

Prospettive sul calcolo scientifico (distribuito)

Davide Salomoni

INFN-CNAF

Seminario Fondazione Occhialini

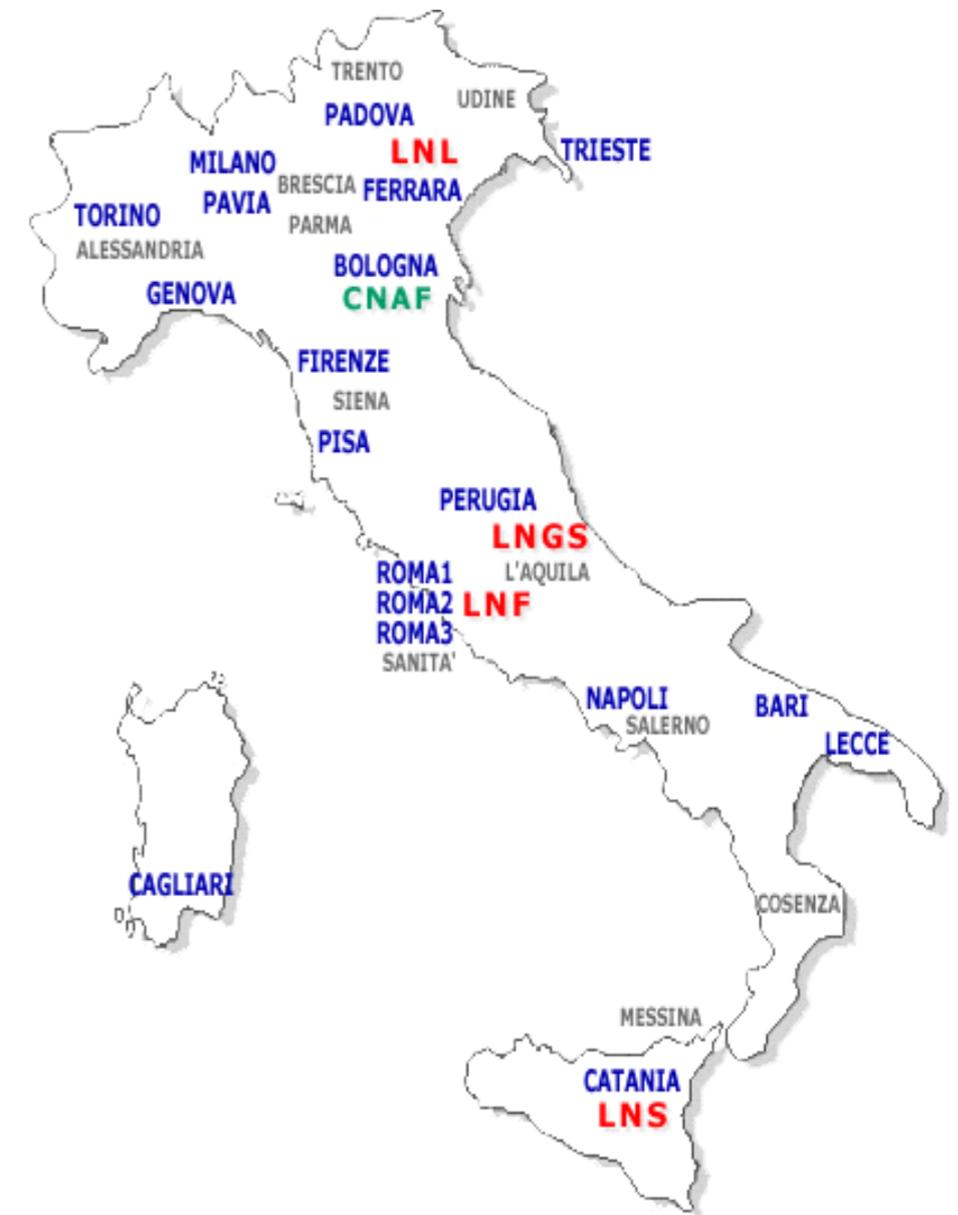
20 aprile 2011

Il mio background

- Laurea in Fisica all'Università di Bologna
- 8 anni all'Istituto Nazionale di Fisica Nucleare (INFN)
 - Occupandomi soprattutto di ricerca e applicazioni di reti di calcolatori applicate alla fisica
- 2 anni allo Stanford Linear Accelerator Center, USA
 - Coordinatore del gruppo reti
- 5 anni in Olanda tra pubblico e privato
 - Ricerca e implementazione di infrastrutture distribuite
- Dal 2006 di nuovo all'INFN a Bologna
 - Coordinatore del calcolo al Centro Nazionale per la Ricerca e Sviluppo nelle Tecnologie Informatiche e Telematiche dell'INFN (CNAF)

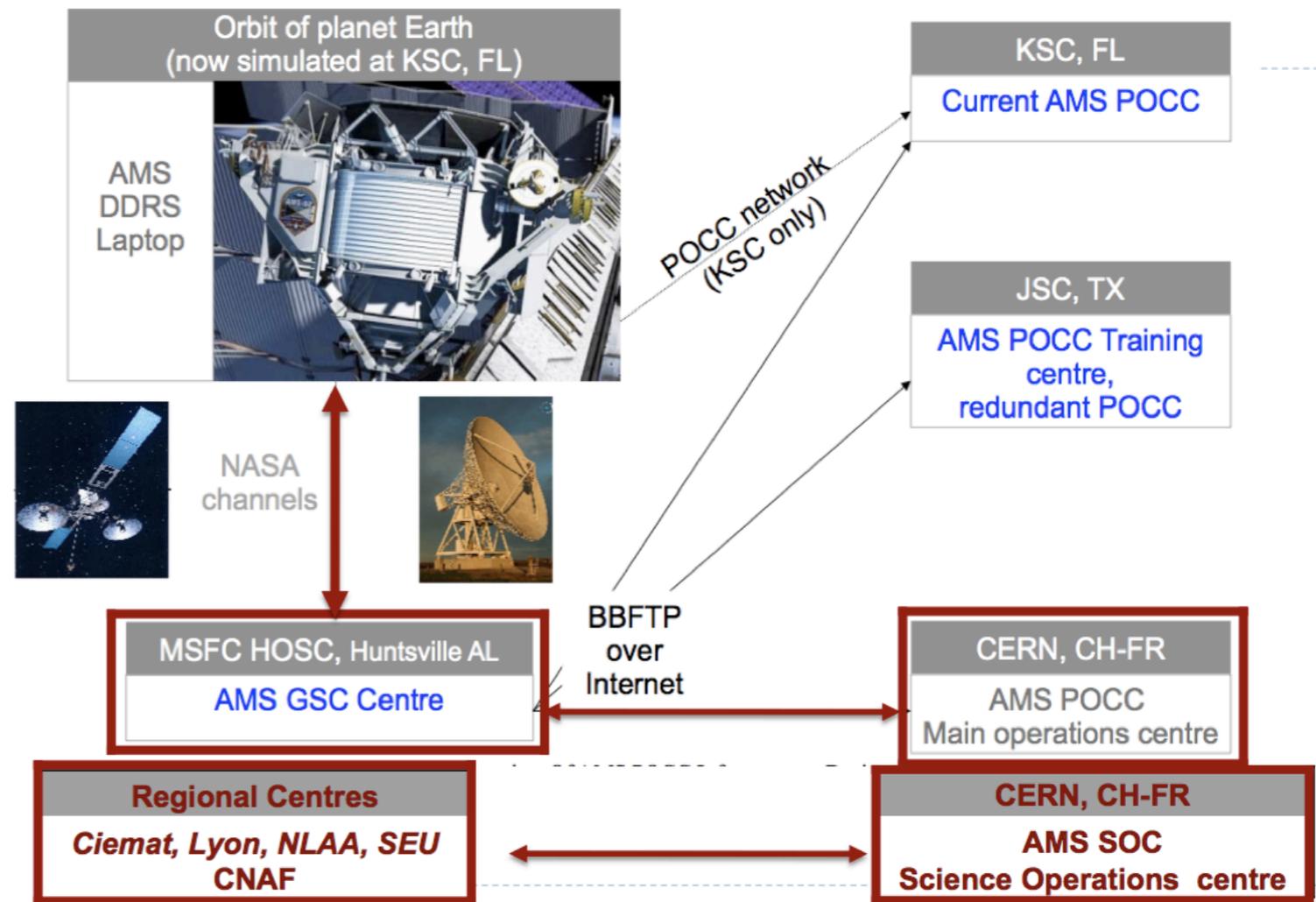
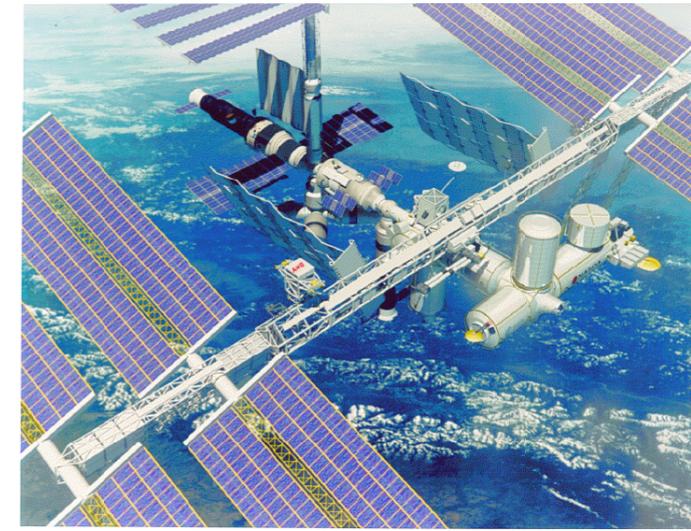
L'Istituto Nazionale di Fisica Nucleare

- L'INFN è un ente di ricerca che ha come missione fondamentale lo studio dei costituenti fondamentali della materia.
- A questo scopo l'Istituto conduce da oltre 50 anni ricerche teoriche e sperimentali nel campo della fisica nucleare, subnucleare e astroparticellare nonché ricerca e sviluppo tecnologico pertinenti all'attività in tali settori.
- L'INFN è oggi organizzato in 20 dipartimenti presenti in quasi tutte le regioni italiane. Ci sono poi 11 gruppi di ricerca collegati, 4 laboratori nazionali (Frascati, Legnaro, Catania e sotto il massiccio del Gran Sasso) e **1 centro di calcolo nazionale** (CNAF, a Bologna).
- L'INFN ha attualmente circa 2000 dipendenti, ai quali vanno aggiunti circa 2450 ricercatori e professori universitari e circa 1300 studenti universitari e assegnisti.



