

Cloud Computing e Open Data

Tecnologie per la PA

Domenico Talia

Università della Calabria

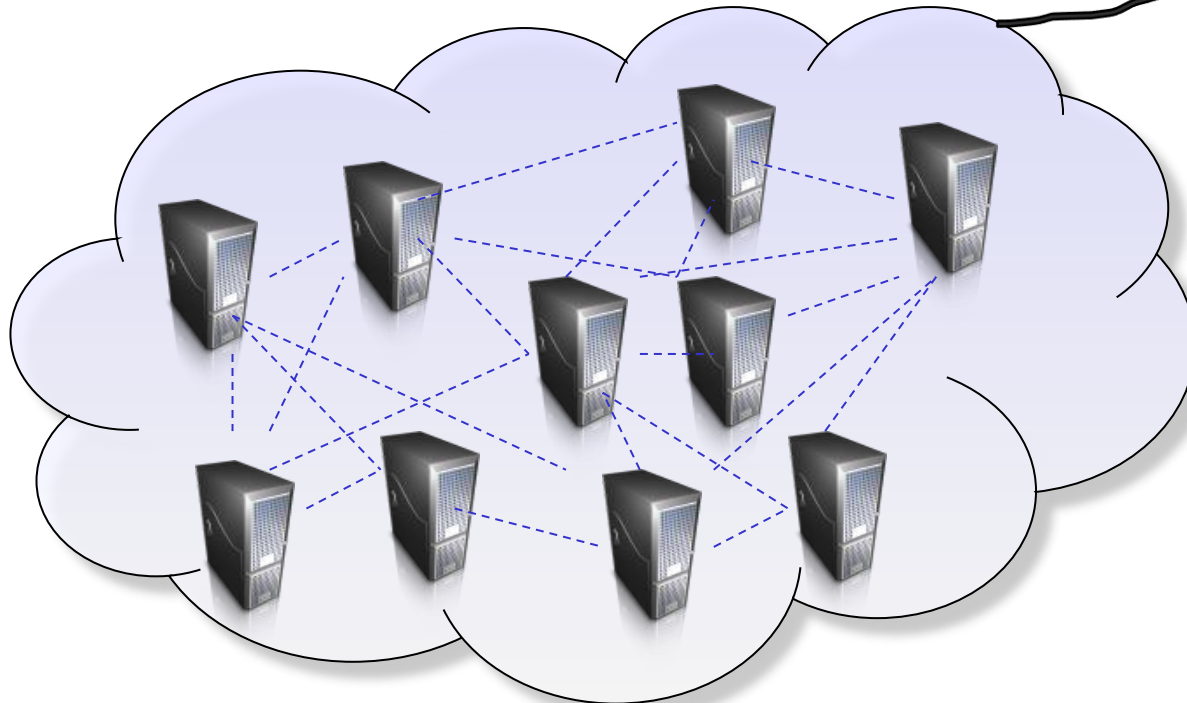
talia@dimes.unical.it



Introduzione al Cloud Computing



- **Cloud computing:** un **modello** di elaborazione distribuita basato su una infrastruttura che offre collezioni di risorse hardware e software accessibili attraverso Internet.



Introduzione al Cloud Computing



- Il termine Cloud Computing nasce dalla nuvoletta usata per rappresentare Internet e da “telecom cloud”.
- I sistemi Cloud **estendono le funzionalità di Internet** offrendo **tutte le operazioni** che si possono eseguire su un computer.
- Gli utenti di un sistema Cloud non hanno necessità di avere diretta conoscenza o controllo sulla infrastruttura tecnologica (hardware/software) che supporta i loro usi.

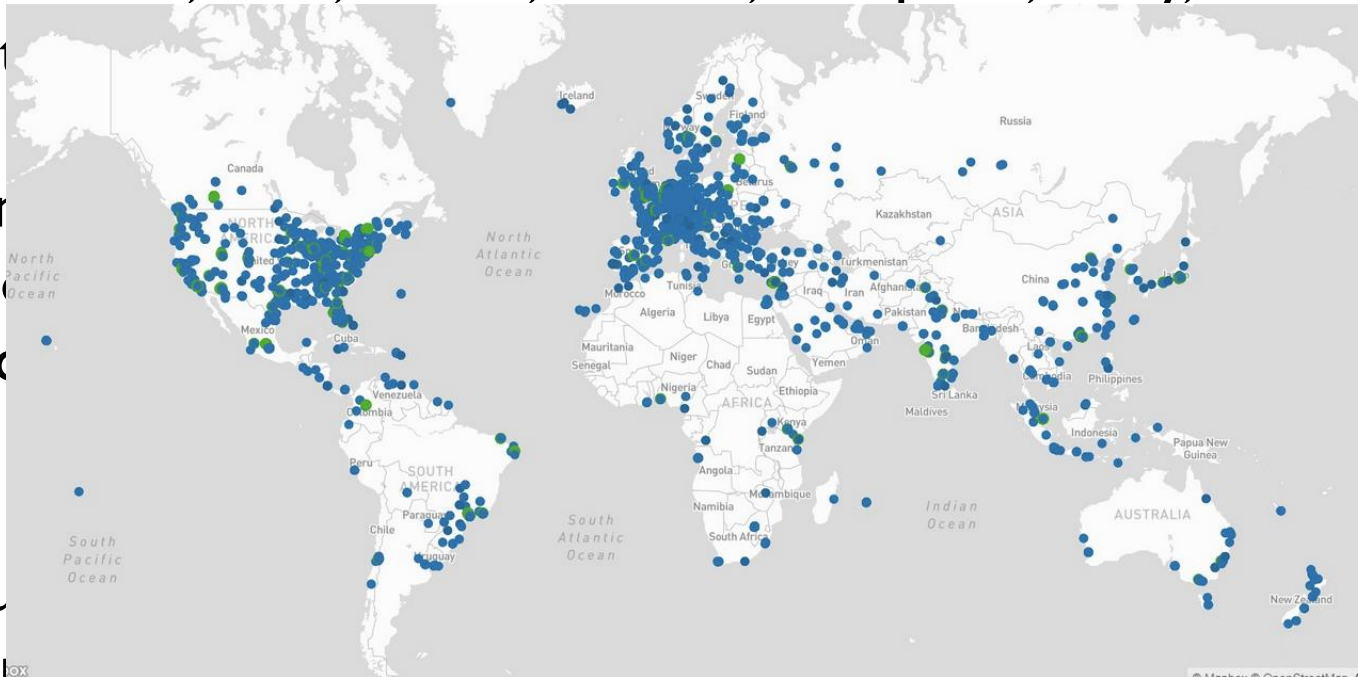
Introduzione al Cloud Computing



- Quando dal PC o dallo smartphone usiamo i servizi di Google, Facebook, TIM, Gmail, Netflix, Dropbox, eBay, e molti altri sistemi

- Il nostro computer o smartphone
- Facebook

- Le nostre attività quotidiane
- Diamo un'occhiata alle nostre foto
- Usiamo altri programmi del mondo senza che noi sappiamo dove sono e come sono fatti.



rete ad Google,

ornia, a
e tante
siano e

Introduzione al Cloud Computing



- Il Cloud computing è un modello di elaborazione in cui le risorse (hardware/software)
 - Sono scalabili dinamicamente e
 - Sono virtualizzate
 - Sono fornite con un'interfaccia a servizi (*as a service*) su Internet.
- Le risorse che un sistema Cloud mette a disposizione degli utenti e delle applicazioni sono CPU, memorie, reti, sistemi operativi, middleware, applicazioni.

Cloud Computing: Definizione del NIST

- Il NIST (National Institute of Standards and Technology) da la seguente definizione di Cloud Computing:
*“Cloud computing is a model for enabling convenient, **on-demand network access to a shared pool of configurable computing resources** (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”*
- Le sue caratteristiche principali sono:
On-demand self-service, Broad network access, Resource pooling, Rapid elasticity, Measured Service.

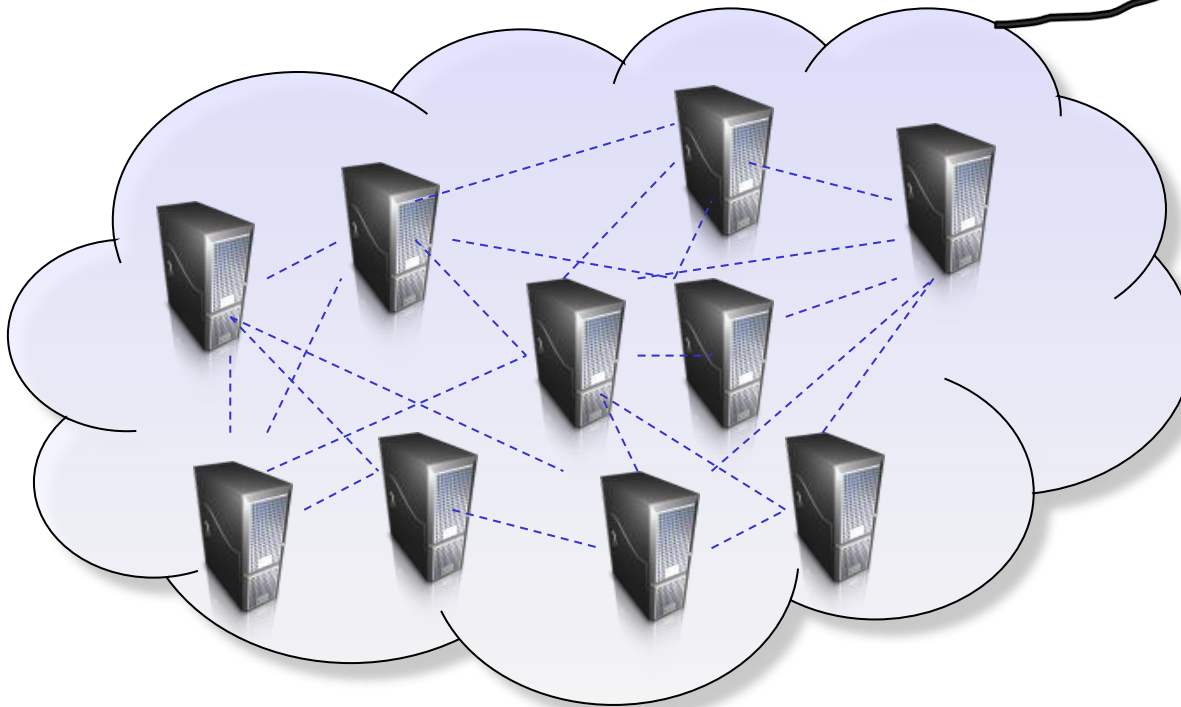
Cloud Computing: modello di costo

I sistemi Cloud permettono di:

