessivo di energia, che include gli impianti di illuminazione, raffreddamento, sistemi UPS e altri servizi accessori. Nello scenario

---

<sup>84</sup> *Coefficient of Performance* (COP): rappresenta l'efficienza del sistema. Nel caso del condizionamento è il rapporto tra l'energia termica rimossa e l'energia elettrica consumata.

ideale tutta l'energia consumata è usata esclusivamente per alimentare le attrezzature informatiche e il PUE vale 1. In genere si considera buono un PUE inferiore a 2, mentre un valore pari a 2 indica che una parte significativa dell'energia totale è usata per scopi diversi dai dispositivi IT e che per ogni watt di energia informatica utilizzata per le attrezzature informatiche occorre un watt aggiuntivo per raffreddamento e distribuzione dell'energia.

In conclusione, si riporta nella Tabella 20 il calcolo dei costi energetici sulla base delle tariffe in convenzione Consip, aggiornate a ottobre 2023. Per semplificare la ripartizione del consumo nelle fasce orarie si sono considerati il sabato in fascia 2, la domenica in fascia 3 e tutti gli altri giorni in fascia 1. L'importo risulta sovrastimato ma in tal modo si compensano eventuali eventi sporadici di picchi di carico o consumi.

Tabella 20. Analisi economica: costo annuale per consumi elettrici

Fascia	Costo ott-2023 [€/MWh]	Numero giorni	Costo annuale
F1	153,74 €	261	42.327 €
F2	157,81 €	52	8.656 €
F3	128,26 €	52	7.035 €
<b>Totale</b>			<b>58.018 €</b>

### 6.1.7 Costi del personale

Nel modello di stima dei costi, considerate le attività di configurazione iniziale, manutenzione periodica, risoluzione dei problemi e eventuali ottimizzazioni, si ritengono necessarie almeno 4 risorse con profilo amministratore di sistema e specialista cloud. Tali figure sarebbero da collocare a tempo pieno sulla gestione OpenStack.

L'*Ufficio Trattamenti Economici* del Politecnico di Torino rende disponibili nell'area Intranet le tabelle con i costi del personale tecnico amministrativo dell'Ateneo, usate per le rendicontazioni dei progetti di ricerca. Si ipotizzano due figure inquadrare in categoria C1 e due in categoria D1<sup>85</sup>. I costi risultanti sono calcolati nella Tabella 21.

Tabella 21. Costi on-premise: personale tecnico

Categoria	Costo lordo annuale	Num	Totale annuale
C1	33.656,62 €	2	67.313,24 €
D1	39.839,89 €	2	79.679,78 €
<b>Totale</b>			<b>146.993,02 €</b>

<sup>85</sup> Nella classificazione del personale universitario prevista dall'attuale contratto, la categoria C è caratterizzata da un minore grado di autonomia e di responsabilità; il titolo di studio richiesto per l'accesso esterno è il diploma di scuola secondaria di secondo grado. La categoria D ha un livello maggiore di progettualità e per l'accesso esterno è richiesto il diploma di laurea.

### 6.1.8 Riepilogo dei costi on-premise

In questa sezione si riportano le tabelle di riepilogo con la sintesi dei costi per un'infrastruttura on-premise.

La Tabella 22 mostra i costi totali per le apparecchiature hardware. Ai fini del calcolo del TCO, si considera l'utilizzo di un orizzonte temporale di 5 anni, periodo per il quale è stata anche selezionata l'estensione della garanzia per le forniture Consip.

Nella Tabella 23 vengono calcolati i costi complessivi di apparecchiature, consumi energetici, personale e licenze rapportati ad un anno di esercizio.

Tabella 22. Costi on-premise: apparecchiature hardware

Apparato	Descrizione	Quantità	Prezzo	Totale
Server Dell Poweredge R940 14HDD,1TB RAM	Compute/storage node	32	14.705 €	470.560 €
Server Dell Poweredge R940 2HDD, 1TB RAM	Controller e network node	5	13.050 €	65.250 €
Server Dell Poweredge R940 2HDD, 512GB RAM	Management node	2	8.832 €	17.664 €
Rack, PDU e KVM	Rack e accessori	4	4.864 €	19.456 €
Switch Juniper EX2300-24P	Switch 1GbE	4	654 €	2.616 €
Switch Juniper EX4650-48Y e transceiver	Switch 10GbE (leaf)	4	6.453 €	25.812 €
Switch C9500X-28C8D e transceiver	Switch 40GbE (spine)	2	38.000 €	76.000 €
Cablaggi e servizi	Cavi rame, fibre e posa	1	1.256 €	1.256 €
<b>Totale IVA esclusa</b>				<b>678.614 €</b>
<b>Totale IVA inclusa</b>				<b>827.909 €</b>

Tabella 23. Analisi economica: TCO on-premise su 1 e 5 anni di esercizio

Descrizione	Totale 5 anni	Totale 1 anno
Apparecchiature hardware (server, switch, rack)	827.910 €	165.582 €
Licenze software	61.490 €	12.298 €
Consumi elettrici	290.090 €	58.018 €
Personale	734.965 €	146.993 €
<b>Totale complessivo</b>	<b>1.914.460 €</b>	<b>382.892 €</b>

## 6.2 Infrastruttura su cloud pubblico

Tra i fornitori di Public cloud, si sceglie Microsoft che proprio lo scorso giugno 2023 ha inaugurato la prima regione cloud in Italia, basata su tre data center in Lombardia<sup>86</sup>.

### 6.2.1 Modello di tariffazione per VM Azure

Il servizio IaaS Azure propone due modalità di acquisto per le VM.

- Opzione *pay as you go*. È la modalità classica del cloud pubblico in cui le risorse vengono addebitate in base all'effettivo consumo. È un modello flessibile, ideale per carichi di lavoro variabili o poco prevedibili. È possibile aumentare o ridurre le risorse in base alle esigenze, ma in caso di utilizzo continuativo nel lungo periodo espone a costi potenzialmente più alti.
- Opzione *reserved*. Questa modalità prevede la sottoscrizione di un impegno per l'uso delle risorse a lungo termine (1 anno o 3 anni), con uno sconto significativo rispetto alle tariffe a consumo. È un modello adatto a carichi di lavoro stabili e a lungo periodo, ma è meno flessibile in quanto non prevede variazioni.

Virtual Machines in Azure Your options on selecting the right VMs for your workloads		
Tipo	Dimensioni	Descrizione
Utilizzo generico	B, Dsv3, Dv3, Dasv4, Dav4, Dsv2, Dv2, Av2, DC, DCv2, Dpdsv5, Dpldsv5, Dpsv5, Dplsv5, Dv4, Dsv4, Ddv4, Ddsv4, Dv5, Dsv5, Ddv5, Ddsv5, Dasv5, Dadsv5, DCasv5, DCadsv5	Rapporto equilibrato tra CPU e memoria. Soluzione ideale per test e sviluppo, database medio-piccoli e server Web con traffico da medio a ridotto.
Con ottimizzazione per il calcolo	F, Fs, Fsv2, FX	Rapporto elevato tra CPU e memoria. Soluzione idonea per server Web con livelli medi di traffico, dispositivi di rete, processi batch e server applicazioni.
Ottimizzate per la memoria	Esv3, Ev3, Easv4, Eav4, Epdsv5, Epsv5, Ev4, Esv4, Edv4, Edsv4, Ev5, Esv5, Edv5, Edsv5, Easv5, Eadsv5, Mv2, M, Dsv2, Dv2, ECasv5, ECadsv5	Rapporto elevato tra memoria e CPU. Soluzione ideale per server di database relazionali, cache medio-grandi e analisi in memoria.
Con ottimizzazione per l'archiviazione	Lsv2, Lsv3, Lasv3	I/O e velocità effettiva del disco elevati ideali per i database NoSQL, SQL e Big Data, data warehousing e database transazionali di grandi dimensioni.
GPU	NC, NCv2, NCv3, NCasT4_v3, NC A100 v4, ND, NDv2, NGads V620, NV, NVv3, NVv4, NDasrA100_v4, NDm_A100_v4	Macchine virtuali specializzate ottimizzate per livelli intensivi di rendering della grafica e modifica di video, nonché per il training e l'inferenza dei modelli (ND) con apprendimento profondo. disponibili con GPU singole o più GPU.
High Performance Computing (HPC)	HB, HBv2, HBv3, HBv4, HC, HX	Le nostre macchine virtuali con CPU più veloci e potenti, con interfacce di rete ad alta velocità effettiva facoltative (RDMA).

