



Introduzione al Cloud

Quest'opera è distribuita con Licenza Creative Commons
Attribuzione - Condividi allo stesso modo 4.0 Internazionale



Agenda

- **Il Cloud computing**
- Cloud computing e/o Virtualizzazione?
- Strategie di adozione del Cloud

Cloud computing

- La definizione classica di riferimento è quella del National Institute of Standards and Technology (NIST) USA (<http://goo.gl/eBGBk>)
- In sintesi il Cloud computing si occupa di:

Fornitura di tecnologia di informazione e comunicazione (ICT) come servizio

Caratteristiche del Cloud

- **Self-service, on-demand**
 - Il cliente chiede autonomamente ciò che gli serve, quando gli serve (e sperabilmente lo ottiene).
- **Accesso attraverso la rete**
 - Assume che una rete (Internet o intranet) sia disponibile, normalmente a banda larga.
- **Pool di risorse**
 - L'utente non si preoccupa di conoscere i dettagli delle risorse, che sono gestiti dai Cloud resource provider.
- **Elasticità**
 - Il servizio Cloud può scalare rapidamente come dimensioni a seconda delle necessità del cliente.
- **Pagamento a consumo**
 - Il cliente paga solo per ciò che usa.

Una analogia: l'autonoleggio

- Self-service, on-demand
 - Prenotazione telefonica oppure online
- Rete
 - Estesa rete di autonoleggi in tutto il mondo
- Pool di risorse
 - Pensa l'autonoleggio a gestire e sapere quante macchine gli servono
- Elasticità
 - Il numero di auto disponibili normalmente varia a seconda della richiesta
- Pagamento a consumo
 - Il cliente paga per il tempo in cui usa l'auto (e non pensa ad assicurazione, gomme, etc.)