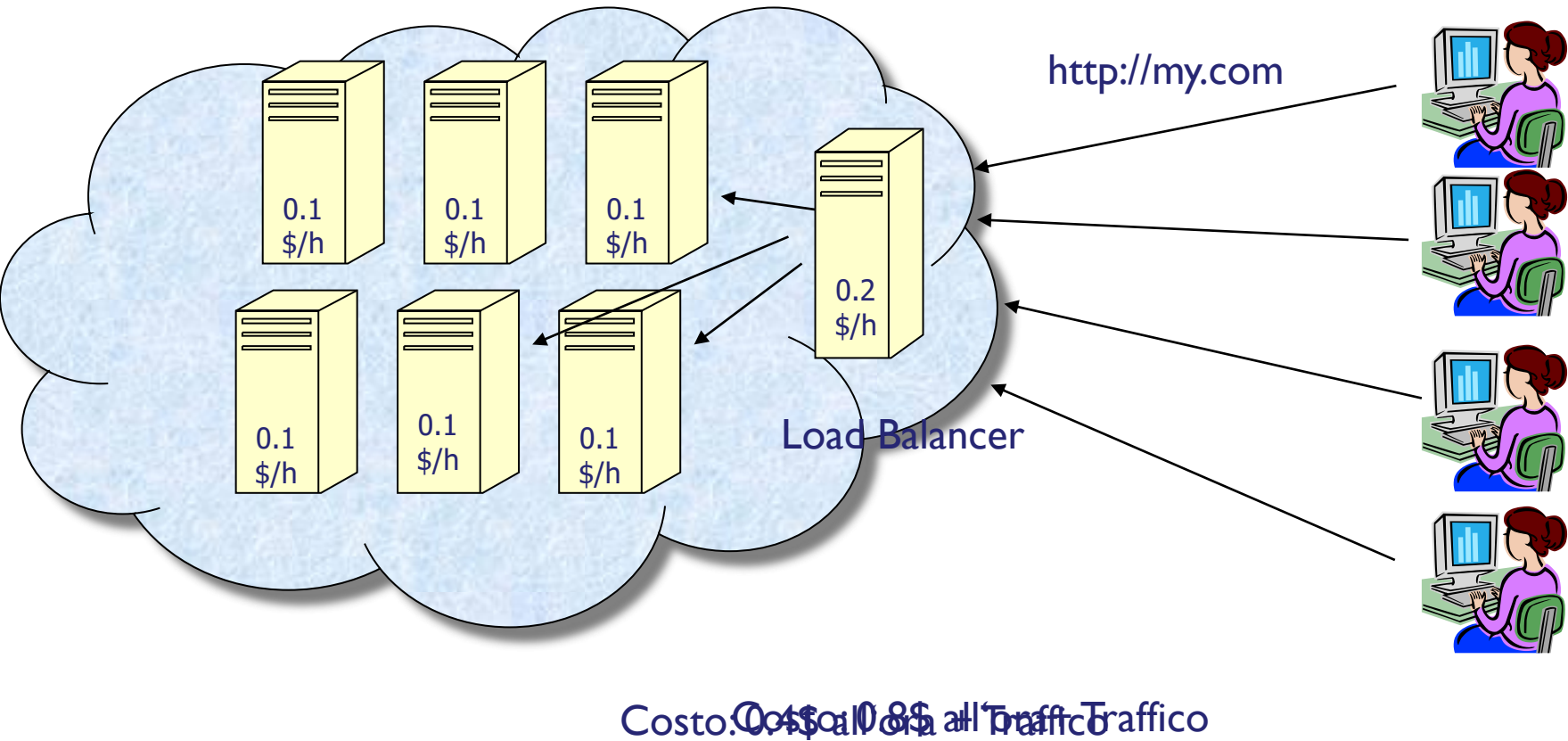
- **Affittare** i servizi di elaborazione da un fornitore.
- Aumentare e diminuire **dinamicamente** le unità di elaborazione/memoria.
- **Pagare** solamente per **le risorse** che sono effettivamente **utilizzate**
 - ore CPU, banda di rete, quantità di storage, software, applicazioni, ...

Cloud Computing: modello di uso

- Un utente usa la risorse delle macchine di un sistema Cloud senza conoscerne le caratteristiche di dettaglio e senza occuparsi della loro gestione.

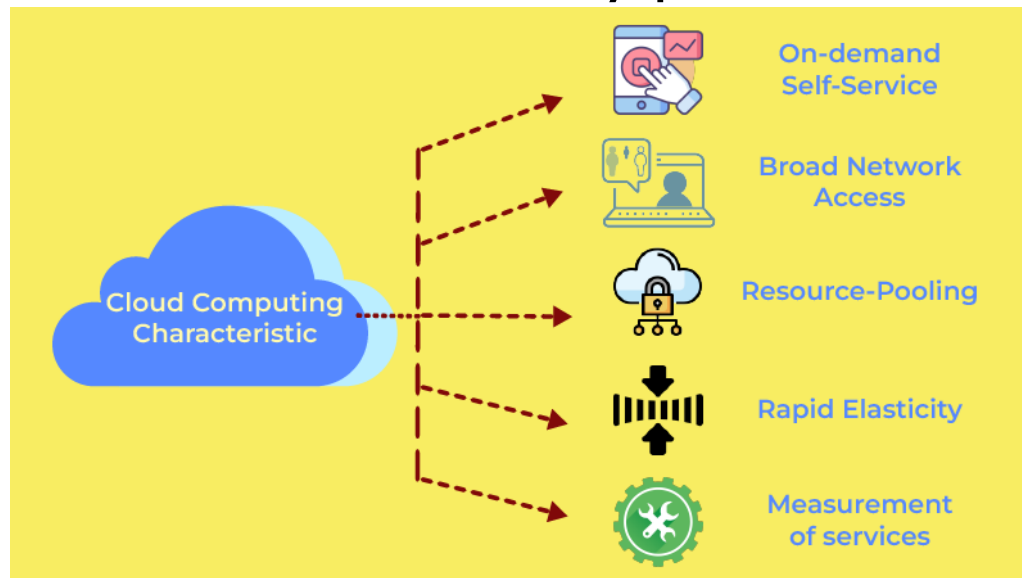


Cloud Computing: scalabilità



Caratteristiche dei Sistemi Cloud

- Interfaccia a servizi
- On-demand
- Accesso Ubiquo tramite Internet
- Aggregazione di risorse
- location independent
- Elasticità
- Pay per use