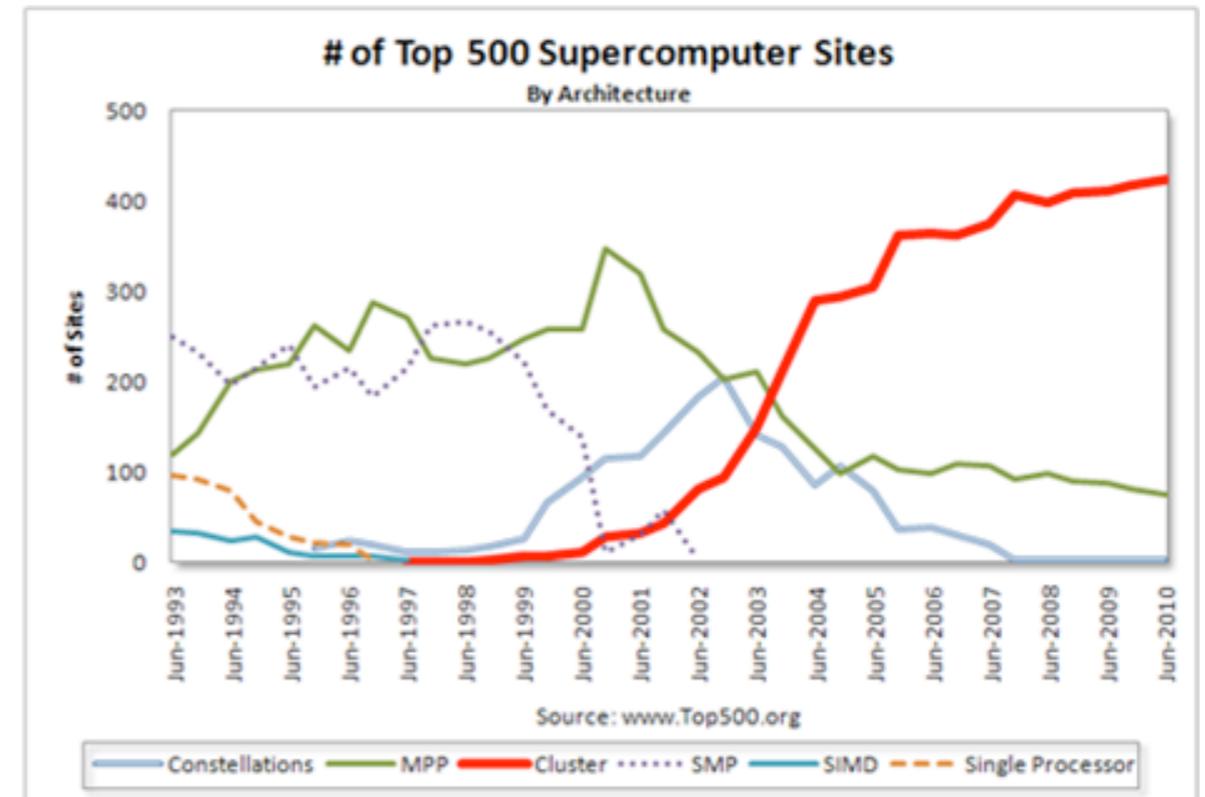
Perché distribuire il calcolo?

- **Geografia:** produco dati in un posto, ne ho bisogno altrove
 - Esempio: AMS, spettrometro magnetico per lo studio della radiazione cosmica, a bordo della ISS
 - In partenza con il penultimo Shuttle, 29/4/2011
 - L'Italia contribuisce per circa il 20% della collaborazione internazionale
 - 17 anni, 16 nazioni, 60 istituti, 600 fisici/ingegneri



Perché distribuire il calcolo?

- **Costo:** il costo di un cluster di molti computer economici può essere minore del costo di un singolo supercomputer



- **Affidabilità:**
evitare single point of failure

- **Scalabilità /
espandibilità:**
architettura modulare

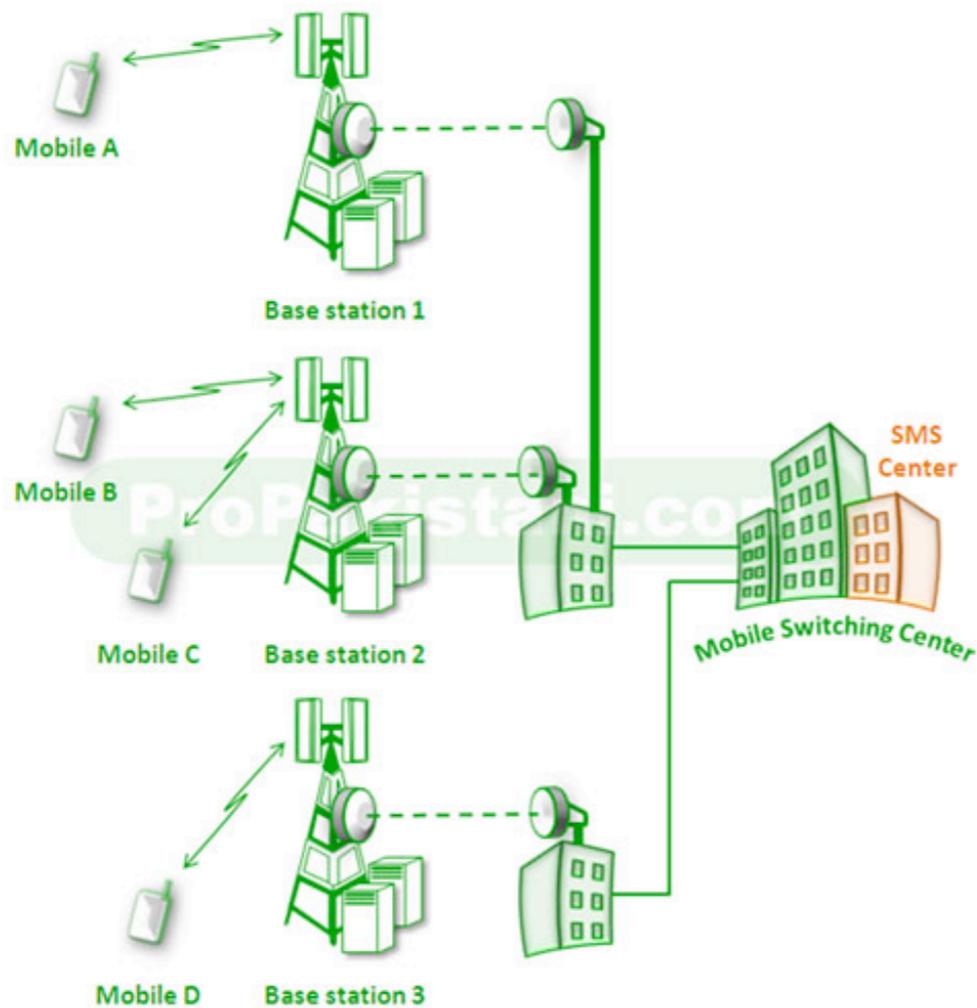


(intermezzo)



Esempi di calcolo distribuito

■ SMS



■ Internet, WWW

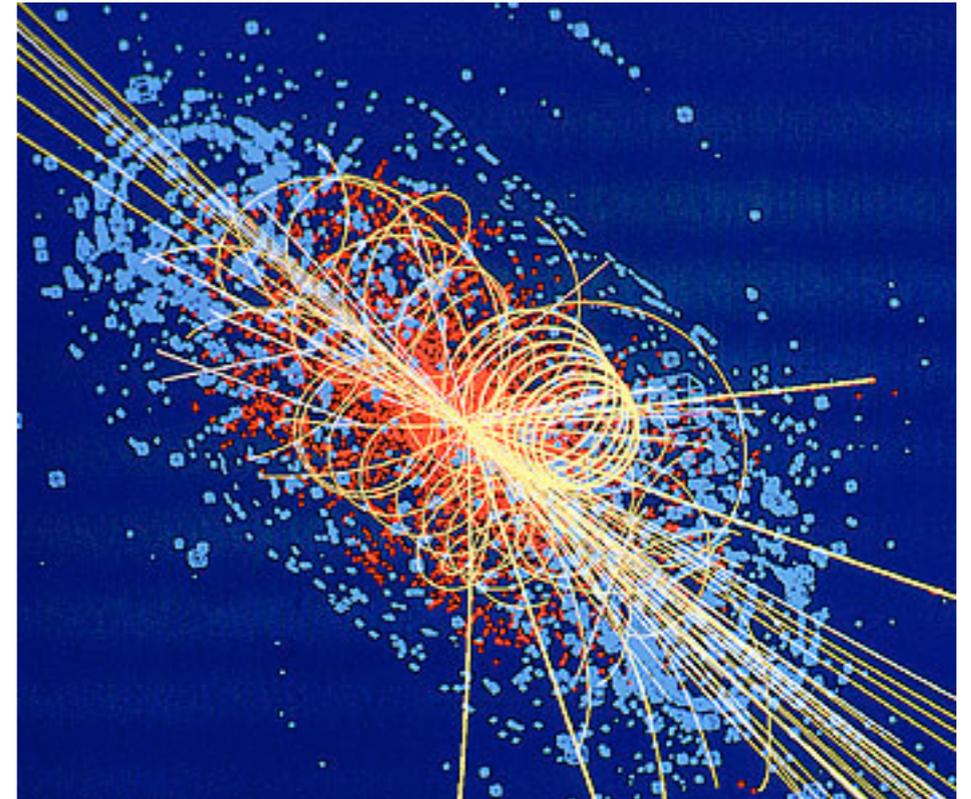


■ Multiplayer online games



Problemi scientifici di grandi dimensioni

- **Calcolo per esperimenti di fisica**
- Oggi la maggior parte dei dati sperimentali prodotti dai rivelatori consistono in collezioni di **immagini digitali** con informazioni spazio-temporali
- Tali immagini o “eventi” indipendenti si prestano bene ad una **elaborazione parallela** debolmente accoppiata
- **Grandi moli di dati** stoccati in modalità per lo più a sola lettura
- **Scalabilità** con l’aumento del numero dei processori coinvolti nell’elaborazione



Calcolo per un esperimento

- Nel campo della fisica sperimentale, largo uso di risorse di calcolo per:
 - **Acquisizione dati**
 - **Simulazione** con tecniche Monte Carlo di apparati sperimentali
 - In quantità spesso > dati sperimentali
 - **Ricostruzione e selezione** degli eventi prodotti da apparati sperimentali e da simulazioni
 - **Analisi** finale delle interazioni elementari



Il Large Hadron Collider (LHC)

■ Numeri

- Collisore protone-protone da **7+7 TeV**
(costruzione approvata nel 1995)
- **27 km** tunnel
 - dipoli superconduttori
 - nuova scala di energia e di risoluzione
(10^{-20} m)
- Fiotti di protoni che si scontrano ogni
25 ns
- Costo circa **3.8 G€**
 - 1 Km di alta velocità: 62 M€, TAV MI-
TO (125 Km): 7.8 G€
 - 1 anno di Formula 1 Toyota 2005: 500
M€ (più di un grande esperimento al
CERN)
 - Costo Parlamento italiano 2009: circa
2.4 G€ (3.5 x budget del CERN)
- Prime collisioni registrate nel
dicembre 2009