Economy



Compact



Intermediate



Full Size



Premium



Luxury



Minivan



Convertible

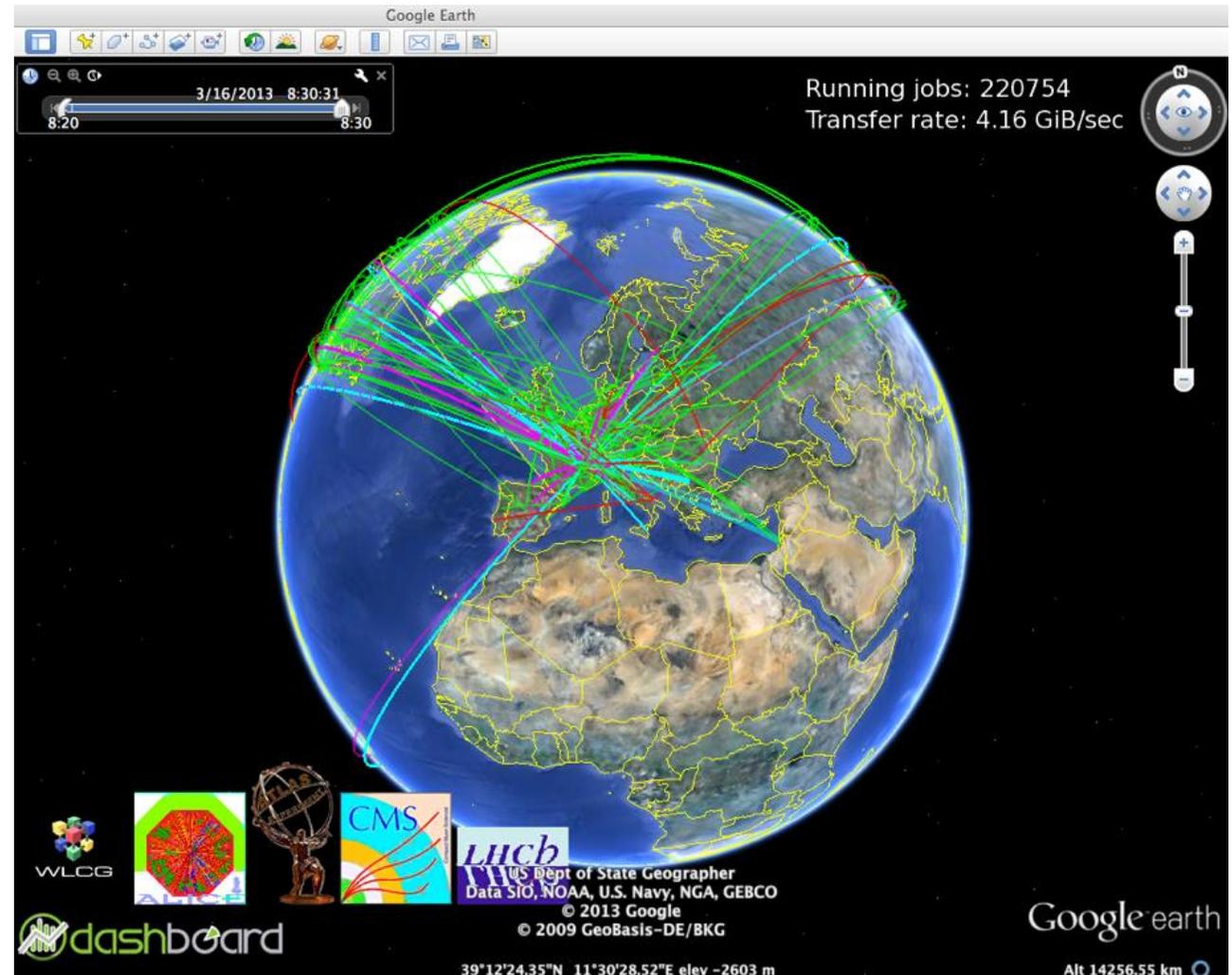


Premium SUV

Fonte: <http://goo.gl/cEa8M>

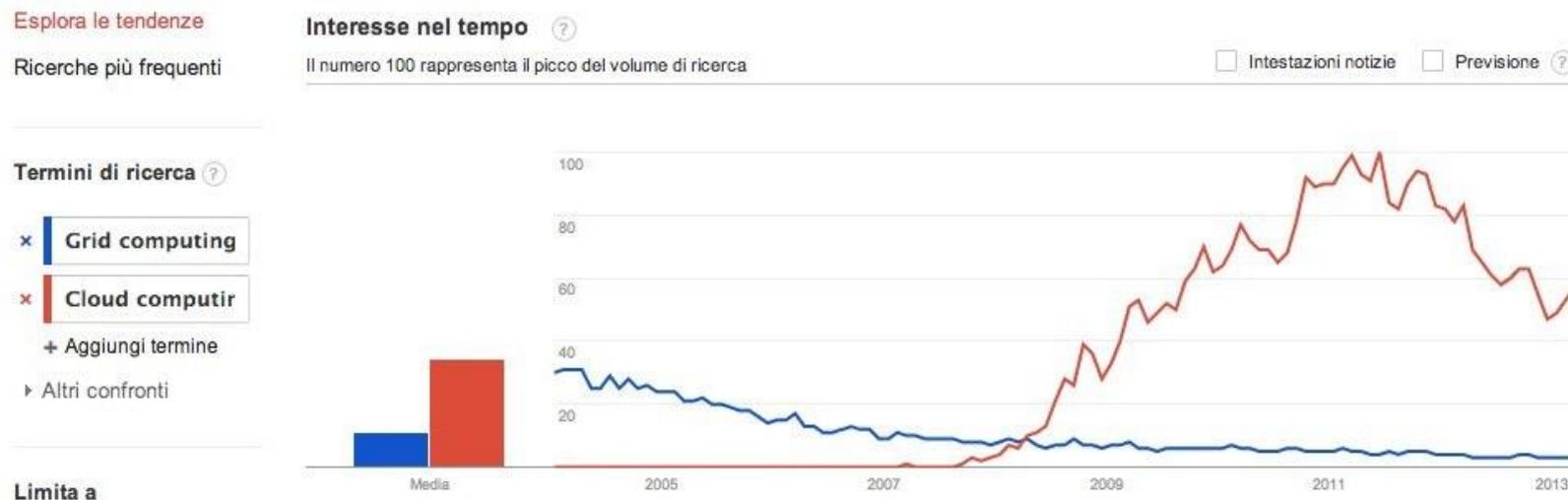
Che cosa c'è di nuovo?

- La fornitura di risorse in modo distribuito, in forme *simili*, avviene da molti anni
- Un esempio di successo nel mondo scientifico è rappresentato dal **Grid Computing**. Cf. <http://goo.gl/i5Rkt>:
 - Stato in tempo reale dei job di calcolo (Computing) e del trasferimento dati (Storage) legati agli esperimenti LHC al CERN



Tuttavia...

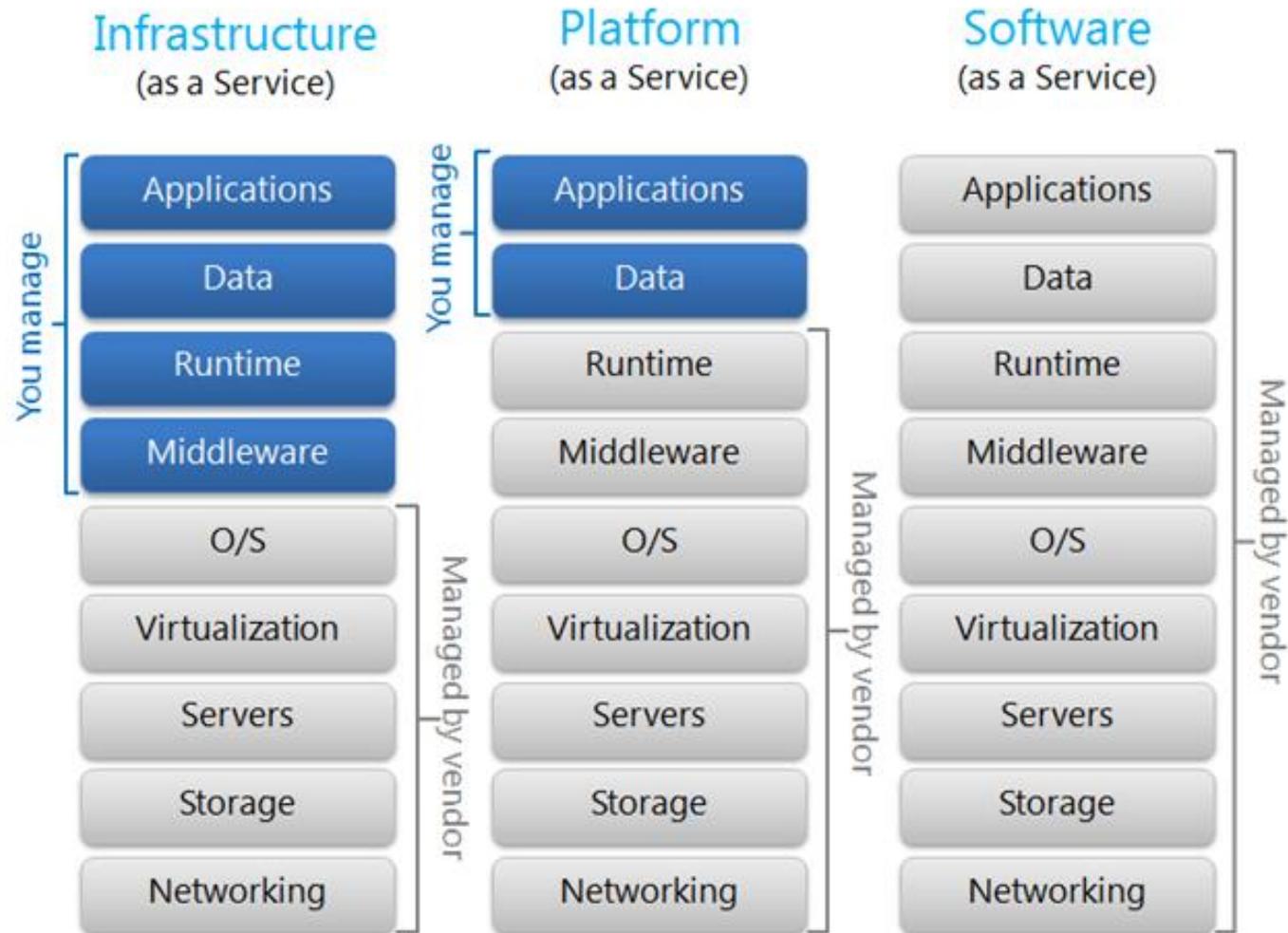
- Il Grid Computing, di enorme successo nel suo campo, non ha mai avuto significativa diffusione al di fuori di grandi collaborazioni scientifiche
- I trends (<http://www.google.com/trends/>): Grid Computing vs. Cloud Computing



Il focus sul “service”

- Abbiamo visto che nella definizione di Cloud computing (“Fornitura di tecnologia di informazione e comunicazione come servizio”) il ***servizio*** nei confronti del cliente è parte essenziale.
- Il Cloud computing si può modellare infatti intorno a *servizi* legati principalmente a
 - Infrastruttura (**IaaS** → Infrastructure as a Service)
 - Piattaforma (**PaaS** → Platform as a Service)
 - Software (**SaaS** → Software as a Service)