Classi di Sistemi Cloud

- **Private cloud**

- Di proprietà (o in affitto) di un'azienda

- **Community cloud**

- Infrastruttura condivisa da una comunità di utenti

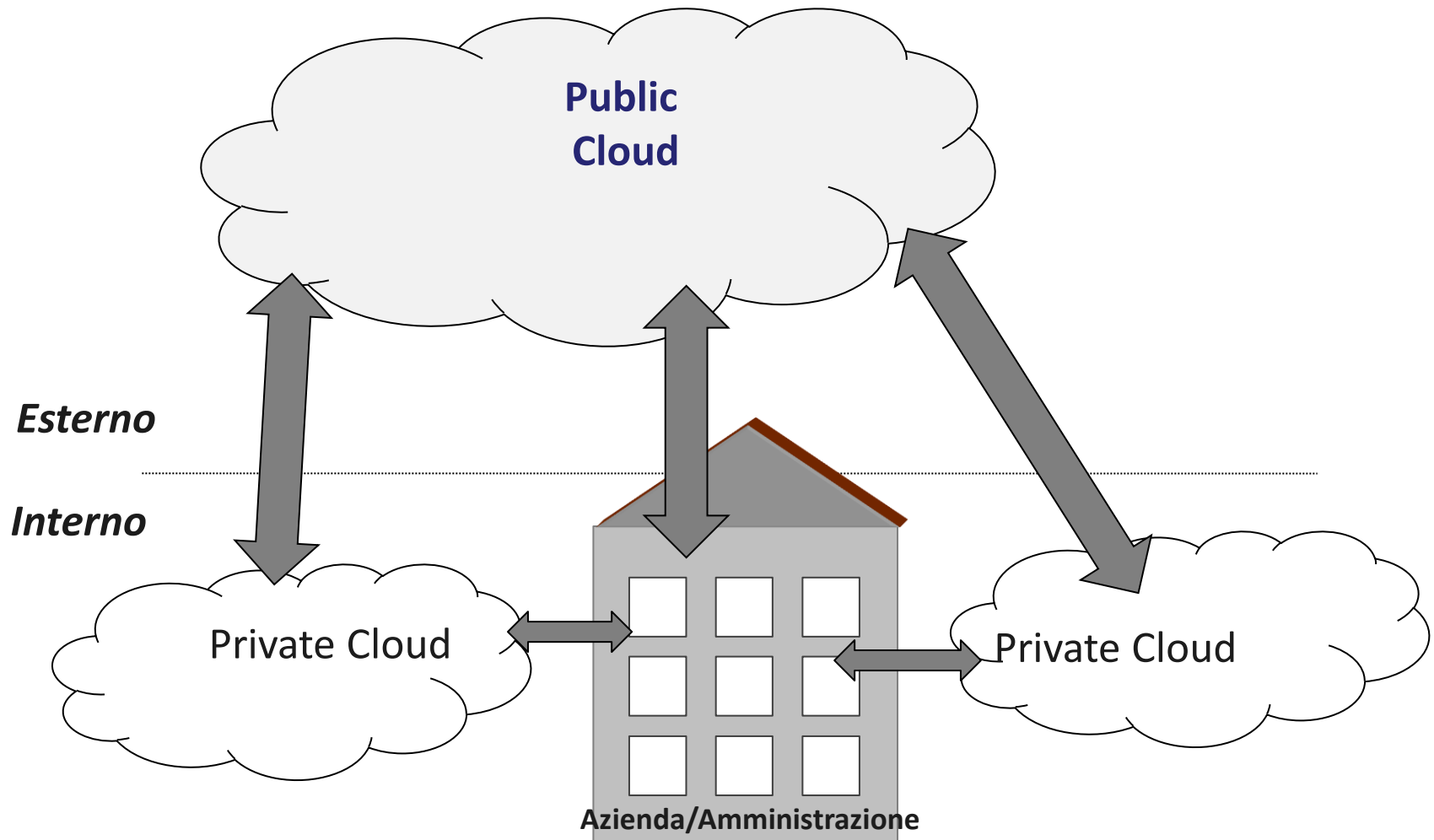
- **Public cloud**

- In vendita al pubblico, infrastruttura di grande dimensione

- **Hybrid cloud**

- Composizione di due o più cloud

Classi di Sistemi Cloud



Modelli a Servizi

- **Software as a Service (SaaS)**

- Software è fornito agli utenti “On-demand”
- Non serve acquistare le licenze
- “Utility-based” computing

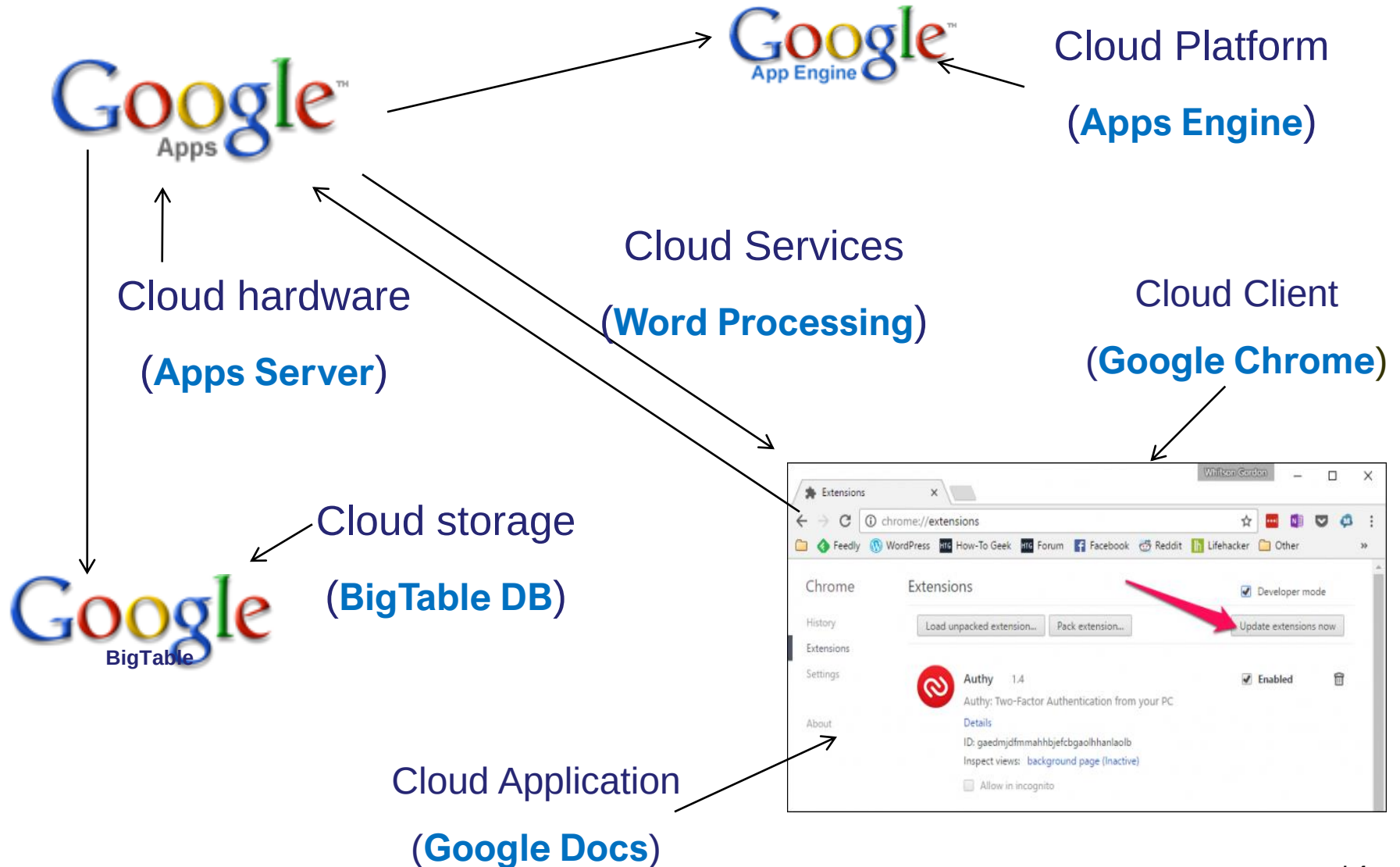
- **Platform as a Service (PaaS)**

- L’ ambiente per lo sviluppo di Cloud applications è fornito secondo il modello “Software as a Service”

- **Infrastructure as a Service (IaaS)**

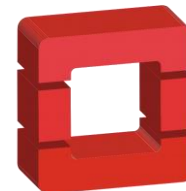
- Una infrastruttura di elaborazione (hardware) viene fornita come un servizio

Un Esempio: Alcuni servizi cloud di Google

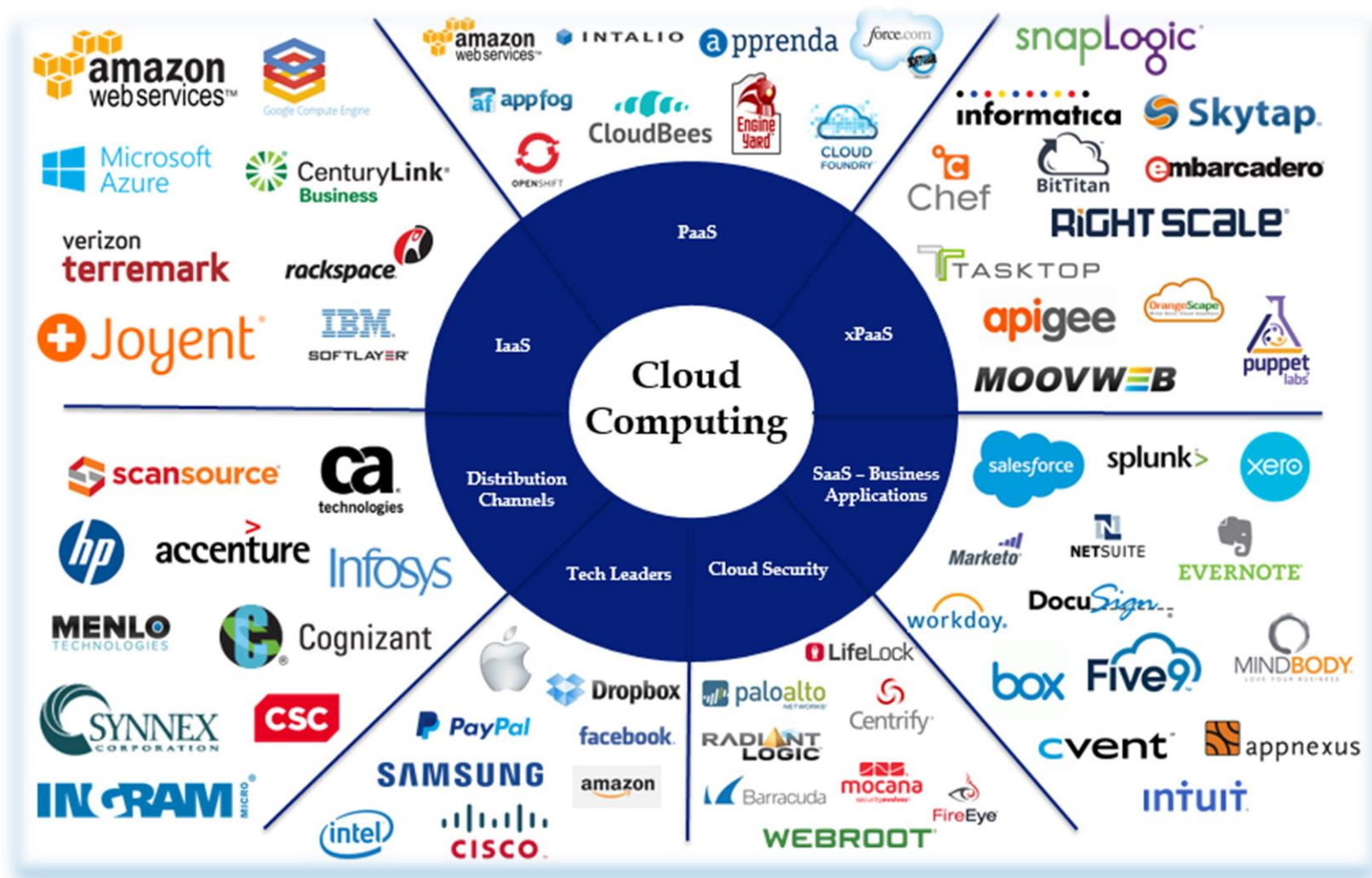


Sistemi Cloud Commerciali e OS

- Amazon Elastic Compute Cloud (EC2)
- Amazon Simple Storage Service (S3)
- Google App Engine
- Microsoft Windows Azure
- Eucalyptus
- OpenStack