Figura 51. Cloud Azure: catalogo delle VM.

<sup>86</sup> Link: [https://www.repubblica.it/tecnologia/2023/06/05/news/cloud\\_region\\_microsoft\\_italia-403274136](https://www.repubblica.it/tecnologia/2023/06/05/news/cloud_region_microsoft_italia-403274136)



Fonte [91]

In merito al listino prezzi, l'accordo Microsoft-CRUI garantisce uno sconto dell'8% sui servizi cloud di tipo *pay as you go*, mentre non sono previste riduzioni sulle tariffe riservate a 1 o 3 anni.

Per la scelta delle VM, il catalogo Azure è composto da una gamma di profili diversi, definiti serie, ciascuna con specifiche caratteristiche di ottimizzazione come illustra la Figura 51. Si prendono in esame la Serie Bs e la D. La *Serie Bs* offre VM economiche, progettate per fornire opzioni di calcolo a basso costo e adatte a carichi di lavoro di tipo generale. L'uso della CPU è basso o moderato, ma possono essere gestite situazioni di picco (*burst*) inferiori ai 30 minuti. Anche le VM della *Serie D* sono considerate di calcolo per uso generale, ma offrono processori più veloci, un elevato rapporto memoria/core e eventuali unità SSD per dischi temporanei [91]. All'interno di ciascuna serie sono previsti diversi tagli, in maniera simile ai *flavor* OpenStack, che raggruppano le opzioni in base alla quantità di vCPU, RAM e disco. Si assume come tipologia una VM con la dimensione media delle VM nel cluster Polito pari a 16 vCPU e 64GB di RAM (riferimento al paragrafo 5.3.2).

La Tabella 24 riporta i prezzi per le due tipologie di VM in funzione dei modelli di prezzo e del sistema operativo.

Tabella 24. Cloud Azure: esempio tariffazione VM

<b>Serie VM con s.o. Linux</b>	<b>Modello</b>	<b>Costo €/ora</b>	<b>Sconto CRUI</b>	<b>Costo €/mese</b>
Bs 16ms, 16cores 64RAM 128GB temporary storage	Pay as you go	0,7300 €	8%	490,27 €
	Riservato 1 anno	0,4243 €		309,74 €
	Riservato 3 anni	0,2737 €		199,80 €
D16a v4, 16vCPU 64RAM 400GB temporary storage	Pay as you go	0,8470 €	8%	568,85 €
	Riservato 1 anno	0,4884 €		356,53 €
	Riservato 3 anni	0,3177 €		231,92 €
<b>Serie VM con s.o. Windows</b>	<b>Modello</b>	<b>Costo €/ora</b>	<b>Sconto CRUI</b>	<b>Costo €/mese</b>
Bs 16ms, 16cores 64RAM 128GB temporary storage	Pay as you go	0,7865 €	14%	528,21 €
	Riservato 1 anno	0,4848 €		353,90 €
	Riservato 3 anni	0,3341 €		243,89 €

Sul sito Microsoft Azure è presente un calcolatore<sup>87</sup> per esplorare i prezzi e comporre un preventivo personalizzato, come riportato in Figura 52, inserendo le proprie esigenze:

- zona geografica;
- sistema operativo;
- serie delle VM e profilo delle risorse;
- modello di tariffazione.

The screenshot shows the Azure Virtual Machines pricing calculator interface. At the top, there's a search bar for 'Virtual Machines' and the Azure logo. Below that, a promotional banner offers €200 credit. The main configuration area includes dropdown menus for Region (Italy North), Operating system (Linux), Type (Ubuntu), and Tier (Standard). Below these are Category (General purpose) and Instance Series (All). A specific instance is selected: 'D16 v3: 16 vCPUs, 64 GB RAM, 400 GB Temporary storage, €0.318/hour'. A search bar shows '135' results. The 'Savings Options' section includes a 'Learn more' button. The 'Compute (D16 v3)' section lists options: 'Pay as you go', 'Savings plan' (1 year and 3 year), and 'Reserved instances' (1 year and 3 year reserved). The 'Compute payment options' dropdown is set to 'Upfront'. At the bottom, the total price is displayed as €31,300.50, with a breakdown of average per month costs and upfront charges.

Figura 52. Cloud Azure: calcolatore prezzi VM.

Fonte [91]

<sup>87</sup> Link: <https://azure.microsoft.com/pricing/calculator>

## 6.2.2 Servizi a costi aggiuntivi

Nonostante il calcolatore web dei prezzi sia intuitivo, alcune opzioni non sono di immediata comprensione e richiedono ulteriori approfondimenti per effettuare una scelta di acquisto consapevole.

Rientrano tra queste la selezione del disco (oltre a quello temporaneo incluso nel profilo base), l'attivazione di snapshot, le eventuali transazioni storage e le opzioni sul traffico di rete (Figura 53).

The screenshot displays the 'Managed Disks' section of the Azure pricing calculator, showing various cost components and their calculations:

- Managed Disks:** Total cost €2.89. Tier: Standard HDD. Disk Size: S6: 64 GiB, €2.843/month. 1 Disk at €2.84 Per month = €2.84.
- Snapshot:** Redundancy: LRS. 1 Size of snapshot (GB) at €0.0473 Per GB = €0.05.
- Storage transactions:** 0 Transaction units (10,000 transactions) at €0.0005 Per unit = €0.00.
- Bandwidth:** Total cost €0.00. Data Transfer Type: Inter Region. Source Region: Italy North. Destination Region: West Europe. Outbound Data Transfer: 5 GB. A tooltip states: "The first 5 GB/Month of Inter Region data transfer and the first 100 GB/Month of Internet Egress data transfer are free in each zone."

Figura 53. Cloud Azure: costi aggiuntivi da considerare.

Fonte [91]

Si descrivono di seguito alcuni servizi aggiuntivi che concorrono all'incremento dei costi.

- *Storage persistente.* L'opzione *Temporary Storage* inclusa nei profili delle VM è uno spazio disco che viene cancellato quando la VM viene deallocata oppure riavviata. Se si desidera uno storage a blocchi persistente, simile al concetto di disco fisico locale, è necessario aggiungere l'opzione *Managed Disks* che viene calcolata a parte e tariffata su base mensile.
- *Snapshot dischi.* Su un Managed Disk è possibile creare una snapshot, cioè una copia istantanea del contenuto del disco ai fini di backup. Le snapshot sono fatturate a parte in base alla dimensione effettiva dei dati e non dell'intero disco.
- *Transazioni storage.* Sono le operazioni di lettura e scrittura eseguite sullo storage Azure; è considerata transazione ogni chiamata REST eseguita verso lo storage. Non sono calcolabili in modo immediato e comportano addebiti aggiuntivi se superano il valore di 10.000. Possono incidere se si usano servizi Azure storage (archiviazione File, storage a oggetti Blob); nel caso di VM dovrebbero essere limitate alle operazioni su eventuali Managed Disk: ad esempio la creazione di snapshot genera *storage transactions*.
- *Larghezza di banda.* Si riferisce al trasferimento di dati all'interno e all'esterno dei data center Azure e tra data center diversi. Non è soggetto a tariffazione il flusso di dati in ingresso (*inbound*), cioè il traffico in entrata da Internet. Viene invece addebitato il traffico in uscita dalla rete Azure (*outbound*) che eccede due soglie. Il primo limite di 5GB al mese riguarda il traffico tra regioni, il secondo di 100GB al mese è relativo al traffico dati in uscita verso Internet da ogni zona.
- *Servizi di rete.* Hanno un costo supplementare eventuali servizi di rete che consentono la connettività verso la propria rete aziendale (ad es. VPN o Azure *Express Route*), oltre che la dotazione di indirizzi IP pubblici e IP riservati.