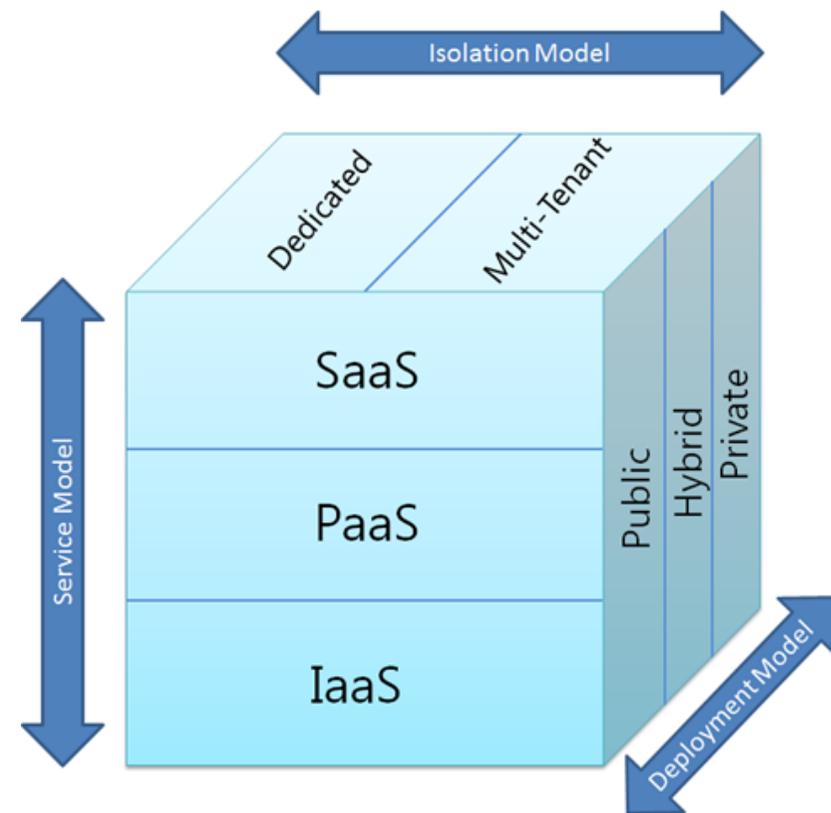
Chi fa cosa?



Fonte: <http://goo.gl/1jmkR>

Aggiungiamo dimensioni

- Oltre i modelli di *servizio*, parti importanti per definire e capire il Cloud computing sono i modelli di:
 - ***deployment*** (dove distribuisco i servizi)
 - ***isolamento*** (come isolo i servizi)



Fonte: <http://goo.gl/1jmkR>

Deployment: i “tipi di Cloud”

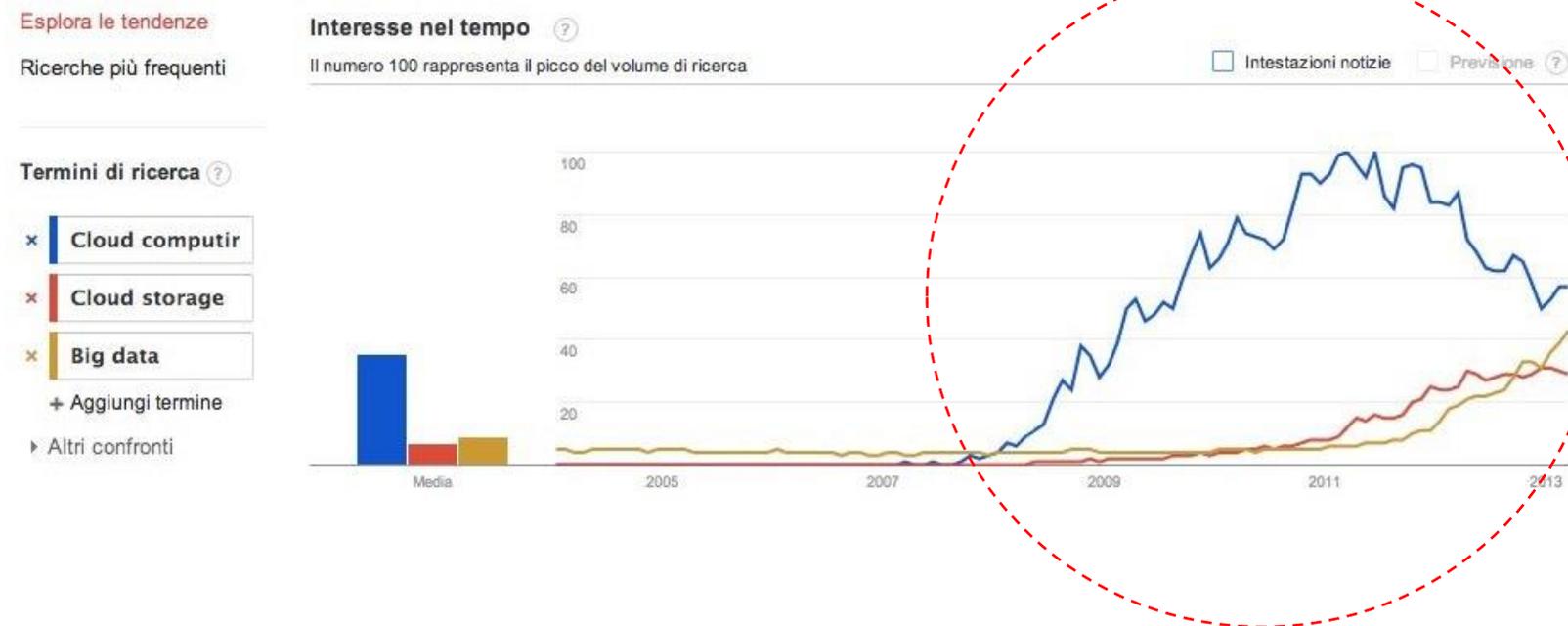
- **Cloud privata:**
 - L’infrastruttura viene fornita per un *uso* esclusivo da parte di una singola organizzazione. La gestione, l’operazione, la proprietà, la dislocazione della Cloud privata tuttavia può essere anche indipendente dall’organizzazione che la usa.
- **Cloud di comunità (Community Cloud):**
 - L’infrastruttura è disponibile ad una comunità di organizzazioni che hanno uno scopo comune (ad esempio missione, requisiti di sicurezza, conformità a regole comuni, etc.)
- **Cloud pubblica:**
 - L’infrastruttura è disponibile in generale al pubblico. La gestione può essere pubblica o privata. La dislocazione è presso il fornitore di servizi.
- **Cloud ibrida:**
 - L’infrastruttura è una combinazione di due o più infrastrutture Cloud (private, di comunità o pubbliche) che sono collegate in modo da garantire forme di portabilità ad esempio di dati o applicazioni.