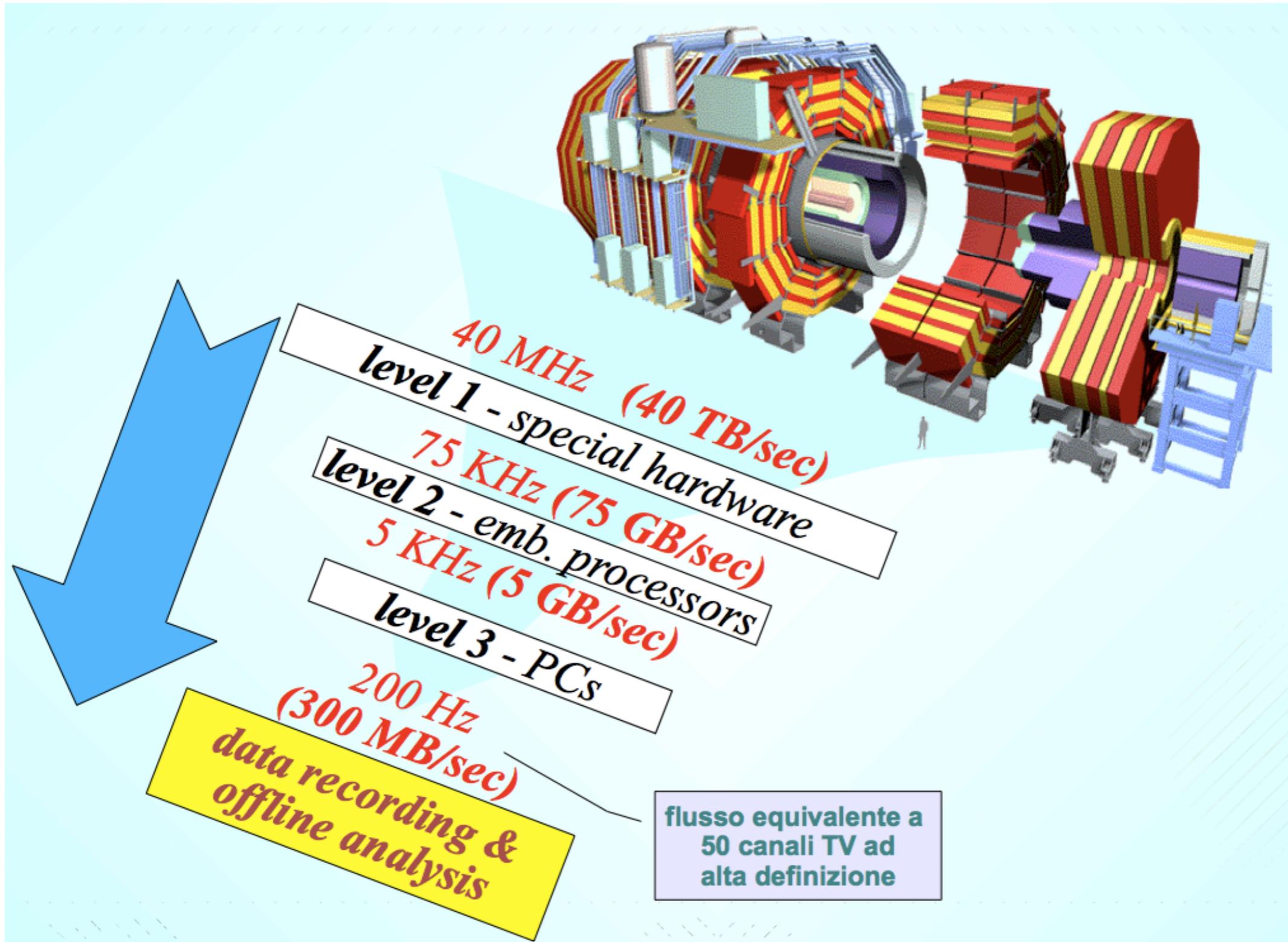


(intermezzo)

- Incidente a LHC, 19 settembre 2008
 - Problema a una connessione elettrica tra due magneti (LHC ne ha 9593) durante un test a 5 TeV

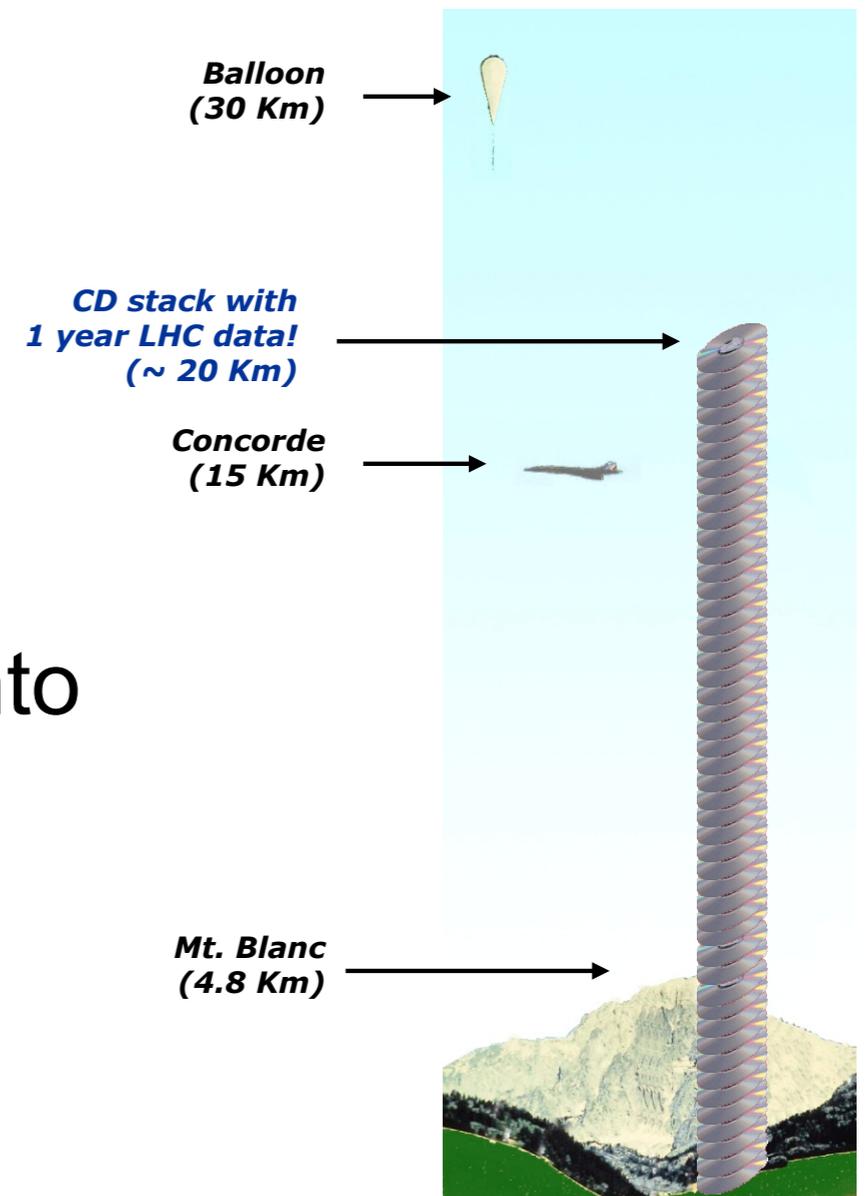


Il flusso dei dati in LHC

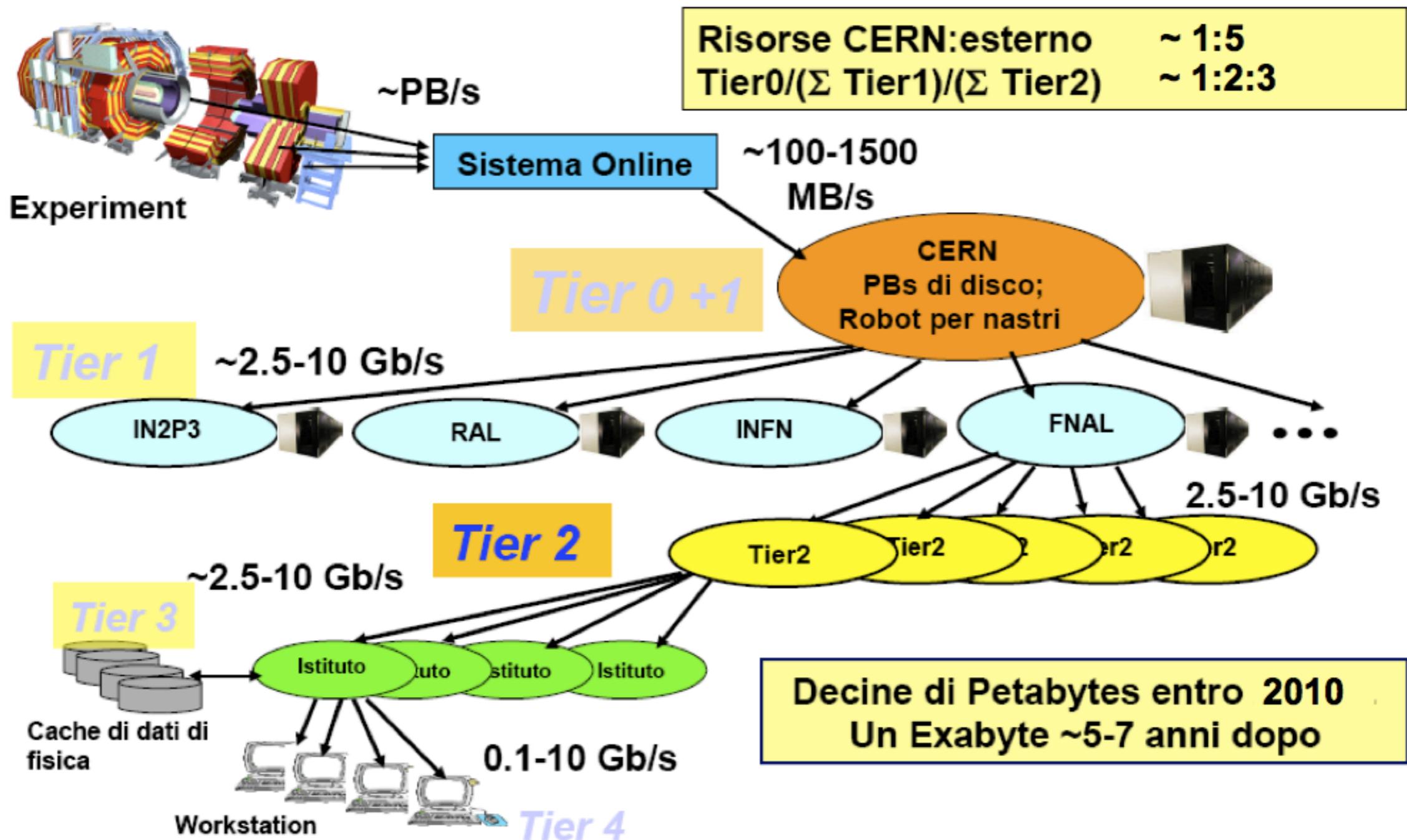


Le dimensioni del problema

- **Quantità di dati** accumulata da ogni esperimento ad LHC:
 - ~ 1-2 MB / evento
 - ~ 0.1-1 GB/s
 - durata presa dati ~ 10^7 s / anno
 - totale: ~ 1-4 PB / anno per esperimento
(1 PB = 1 Petabyte = 1 milione di Gigabyte)
- **Risorse** necessarie:
 - stoccaggio dati: **10-15 PB / anno**
 - CPU: ~ **10^5 processori** attuali