Inter Region					
<b>Intra-continental Data Transfer</b>			<b>Price</b>		
Between regions within North America			€0.019 per GB		
Between regions within Europe					
<b>Inter-continental Data Transfer</b>			<b>Price</b>		
From North America to other continents			€0.048 per GB		
From Europe to other continents					
From Asia to other continents			€0.076 per GB		
From Oceania to other continents					
From Africa to other continents					
<b>Internet Egress (routed via Microsoft Premium Global Network)</b>					
Source Continent	First 100GB / Month	Next 10TB / Month	Next 40TB / Month	Next 100TB / Month	Next 350TB / Month
From North America, Europe to any destination	Free	€0.0823 per GB	€0.0785 per GB	€0.0662 per GB	€0.0473 per GB
From Asia (China excluded), Australia, MEA to any destination	Free	€0.1135 per GB	€0.0804 per GB	€0.0776 per GB	€0.0757 per GB

Figura 54. Cloud Azure: costi per traffico di rete in uscita.

Fonte [91]

### 6.3 Scenari di confronto

In questa sezione si definiscono gli scenari usati per la comparazione dei costi tra l'infrastruttura on-premise e l'esecuzione delle VM nel cloud Azure

#### 6.3.1 Scenario 1

Il punto di partenza di questa analisi è la quantità di risorse offerte dall'infrastruttura on-premise delineata nel paragrafo 6.1.2. Al netto di una quota di overhead riservata agli hypervisor, vi sono a disposizione un totale di 5.530 vCPU e 29TB di RAM. Si tralasciano in questa sede le risorse storage in quanto sono soggette a valutazioni diverse per la ridondanza e comunque disponibili in misura sufficiente per gli scopi.

Si ipotizza che l'infrastruttura on-premise possa operare con diverse percentuali di carico, prendendo come riferimento la VM media nel cluster Polito (16 vCPU e 64GB di RAM, si veda 5.3.2) e si calcola un preventivo Azure per la stessa quantità di VM. I valori risultanti da tali ragionamenti sono indicati in Tabella 25.

Tabella 25. Analisi economica: ipotesi di carico on-premise e cloud

Infrastruttura on-premise		Tot Core (hypertthread)	Tot RAM [TB]
Risorse lorde		6.144	32
Risorse nette (10% overhead per hypervisor)		5.530	28,80

Ipotesi Carico	Tot Core usati	Tot RAM usata [TB]	Num VM medie (16 vCPU e 64GB RAM)	Num VM Azure
50%	2.765	14,40	172	146 Linux + 26 Windows
70%	3.871	20,16	242	206 Linux + 36 Windows
90%	4.977	25,92	311	264 Linux + 47 Windows

La stima sul calcolatore Azure segue i seguenti criteri.

- scelta VM della serie Bs, più economica rispetto alla serie D (si rimanda alla Tabella 24 per i dettagli);
- zona geografica per la collocazione: Italia nord;
- modalità di tariffazione riservata a 3 anni con pagamento anticipato, che risulta la più conveniente economicamente;
- storage: ogni VM viene dotata di un disco persistente di 64GB, oltre al disco temporaneo compreso di default;
- licenza del sistema operativo inclusa per le VM Windows (circa 15% del totale);
- opzioni: non si aggiungono transazioni storage e si mantengono i valori di default per il traffico di rete in uscita.

La quotazione totale per le VM in Azure viene riportata nella Tabella 26. La modalità prescelta prevede il pagamento anticipato, tuttavia Microsoft mantiene separato il costo dei dischi persistenti, fatturato su base mensile.

Tabella 26. Cloud Azure: scenario 1, preventivo costi VM

Descrizione VM e servizi	Costi mensili	Costi anticipati
146 B16ms (16 Cores, 64 GB RAM) (3 year res.), Linux; 146 managed disks S6; Inter Region transfer type, 5 GB outbound data transfer from Italy North to West Europe	415,13 €	1.049.837,68 €
26 B16ms (16 Cores, 64 GB RAM) (3 year res.), Windows (License included); 26 managed disks S6; Inter Region transfer type, 5 GB outbound data transfer from Italy North to West Europe	1.222,16 €	186.957,40 €
<b>Totale IVA esclusa</b>	<b>1.637,29 €</b>	<b>1.236.795,08 €</b>
<b>Totale IVA inclusa</b>	<b>1.997,49 €</b>	<b>1.508.890,00 €</b>

Nonostante l'orizzonte temporale per l'infrastruttura on-premise sia distribuito su 5 anni, si mantiene il confronto tra le due soluzioni su base annuale e su 3 anni, in quanto per il cloud Azure la proiezione dei costi di ulteriori 2 anni non sarebbe affidabile. La Tabella 27 presenta i risultati.

Tabella 27. Analisi economica: confronto costi on-premise e cloud

Periodo	On-premise	Cloud Azure	Cloud Azure - On-premise
1 anno	382.892 €	526.933 €	144.041 € (+38%)
3 anni	1.148.676 €	1.580.800 €	432.124 € (+38%)

Considerando solo l'esecuzione delle VM, la soluzione cloud Azure costa il 38% in più rispetto a una potenziale infrastruttura on-premise. Il divario aumenterebbe ulteriormente se si aggiungessero anche i costi relativi al personale interno addetto alle operazioni sulla sottoscrizione Azure nonché eventuali costi cloud fatturati a consumo come traffico in uscita, snapshot, servizi di rete. Visti gli esiti dei calcoli, non si ritiene utile proseguire con il confronto degli scenari di carico dell'infrastruttura on-premise al 70% e 90% presentati in Tabella 25.

### 6.3.2 Scenario 2

In questo scenario si propone un ulteriore confronto tra le due soluzioni.

A partire dal costo annuale dell'infrastruttura on-premise, calcolato nel paragrafo 6.1.8 e riportato nella Tabella 23, si determina il numero di VM che si potrebbero acquistare nel cloud Azure a parità di esborso. I risultati emersi dallo scenario precedente forniscono già una previsione, ma si riporta l'analisi per valutare un eventuale punto di pareggio annuale.

Il TCO on-premise per un anno esercizio ammonta a 382.892 €. Con tale cifra si possono acquisire due opzioni di servizio su Azure. Con l'opzione di risorse riservate a 1 anno si possono eseguire 73 VM Linux e 10 VM Windows, mentre con l'opzione di risorse riservate a 3 anni si possono ottenere 111 VM Linux e 15 VM Windows. Come nei casi precedenti, si mantiene la selezione di VM della serie più economica Bs equipaggiate con 16 core, 64GB di RAM e disco persistente 64GB. I risultati sono descritti nella Tabella 28.

Tabella 28. Analisi economica: punto di pareggio Cloud e TCO on-premise

Descrizione VM e servizi	Costi mensili	Costi anticipati	Costo 1 anno
73 VM Linux + 10 VM Windows Serie Bs 16 Core, 64GB RAM, disco persistente 64GB Opzione riservata 1 anno	826,71 €	376.364,21 €	386.284,73 €
111 VM Linux + 15 VM Windows Serie Bs 16 Core, 64GB RAM, disco persistente 64GB Opzione riservata 3 anni	1.245,26 €	1.105.349,65 €	383.393,00 €

Come è da aspettarsi, il caso più favorevole si manifesta con l'impegno contrattuale a 3 anni, che tuttavia richiede di anticipare un importo consistente. In ogni caso, in entrambi i casi il numero di VM sul cloud Azure è inferiore alle VM che potrebbero essere eseguite sull'infrastruttura on-premise a metà del suo carico massimo.

### 6.3.3 Osservazioni finali

I due scenari precedenti convergono sulle conclusioni analoghe che, per il tipo di carico di VM esaminato, è economicamente più conveniente l'esecuzione su un'infrastruttura on-premise piuttosto che sul cloud Azure.

Dalla Tabella 27 emerge che la soluzione cloud Azure costa il 38% in più rispetto a una potenziale infrastruttura on-premise, utilizzando le VM della serie più economica e senza considerare spese aggiuntive per eventuali costi a consumo e per il personale interno che gestisca le risorse cloud.