Isolamento

- I modelli di **isolamento** nel Cloud (spesso ignorati) sono importanti e si dividono in:
 - Infrastrutture dedicate
 - Infrastrutture “multi-tenant” (con diversi [tipi di] clienti)
- Il tipo di isolamento è importante per molti aspetti, come:
 - Segmentazione delle risorse
 - Protezione dei dati
 - Sicurezza delle applicazioni
 - Auditing
 - Disaster recovery

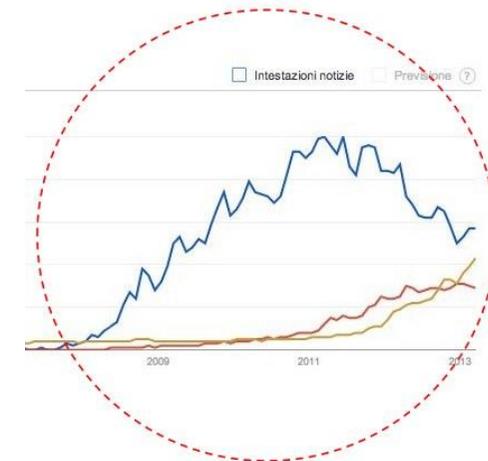
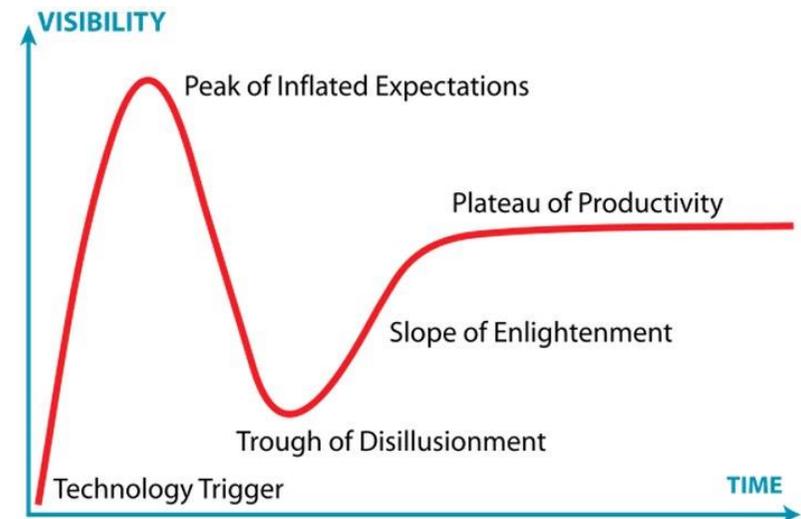
Ma la moda è già passata?

- Cf. i trend di Google per i termini “Cloud computing”, “Cloud storage”, “Big data”



Dunque il Cloud è una moda?

- In molte tecnologie si passa attraverso varie fasi di “hype” (di moda). Possiamo distinguere (cf. <http://goo.gl/1qaC>) spesso tra:
 - Inizio della tecnologia
 - Picco di aspettative
 - Disillusione (aspettative fallite)
 - Crescita di maturazione
 - Altopiano di produttività



Agenda

- Il Cloud computing
- **Cloud computing e/o Virtualizzazione?**
- Strategie di adozione del Cloud

Virtualizzazione

- Il Cloud computing può anche essere fornito *senza* l'utilizzo di tecnologie di virtualizzazione.
 - Spesso, tuttavia, l'utilizzo di tecnologie di virtualizzazione consente di ridurre i costi operativi e in conto capitale.
 - Essere in grado di fornire molto rapidamente delle macchine virtuali non è comunque efficiente, se ad esempio servono diversi mesi per acquisire ed installare gli host fisici che ospiteranno le macchine virtuali.
 - Inoltre, il tempo impiegato per la fornitura o gestione dello strato di virtualizzazione è recuperato dai risparmi associati al non dover dedicare dei server fisici?
 - Importanza di avere tool di installazione, monitoring e accounting il più possibile automatizzati.
- Ma che cosa si intende con *virtualizzazione*?

Tipi di virtualizzazione

- Quando i sistemi virtuali usano un insieme di istruzioni di linguaggio macchina (instruction set) *differente* da quello disponibile sul sistema ospitante, si parla di **emulazione**.
 - Esempio: emulare ARM o PowerPC su piattaforma Intel.
- Quando i sistemi virtuali usano *lo stesso* instruction set del sistema ospitante, si parla propriamente di **virtualizzazione**.
 - A sua volta distinta tra *virtualizzazione completa* e *para-virtualizzazione*.

Vantaggi della virtualizzazione (1/2)

- Consolidamento di server.
 - Molte VM su una macchina fisica.
 - Riduce i costi di acquisizione dell'hardware e può semplificare operazioni come monitoring e amministrazione.
- Isolamento (*sandboxing*).
 - Isolamento delle applicazioni.
 - Sviluppo, test e debug di codice.
 - Creazione di ambienti dedicati per applicazioni legacy.
- Creazione di VM *on-demand*.

Vantaggi della virtualizzazione (2/2)

- Disaccoppiamento dell'hardware dal software.
 - Suspend/Resume delle VM.
 - Migrazione di VM tra server fisici (con vari gradi di sofisticazione).
- Test di nuove versioni di sistema operativo, applicazioni.
 - O di vecchie versioni: data preservation.
- Emulazione di hardware non presente sulla macchina fisica.
- Esecuzione di applicazioni non compatibili con il sistema operativo della macchina fisica.

Svantaggi della virtualizzazione

- Sicurezza.
 - Si mettono sullo stesso hardware diversi sistemi operativi, gestiti da un software – aumento della probabilità di bug e aumento del numero dei cosiddetti *attack vectors*:
 - VM-to-VM network attacks.
 - VM-to-HV (KVM o XEN). KVM è un modulo del kernel Linux. Xen è direttamente connesso all'hardware. → tutto può diventare compromesso.
 - VM-to-QEMU. QEMU è un software grande e complesso. Se viene compromesso, l'attacco può arrivare al sistema operativo.
- Performance.
 - Overhead per la macchina fisica.
 - Minori prestazioni per la VM, soprattutto nell'I/O.