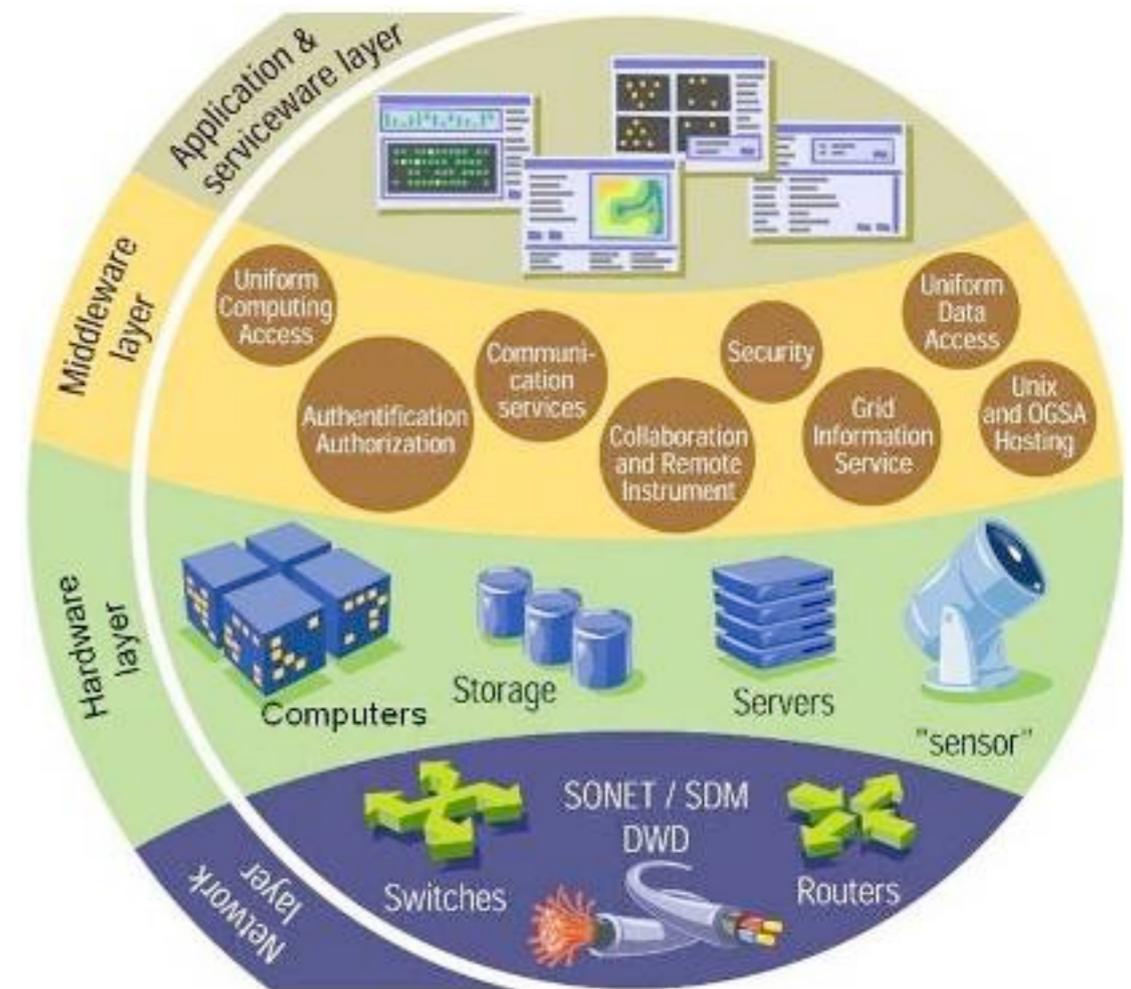


Soluzione distribuita del problema del calcolo a LHC



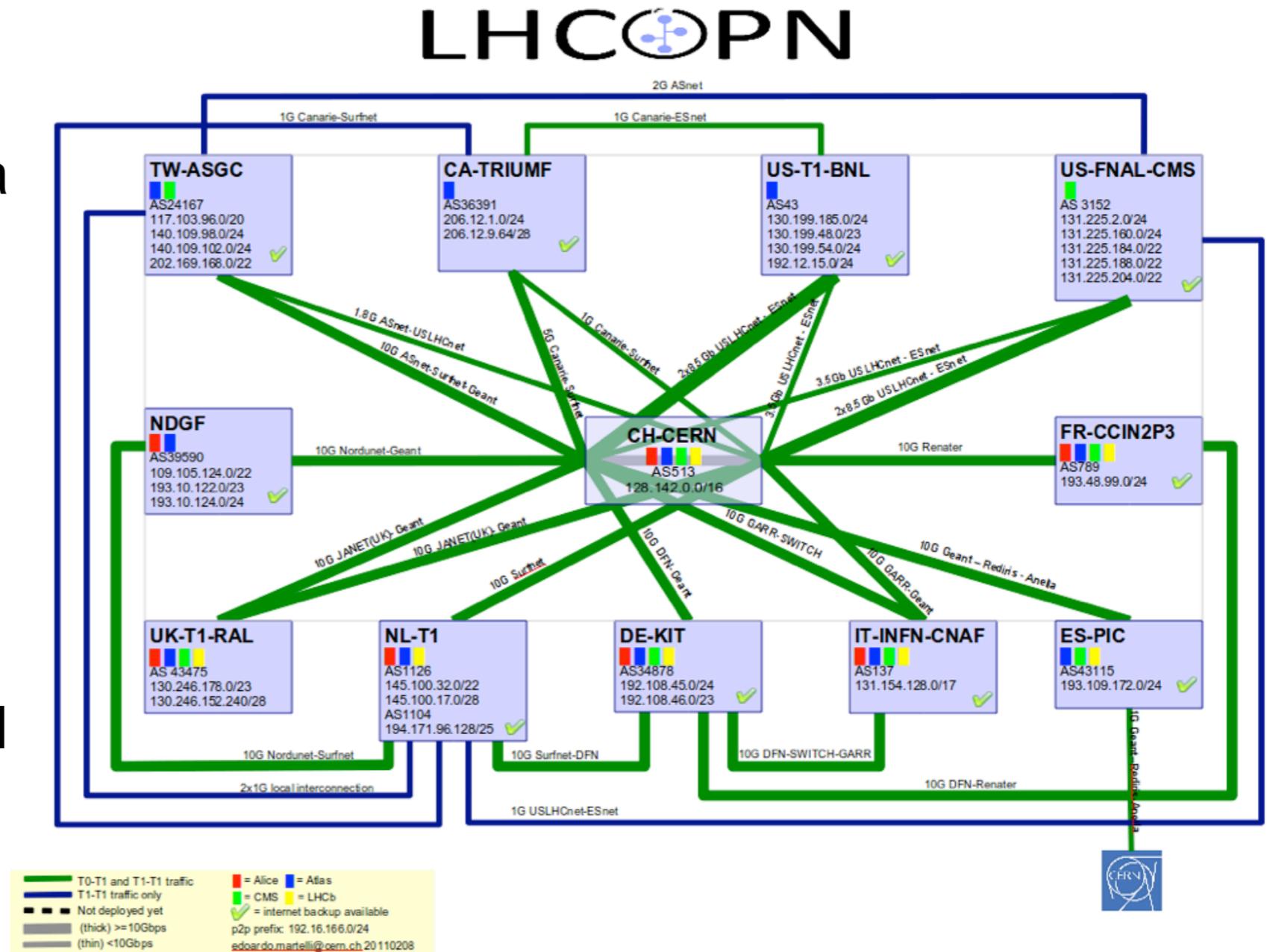
Sì, ma come?

- Non esistono (*in pratica*) tool o servizi commerciali in grado di soddisfare le richieste computazionali di LHC
- 4 layer “tradizionali”:
 - Networking (data transfer)
 - Hardware (computer, storage)
 - “Middleware”
 - Software di fisica
- Aggiungeremo più avanti la dimensione “virtualizzazione”



Data Transfer

- La rete “Internet” standard non è adatta ai trasferimenti di dati perché troppo limitata come banda
- I maggiori centri (“Tier-1”) partecipanti a LHC si sono dotati di una **rete dedicata** chiamata LHC Optical Private Network, basata su link multipli a 10 Gbit/s



Ma come interconnettere i dati, le risorse, le persone, il software?

- **Grid Computing**: negli ultimi 10 anni la comunità internazionale della fisica ha contribuito sostanzialmente alla scrittura di “middleware”, definendo *de facto* il Grid Computing, che rendesse possibile:
 - **condivisione** in modo efficiente e sicuro di risorse di calcolo appartenente a istituzioni diverse in tutto il mondo
 - **sottomissione di elaborazione dati** attraverso un sistema unico
 - **sicurezza dell'accesso e tracciabilità** attraverso l'uso di certificati
 - **gestione su catalogo** dei dati e loro replica

Funziona?



- Google Earth:
<http://dashb-earth.cern.ch/dashboard/dashb-earth-all.kmz>
- Mappa “live” del funzionamento della Computing Grid di LHC
 - 20/4/2011 ore 00:20
 - A quell’ora, circa **115.000 job in esecuzione**
 - Media di circa **1.000.000 di job eseguiti al giorno**
- LHC è in funzione 24h/24, con una previsione di **presa dati ininterrotta fino alla fine del 2012**

In Italia...

- INFN partecipa a tutti i **quattro grandi esperimenti** a LHC
- Nonostante la partecipazione italiana nei vari esperimenti sia inferiore al 20%, **TUTTE le collaborazioni** sono guidate oggi da colleghi italiani
- L'attuale **Direttore della Divisione Ricerca e Calcolo del CERN** è Sergio Bertolucci, ex Direttore di uno dei Laboratori Nazionali dell'INFN

Calcolo INFN per LHC (e non solo)

- **1 centro Tier-1** al CNAF di Bologna
 - **4 esperimenti LHC** ma anche calcolo e storage per altri 16 esperimenti internazionali, tra cui:
 - **Virgo**: interferometro gravitazionale
 - **Fermi**: telescopio per raggi gamma in orbita intorno alla terra
 - **AMS**: (vedi sopra)
 - **CDF, BaBar**: esperimenti di fisica delle particelle basati negli USA
 - **Auger**: osservatorio per raggi cosmici in Argentina distribuito su una superficie di 3000 km²
- **10 centri Tier-2**



Il Tier-1 di Bologna

- Presso INFN-CNAF
 - 60 persone (23 permanenti)
- Circa 8.000 “slot” di calcolo (“CPU cores”), upgrade a 10.000 in corso
 - Completamente occupati 24h/24, con 6-8.000 job sempre “in coda”
- 8 PB di spazio disco, 10 PB di spazio su nastro con libreria robotizzata
- Sala calcolo di 1200 m²
- Potenza max dissipabile 1.4 MW
 - Consumo massimo previsto 2.6 MW (come una cittadina di 25.000 persone!)
 - 2+1 trasformatori elettrici da 2.5 MVA ciascuno, 2 UPS dinamici, gruppo di continuità, 5+2 chiller per il raffreddamento



(intermezzo)



Tornando alle infrastrutture distribuite...