Dalla Tabella 28 si evince che il punto di pareggio tra le due soluzioni per un anno di esercizio si ottiene con un contratto Azure di tipo triennale e con un numero di VM in cloud inferiore rispetto alle VM che si potrebbero eseguire on-premise utilizzando il 50% delle risorse hardware. In particolare, a parità di spesa, on-premise è possibile creare il 37% di VM in più.

In generale, le due soluzioni sono accomunate da un elevato esborso iniziale, che si verifica sia per l'eventuale contratto triennale Microsoft (1.105.349,65€, si veda Tabella 28) sia per le spese di acquisto dell'hardware (827.910€, si veda Tabella 22).

Il cloud potrebbe essere utilizzato come banco di prova in contesti di applicazione sporadici. Potrebbe essere indicato nel caso sia necessario eseguire carichi di breve durata ma con elevate esigenze di prestazioni, ad esempio per simulazioni che richiedono l'esecuzione di batterie di processi worker che non hanno bisogno di disco persistente. In tali scenari, è possibile valutare l'opzione di fatturazione a consumo nel cloud, consentendo poi di dimensionare il carico di lavoro e determinare se posizionare eventualmente i processi sul cluster on-premise.

L'attivazione di una sottoscrizione di test su Microsoft Azure potrebbe risultare vantaggiosa per esaminare alcuni costi difficili da stimare preventivamente, come le transazioni di storage o i costi associati al traffico di rete in uscita. Spesso, quando si richiede una VM, l'utente potrebbe non avere una chiara comprensione del tipo di carico di lavoro che sarà eseguito. In alcuni casi sono presenti specifiche richieste per l'installazione di software personalizzato o requisiti provenienti da terze parti, come consulenti esterni o aziende. Di conseguenza, il docente o il ricercatore potrebbe avere poche informazioni sulla natura del carico di lavoro e una conoscenza limitata del comportamento della VM nella rete. La sottoscrizione di test fornirebbe quindi una modalità efficace per acquisire maggiori informazioni e visibilità su tali aspetti.

L'utilizzo di un cloud interno offre ulteriori vantaggi. Molti servizi che potrebbero comportare costi aggiuntivi in un cloud pubblico sono inclusi nel cloud privato Polito: tra questi l'uso di indirizzi IP pubblici per le VM e la possibilità di effettuare snapshot dei volumi per scopi di backup o ripristino delle VM. Dal punto di vista della rete non vi sono costi o limitazioni relativi al traffico ed è possibile usufruire di connessioni ad alta velocità grazie ai collegamenti di ateneo alla rete GARR.

In sintesi, l'infrastruttura on-premise oltre alla convenienza economica offre controllo diretto e maggiori personalizzazioni, mentre il cloud Azure offre agilità immediata e, se usato in modalità a consumo può consentire una componente di elasticità con una modalità di pagamento più flessibile.

In merito all'implementazione di un cloud privato per un'università pubblica come il Politecnico di Torino, ulteriori considerazioni si svincolano dall'ambito economico e riguardano temi più ampi come la sovranità digitale, la rispondenza a eventuali normative, il rischio di lock-in che vengono affrontati nel capitolo delle conclusioni.



# Capitolo 7.

## Conclusioni e spunti futuri

L'obiettivo di analizzare l'implementazione del cluster OpenStack come regione del modello federato GARR presso il Politecnico di Torino si è concretizzato nella creazione di un percorso stimolante tra molteplici argomenti collegati. Questo lavoro ha infatti rappresentato l'occasione per guardare alle attività svolte in un modo organico, senza limitarsi alla realizzazione tecnica e ai suoi effetti.

L'esame del contesto teorico e della diffusione del cloud computing ha consentito di acquisire consapevolezza sull'evoluzione di questa tecnologia e contestualizzarla nello scenario normativo e organizzativo di un'università pubblica. La fase di piena maturità e la trasformazione innescata dai nuovi modelli hanno infatti raggiunto anche il settore pubblico che ha investito il cloud di un ruolo abilitante; non a caso molti paesi hanno inserito la politica *cloud first* nelle proprie linee di indirizzo.

Il cloud ha da tempo attirato anche l'interesse del mondo scientifico sia perché non tutte le organizzazioni hanno fondi e competenze per gestire dei propri data center sia perché alcune caratteristiche del modello cloud sono particolarmente adatte a svolgere attività computazionali che richiedono un numero di risorse storage e di elaborazione sempre crescente. Le università e gli enti di ricerca pubblici si collocano giuridicamente come PA e si trovano a dover bilanciare il rispetto dei vincoli normativi e la gestione delle esigenze meno standardizzate e peculiari del mondo accademico e della ricerca. Inoltre, in deroga alle disposizioni AgID ed esclusivamente per le necessità legate alla ricerca, hanno la facoltà di mantenere e far evolvere i propri data center on-premise.

Il mercato dei servizi cloud è in crescita, ma meno del 15% dei provider è europeo: ciò porta a interrogarsi sulla dipendenza tecnologica dell'Europa dalle multinazionali statunitensi, fornitrici di servizi anche attraverso i contratti di appalti pubblici. Le direttive della strategia digitale della Commissione europea, pur riconoscendo il valore del cloud quale tecnologia abilitante, sono focalizzate sui temi della sovranità digitale e sulle politiche di standardizzazione e interoperabilità. In questo contesto, le reti europee di ricerca tra cui GARR, si stanno da tempo muovendo per valutare soluzioni cloud che possano armonizzare le necessità di indipendenza, sicurezza e controllo mantenendo al contempo alti standard tecnologici.

In questo scenario si colloca la realizzazione del cloud privato OpenStack all'interno del Politecnico di Torino, che è stato ed è tutt'ora un progetto stimolante e ambizioso.

L'adesione al modello cloud GARR ha permesso di implementare un'architettura replicabile, progettata per esigenze comuni, basata interamente su software open source. Oltre all'accrescimento di competenze tecniche per tutti i partecipanti, la ricaduta positiva per l'Ateneo è stata l'ampliamento di servizi verso la comunità della ricerca, a cui è offerta la possibilità di disporre di ampie risorse computazionali talvolta difficilmente acquisibili e gestibili a livello di singolo gruppo di ricerca o di dipartimento.

Pur mantenendo l'autonomia sulle proprie risorse, le caratteristiche del modello GARR quali l'organizzazione della struttura federata e l'autenticazione centralizzata aprono a opportunità di collaborazione anche in termini di scambio di risorse, consentendo a gruppi di ricerca di mettere a fattore comune eventuali risorse private.

In aggiunta alla debita panoramica sugli aspetti tecnologici e sulle attività di installazione e configurazione, questo lavoro ha avviato lo spunto per interrogarsi sulla diffusione e le modalità di utilizzo delle risorse a disposizione, nonché sul lato economico del cloud privato Polito. Per delineare un'analisi economica di sintesi, è stato definito un modello di alto livello simulando l'implementazione di una nuova infrastruttura on-premise con risorse paragonabili a quelle dell'attuale cluster OpenStack.

Il presupposto di partenza è che il locale data center fosse esistente, operativo e attrezzato con gruppi di continuità e impianto di raffrescamento. Sono stati presi in considerazione i costi per l'acquisto delle apparecchiature hardware, manutenzione, contratti di assistenza, servizi accessori, consumi energetici, eventuali licenze software e costi relativi al personale IT addetto alla gestione.

La stima del TCO ottenuto è stata confrontata con i costi per l'esecuzione di un numero paragonabile di VM in un cloud pubblico, scegliendo come fornitore Microsoft Azure che ha inaugurato lo scorso giugno 2023 la prima regione cloud in Italia, disposta su tre data center in Lombardia. Gli scenari esaminati mostrano che, per il tipo di carico di lavoro, è economicamente più conveniente l'esecuzione su un'infrastruttura on-premise piuttosto che sul cloud Azure. In particolare la soluzione cloud Azure costa il 38% in più rispetto a una potenziale infrastruttura on-premise, utilizzando le VM della serie più economica e senza considerare spese aggiuntive per eventuali costi a consumo e per il personale interno che gestisca le risorse cloud.