Una (una!) slide sulla security

- **La sicurezza nell'utilizzo della virtualizzazione** non è un argomento da prendere alla leggera. A parte problemi configurazione legati al corretto isolamento delle reti, questi sono alcuni esempi recenti di exploit di sicurezza:
 - CloudBurst, 2008 - VMware, result: full breakout.
 - Xen Ownage Trilogy, 2011 - Xen, result: full breakout.
 - VirtuNoid, 2011 - KVM, result: full breakout.
 - SYSRET-64, 2012 - Xen, result: full breakout.
 - VMDK Has Left The Building, 2012 - VMware, result: data leakage, loss.
 - KVM IOAPIC, SET MSR, TIME, 2013 - KVM, result: DoS, potential breakout.
- Un testo utile (originato da OpenStack ma con molti concetti validi in modo generale) è la **OpenStack Security Guide**, disponibile su <http://goo.gl/ibkNsD>.
- Esiste un impatto (positivo) di sistemi efficienti di provisioning sulla security.

Riassumendo: virtualizzazione vs. Cloud computing

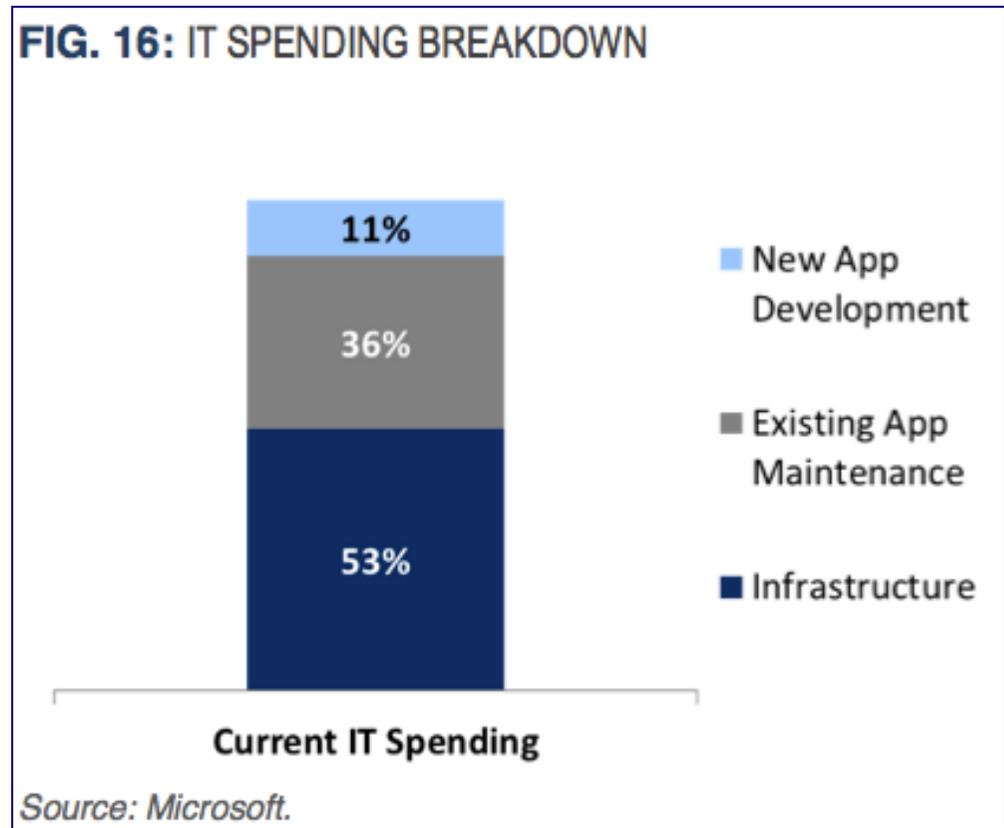
- **Installazione/reinstallazione** di server o di applicazioni su VM di per sé ***non è Cloud computing***.
- Verificare con le **5 caratteristiche del Cloud** mostrate precedentemente:
 - Self-service, on-demand → **NO** (tipicamente è un dipartimento IT che fornisce le VM)
 - Accesso attraverso la rete → **NO** (deployment limitato a “internal customers”)
 - Pool di risorse → **Sì**
 - Elasticità → **NO** (tipicamente è un dipartimento IT che deve installare sistema operativo + software, e non necessariamente in modo scalabile)
 - Pagamento a consumo → **NO** (spesso il billing non viene fatto a consumo ma in modo tradizionale)

Agenda

- Il Cloud computing
- Cloud computing e/o Virtualizzazione?
- **Strategie di adozione del Cloud**

Vantaggi e svantaggi

- A partire dalle caratteristiche principali del Cloud si possono derivare un certo numero di liste di vantaggi e svantaggi.
- Nelle slides seguenti compareremo le caratteristiche principali del Cloud con quelle di infrastrutture tradizionali.



Riduzione dei costi

- Un resource center a larga scala consente di **abbassare i costi per singolo server**.
 - Per il consumo elettrico (un grande centro è probabile che paghi l'energia meno di un centro piccolo).
 - Il costo di manodopera viene ammortizzato su più risorse.
 - Il costo di acquisizione dell'hardware è più basso grazie alla scala.
- Più clienti ci sono, meno incidono **i costi di gestione per cliente**.
- L'aggregazione delle risorse porta ad un loro **più efficiente utilizzo**.
 - E ad una diminuzione della complessità relativa.
 - (con una buona progettazione ed organizzazione)

Flessibilità e scalabilità

- **Self-service provisioning**
 - Da comparare con i tempi di procurement delle risorse in un data center tradizionale.
- **Scale-out**
 - In caso di necessità è possibile aumentare risorse attraverso altri Cloud (anziché comperare nuove risorse).
 - Load Balancer as a Service (ad esempio in OpenStack) – anziché configurare / acquistare dei load balancer proprietari.
- **Mitigazione delle incertezze di crescita**
 - Spesso i pattern di crescita di un business non sono noti; con infrastrutture tradizionali questo può portare ad un overprovisioning eccessivo.
- **Modulazione flessibile del tipo di workload**
 - Con il Cloud è possibile decidere di cambiare il tipo di workload, (es. CPU-intensive vs. I/O-intensive vs. HPC) senza necessità di dedicare risorse o prendere decisioni tecnologiche troppo in anticipo.

Collaborazione, opportunità

- Soprattutto nel caso di adozione di soluzioni di tipo SaaS (es. Google Apps, Dropbox), il Cloud fornisce **accesso a risorse in maniera ubiqua su qualunque tipologia di computer**.
 - Comprese piattaforme mobili, Internet appliances.
 - Senza richiesta di installazione di applicazioni, software, sistemi operativi specifici, licenze ad-hoc.