- Una definizione generale: una **infrastruttura di calcolo distribuita** tipicamente deve:
 - fornire soluzioni per la **scoperta, l'uso e il controllo delle risorse**
 - onorare dei **Service Level Agreements** (contratti sul livello del servizio)
 - assicurarsi che vengano prese opportune **misure di sicurezza**, ad esempio per quanto riguarda la **autenticazione** (chi sei?) e la **autorizzazione** (che cosa puoi fare?)
- Questo è quanto negli ultimi 10 anni il Grid Computing ha cercato di implementare con grandi progetti come il calcolo per LHC
- Sondaggio: chi ha sentito parlare di **Cloud Computing**?

Cloud Computing in una frase (lunga)

- Un *modello* per un **accesso ubiquo, facile, on-demand, remoto** (che utilizzi la rete)...
- ... a un insieme di risorse condivise che possono essere **configurabili dall'utente**...
- ... e che possono essere **rapidamente fornite e rilasciate** con uno sforzo minimo da parte dell'utente o del fornitore di servizi.



(NIST, Working Definition of Cloud Computing)

Grid Computing, dal punto di vista dell'utente

■ Come:

- Occorre fare parte di una **Virtual Organization**. Se questa non esiste, occorre crearla.
- Alla “Grid” si accede attraverso una **User Interface**, autenticandosi con un **certificato digitale**.
- **Le richieste dei miei “job”** di calcolo si specificano attraverso un Job Description Language.
- Le richieste vengono **validate sulle risorse disponibili** nella Grid. Se esistono risorse compatibili alle richieste, prima o poi il mio job andrà in esecuzione da qualche parte.
- La Grid mi dà la possibilità di **controllare** lo stato e **ottenere** l'output dei miei job, di trovare, recuperare e salvare dati.

■ Architettura:

- Enfasi sulla **condivisione delle risorse** a un livello di organizzazione (virtuale).
- Sostanzialmente adottata da **comunità scientifiche**. Pochissimo interesse da parte dell'industria.
- Tipicamente pensata per l'esecuzione di **job non interattivi** (“batch”) e non dinamici.

Cloud Computing, dal punto di vista dell'utente

■ Come:

- Devo identificare un **Service Provider**.
 - Ad esempio, Amazon (Elastic Compute Cloud, o EC2), Google, GoGrid, Microsoft
- Posso allocare una (apparente) **quantità infinita di risorse**, tipicamente attraverso una interfaccia Web.
- Il modello di **pagamento è a consumo** (“pay-as-you-go”).

■ Architettura:

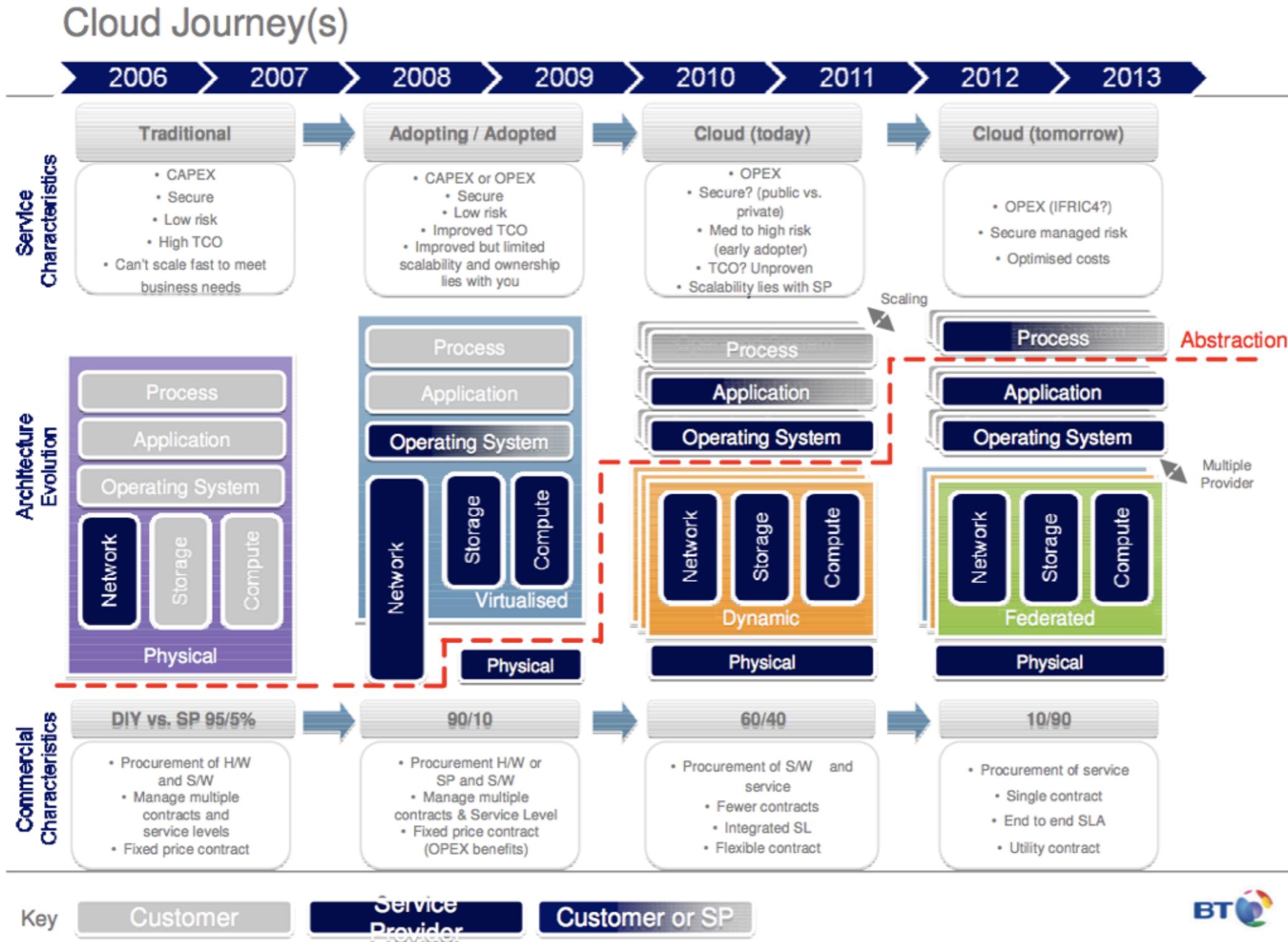
- Enfasi sulla **facilità di accesso** alle risorse per *utenti individuali*.
- Iniziativa partita dal **settore privato**, con ottima accoglienza (in alcuni settori).
- **Diversi livelli di astrazione** sono possibili: posso vedere le risorse come “Infrastructure as a Service” (chiedo una intera infrastruttura), come “Platform as a Service” (ad esempio un ambiente di sviluppo, 1-10-100 core o altro), etc.

Schematizzando

- Cloud: **fornitura di risorse**
- Grid: **federazione di risorse**

- Osservazioni:
 - Il calcolo scientifico (specie se distribuito internazionalmente) è normalmente basato sulla **collaborazione e sulla condivisione** di risorse umane, tecnologiche, di metodi, dati e risultati.
 - Qualunque modello di calcolo scientifico distribuito dovrebbe dunque **evitare modelli proprietari di software e di accesso alle risorse, nonché limitazioni di proprietà intellettuale.**

L'evoluzione delle Cloud



Tuttavia c'è Cloud e Cloud

- Posso personalizzare l'ambiente!
- Accesso on-demand!
- Risorse infinite!
- Gestione dei costi e della crescita ottimale!
- Meno CO₂!



- Ma devo installarmi da solo il sistema operativo?
- Quanto ci vuole ad avere 5000 nodi? Quanto tempo ci metteranno i miei calcoli?
- Ma queste risorse ogni tanto spariscono, il servizio non è affidabile!
- Costa troppo! E se Amazon poi alza i prezzi?
- In quale continente si trovano i miei dati? Chi vi ha accesso?

Alcuni problemi del Cloud Computing

■ Sicurezza

- Diversi report indicano come **controllo e fiducia** siano i maggiori ostacoli all'adozione di tecnologie cloud.

■ Autorità

- Regolamentazione legale**, ad esempio della privacy.

■ Performance

- Quali sono gli **impatti della virtualizzazione** delle risorse?
- Ci sono **livelli di servizio garantiti**? (SLA)

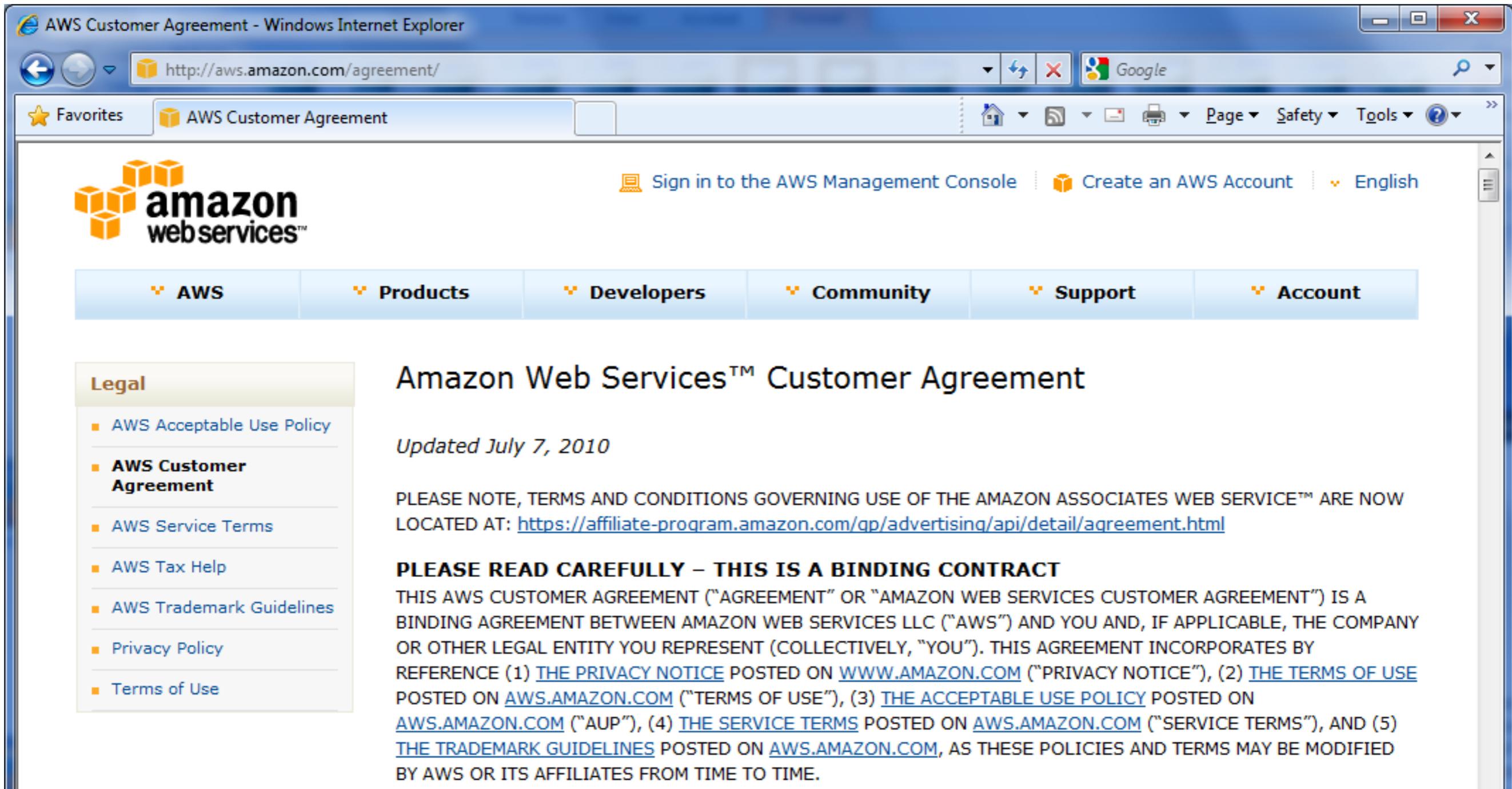
■ Portabilità delle applicazioni

- Cosa serve perché la mia **applicazione** possa funzionare “in the cloud”?
- Problema di legarsi a un **unico fornitore**?