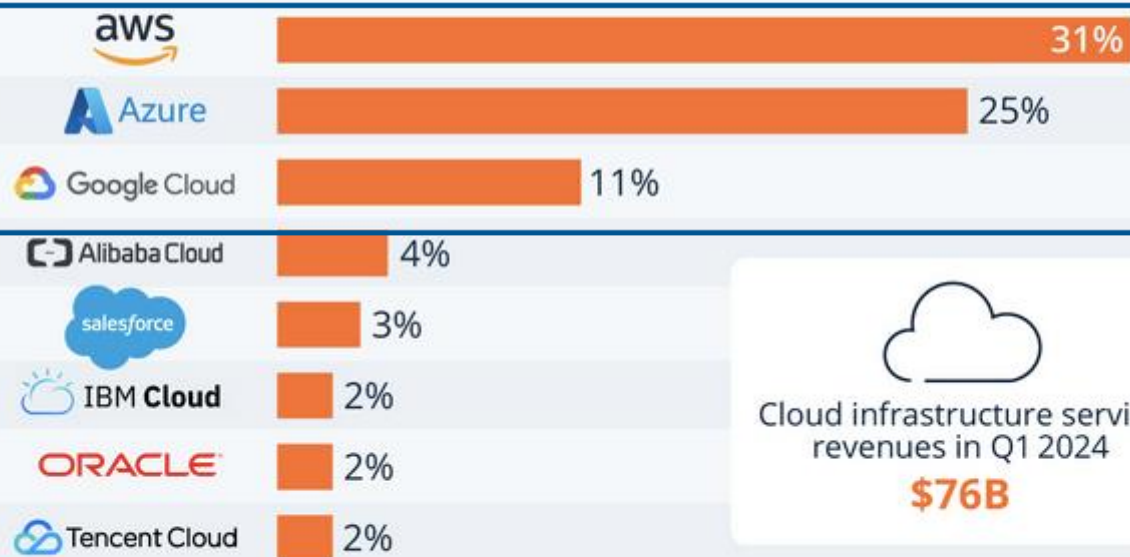
Alcuni Providers di Sistemi Cloud



Top Providers di Sistemi Cloud

Amazon Maintains Cloud Lead as Microsoft Edges Closer

Worldwide market share of leading cloud infrastructure service providers in Q1 2024*



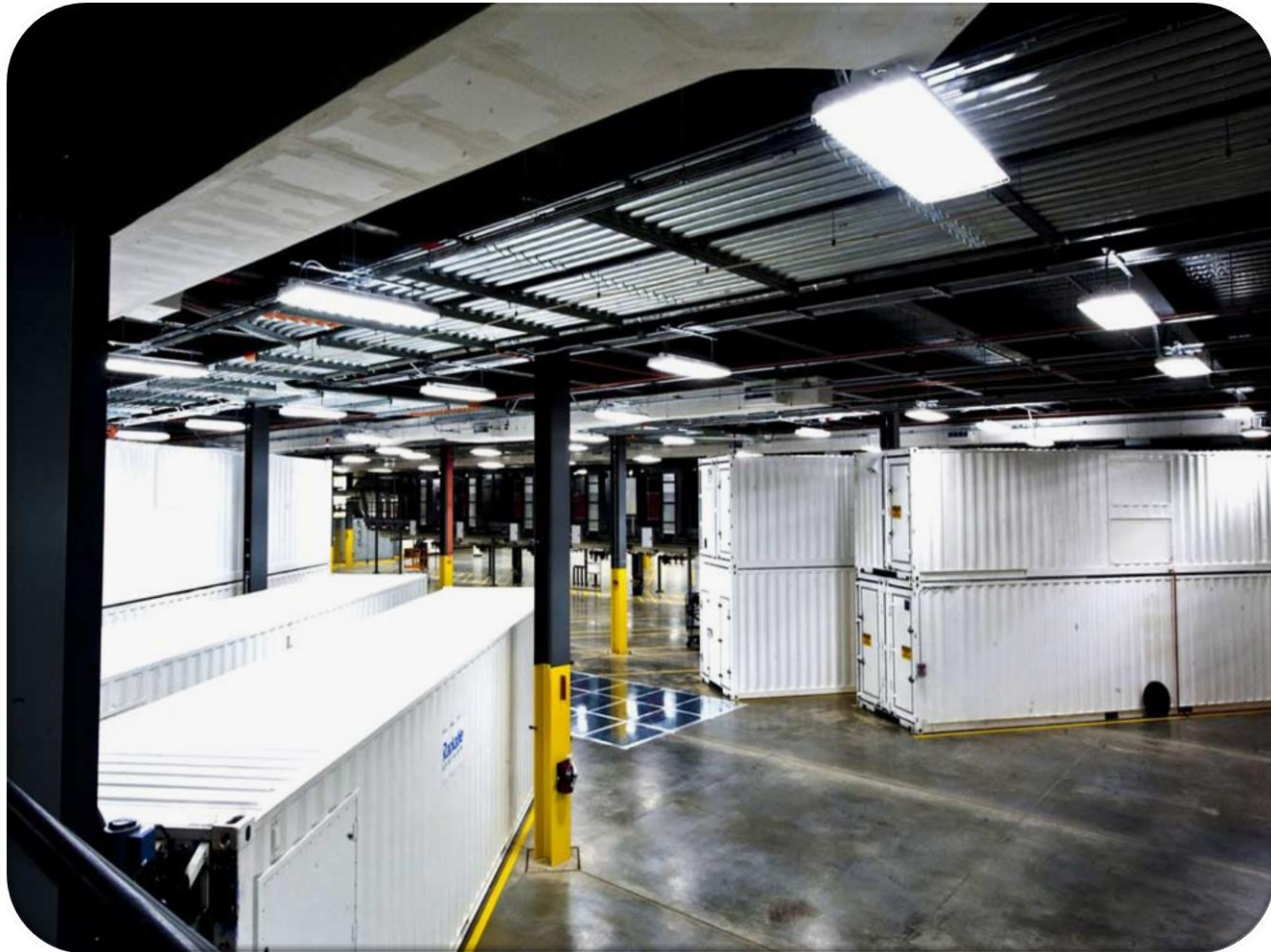
Cloud infrastructure service
revenues in Q1 2024
\$76B

* Includes platform as a service (PaaS) and infrastructure as a service (IaaS)
as well as hosted private cloud services