Analizzando i dati per un anno di esercizio si può determinare che il punto di pareggio tra le due soluzioni si ottiene con un contratto Azure di tipo triennale e con un numero di VM in cloud inferiore rispetto alle VM che si potrebbero eseguire on-premise utilizzando il 50% delle risorse hardware. In particolare, a parità di esborso economico, on-premise è possibile creare il 37% di VM in più.

Se flessibilità e scalabilità sono caratteristiche indubbie del cloud pubblico, non si può dire lo stesso in merito alla convenienza economica, sebbene fosse uno dei vantaggi

decantati con l'avvento della tecnologia. Questo aspetto pone un interrogativo in merito alle disposizioni normative della Strategia Cloud Italia che prevedono per tutte le PA italiane il piano di migrazione delle proprie risorse locali verso i cloud accreditati. Accomunati dallo stesso destino sono sia gli enti di piccole dimensioni, sia amministrazioni più strutturate che avrebbero con capacità tecniche e organizzative per mantenere i servizi on-premise.

Le considerazioni emerse non significano necessariamente che il cloud pubblico sia da evitare. L'attivazione di una sottoscrizione di test su Microsoft Azure potrebbe risultare vantaggiosa per esaminare alcuni costi difficili da stimare preventivamente, oppure per valutare le prestazioni di determinate VM. In generale il cloud pubblico potrebbe essere utilizzato come banco di prova in contesti di applicazione sporadici, ad esempio nel caso sia necessario eseguire carichi di breve durata ma con elevate esigenze di prestazioni, ad esempio per simulazioni che richiedono l'esecuzione di batterie di processi worker senza la necessità di disco persistente. In tali scenari, sarebbe possibile valutare l'opzione di fatturazione a consumo, consentendo poi di dimensionare il carico di lavoro e determinare se posizionare eventualmente i processi sul cluster on-premise.

Per un'università pubblica come il Politecnico di Torino, si ritiene vantaggioso poter contare su un cloud privato interno. In primo luogo vi è una maggiore flessibilità per la personalizzazione delle VM e possono essere definiti progetti su misura che servano obiettivi specifici, in contrapposizione alle proposte precostruite dei cloud commerciali. Inoltre sono disponibili molti servizi che comporterebbero costi aggiuntivi in un cloud pubblico, come l'uso di indirizzi IP pubblici e la possibilità di effettuare snapshot dei volumi per scopi di backup o ripristino di VM.

Dal punto di vista della rete non vi sono costi o limitazioni relativi al traffico, come invece avviene su Azure, ed è possibile usufruire di connessioni ad alta velocità grazie ai collegamenti di ateneo alla rete GARR, rendendo la piattaforma adatta alle applicazioni che richiedono basse latenze.

La scelta di soluzioni open source e la collocazione on-premise contribuiscono inoltre a minimizzare il rischio di vendor lock-in. Un ulteriore aspetto decisamente rilevante, soprattutto in tema di sovranità digitale e di indipendenza della ricerca, è che il Private cloud consente di mantenere la sovranità su dati e tecnologie software, grazie alla collocazione delle infrastrutture in un data center gestito in maniera diretta. Si ottiene pertanto un maggiore controllo sull'intero ciclo di vita dei dati, anche nel rispetto di vincoli legislativi come ad esempio il GDPR. Non da ultimo il cloud privato permette di creare e mantenere le competenze tecnologiche all'interno dell'ateneo e della comunità, favorendo lo scambio e la collaborazione con altre reti della ricerca.

Ciò è quanto in effetti è avvenuto con il progetto di realizzazione del cloud privato OpenStack all'interno del Politecnico di Torino, che ha portato ad una crescita di competenze tecniche per tutti i partecipanti. Al momento il cluster è operativo e offre servizi IaaS e un virtual data center per l'esecuzione di VM. In futuro sul piano organizzativo, nell'ottica di una maggiore diffusione del servizio, sarebbe opportuno perfezionare le strategie di comunicazione in ateneo e definire delle policy per regolare l'utilizzo del cloud. Per i gruppi di ricerca che lavorano su progetti, visto anche l'aumento delle attività legate al PNRR, sarebbe molto utile implementare un sistema di contabilizzazione che permetta di inserire i costi delle VM impiegate nelle rendicontazioni dei progetti, in modo da realizzare una compartecipazione alle spese di gestione dell'infrastruttura.

Dal punto di vista tecnico, OpenStack è un sistema estremamente complesso e richiede competenze avanzate di tipo trasversale. Da un lato, in vista della crescente complessità hardware e software, si rivela fondamentale ottimizzare le attività di gestione puntando su eventuali nuovi strumenti che forniscano maggiore automazione e capacità di monitoraggio. Dall'altro lato è necessario poter contare su un team dedicato e adeguatamente preparato, rafforzando le competenze interne attraverso iniziative di formazione sia avviando strategie per l'organizzazione e il reclutamento di nuovo personale.

In conclusione, riportando il focus su un ambito più generale, è indubbio che l'adesione al cloud pubblico ha il vantaggio innegabile di poter disporre di ampie risorse computazionali e di storage, senza investimenti di capitale, molto conveniente per poter avviare le proprie attività, come nello scenario di piccole startup. Tuttavia è ormai assodato che non tutti i carichi di lavoro di dati e applicazioni sono adatti al cloud pubblico e gioca un ruolo sempre più importante l'analisi su dove e come eseguire i servizi. La maggior parte delle aziende non cerca più una risposta unica e onnicomprensiva alle proprie esigenze IT quanto piuttosto una struttura che possa soddisfare i requisiti di costo, prestazioni e controllo dei diversi carichi di lavoro. I servizi caratterizzati da modelli d'uso noti con consumi di risorse stabili e prevedibili, o che risentono in caso di latenze eccessive possono essere collocati su cloud privati, lasciando invece sul cloud pubblico le applicazioni contraddistinte da picchi di utilizzo e traffico imprevedibile che richiedono scalabilità.

Secondo quanto presentato dai risultati della Ricerca dell'Osservatorio Cloud Transformation del Politecnico di Milano, discussi nel convegno dello scorso 5 ottobre 2023<sup>88</sup>, molte organizzazioni stanno raggiungendo una solida consapevolezza del cloud computing e puntano pertanto agli ambienti ibridi per il loro IT. Gli analisti prevedono

---

<sup>88</sup> Convegno Cloud Transformation 2023. Link: [https://blog.osservatori.net/it\\_it/eventi-cloud-transformation](https://blog.osservatori.net/it_it/eventi-cloud-transformation)

che a tendere il cloud pubblico coesisterà con le soluzioni private on-premise/colocation. Dello stesso avviso sono stati alcuni interventi durante l'annuale workshop della comunità GARR<sup>89</sup> il 2023 svoltosi a Roma dall'8 al 10 novembre 2023.

Come avviene in molteplici settori, la chiave del successo risiede nell'integrazione tra diverse soluzioni: in questo caso mantenendo le caratteristiche più apprezzate del modello operativo cloud, l'agilità e l'erogazione efficiente dei servizi, la direzione è sicuramente verso un'infrastruttura IT ibrida o in un ambiente ibrido/multi-cloud.

---

<sup>89</sup> Massimo Carboni, "Evoluzione del modello cloud GARR". Link: <https://ws23.garr.it/>

# Bibliografia

- [1] R. Chellappa, «Intermediaries in cloud-computing: A new computing paradigm,» in *INFORMS Annual Meeting*, Dallas, Texas, 1997.
- [2] P. Mell e T. Grance, «The NIST Definition of Cloud Computing,» 2011.
- [3] S. Bhowmik, *Cloud Computing*, Cambridge University Press, 2017, pp. Cloud Deployment Models 66-70.
- [4] M. J. Kavis, *Architecting The Cloud Design Decisions for Cloud Computing Service Models (SaaS, PaaS, and IaaS)*, Wiley, 2014.
- [5] H. B. Patel e N. Kansara, «Cloud Computing Deployment Models: A Comparative Study,» *International Journal of Innovative Research in Computer Science & Technology*, vol. 9, marzo 2021.
- [6] H. Al-Baroodi, S. Manickam e M. Aboalmaaly, «The classification and arts of open source cloud computing: a review,» *Advances in Information Sciences & Service Sciences*, vol. 5, 1 gennaio 2013.
- [7] A. Barron, «Pizza as a Service,» 30 luglio 2014. [Online]. Available: <https://www.linkedin.com/pulse/20140730172610-9679881-pizza-as-a-service>. [Consultato il giorno 20 novembre 2023].
- [8] The Linux Foundation, «Kubernetes Documentation,» 19 settembre 2023. [Online]. Available: <https://kubernetes.io/it/docs/concepts/overview/what-is-kubernetes/>. [Consultato il giorno 20 novembre 2023].
- [9] Eurostat, European Union, «Cloud computing - statistics on the use by enterprises,» Eurostat, 2022.
- [10] Istat - Istituto nazionale di statistica, «Rapporto annuale 2021,» Roma.
- [11] B. Balabio e D. Gatti, «Mercato Cloud Computing 2022 in crescita nonostante l'instabilità geopolitica ed economica,» 6 ottobre 2022. [Online]. Available: <https://www.osservatori.net/it/ricerche/comunicati-stampa/cloud-italia-mercato>. [Consultato il giorno 30 ottobre 2023].