Opportunità di business

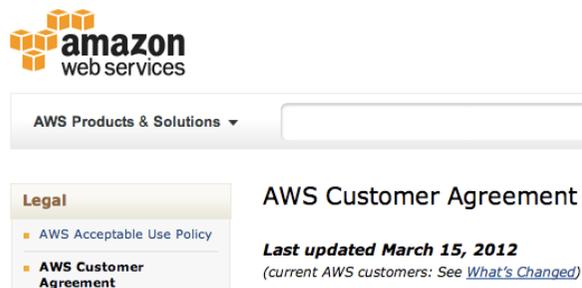
- Dal punto di vista del business, il Cloud computing offre **opportunità di profitto** non legate alla semplice rivendita di software sviluppati da altri.
 - Ad esempio nei campi di sviluppo tecnologico, integrazione, training e supporto.
 - Il modello Open Source è in questo caso particolarmente adatto (cf. ditte come Red Hat, Canonical, Mirantis, ma anche IBM e HP, per esempio).

E ora, gli svantaggi...



Non esclusività

- **Perdita della proprietà del prodotto**
 - Ad esempio, Amazon può sviluppare prodotti in competizione con quanto sviluppate voi stessi su AWS, adottare tecnologie da voi utilizzate, ...
 - ... o assistere terze parti perché sviluppino prodotti in competizione con i vostri.



Un esempio di contratto di un noto Cloud provider (Amazon)

13.3 Independent Contractors; Non-Exclusive Rights. We and you are independent contractors, and neither party, nor any of their respective affiliates, is an agent of the other for any purpose or has the authority to bind the other. Both parties reserve the right (a) to develop or have developed for it products, services, concepts, systems, or techniques that are similar to or compete with the products, services, concepts, systems, or techniques developed or contemplated by the other party and (b) to assist third party developers or systems integrators who may offer products or services which compete with the other party's products or services.

Non disponibilità

- **Rifiuto di responsabilità** in caso di non disponibilità di servizi o dati.
 - Per esempio a causa di power outages, di fallimenti di sistema o di qualunque altra interruzione.
 - O a causa di accesso non autorizzato, alterazioni, cancellazione, distruzione, perdita, incapacità di memorizzazione o altro di dati o di qualunque altro contenuto in AWS.

11. Limitations of Liability.

WE AND OUR AFFILIATES OR LICENSORS WILL NOT BE LIABLE TO YOU FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR EXEMPLARY DAMAGES (INCLUDING DAMAGES FOR LOSS OF PROFITS, GOODWILL, USE, OR DATA), EVEN IF A PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. FURTHER, NEITHER WE NOR ANY OF OUR AFFILIATES OR LICENSORS WILL BE RESPONSIBLE FOR ANY COMPENSATION, REIMBURSEMENT, OR DAMAGES ARISING IN CONNECTION WITH: (A) YOUR INABILITY TO USE THE SERVICES, INCLUDING AS A RESULT OF ANY (I) TERMINATION OR SUSPENSION OF THIS AGREEMENT OR YOUR USE OF OR ACCESS TO THE SERVICE OFFERINGS, (II) OUR DISCONTINUATION OF ANY OR ALL OF THE SERVICE OFFERINGS, OR, (III) WITHOUT LIMITING ANY OBLIGATIONS UNDER THE SLAS, ANY UNANTICIPATED OR UNSCHEDULED DOWNTIME OF ALL OR A PORTION OF THE SERVICES FOR ANY REASON, INCLUDING AS A RESULT OF POWER OUTAGES, SYSTEM FAILURES OR OTHER INTERRUPTIONS; (B) THE COST OF PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; (c) ANY INVESTMENTS, EXPENDITURES, OR COMMITMENTS BY YOU IN CONNECTION WITH THIS AGREEMENT OR YOUR USE OF OR ACCESS TO THE SERVICE OFFERINGS; OR (D) ANY UNAUTHORIZED ACCESS TO, ALTERATION OF, OR THE DELETION, DESTRUCTION, DAMAGE, LOSS OR FAILURE TO STORE ANY OF YOUR CONTENT OR OTHER DATA. IN ANY CASE, OUR AND OUR AFFILIATES' AND LICENSORS' AGGREGATE LIABILITY UNDER THIS AGREEMENT WILL BE LIMITED TO THE AMOUNT YOU ACTUALLY PAY US UNDER THIS AGREEMENT FOR THE SERVICE THAT GAVE RISE TO THE CLAIM DURING THE 12 MONTHS PRECEDING THE CLAIM.

Perdita di garanzie

- **Nessuna garanzia** che il servizio offerto sia “ininterrotto, senza errori, privo di componenti pericolosi”.
 - Né che quanto contenuto in AWS sia sicuro, che non venga perso o danneggiato.
 - Cf. il caso in cui io decida di utilizzare Amazon per memorizzare i miei dati (magari qualche PB...)

10. Disclaimers.

THE SERVICE OFFERINGS ARE PROVIDED “AS IS.” WE AND OUR AFFILIATES AND LICENSORS MAKE NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, WHETHER EXPRESS, IMPLIED, STATUTORY OR OTHERWISE REGARDING THE SERVICE OFFERINGS OR THE THIRD PARTY CONTENT, INCLUDING ANY WARRANTY THAT THE SERVICE OFFERINGS OR THIRD PARTY CONTENT WILL BE UNINTERRUPTED, ERROR FREE OR FREE OF HARMFUL COMPONENTS, OR THAT ANY CONTENT, INCLUDING YOUR CONTENT OR THE THIRD PARTY CONTENT, WILL BE SECURE OR NOT OTHERWISE LOST OR DAMAGED. EXCEPT TO THE EXTENT PROHIBITED BY LAW, WE AND OUR AFFILIATES AND LICENSORS DISCLAIM ALL WARRANTIES, INCLUDING ANY IMPLIED WARRANTIES OF MERCHANTABILITY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, OR QUIET ENJOYMENT, AND ANY WARRANTIES ARISING OUT OF ANY COURSE OF DEALING OR USAGE OF TRADE.

Rimane la responsabilità

- **E' responsabilità dell'utente assicurarsi della sicurezza, della protezione e *del backup* di quanto memorizzato in Amazon (e di nuovo, se fosse qualche PB?), eventualmente utilizzando tecnologie di crittografia.**

4.2 Other Security and Backup. You are responsible for properly configuring and using the Service Offerings and taking your own steps to maintain appropriate security, protection and backup of Your Content, which may include the use of encryption technology to protect Your Content from unauthorized access and routine archiving Your Content. AWS log-in credentials and private keys generated by the Services are for your internal use only and you may not sell, transfer or sublicense them to any other entity or person, except that you may disclose your private key to your agents and subcontractors performing work on your behalf.

Proprietà / privacy dei dati?

- Quando viene cancellato un contratto con un Cloud provider come possiamo essere sicuri che anche tutti i **propri dati vengano rimossi**?
- E come evitare il cosiddetto ***vendor lock-in***?
- Ma dove stanno i miei dati? E per quanto riguarda il ***tapping***?