■ Costo

Una veloce occhiata ad alcune considerazioni di sicurezza

Una veloce occhiata ad alcune considerazioni di sicurezza

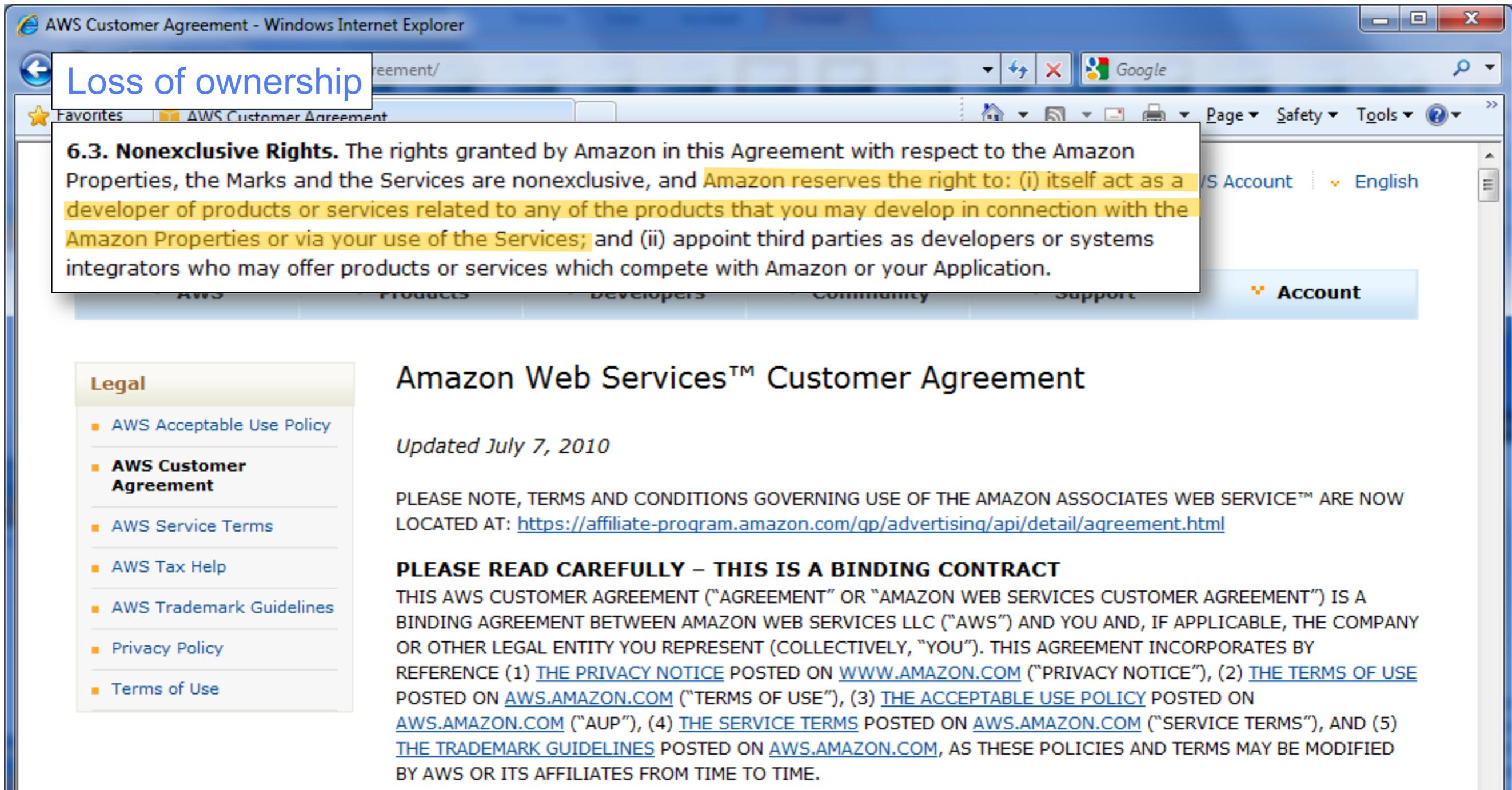


The screenshot shows a Windows Internet Explorer browser window displaying the AWS Customer Agreement page. The address bar shows the URL <http://aws.amazon.com/agreement/>. The page features the Amazon Web Services logo and navigation links for 'Sign in to the AWS Management Console', 'Create an AWS Account', and 'English'. A horizontal menu contains links for 'AWS', 'Products', 'Developers', 'Community', 'Support', and 'Account'. On the left, a 'Legal' sidebar lists various policies, with 'AWS Customer Agreement' highlighted. The main content area is titled 'Amazon Web Services™ Customer Agreement' and includes a date update of 'Updated July 7, 2010'. A prominent warning states: 'PLEASE READ CAREFULLY – THIS IS A BINDING CONTRACT'. The text explains that the agreement is a binding contract between AWS and the user, incorporating several other policies: (1) The Privacy Notice, (2) The Terms of Use, (3) The Acceptable Use Policy, (4) The Service Terms, and (5) The Trademark Guidelines, all posted on aws.amazon.com. It notes that these policies and terms may be modified by AWS or its affiliates over time.

S. Bradshaw, C. Millard, I. Walden

“Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services”

Una veloce occhiata ad alcune considerazioni di sicurezza



Loss of ownership

6.3. Nonexclusive Rights. The rights granted by Amazon in this Agreement with respect to the Amazon Properties, the Marks and the Services are nonexclusive, and Amazon reserves the right to: (i) itself act as a developer of products or services related to any of the products that you may develop in connection with the Amazon Properties or via your use of the Services; and (ii) appoint third parties as developers or systems integrators who may offer products or services which compete with Amazon or your Application.

Legal

- AWS Acceptable Use Policy
- **AWS Customer Agreement**
- AWS Service Terms
- AWS Tax Help
- AWS Trademark Guidelines
- Privacy Policy
- Terms of Use

Amazon Web Services™ Customer Agreement

Updated July 7, 2010

PLEASE NOTE, TERMS AND CONDITIONS GOVERNING USE OF THE AMAZON ASSOCIATES WEB SERVICE™ ARE NOW LOCATED AT: <https://affiliate-program.amazon.com/gp/advertising/api/detail/agreement.html>

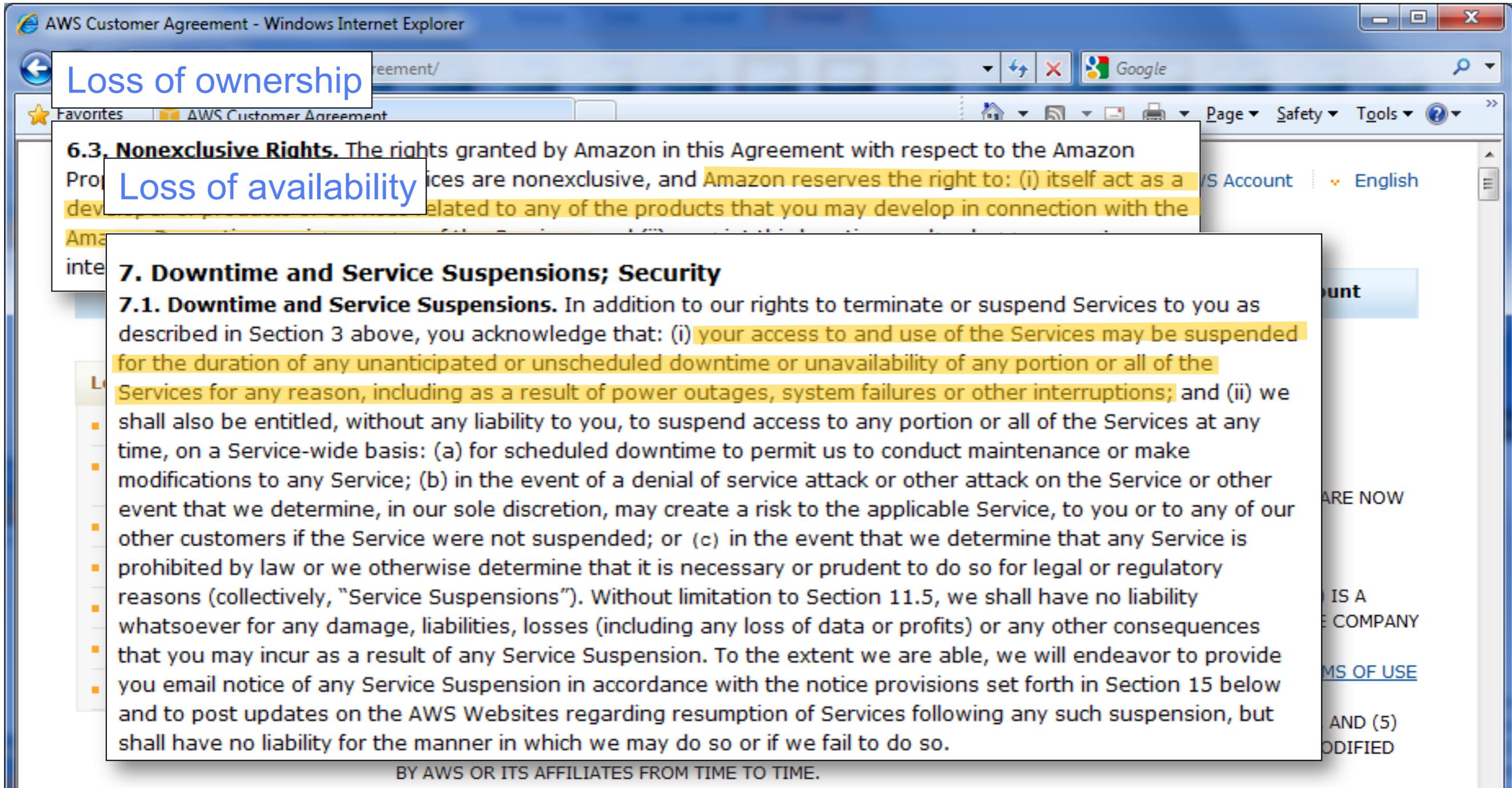
PLEASE READ CAREFULLY – THIS IS A BINDING CONTRACT