Source: Synergy Research Group

Windows Azure: Un esempio di cloud pubblico

- L'infrastruttura di un data center Azure



Windows Azure: Un esempio di cloud pubblico

- Microsoft Data Center di 4 generazione

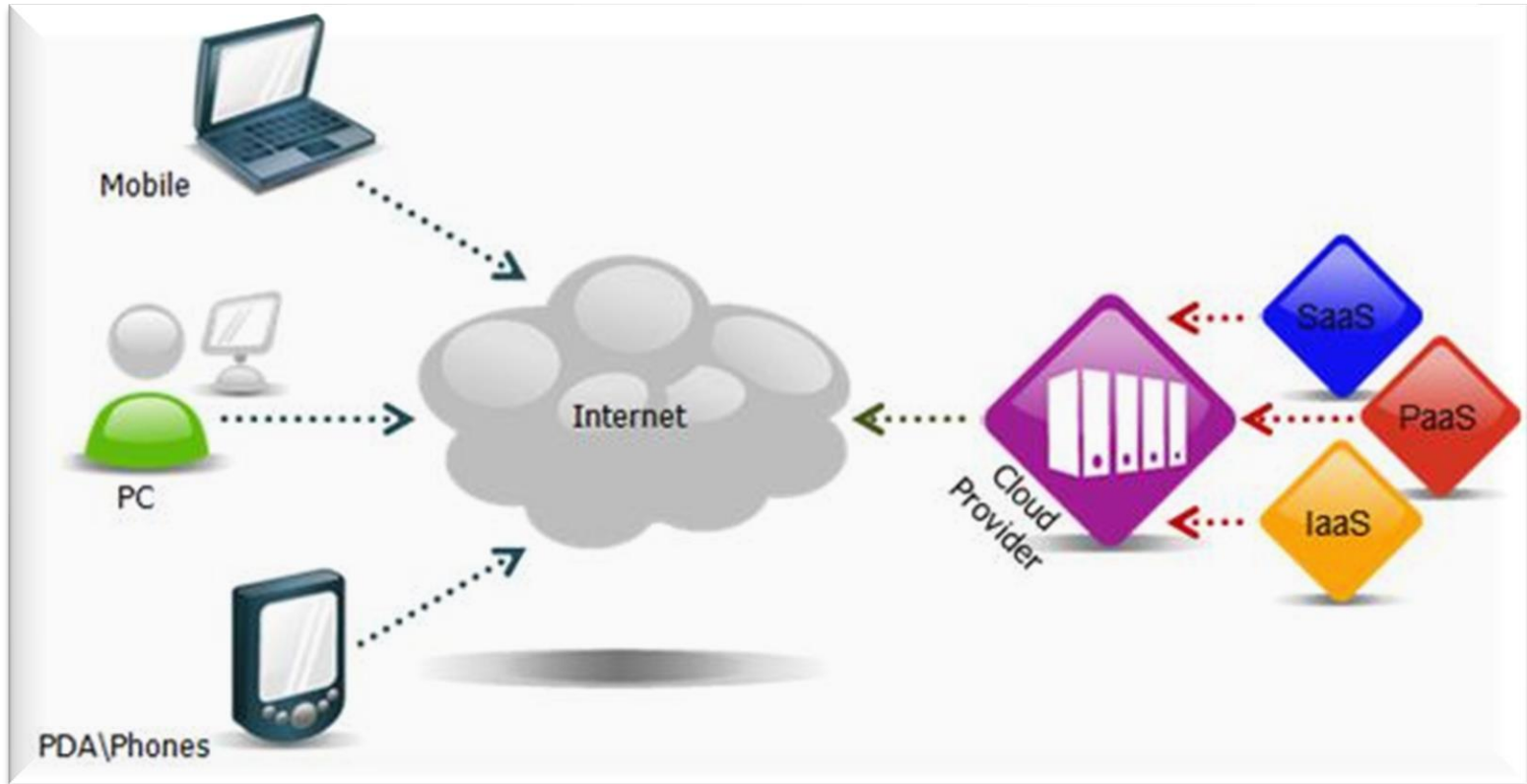


Sistemi Cloud Open Source

- Eucalyptus (www.eucalyptus.com)
- Red Hat Cloud (www.redhat.com/solutions/cloud)
- Cloudera (<http://www.cloudera.com>)
- OpenNebula (www.opennebula.org)
- Puppet (projects.puppetlabs.com/projects/puppet)
- Enomaly (www.enomaly.com)
- Zoho (www.zoho.com)
- Joyent (www.joyent.com)
- Reservoir (www.reservoir-fp7.eu)
- OpenStack (www.openstack.org)

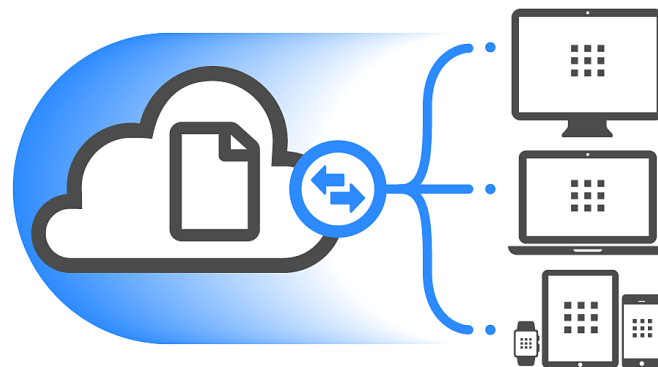
Cloud Computing fornisce Servizi Ubiqui a Client Ubiqui

Dispositivi mobili ed ubiqui possono essere “punti di accesso” (client) verso servizi ed applicazioni complesse fornite da sistemi di Cloud Computing



Archiviazione documentale in Cloud

- Consente di archiviare, velocemente ogni tipo di documento (elettronico, cartaceo, mail) nella infrastruttura Cloud.
- L'archiviazione digitale documentale in Cloud favorisce la catalogazione, la gestione, la condivisione e il controllo dei flussi documentali all'interno o verso l'esterno.
- Queste operazioni sono realizzate come servizi nelle piattaforme Cloud e sono accedibili da qualunque luogo.
- Permette la gestione multi-organizzazione attraverso la suddivisione di più archivi digitali.



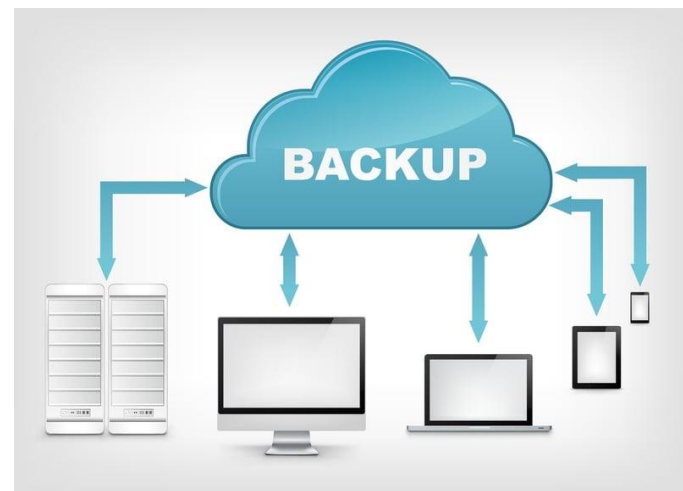
Archiviazione documentale in Cloud

- Archiviazione dei files on-line su una sorta di "hard disk virtuale" (offerto dal cloud).
- Elaborazione locale dei dati tramite cloud services.
- Accesso ai files, da qualunque luogo via app o via web (senza necessità di scaricare alcun software sul proprio computer, tablet, smartphone).



Archiviazione documentale in Cloud

- DB e files sono conservati e sottoposti a backup.
- I documenti sono condivisibili con altri soggetti e in modo controllato dall'amministratore.
- Possibilità di generare più archivi digitali da associare ai diversi utenti.
- Necessità di protezione dei documenti ai fini della riservatezza (cifratura).



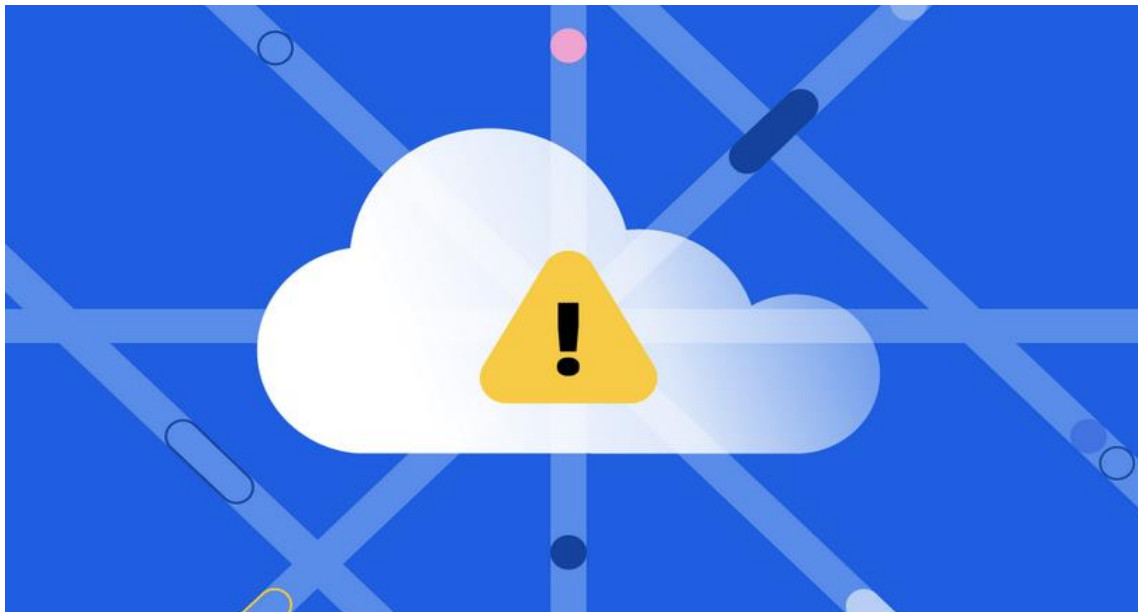
Criticità del Cloud computing

- Poiché i sistemi cloud non permettono agli utenti di possedere i dispositivi di storage dei propri dati, lasciano la responsabilità e il controllo dei dati nella mani dei provider.
- Gli utenti dipendono dai provider dei sistemi di Cloud computing per l'uso che ne possono fare.
- Gli utenti non hanno la libertà di installare su un sistema Cloud nuove applicazioni senza l'approvazione degli amministratori del sistema.



Criticità del Cloud computing

- Privacy e sicurezza dei dati e delle applicazioni degli utenti non sono a rischio zero.
- Un guasto di un sistema Cloud di grandi dimensioni può coinvolgere migliaia di utenti, molte applicazioni ed enormi quantità di dati.



Cloud computing elements

- Numerosi elementi hardware/software sono necessari per rendere funzionante un sistema di Cloud computing:
 - Cloud applications
 - Cloud clients
 - Cloud hardware
 - Cloud platform
 - Cloud services
 - Cloud storage.
- Tutti questi possono essere offerti “as a Service” in un Cloud.