- [12] P. Todorovich, «Il cloud ormai vicino alla piena maturità,» 27 agosto 2023. [Online]. Available: <https://www.zerounoweb.it/cloud-computing/il-cloud-ormai-vicino-alla-piena-maturita/>. [Consultato il giorno 30 ottobre 2023].
- [13] S. Moore, «Gartner says more than half of enterprise IT spending in key market segments will shift to the cloud by 2025,» 9 febbraio 2022. [Online]. Available: <https://www.gartner.com/en/newsroom/press-releases/2022-02-09-gartner-says-more-than-half-of-enterprise-it-spending>. [Consultato il giorno 30 ottobre 2023].
- [14] A. Khajeh-Hosseini, D. Greenwood e I. Sommerville, «Cloud migration: A case study of migrating an enterprise IT system to IaaS,» in *IEEE 3rd International Conference on Cloud Computing*, Miami, FL, USA, 2010.
- [15] M. F. Gholami, F. Daneshgar, G. Low e G. Beydoun, «Cloud migration process. A survey, evaluation framework and open challenges,» *The Journal of Systems and Software*, vol. 120, 28 giugno 2016.
- [16] P. Scandurra, G. Psaila, R. Capilla e R. Mirandola, «Challenges and assessment in migrating IT legacy applications to the cloud,» in *MESOCA 2015: 9th IEEE Symposium on the Maintenance and Evolution of Service-Oriented Systems and Cloud-Based Environments*, Brema, Germania, 2015.
- [17] A. Alkhalil, R. Sahandi e D. John, «An exploration of the determinants for decision to migrate existing resources to cloud computing using an integrated TOE-DOI model,» *Journal of Cloud Computing*, vol. 6, n. 1, pp. 1-20, dicembre 2017.
- [18] N. Ahmad, Q. N. Naveed e N. Hoda, «Strategy and procedures for Migration to the Cloud Computing,» in *IEEE 5th International Conference on Engineering Technologies & Applied Sciences*, Bangkok Thailand, 2018.
- [19] D. S. Linthicum, «Cloud-Native Applications and Cloud Migration: The Good, the Bad, and the Points between,» *IEEE Cloud Computing*, vol. 4, n. 5, pp. 12 - 14, 1 dicembre 2017.
- [20] Y.-C. Lee e A. Zomaya, «Energy efficient utilization of resources in cloud computing systems,» *The Journal of Supercomputing*, vol. marzo, n. 2, pp. 268-280, 2012.
- [21] F. D. Rossi, M. G. Xavier, C. A. De Rose, R. N. Calheiros e R. Buyya, «E-co: Performance-aware energy-efficient cloud data center orchestration,» *Journal of Network and Computer Applications*, vol. 78, pp. 83-96, 15 gennaio 2017.
- [22] A. Strunk, «A lightweight model for estimating energy cost of live migration of virtual machines,» in *IEEE Sixth International Conference on Cloud Computing*, Santa Clara, CA, USA, giugno 2013.

- [23] K. Rybina e A. Schill, «Estimating energy consumption during live migration of virtual machines,» in *IEEE BlackSeaCom, International Black Sea Conference on Communications and Networking*, Varna, Bulgaria, 06-09 giugno 2016.
- [24] N. S. More e R. B. Ingle, «Research issues for energy-efficient cloud computing,» in *2nd ICICC, Intelligent Computing and Information and Communication Conference*, Pune, India, 2-4 agosto 2017.
- [25] A. Vafamehr e M. E. Khodayar, «Energy-aware cloud computing,» *The Electricity Journal*, vol. 31, n. 2, pp. 40-49, marzo 2018.
- [26] Z. Pantić e M. Babar, «Guidelines for Building a Private Cloud Infrastructure,» Copenhagen, Denmark, 2012.
- [27] N. A. Salleh, H. Hussin, M. A. Suhaimi e A. Md Ali, «A systematic literature review of cloud computing adoption and impacts among small medium enterprises (SMEs),» in *International Conference on Information and Communication Technology for the Muslim World (ICT4M)*, Kuala Lumpur, Malaysia, 2018.
- [28] O. Jayeola, S. Sidek, A. A. Rahman, A. S. B. Mahomed e J. Hu, «Cloud Computing adoption in small and medium enterprises (SMEs): a systematic literature review and directions for future research,» *International Journal of Business and Society*, vol. 23, n. 1, pp. 226-243, 31 marzo 2022.
- [29] B. Bashari Rad, T. Diaby e M. Ehsan Rana, «Cloud computing adoption: A short review of issues and challenges,» in *ICEEG 2017: Proceedings of the 1st International Conference on E-commerce, E-Business and E-Government*, Turku Finland, 2017.
- [30] M. A. Suhaimi, H. Hussin, A. M. Ali e N. H. M. Puad, «A Review on Cloud Computing Adoption from the Perspectives of Providers and Clients,» in *The 6th International Conference on Cyber and IT Service Management (CITSM 2018)*, Medan, Indonesia, 2018.
- [31] D. Catteddu e G. Hogben, «Cloud Computing Risk Assessment,» European Network and Information Security Agency (ENISA), 2009.
- [32] J.-M. Brook, A. Stone Getsin e M. Roza, «Top Threats to Cloud Computing - Pandemic Eleven,» Cloud Security Alliance (CSA), 2022.
- [33] M. Sullivan, P. Kamleshkumar Kishnani, J. Mariani e M. Jackson, «Don't just adopt cloud computing, adapt to it,» Deloitte Consulting, 2022.
- [34] D. C. Wyld, «The cloudy future of government it: Cloud Computing and the public sector around the world,» *International Journal of Web & Semantic Technology (IJWesT)*, vol. 1, n. 1, gennaio 2010.



- [35] H. Sallehudin, R. Che Razak, M. Ismail, A. F. Md Fadzil e R. Baker, «Cloud Computing Implementation in The Public Sector: Factors and Impact,» 2018.
- [36] T. Branco, I. Bianchi e F. de Sá-Soares, «Cloud Computing Adoption in the Government Sector in Brazil: An Exploratory Study with Recommendations from IT Managers,» 2019.
- [37] A. Dahiru e H. Abubakar, «Cloud Computing Adoption: A Cross-Continent Overview of Challenges,» *Nigerian Journal of Basic and Applied Sciences*, vol. 25, n. 1, 2018.
- [38] E. Cidres, A. Vasconcelos e F. Leiteao, «Cloud calculator: A cloud assessment tool for the public administration,» in *21st Annual International Conference on Digital Government Research*, Seoul, Republic of Korea, giugno 2020.
- [39] K. E. Ali, S. A. Mazen e E. Hassanein, «A proposed hybrid model for adopting cloud computing in e-government,» *Future Computing and Informatics Journal*, vol. 3, n. 2, pp. 286-295, dicembre 2018.
- [40] S. Jones, Z. Irani, U. Sivarajah e P. E. Love, «Risks and rewards of cloud computing in the UK public sector: A reflection on three Organisational case studies,» *Information Systems Frontiers*, vol. 21, n. 2, pp. 359-382, 2019.
- [41] M. Draoli, «Public sector study. FP7-610650 Cloud for Europe,» Commissione europea. Servizio Comunitario di Informazione in materia di Ricerca e Sviluppo (CORDIS), 2014.
- [42] A.-M. Zaharia-Rădulescu e I. Radu, «Cloud computing and public administration: approaches in several European countries,» in *11th ICBE: International Conference on Business Excellence*, Bucharest, Romania, marzo 2017.
- [43] J. L. Marín, «European cloud services: Big Tech's profits grow thanks to public sector contracts,» *El Orden Mundial*, 2022.
- [44] A. Fuller, «How the Public Cloud has transformed university IT,» 17 gennaio 2023. [Online]. Available: <https://kion.io/resources/how-the-public-cloud-has-transformed-university-it>. [Consultato il giorno 30 ottobre 2023].
- [45] M. Mircea e A. Andreescu, «Using Cloud Computing in higher education. A strategy to improve agility in the current financial crisis,» *Communications of the IBIMA*, vol. 2011, 2011.
- [46] H. Aydin, «A Study of Cloud Computing Adoption in Universities as a Guideline to Cloud Migration,» *SAGE Open*, vol. 11, n. 3, pp. 1-14, luglio-settembre 2021.