THIS AWS CUSTOMER AGREEMENT ("AGREEMENT" OR "AMAZON WEB SERVICES CUSTOMER AGREEMENT") IS A BINDING AGREEMENT BETWEEN AMAZON WEB SERVICES LLC ("AWS") AND YOU AND, IF APPLICABLE, THE COMPANY OR OTHER LEGAL ENTITY YOU REPRESENT (COLLECTIVELY, "YOU"). THIS AGREEMENT INCORPORATES BY REFERENCE (1) [THE PRIVACY NOTICE](#) POSTED ON WWW.AMAZON.COM ("PRIVACY NOTICE"), (2) [THE TERMS OF USE](#) POSTED ON AWS.AMAZON.COM ("TERMS OF USE"), (3) [THE ACCEPTABLE USE POLICY](#) POSTED ON AWS.AMAZON.COM ("AUP"), (4) [THE SERVICE TERMS](#) POSTED ON AWS.AMAZON.COM ("SERVICE TERMS"), AND (5) [THE TRADEMARK GUIDELINES](#) POSTED ON AWS.AMAZON.COM, AS THESE POLICIES AND TERMS MAY BE MODIFIED BY AWS OR ITS AFFILIATES FROM TIME TO TIME.

S. Bradshaw, C. Millard, I. Walden

"Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services"

Una veloce occhiata ad alcune considerazioni di sicurezza



The screenshot shows a Windows Internet Explorer browser window displaying the AWS Customer Agreement. The address bar shows 'AWS Customer Agreement - Windows Internet Explorer'. The page content includes several sections with callouts:

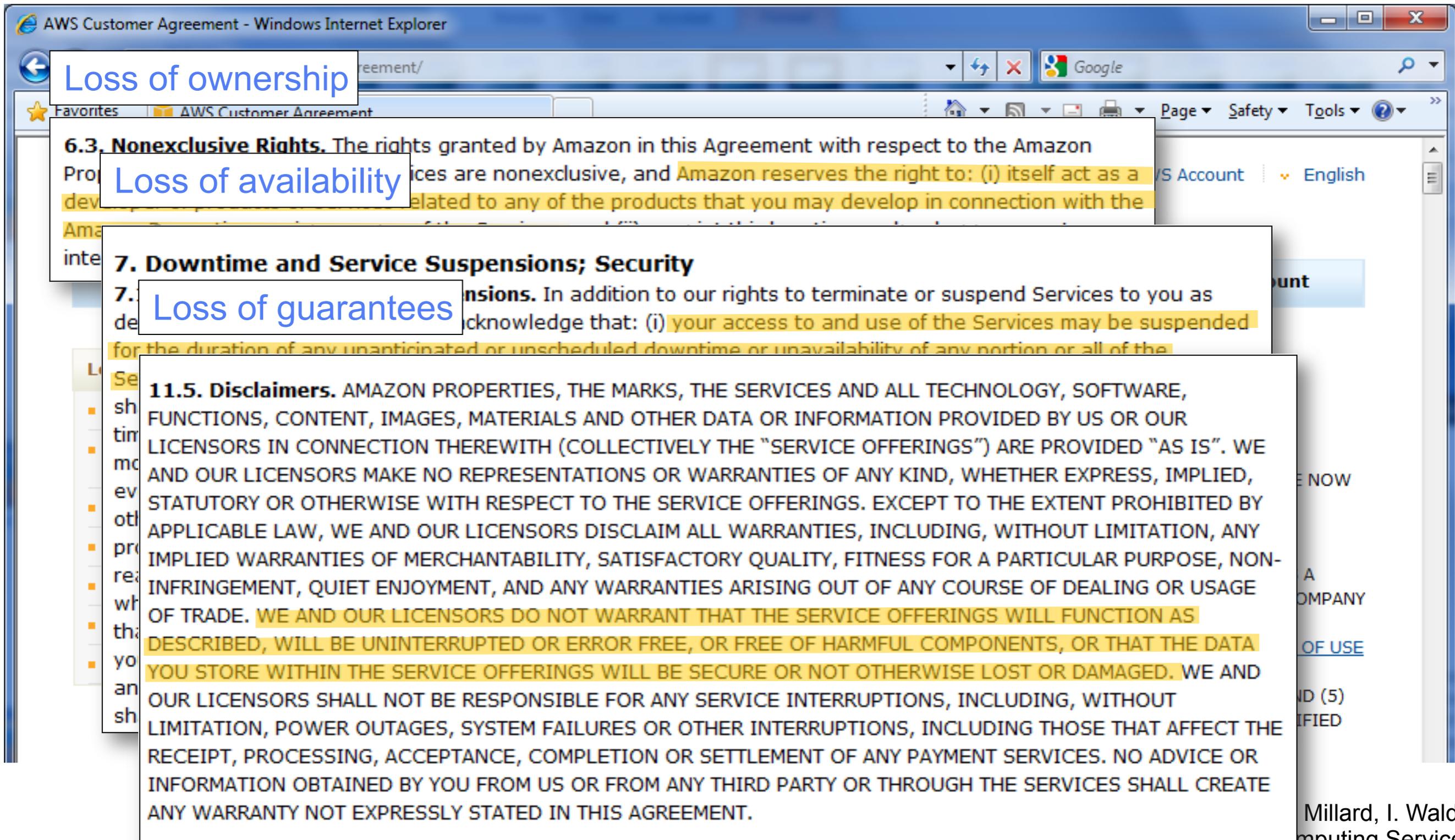
- Loss of ownership:** A callout box pointing to the text 'Amazon reserves the right to: (i) itself act as a provider of services related to any of the products that you may develop in connection with the Amazon Services'.
- Loss of availability:** A callout box pointing to the text 'your access to and use of the Services may be suspended for the duration of any unanticipated or unscheduled downtime or unavailability of any portion or all of the Services for any reason, including as a result of power outages, system failures or other interruptions'.
- 7. Downtime and Service Suspensions; Security:** A large callout box containing the following text:

7.1. Downtime and Service Suspensions. In addition to our rights to terminate or suspend Services to you as described in Section 3 above, you acknowledge that: (i) your access to and use of the Services may be suspended for the duration of any unanticipated or unscheduled downtime or unavailability of any portion or all of the Services for any reason, including as a result of power outages, system failures or other interruptions; and (ii) we shall also be entitled, without any liability to you, to suspend access to any portion or all of the Services at any time, on a Service-wide basis: (a) for scheduled downtime to permit us to conduct maintenance or make modifications to any Service; (b) in the event of a denial of service attack or other attack on the Service or other event that we determine, in our sole discretion, may create a risk to the applicable Service, to you or to any of our other customers if the Service were not suspended; or (c) in the event that we determine that any Service is prohibited by law or we otherwise determine that it is necessary or prudent to do so for legal or regulatory reasons (collectively, "Service Suspensions"). Without limitation to Section 11.5, we shall have no liability whatsoever for any damage, liabilities, losses (including any loss of data or profits) or any other consequences that you may incur as a result of any Service Suspension. To the extent we are able, we will endeavor to provide you email notice of any Service Suspension in accordance with the notice provisions set forth in Section 15 below and to post updates on the AWS Websites regarding resumption of Services following any such suspension, but shall have no liability for the manner in which we may do so or if we fail to do so.

S. Bradshaw, C. Millard, I. Walden

"Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services"

Una veloce occhiata ad alcune considerazioni di sicurezza



Loss of ownership

6.3. Nonexclusive Rights. The rights granted by Amazon in this Agreement with respect to the Amazon Properties are nonexclusive, and Amazon reserves the right to: (i) itself act as a developer of products related to any of the products that you may develop in connection with the Amazon Properties.

Loss of availability

7. Downtime and Service Suspensions; Security

7. Service Suspensions. In addition to our rights to terminate or suspend Services to you as described in this Agreement, you acknowledge that: (i) your access to and use of the Services may be suspended for the duration of any unanticipated or unscheduled downtime or unavailability of any portion or all of the Amazon Properties.