Modelli di Cloud services

SaaS

Software as a Service

PaaS

Platform as a Service

IaaS

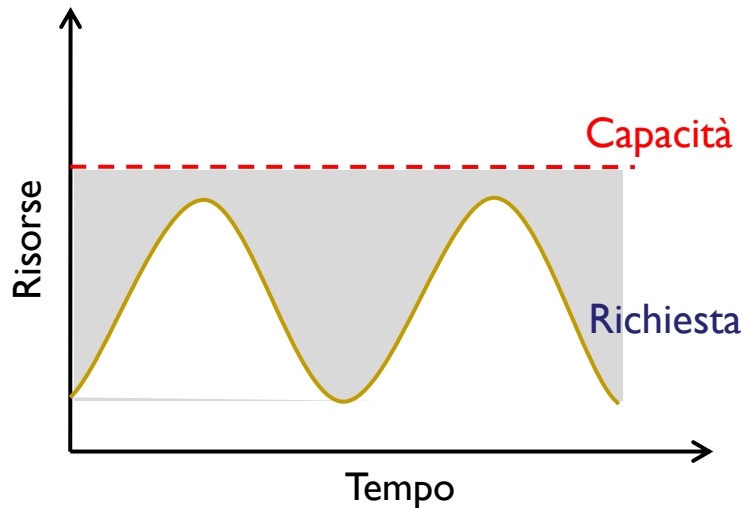
Infrastructure as a Service

Modelli di Cloud Services

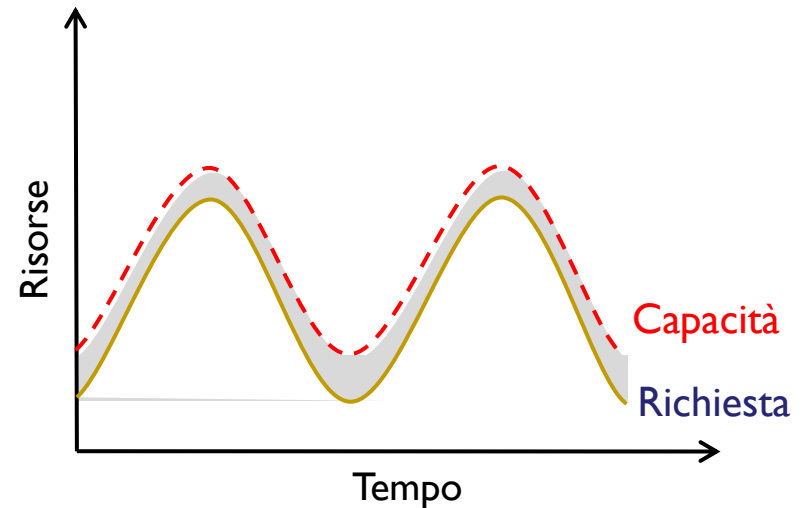
- I modelli di uso service oriented per un sistema Cloud sono i seguenti:
 - Software as a Service (SaaS)
 - Software è fornito agli utenti “On-demand”
 - Non serve acquistare le licenze
 - Platform as a Service (PaaS)
 - L’ ambiente per lo sviluppo di Cloud applications è fornito “as a Service”
 - Infrastructure as a Service (IaaS)
 - Un sistema di elaborazione completo (hardware) viene fornito come un servizio

Modelli di uso di sistemi Cloud

- Modello di account “Pay per use” invece di “provisioning for peak”. Si paga per i servizi richiesti e utilizzati



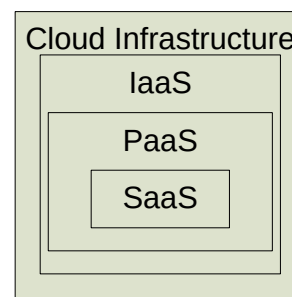
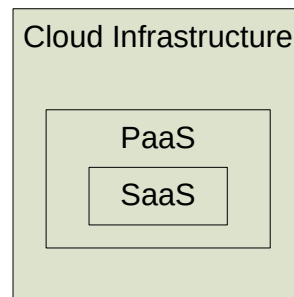
Static data center



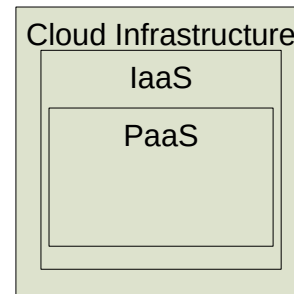
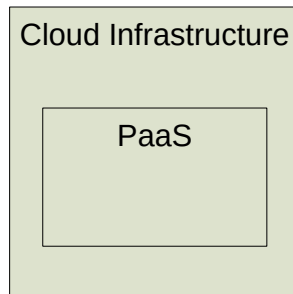
Data center in the cloud

 Risorse inutilizzate

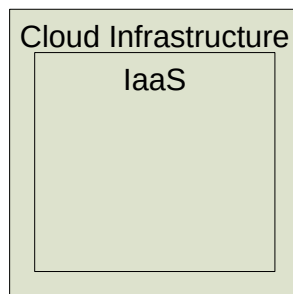
Architettura Cloud a Servizi: diverse possibilità



Software as a Service
(SaaS)
Architectures



Platform as a Service (PaaS)
Architectures



Infrastructure as a Service (IaaS)
Architectures

Software as a Service (SaaS)

■ Problema:

- Software proprietario che viene installato su un computer non può essere installato su altri computer senza pagare altre licenze.
- E' possibile l'accesso remoto ma con problemi di uptime, security...
- Invece il software viene fornito come *cloud application* che può essere acceduto da qualsiasi *cloud client* (es: Web browser)

■ Soluzione: Software as a Service

- Il software commerciale può essere configurato ed usato su una macchina e usato da molti client connessi al server tramite la rete.

Platform as a Service (PaaS)

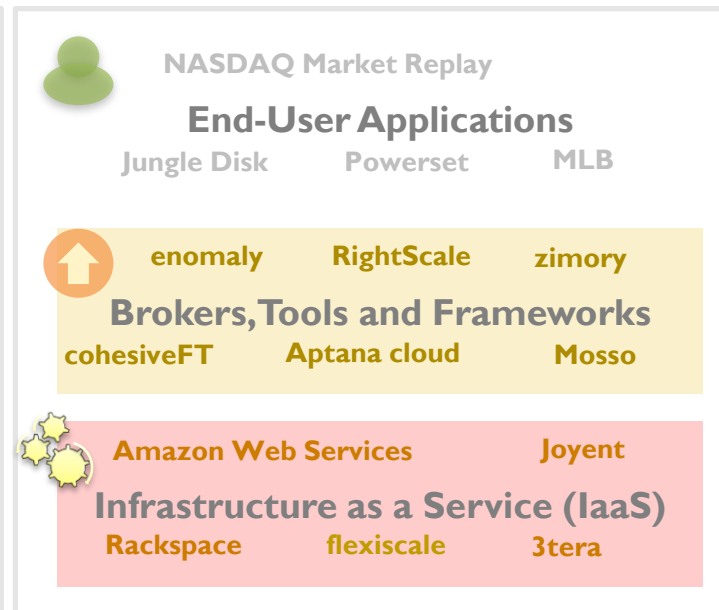
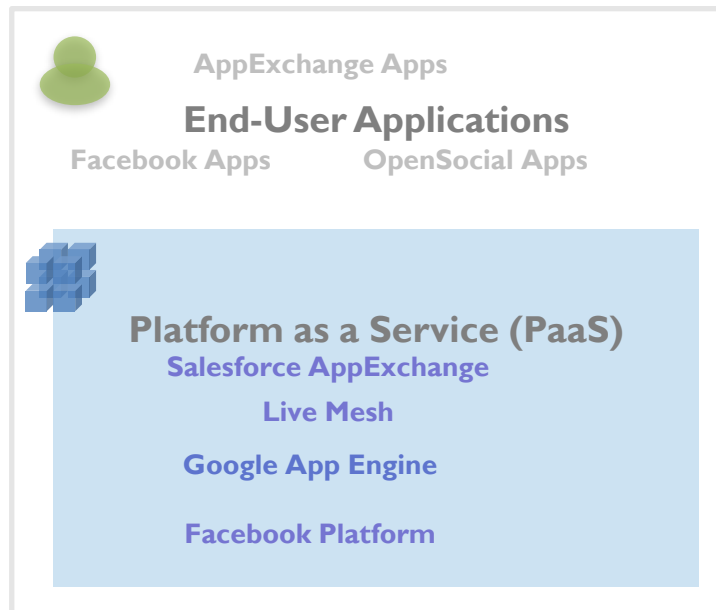
- Implementa il “Software Lifecycle” su un Cloud.
- Come funziona?
 - I server Cloud permettono l’hosting di applicazioni durante il loro sviluppo (*applications in progress*)
 - Interfacce utente nel web browser per poterle usare
 - Permette la valutazione del software in situazioni di carico
 - Integrazione automatica con altre sorgenti dati disponibili sul Cloud
 - Molti strumenti di sviluppo disponibili

Infrastructure as a Service (IaaS)

- Components of Infrastructure as a Service:
 - Computing hardware
 - Es. Storage Space per gli utenti, Processing power di grandi prestazioni
 - Virtualizzazione
 - Es. VM Ware, VirtualBox
 - Networking
 - Security, elevata velocità di comunicazione, server di rete
 - Internet Access
 - Connette applicazioni client a server “anywhere”
 - Utility Computing
 - Pagamento ad ora, gigabyte, processor load
 - Contratti
 - Service level specifici (richiesti dagli utenti).