- [47] O. K. Jasim Mohammad, M. Alwan e O. Abduljabbar, «Campus Cloud Computing for Universities: State-of-the-Art,» 28 marzo 2022.
- [48] A. R. Alzoubaidi, «Multi-Campus universities Private-Cloud migration infrastructure,» *International Journal on Cloud Computing: Services and Architecture*, vol. 6, n. 3, giugno 2016.
- [49] N. Kyriakou, Z. Lachana, D. N. Skoutas, C. Skianis e Y. Charalabidis, «Achieving Seamless Migration to Private-Cloud Infrastructure for Multi-Campus Universities,» 2023.
- [50] N. Sultan, «Cloud computing for education: A new dawn?,» *International Journal of Information Management*, vol. 30, n. 2, pp. 109-116, aprile 2010.
- [51] K. Marken e L. Jain, «Deployment of Private Cloud in Campus Wide Network,» *International Journal for Research in Applied Science and Engineering Technology*, vol. 8, n. 7, luglio 2020.
- [52] M. B. Ali, «Multiple Perspective of Cloud Computing Adoption Determinants in Higher Education a Systematic Review,» *International Journal of Cloud Applications and Computing*, vol. 9, n. 3, pp. 89-109, dicembre 2019.
- [53] K. Jewargi, «Public cloud to cloud repatriation trend,» *Scholars Journal of Engineering and Technology*, vol. 11, n. 1, 2023.
- [54] M. Bailey e M. Eastwood, «Cloud Repatriation Accelerates in a Multi-Cloud World,» IDC, 2018.
- [55] M. Posey, D. Thompson e P. Schweizer, «Cloud repatriation: What it is, what it isn't, and why it's not going away,» S&P Global Market Intelligence, 2021.
- [56] S. Wang e M. Casado, «The Cost of Cloud, a Trillion Dollar Paradox,» 27 marzo 2021. [Online]. Available: <https://a16z.com/the-cost-of-cloud-a-trillion-dollar-paradox/>. [Consultato il giorno 30 ottobre 2023].
- [57] D. Parmar, «Step-By-Step Guide to Building a Private Cloud. A private cloud isn't just an engineering project.,» 19 Aprile 2022. [Online]. Available: <https://www.nutanix.com/theforecastbynutanix/technology/step-by-step-guide-to-building-a-private-cloud>. [Consultato il giorno 21 novembre 2023].
- [58] E. Montagnani, «Le pubbliche amministrazioni nell'era delle tecnologie cloud ed edge computing tra opportunità e rischi: il Piano Nazionale di Ripresa e Resilienza e le comunità digitali,» *Rivista italiana di informatica e diritto*, vol. 4, n. 1, pp. 189-200, aprile 2022.
- [59] Agenzia per l'Italia Digitale, «Censimento patrimonio ICT della PA. Sintesi dei dati rilevati,» Roma, 2020.

- [60] G. Attardi, «Agenda Digitale. Cloud per la ricerca: i vantaggi di un approccio federato,» 19 luglio 2017. [Online]. Available: <https://www.agendadigitale.eu/infrastrutture/cloud-per-la-ricerca-i-vantaggi-di-un-approccio-federato/>. [Consultato il giorno 30 ottobre 2023].
- [61] GÉANT SIG-CISS (Cloudy Interoperable Software Stacks), «Community Clouds for European NRENs,» GÉANT SIG-CISS, novembre 2021.
- [62] G. Attardi, A. Barchiesi, A. Colla, F. Galeazzi, G. Marzulli e M. Reale, «Declarative Modeling for Building a Cloud Federation and Cloud Applications,» 2017.
- [63] V. Ardizzone, G. Attardi, A. Barchiesi, A. Colla, M. Di Fazio, R. Di Lallo, F. Galeazzi, M. Lorini, D. Passalacqua, C. Pisa, M. Reale, F. Ruggieri e D. Vagheti, «GARR Cloud: a Federated and Replicable Model for the Provisioning of Community Cloud Services,» Rome, Italy, 2022.
- [64] A. Colla, «La Cloud GARR. Modulo 1,» in *Webinar corso di aggiornamento APM GARR*, Roma, 2 novembre 2022.
- [65] GARR - Dipartimento Calcolo e Storage Distribuito, «Reference Architecture for the GARR,» GARR, Roma, 2016.
- [66] F. Galeazzi e D. Passalacqua, «La Cloud GARR. Modulo 2,» in *Webinar corso di aggiornamento APM GARR*, Roma, 9 novembre 2022.
- [67] Canonical Ltd., «MAAS Documentation. Overview of MAAS,» 10 ottobre 2023. [Online]. Available: <https://maas.io/docs/overview-of-maas>. [Consultato il giorno 2 novembre 2023].
- [68] Canonical Ltd., «Juju documentation. Juju Reference,» Canonical Ltd., 1 ottobre 2023. [Online]. Available: <https://juju.is/docs/juju/reference>. [Consultato il giorno 14 novembre 2023].
- [69] Red Hat, Inc., «Ceph Storage 6. Storage Strategies Guide,» 12 ottobre 2023. [Online]. Available: [https://access.redhat.com/documentation/en-us/red\\_hat\\_ceph\\_storage/6/html/storage\\_strategies\\_guide](https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/6/html/storage_strategies_guide). [Consultato il giorno 2 novembre 2023].
- [70] A. Manacero, R. Spolon, N. G. Bachiega e H. Martins, «Open Source Cloud Computing: Characteristics and an Overview,» in *International Conference on Parallel and Distributed Processing Techniques and Applications*, Las Vegas, USA, 2013.
- [71] G. Raj, C. Kapoor e D. Singh Dr., «Comparative cloud deployment and service orchestration process using juju charms,» *International Journal of Engineering and Technology*, vol. 5, n. 2, pp. 1412-1418, 2013.