NSA infiltrates links to Yahoo, Google data centers worldwide, Snowden documents say

Infine, il grande equivoco

- **La capacità non è infinita** (nonostante questo sia un assioma del Cloud).
 - Dunque, le risorse potrebbero essere non disponibili quando servono, o potrebbero non avere le caratteristiche che servono.



Ma come si usa il Cloud?

- Una domanda essenziale è, dati i casi d'uso discussi in precedenza (che presuppongono spesso indipendenza dal vendor, aderenza agli standard, risparmi economici, etc.):

Come migrare verso il Cloud computing?

Migrazione di applicazioni a infrastrutture Cloud

- **Obiettivo:** passaggio di una applicazione da un data center esistente a una Cloud di riferimento.
- Quali sono i **fattori tecnici e di business** che portano alla migrazione? Tipicamente:
 - Riduzione dei costi → resource pooling, pay-per-use
 - “Business agility” → semplificazione del deployment
 - Risparmi gestionali → performance (es. piattaforme performanti, auto-scaling), facilità di gestione (es. delega di responsabilità operative)
- **Cloud pubbliche o private?** Vale quanto detto sopra, in particolare:
 - Traffico a livello WAN? (tipicamente costoso)
 - Sicurezza?
 - Integrazione con altre applicazioni di tipo *legacy*? (ad es. forte accoppiamento con applicazioni che girano su AS400 o simili)

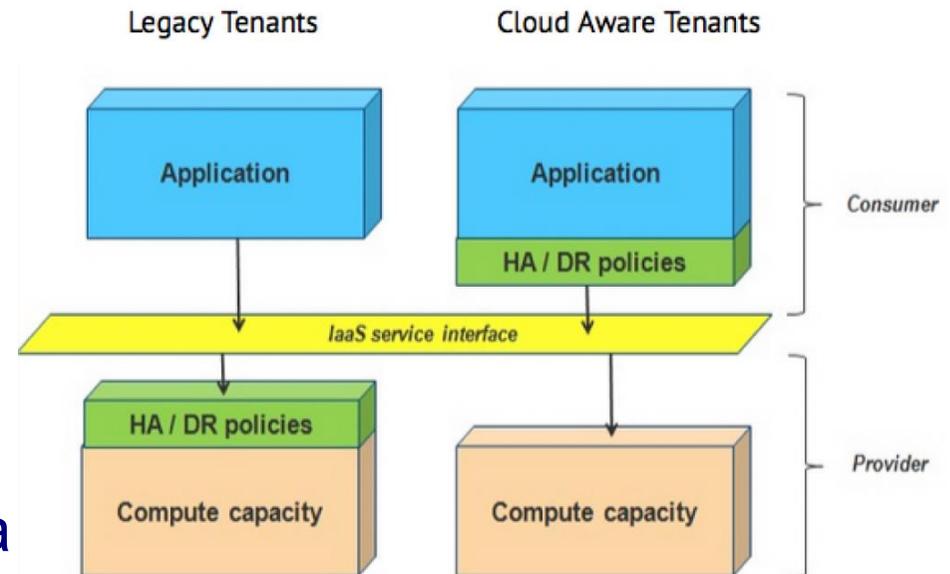
La mia applicazione è “cloud-friendly”?

- **Applicazioni “cloud-aware”:**

- Distribuite
- Stateless
- Fail-over in the app
- Scaling in the app

- **Applicazioni “legacy”:**

- Client-server
- Monolitiche, senza scalabilità orizzontale
- Fail-over nell’infrastruttura
- Scaling nell’infrastruttura



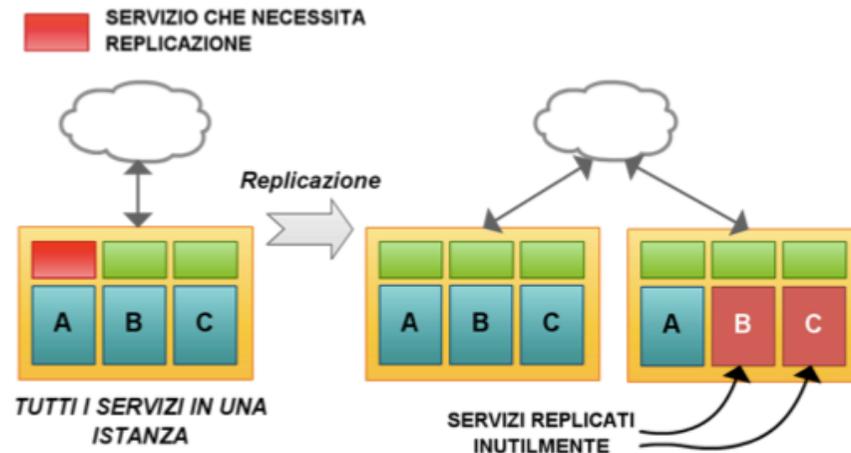
Fonte: VMware

Scomposizione dei servizi

Necessario replicare solamente i componenti che ne hanno reale necessità

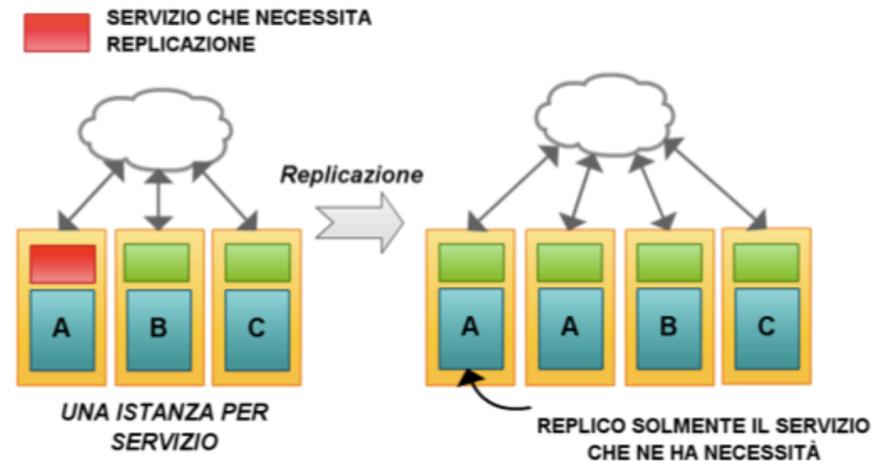
Tutti i servizi in una unica istanza

- *Per scalare un servizio devo replicare l'istanza che fornisce tutti i servizi*



Una servizio per istanza

- *Per scalare un servizio posso replicare solamente l'istanza che lo fornisce*



Fonte: F.Cacco, Tesi di laurea in Informatica, UniPD, Ottobre 2013

Il Cloud, dunque...