Ecosistemi Cloud

- End-user applications, Software-as-a-Service (SaaS)
- Platform-as-a-Service (PaaS)
- Infrastructure-as-a-Service (IaaS)



Software delivery model

SaaS

- Nessun hardware o software da gestire
- Servizi erogati attraverso un browser/app
- Uso del servizio “on demand” da parte dei clienti
- Scalabilità istantanea

Esempi

SaaS

- Il Customer Relationship Management (CRM) è sottodimensionato o non lo si vuole gestire in proprio. Si può usare un provider come Salesforce.com
- Il server di posta è lento. Si può usare un mail server su Cloud che non ha problemi di dimensionamento.

Platform delivery model

PaaS

- Le piattaforme di sviluppo sono costruite su infrastrutture che possono essere costose.
- È difficile stimare la domanda di uso.
- La gestione delle piattaforme proprietarie è costosa.

Esempi

PaaS

- Bisogna utilizzare un file grande (10Mb) e renderlo disponibile a 20.000 utenti per 2 mesi. Usa Cloud Front di Amazon.
- Vuoi fornire storage services per grandi file a molti clienti. Usa Amazon S3 o Microsoft Azure.

Infrastructure delivery model

IaaS

- Un ambiente di virtualizzazione di infrastrutture di elaborazione scalabili.
- Integrazione di risorse di calcolo, memorizzazione e comunicazione.
- Macchine virtuali diverse e flessibili.

Esempi

IaaS

- Vuoi eseguire un numero elevato di job e non hai una macchina parallela. Usa Amazon EC2, Azure o Flexiscale.
- Vuoi ospitare un portale web per un periodo di tempo limitato. Usa Flexiscale.

Elementi comuni

SaaS

- Pay per use
- Scalabilità istantanea
- Adattabilità
- Affidabilità
- API

PaaS

IaaS

Attenzione a:

- Sicurezza e Privacy

Vantaggi

SaaS

- Politiche di aggiornamento automatico

PaaS

- Virtualizzazione per abbassare i costi di uso

IaaS

- Economia di scala

Vantaggi

SaaS

- Costi ridotti di ownership
- Responsabilità di gestione dell'infrastruttura molto ridotta o nulla

PaaS

- Gestione di sovraccarichi improvvisi di CPU e/o storage

IaaS

- Più veloce configurazione/istallazione di applicazioni

Criteri di Scelta

1. Quali servizi?
2. Quanto sicuri?
3. Dove si trova?
4. I costi?
5. Quanto downtime?
6. Come accedo?
7. Import/export dati?
8. Scalabile?
9. Supporto al cliente?
10. Quale accoglienza?



Sicurezza

Principali ambiti sensibili:

1. **Emissione dei dati (processo interno)**
2. **Trasmissione dei dati (connessione)**
3. **Trattamento dei dati (processo sia interno che esterno)**
4. **Conservazione dei dati (processo esterno)**
5. **Salvataggio dei dati (processo esterno)**



In prima istanza notiamo subito che CC non significa esternalizzare tutti i rischi. Alcune minacce rimangono relative a processi interni.



Cos'è il rischio ?

La probabilità che la minaccia percorra la vulnerabilità (per il valore dell'asset)

Immagini tratte da <http://it.123rf.com>

Sicurezza

Il Garante consiglia (tratto dalle Schede di documentazione "Cloud computing: indicazioni per l'utilizzo consapevole dei rischi") :

1. Ponderare **prioritariamente rischi e benefici** dei servizi offerti
2. Effettuare una verifica in ordine **all'affidabilità del fornitore – Certificazioni.**
3. Privilegiare i servizi che favoriscono la portabilità dei dati
4. Assicurarsi la disponibilità dei dati in caso di necessità
5. Selezionare i dati da inserire nella cloud
6. Non perdere di vista i dati
7. Informarsi su dove risiederanno, concretamente, i dati
8. Attenzione alle **clausole contrattuali**
9. Verificare le politiche di persistenza dei dati legate alla loro conservazione
10. Esigere e adottare opportune cautele per tutelare la confidenzialità dei dati
11. Formare adeguatamente il personale

Cloud e PA

- La **Strategia Cloud Italia** è realizzata dal Dipartimento per la trasformazione digitale e dall'Agenzia per la cybersicurezza nazionale (**ACN**).
- Contiene gli indirizzi per la **migrazione verso il cloud qualificato** dei dati/servizi della Pubblica Amministrazione.
- L'Agenzia per la cybersicurezza nazionale ha adottato il **Regolamento unico per le infrastrutture e i servizi cloud** per la PA. La nuova fase regolatoria, è iniziata il 1° agosto 2024.
- Il regolamento definisce le misure minime che le infrastrutture come i data center e i servizi cloud devono rispettare per supportare i servizi pubblici.

Cloud e PA

- Le soluzioni qualificate da ACN sono esposte nel catalogo delle infrastrutture e dei servizi cloud che contiene schede tecniche per ogni servizio con la tipologia e il relativo livello di qualifica.
- Le amministrazioni possono acquistare le soluzioni cloud qualificate secondo la normativa vigente in tema di procurement delle PA (Codice degli appalti) con gli strumenti delle centrali di committenza (esempio il Mercato elettronico della PA (MEPA) di CONSIP).
- I soggetti privati che vogliono qualificare le soluzioni cloud devono farlo online fornendo tutte le informazioni richieste.

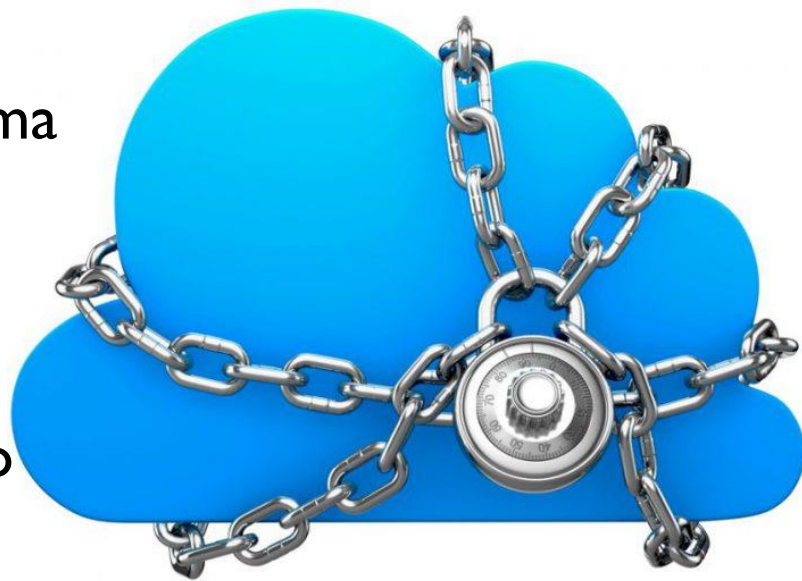
Cloud e PA

- Attenti ai contratti con condizioni predisposte unilateralmente dai fornitori di servizi di cloud computing.
- Il cliente, titolare del trattamento dei dati, deve
 - sapersi tutelare (sapendo dettagliare a livello contrattuale gli obblighi di conformità che intende trasferire all'outsourcer) e
 - adeguare le modalità di controllo con specifici accordi contrattuali.



Cloud e PA

- Un sistema cloud “aperto” può evitare il cosiddetto “**vendor lock-in**”, ovvero, la necessità di prodotti dell'azienda “proprietaria” del sistema per leggere correttamente e completamente i dati conservati.
- I sistemi di cloud computing devono poter essere infatti capaci di scambiarsi dati, (ad es., leggendo e scrivendo sullo stesso file e usando lo stesso protocollo per farlo).



Cloud e (pubblici e private)

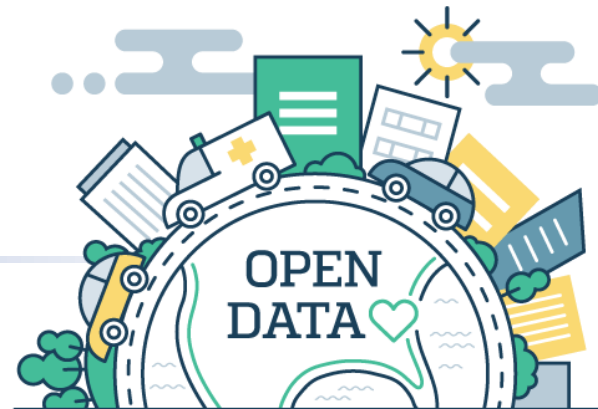
- I vantaggi del cloud computing sono significativi e molti utenti non intendono rinunciarvi.
- Disponibilità, accessibilità e archiviazione dei propri dati e documenti in ogni momento (**everytime**) e da qualunque dispositivo (**everywhere**) collegato alla rete sono elementi chiave.
- Ma ...

Cloud e Utenti (pubblici e private)

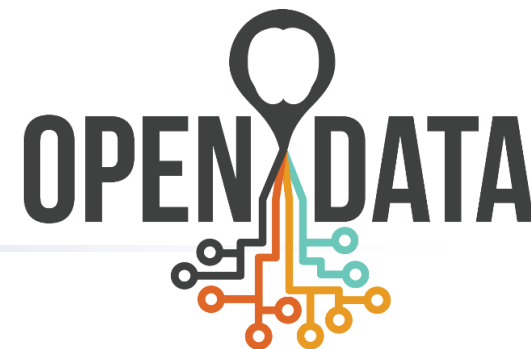
- ... bisogna fare attenzione a verificare l'*outsourcer* di un servizio di storage o applicativo ed evitare generici e standardizzati contratti di servizi di cloud computing che vengono offerti.
- Qualità del servizio, affidabilità, disponibilità, interoperabilità, accessibilità, semplicità d'uso, costi ecc. Vanno tutti valutati per avere benefici ed evitare sorprese.