- [72] R. Kumar, N. Gupta, S. Charu, K. Jain e S. K. Jangir, «Open Source Solution for Cloud Computing Platform Using OpenStack,» *International Journal of Computer Science and Mobile Computing*, vol. 3, n. 5, pp. 89-98, marzo 2014.
- [73] S. Shahzadi, M. Iqbal, Z. U. Qayyum e T. Dagiuklas, «Infrastructure as a service (IaaS): A comparative performance analysis of open-source cloud platforms,» *IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, CAMAD*, Vol. 1 di 22017-June, 6 2017.
- [74] T. Bell, B. Bompastor, S. Bukowiec, J. Castro Leon, M. K. Denis, J. Van Eldik, M. Fermin Lobo, L. Fernandez Alvarez, D. Fernandez Rodriguez, A. Marino, B. Moreira, B. Noel, T. Oulevey, W. Takase, A. Wiebalck e S. Zilli, «Scaling the CERN OpenStack cloud,» *Journal of Physics: Conference Series*, vol. 664, n. 2, 23 dicembre 2015.
- [75] H. T. Ciptaningtyas, R. R. Hariadi, M. Husni, K. Ghazali, R. W. Sholikah e I. M. D. Setyadharma, «OpenStack Implementation using Multinode Deployment Method for Private Cloud Computing Infrastructure,» in *International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Surabaya, Indonesia, 26-27 luglio 2023.
- [76] H. Coullon, C. Perez e D. Pertin, «Production deployment tools for IaaS: An overall model and survey,» in *FiCloud 2017: IEEE 5th International Conference on Future Internet of Things and Cloud*, Prague, Czech Republic, 2017.
- [77] S. Lima, Á. Rocha e L. Roque, «An overview of OpenStack architecture: a message queuing services node,» *Cluster Computing*, vol. 22, pp. 7087-7098, 5 2019.
- [78] Y. Zhang, M. Zhou, A. Mockus e Z. Jin, «Companies' participation in OSS development. An empirical study of OpenStack,» *IEEE Transactions on Software Engineering*, vol. 47, n. 10, pp. 2242-2259, ottobre 2021.
- [79] B. Moreira, «CERN TechBlog. 10 Years of Openstack at CERN - Part 1 and 2,» 11 giugno 2022. [Online]. Available: <https://techblog.web.cern.ch/techblog/post/10-years-of-openstack-at-cern-part-2/>. [Consultato il giorno 25 novembre 2023].
- [80] A. Di Filippo, A. Barchiesi, A. Colla e E. Della Bina, «Cloud multiregion INGV: la Sala Operativa Virtuale,» *Miscellanea INGV. Progetto "Sale Operative Integrate e Rete di monitoraggio del futuro"*, vol. 57, pp. 25-29, dicembre 2020.

- [81] O. Khedher e C. D. Chowdhury, *Mastering OpenStack - Second Edition: Discover Your Complete Guide to Designing, Deploying, and Managing OpenStack-Based Clouds in Mid-to-large IT Infrastructures*, Birmingham, UK: Packt, aprile 2017, p. 470.
- [82] T. Fifield, D. Fleming, A. Gentle e L. Hochstein, *OpenStack Operations Guide*, O'Reilly Media, Inc., 2014.
- [83] K. Tenzer, «OpenStack Multiple Node Configurations,» 26 gennaio 2015. [Online]. Available: <https://keithtenzer.com/openstack/openstack-multiple-node-configurations/>. [Consultato il giorno 21 novembre 2023].
- [84] OpenStack Foundation, «OpenStack reference architecture for 1000 nodes,» OpenStack Foundation, 2016.
- [85] M. Carboni, «Evoluzione del modello cloud GARR,» in *Net Makers 2023*, Roma, novembre 2023.
- [86] Consip S.p.A., «Tecnologie Server 4, lotto 6. Guida alla convenzione,» 15 dicembre 2022.
- [87] Cisco Systems, Inc., «Cisco ACI Multi-tier Architecture White Paper,» Cisco Systems, Inc., San Jose, California, 2023.
- [88] Consip S.p.A., «Reti Locali 8. Guida alla convenzione,» 15 novembre 2023.
- [89] Consip S.p.A., «Energia Elettrica 20 Lotto1. Guida alla convenzione,» 19 dicembre 2022.
- [90] M. Bramucci , D. Di Santo e D. Forni, «Uso Razionale dell'energia nei centri di calcolo. Studio su incarico dell'ENEA,» FIRE. Federazione Italiana per l'uso Razionale dell'Energia, settembre 2010.
- [91] Microsoft Learn, «Documentazione per la creazione e la gestione di macchine virtuali in Azure,» 9 agosto 2023. [Online]. Available: <https://learn.microsoft.com/it-it/azure/virtual-machines/>. [Consultato il giorno 26 novembre 2023].
- [92] Dipartimento per la Trasformazione Digitale, «Strategia Cloud Italia,» 2022. [Online]. Available: <https://docs.italia.it/italia/cloud-italia/>. [Consultato il giorno 30 ottobre 2023].
- [93] Agenzia per l'Italia Digitale, «Piano Triennale per l'informatica nella PA. Aggiornamento 2022-2024,» ottobre 2022. [Online]. Available: <https://www.agid.gov.it/it/agenzia/piano-triennale>. [Consultato il giorno 30 ottobre 2023].
- [94] Agenzia per l'Italia Digitale, «La spesa ICT nella PA italiana 2021 - Principali trend e percorsi in atto,» Roma, 2022.

- [110] D. H. Shin, «User centric cloud service model in public sectors: Policy implications of cloud services,» 2013.
- [111] C. Wu, R. Buyya e K. Ramamohanarao, «Value-based cloud price modeling for segmented business to business market,» 2019.
- [112] R. F. El-Gazzar, «An overview of cloud computing adoption challenges in the norwegian context,» 2014.
- [113] H. Aydin, «A Study of Cloud Computing Adoption in Universities as a Guideline to Cloud Migration,» 2021.
- [114] A. C. M. de Paula e G. d. F. de Carneiro, «A systematic literature review on cloud computing adoption and migration,» 2016.
- [115] M. T. Amron, R. Ibrahim, N. A. A. Bakar e S. Chuprat, «Acceptance of cloud computing in the Malaysian public sector: A proposed model,» 2019.
- [116] O. Mohammad, M. Alwan e O. Abduljabbar, «Campus Cloud Computing for Universities: State-of-the-Art,» 2021.
- [117] M. Janssen e A. Joha, «Challenges for adopting cloud-based software as a service (SAAS) in the public sector,» 2011.
- [118] E. Cidres, A. Vasconcelos e F. Leiteao, «Cloud calculator: A cloud assessment tool for the public administration,» 2020.
- [119] C. R. Alimboyong e M. E. Bucjan, «Cloud computing adoption among state universities and colleges in the Philippines: Issues and challenges,» 2021.
- [120] B. Nicoletti, «Cloud computing and procurement,» 2016.
- [121] A. Kannaki e J. M. Gnanasekar, «Cloud computing overview, security threats and solutions-A survey,» 2016.
- [122] R. P. Z. M. Thomas Erl, Cloud computing: concepts, technology & architecture, Prentice-Hall, 2013, pp. 51-2714-51-2714.
- [123] O. Bushehrian e S. Y. Nabavi, «A new method for migrating legacy applications to the cloud: a finite state process approach,» in *CSSE 2017: International Symposium on Computer Science and Software Engineering Conference*, Shiraz, Iran, 2017.
- [127] R. Nasim e A. J. Kassler, «Deploying OpenStack: Virtual infrastructure or dedicated hardware,» in *COMPSACW 2014: IEEE 38th Annual International Computers, Software and Applications Conference Workshops*, Vasteras, Sweden, 2014.
- [129] A. Dahiru e H. Abubakar, «Cloud Computing adoption: a cross-continent overview of challenges,» *Nigerian Journal of Basic and Applied Sciences*, vol. 25, n. 1, pp. 23-31, giugno 2017.

- [130] J. Fishenden e M. Thompson, «Digital government, open architecture, and innovation: Why public sector it will never be the same again,» *Journal of Public Administration Research and Theory*, vol. 23, n. 4, pp. 977-1004, 6 settembre 2012.
- [131] H. Sallehudin, A. H. M. Aman, R. C. Razak, M. Ismail, N. A. A. Bakar, A. F. M. Fadzil e R. Baker, «Performance and key factors of cloud computing implementation in the public sector,» *International Journal of Business and Society*, vol. 21, n. 1, pp. 134-152, 25 aprile 2021.
- [132] A. Markelov, *Certified OpenStack Administrator Study Guide*, Stockholm, Sweden: Apress, 2016.
- [133] T. Haeberlen, D. Liveri e M. Lakka, «Good Practice Guide for securely deploying Governmental Clouds,» European Union Agency for Network and Information Security (ENISA), Heraklion, Greece, 2013.
- [134] D. W. Chadwick, K. Siu, C. Lee, Y. Fouillat e D. Germonville, «Adding Federated Identity Management to OpenStack,» *Journal of Grid Computing*, vol. 12, n. 1, pp. 3-27, 2014.
- [135] «Storage Strategies Guide,» Red Hat Ceph Storage, 18 ottobre 2023. [Online]. [Consultato il giorno 2 novembre 2023].