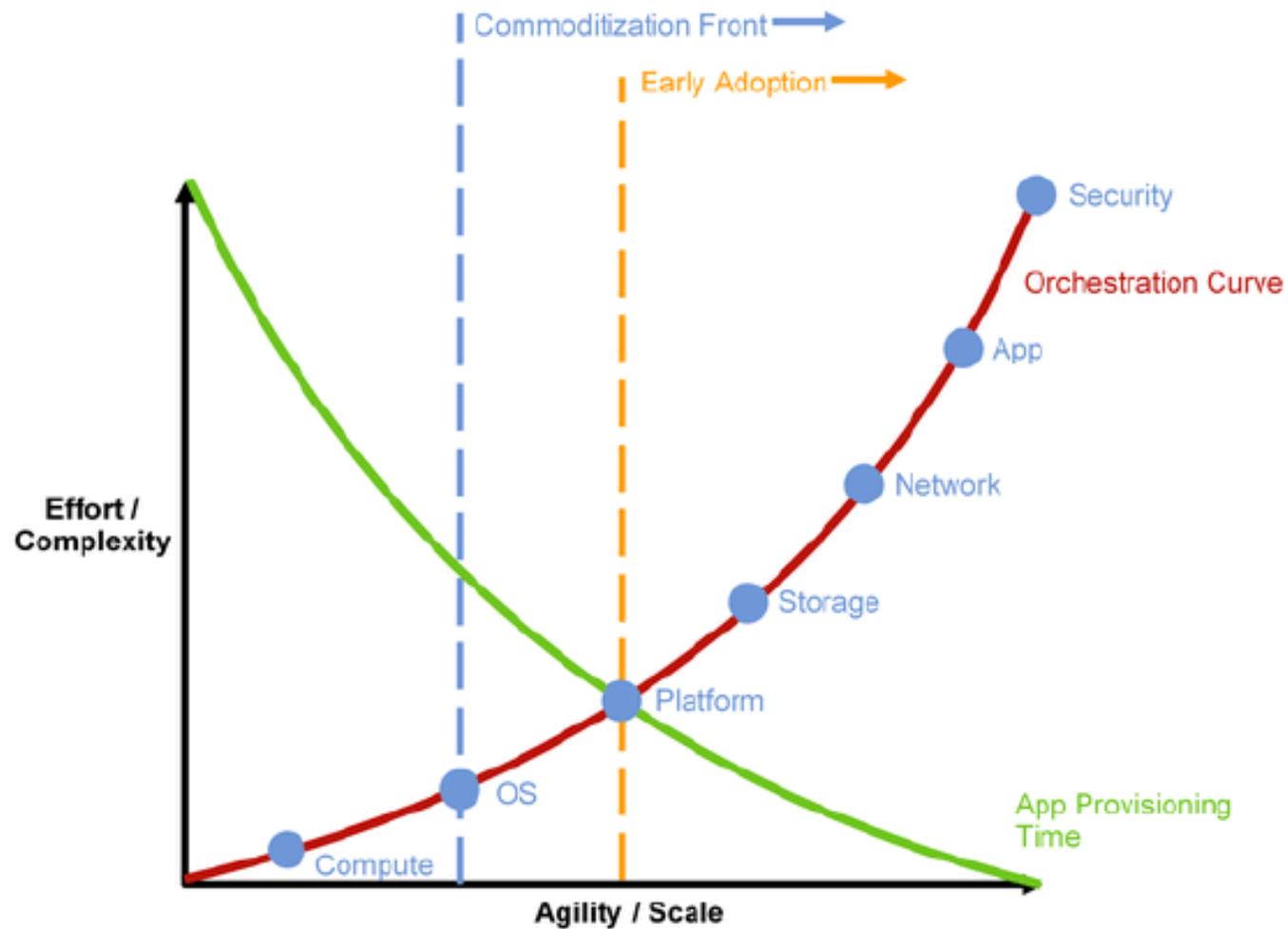
- “Il Cloud computing non è altro che un termine alla moda per indicare il noto concetto di computer collegati via rete. Una nuvola non è che vapore acqueo.”
 - Larry Ellison, co-fondatore e CEO di Oracle Corporation, Settembre 2009
- “Q: qual è la strategia di Oracle sul Cloud computing? A: Oracle ha due obiettivi sul Cloud computing. Il primo è di assicurarsi che sia a livello di una impresa e pronto per l’adozione [...] Il secondo è di supportare sia il Cloud computing pubblico che privato.”
 - Oracle Cloud Computing FAQ, Ottobre 2010
- “Unlike competitors with narrow views of the cloud, Oracle provides the broadest, most complete, and integrated cloud offerings in the industry.”
 - Oracle Cloud Computing, May 2014 (<http://goo.gl/oDVOtA>)
- **“La verità è raramente pura, e mai semplice.”**
 - Oscar Wilde, L’importanza di chiamarsi Ernesto, 1895

Grazie per l'attenzione

per realizzare questa presentazione è stato utilizzato materiale di Paolo Veronesi ed Enrico Fattibene, INFN-CNAF

Maturità

- **Source:** Gartner, Climbing the Cloud Orchestration Curve, Alessandro Perilli, 13 January 2014



Integrazione tra virtualizzazione e Cloud computing

- Nonostante le differenze, ci sono chiaramente **punti di sovrapposizione** tra framework di virtualizzazione e framework di Cloud computing.
- Nel mondo open source, un caso di integrazione è dato dalla crescente **convergenza tra oVirt (<http://www.ovirt.org/>) e OpenStack**.
 - “With [oVirt 3.3], the oVirt community begins integrating oVirt with other important open source projects such as OpenStack and Gluster.” (I. Hein, RedHat)
 - Integrazione delle immagini (Glance) e del networking (Neutron) forniti da OpenStack.
 - Integrazione di cloud-init, il tool standard *de facto* utilizzato per la contestualizzazione di nodi Cloud.

Più risorse per tutti

- In un certo senso, **“democratizzazione” dell’accesso alle risorse.**
 - Ad esempio, una infrastruttura Cloud può facilitare l’accesso e l’utilizzo di risorse da parte di SME (un esempio interessante è il concetto di MCloud B&R).
 - In scala, concetto di equivalenza nell’affitto di 1 macchina per 1000 ore vs. 1000 macchine per 1 ora: possibilità di lavorare a problemi nel passato non affrontabili a causa di limiti di tempo e di costi.
- In generale, il Cloud – con la sua promessa di risorse infinite – può spostare i problemi **da tipologie Batch (accodamento) a tipologie Real-time** o quasi-Real-time.