Open Data: cosa sono?



- Gli **Open Data** sono:
 - ***Dati accessibili a tutti***: informazioni disponibili in formato digitale per il cittadino senza restrizioni e messe a disposizione a titolo gratuito per essere riutilizzate.
 - ***Liberamente utilizzabili***: Possono essere liberamente utilizzati, riutilizzati e ridistribuiti da chiunque, eventualmente citando la fonte e con lo stesso tipo di licenza con cui sono stati rilasciati.
 - ***In formato aperto***: Sono disponibili in un formato utile e modificabile, come CSV, JSON, XML, RDF, che consente a chiunque di leggerli e rielaborarli.
- Gli open data sono un patrimonio informativo comune che può essere utilizzato per creare nuove applicazioni, servizi e conoscenze, favorendo la trasparenza e l'innovazione.



- **Obbligo di pubblicazione:** Le pubbliche amministrazioni sono tenute a pubblicare i propri dati in formati aperti e riutilizzabili, rendendoli accessibili a tutti attraverso portali dedicati come dati.gov.it e data.europa.eu.
- **Licenze aperte:** I dati pubblicati devono essere rilasciati con licenze aperte che ne consentano il libero riutilizzo, la modifica e la ridistribuzione, purché vengano rispettate alcune condizioni.
- **Cataloghi dei dati:** Le amministrazioni devono creare e mantenere aggiornati dei cataloghi che descrivono i dati pubblicati, rendendoli facilmente ricercabili e comprensibili.
- **Qualità dei dati:** I dati pubblicati devono essere di buona qualità, accurati e aggiornati.
- **Valutazione delle performance:** L'apertura dei dati è uno dei parametri di valutazione delle performance delle amministrazioni pubbliche.

Open Data e PA



- Il Piano Triennale per l'informatica nella PA 2024-2026 definisce obiettivi specifici per gli Open Data nella PA, tra questi:
 - la condivisione dei dati tra le pubbliche amministrazioni,
 - il loro riutilizzo da parte di cittadini e imprese,
 - il miglioramento della qualità di dati e metadati e
 - la valorizzazione del patrimonio informativo pubblico.

Open Data e PA

- La costruzione di un'**economia dei dati** è l'obiettivo che l'Unione Europea vuole perseguire attraverso una serie di iniziative di regolazione avviate dal 2020.
- La Strategia europea dei dati ha introdotto la creazione di spazi di dati (***data spaces***) comuni e interoperabili al fine di superare le barriere legali e tecniche alla condivisione dei dati e, quindi, sfruttare l'enorme potenziale dell'innovazione guidata dai dati.

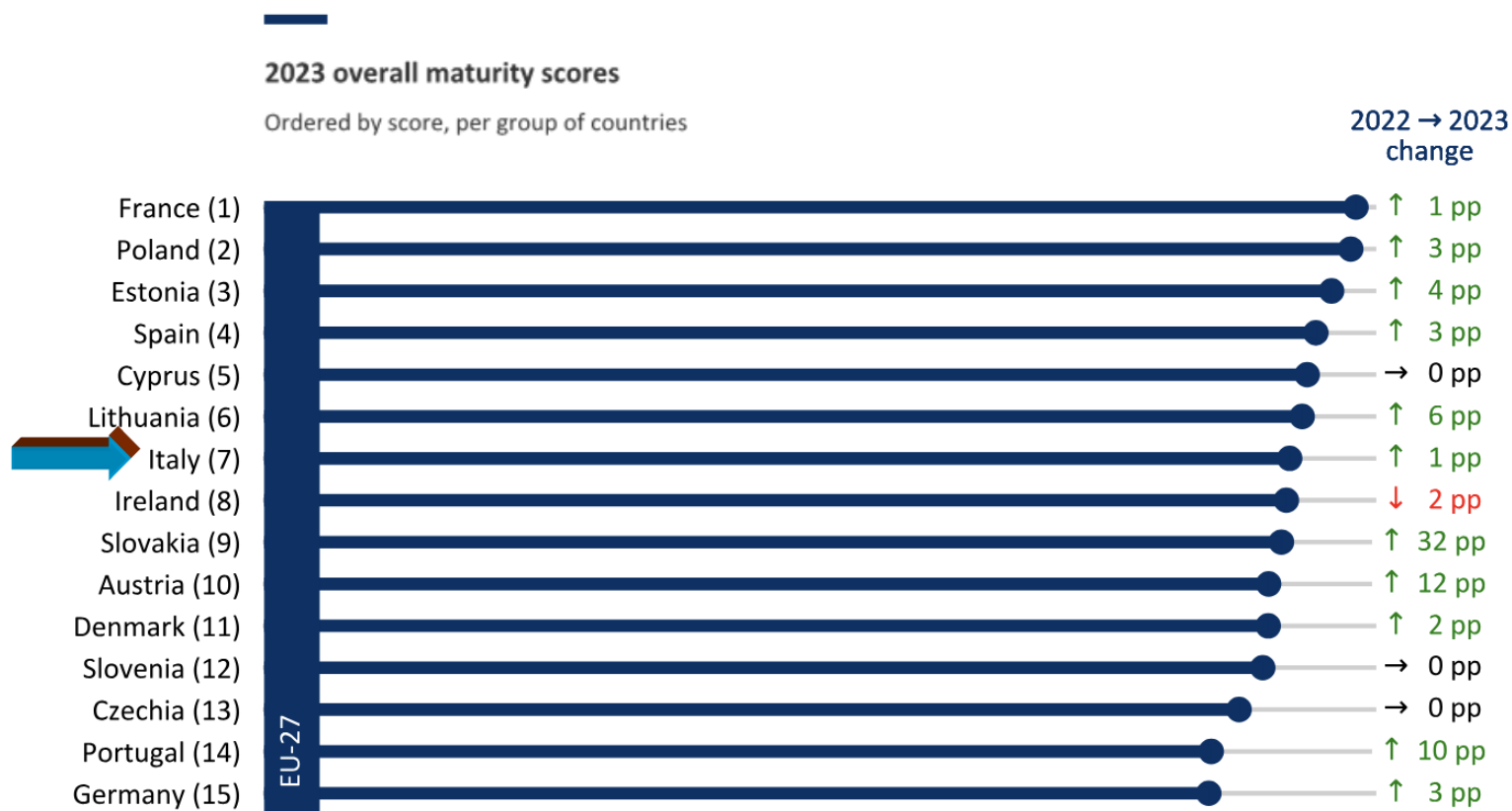


- Con l'adozione dell'atto sulla *governance* dei dati (***Data Governance Act***), sono stati definiti e rafforzati i meccanismi per aumentare la disponibilità dei dati e superare gli ostacoli tecnici al riutilizzo di alcune particolari tipologie di dati.

Open Data in Europa

OPEN DATA MATURITY REPORT – 2023

data.europa.eu



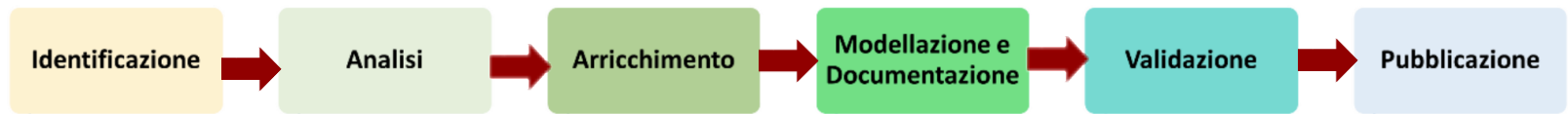
Open Data e PA

- In Italia, il recepimento della Direttiva Europea (UE) 2019/1024 (Direttiva Open Data) sull'apertura dei dati e il riutilizzo dell'informazione del settore pubblico, è avvenuto con il Decreto lgs. n. 200/2021.
- L'obiettivo strategico è perseguito attraverso l'implementazione delle nuove regole tecniche definite con le Linee guida sui dati aperti (Determinazione AGID n. 183/2023).
- L'obiettivo è supportare le pubbliche amministrazioni e gli altri soggetti interessati nel processo di apertura dei dati e favorire l'aumento dell'offerta di dati pubblici preziosi a fini di riutilizzo.



Open Data e PA

- Il processo di formazione e pubblicazione dei dati nelle PA dovrebbe seguire uno schema proposto nelle Linee guida che si compone delle seguenti fasi:
 1. **Identificazione** (ricognizione, analisi dei vincoli, priorità e percorso di apertura dei dati)
 2. **Analisi** (analisi della qualità, bonifica, analisi di processo)
 3. **Arricchimento** (vocabolari controllati, ontologie, mashup e linking nei linked open data - LOD)
 4. **Modellazione e documentazione** (schemi dei dati, modelli dati, conservazione e storicizzazione)
 5. **Validazione** (qualità dei dati)
 6. **Pubblicazione** (meta datazione, politiche di accesso e licenza, modalità di pubblicazione).



Open Data e PA

- Tra i diversi tipi di dati rientrano anche i **dati di elevato valore**, identificati con il Regolamento di esecuzione (UE) 2023/138 sulla base delle sei categorie tematiche stabilite con la Direttiva Open Data.
 - dati geospaziali,
 - dati relativi all'osservazione della Terra e all'ambiente,
 - dati meteorologici,
 - dati statistici,
 - dati relativi alle imprese e alla proprietà delle imprese,
 - dati relativi alla mobilità.
- Nel dicembre 2023 è stata adottata da AGID una specifica Guida operativa sulle serie di dati di elevato valore come documento di orientamento per le PA per guidarle nella implementazione delle disposizioni europee.