Loss of guarantees

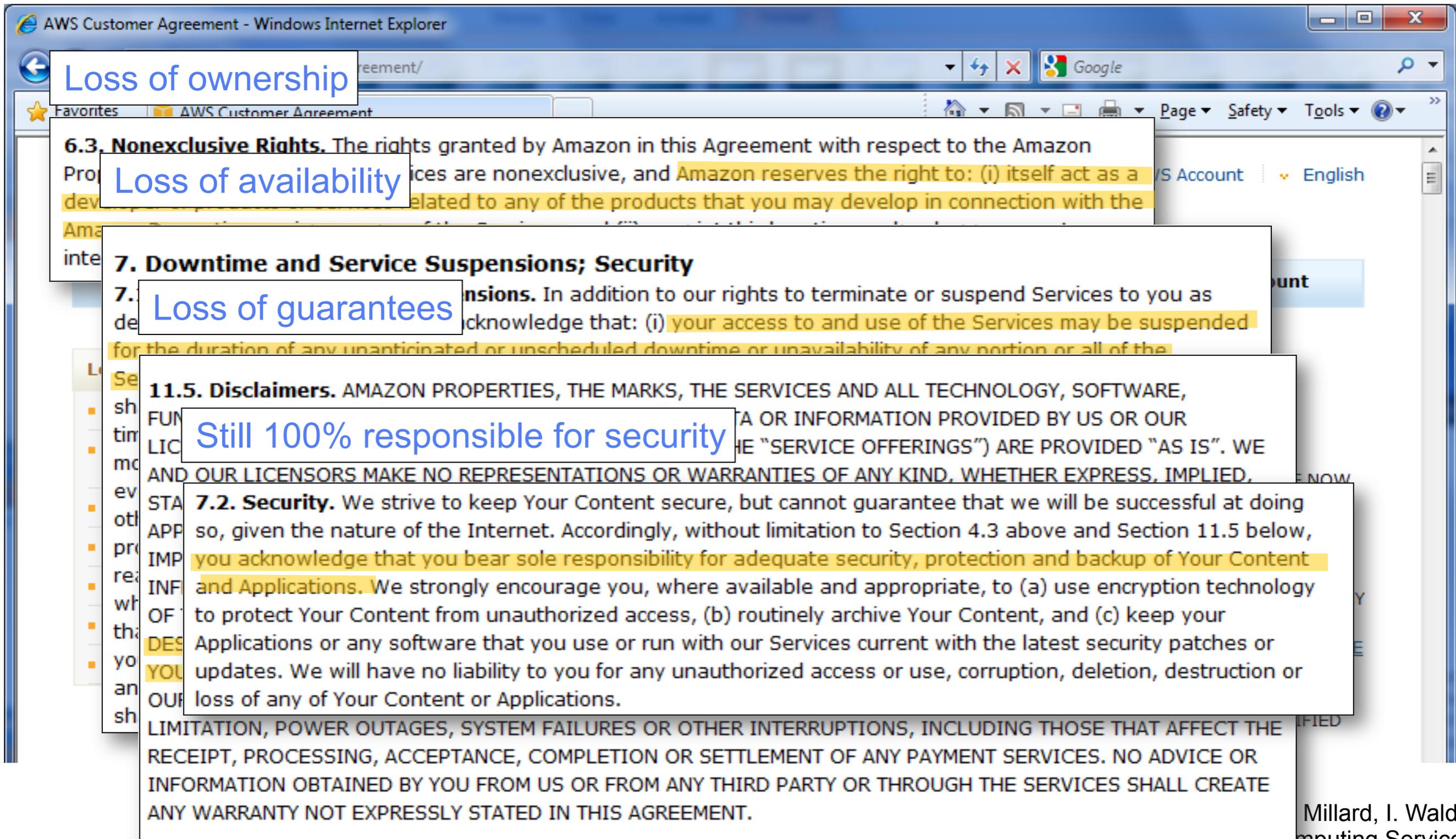
11.5. Disclaimers. AMAZON PROPERTIES, THE MARKS, THE SERVICES AND ALL TECHNOLOGY, SOFTWARE, FUNCTIONS, CONTENT, IMAGES, MATERIALS AND OTHER DATA OR INFORMATION PROVIDED BY US OR OUR LICENSORS IN CONNECTION THEREWITH (COLLECTIVELY THE "SERVICE OFFERINGS") ARE PROVIDED "AS IS". WE AND OUR LICENSORS MAKE NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, WHETHER EXPRESS, IMPLIED, STATUTORY OR OTHERWISE WITH RESPECT TO THE SERVICE OFFERINGS. EXCEPT TO THE EXTENT PROHIBITED BY APPLICABLE LAW, WE AND OUR LICENSORS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, NON-INFRINGEMENT, QUIET ENJOYMENT, AND ANY WARRANTIES ARISING OUT OF ANY COURSE OF DEALING OR USAGE OF TRADE. WE AND OUR LICENSORS DO NOT WARRANT THAT THE SERVICE OFFERINGS WILL FUNCTION AS DESCRIBED, WILL BE UNINTERRUPTED OR ERROR FREE, OR FREE OF HARMFUL COMPONENTS, OR THAT THE DATA YOU STORE WITHIN THE SERVICE OFFERINGS WILL BE SECURE OR NOT OTHERWISE LOST OR DAMAGED. WE AND OUR LICENSORS SHALL NOT BE RESPONSIBLE FOR ANY SERVICE INTERRUPTIONS, INCLUDING, WITHOUT LIMITATION, POWER OUTAGES, SYSTEM FAILURES OR OTHER INTERRUPTIONS, INCLUDING THOSE THAT AFFECT THE RECEIPT, PROCESSING, ACCEPTANCE, COMPLETION OR SETTLEMENT OF ANY PAYMENT SERVICES. NO ADVICE OR INFORMATION OBTAINED BY YOU FROM US OR FROM ANY THIRD PARTY OR THROUGH THE SERVICES SHALL CREATE ANY WARRANTY NOT EXPRESSLY STATED IN THIS AGREEMENT.

Millard, I. Walden

Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services"

31

Una veloce occhiata ad alcune considerazioni di sicurezza



Loss of ownership

6.3. Nonexclusive Rights. The rights granted by Amazon in this Agreement with respect to the Amazon Properties are nonexclusive, and Amazon reserves the right to: (i) itself act as a developer of products related to any of the products that you may develop in connection with the Amazon Properties.

Loss of availability

7. Downtime and Service Suspensions; Security

7.1. Service Suspensions. In addition to our rights to terminate or suspend Services to you as set forth in Section 4.3, you acknowledge that: (i) your access to and use of the Services may be suspended for the duration of any unanticipated or unscheduled downtime or unavailability of any portion or all of the Amazon Properties.

Loss of guarantees

11.5. Disclaimers. AMAZON PROPERTIES, THE MARKS, THE SERVICES AND ALL TECHNOLOGY, SOFTWARE, INFORMATION AND INFORMATION PROVIDED BY US OR OUR LICENSORS (THE "SERVICE OFFERINGS") ARE PROVIDED "AS IS". WE AND OUR LICENSORS MAKE NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, WHETHER EXPRESS, IMPLIED, OR STATUTORY.

Still 100% responsible for security

7.2. Security. We strive to keep Your Content secure, but cannot guarantee that we will be successful at doing so, given the nature of the Internet. Accordingly, without limitation to Section 4.3 above and Section 11.5 below, you acknowledge that you bear sole responsibility for adequate security, protection and backup of Your Content and Applications. We strongly encourage you, where available and appropriate, to (a) use encryption technology to protect Your Content from unauthorized access, (b) routinely archive Your Content, and (c) keep your Applications or any software that you use or run with our Services current with the latest security patches or updates. We will have no liability to you for any unauthorized access or use, corruption, deletion, destruction or loss of any of Your Content or Applications.

LIMITATION, POWER OUTAGES, SYSTEM FAILURES OR OTHER INTERRUPTIONS, INCLUDING THOSE THAT AFFECT THE RECEIPT, PROCESSING, ACCEPTANCE, COMPLETION OR SETTLEMENT OF ANY PAYMENT SERVICES. NO ADVICE OR INFORMATION OBTAINED BY YOU FROM US OR FROM ANY THIRD PARTY OR THROUGH THE SERVICES SHALL CREATE ANY WARRANTY NOT EXPRESSLY STATED IN THIS AGREEMENT.

Millard, I. Walden

Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services"

Una veloce occhiata ad alcune considerazioni di sicurezza

Loss of ownership

Loss of availability

Loss of guarantees

Still 100% responsible

No control, full responsibility!

6.3. **Nonexclusive Rights.** The rights granted by Amazon in this Agreement with respect to the use of the Services are nonexclusive, and Amazon reserves all other rights not expressly granted to you. Amazon is not responsible for any loss of data or other information that you store or transmit using the Services, including any loss of data or other information that you store or transmit using the Services that is caused by any of the products that you use with the Services.

7. **Downtime and Service Suspensions; Security.** In addition to the limitations on availability of the Services set forth in Section 4.3, you acknowledge that the Services may be suspended or interrupted for the duration of any unanticipated or extraordinary circumstances, including but not limited to, natural disasters, war, terrorism, sabotage, government actions, strikes, labor disputes, fire, flood, power outages, telecommunications failures, and other events beyond Amazon's control. Amazon is not responsible for any loss of data or other information that you store or transmit using the Services, including any loss of data or other information that you store or transmit using the Services that is caused by any of the products that you use with the Services.

11.5. **Disclaimers.** AMAZON PROVIDES THE SERVICES, INCLUDING ANY TECHNOLOGY, SOFTWARE, OR CONTENT PROVIDED BY US OR OUR AFFILIATES ("SERVICES") ARE PROVIDED "AS IS". WE MAKE NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, WHETHER EXPRESS, IMPLIED, OR STATUTORY, INCLUDING BUT NOT LIMITED TO, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT. WE DO NOT GUARANTEE THAT WE WILL BE SUCCESSFUL AT DOING ANY OF THE ABOVE. YOUR USE OF THE SERVICES IS LIMITED TO SECTION 4.3 ABOVE AND SECTION 11.5 BELOW. YOU AGREE TO MAINTAIN ADEQUATE SECURITY, PROTECTION AND BACKUP OF YOUR CONTENT AND DATA. WHERE AVAILABLE AND APPROPRIATE, TO (A) USE ENCRYPTION TECHNOLOGY TO PROTECT YOUR CONTENT AND DATA, (B) ROUTINELY ARCHIVE YOUR CONTENT, AND (C) KEEP YOUR CONTENT AND DATA UP TO DATE OR RUN WITH OUR SERVICES CURRENT WITH THE LATEST SECURITY PATCHES OR UPDATES. YOU AGREE TO HOLD US HARMLESS AND TO INDEMNIFY US FOR ANY UNAUTHORIZED ACCESS OR USE, CORRUPTION, DELETION, DESTRUCTION OR LOSS OF YOUR CONTENT AND DATA, INCLUDING ANY CONTENT AND DATA STORED OR TRANSMITTED THROUGH OUR SERVICES OR APPLICATIONS. WE WILL NOT BE RESPONSIBLE FOR ANY SYSTEM FAILURES OR OTHER INTERRUPTIONS, INCLUDING THOSE THAT AFFECT THE AVAILABILITY, ACCEPTANCE, COMPLETION OR SETTLEMENT OF ANY PAYMENT SERVICES. NO ADVICE OR INFORMATION OBTAINED BY YOU FROM US OR FROM ANY THIRD PARTY OR THROUGH THE SERVICES SHALL CREATE A WARRANTY NOT EXPRESSLY STATED IN THIS AGREEMENT.



Millard, I. Walden

Contracts for Clouds: Comparison and Analysis of the Terms and Conditions for Cloud Computing Services"

Ma ha senso usare il Cloud Computing in ambito scientifico?

- A volte sembra di sì. Ad esempio il **progetto Gaia dell'ESA**, che ha come obiettivo il rilevamento di circa un miliardo di stelle per ottenere una precisa mappa tridimensionale della nostra galassia ha fatto questa analisi:
 - For the full 1 billion star project numbers [the Gaia Science Operations Development Team] calculated that they will analyze 100 million primary stars, plus 6 years of data, which will require a total of 16,200 hours of a 20-node EC2 cluster. That's an **estimated total computing cost of 344,000 Euros**. By comparison, **an in-house solution would cost roughly 720,000 EUR** (at today's prices) – which doesn't include electricity or storage or sys-admin costs. (Storage alone would be an additional 100,000 EUR.)

Però...

- L'esperienza decennale sul Grid Computing da parte di grandi collaborazioni mostra che:
 - **L'interoperabilità di provider** cloud è essenziale, ed è ancora da provare.
 - Fare l'"outsourcing" delle proprie risorse a una ditta esterna non è convincente per **motivi politici, legali o di sicurezza**
 - Come si **interconnettono** i componenti di calcolo, di storage e di rete?
 - Ci sono voluti anni per ottimizzare i calcoli per LHC e in particolare l'interazione tra CPU e storage. E' dubbio che un Cloud computing generico possa ottenere facilmente lo stesso risultato.
 - A meno probabilmente di non fare esplodere i costi.



Tecnicamente, qual è il flusso di lavoro di una analisi scientifica?

- Nella analisi di un esperimento di fisica, normalmente un utente applica i seguenti passi:
 - **quale analisi** voglio compiere per la mia ricerca?
 - **sviluppo il codice** per la mia analisi (normalmente una macro, un plugin di qualche componente software già pronto)
 - Mi serve sapere **di quali dati ho bisogno**:
 - Quali files contengono le informazioni che mi servono? Spesso la risposta è data da un catalogo generale o di esperimento.
 - Chiedo a “qualcosa” di eseguire la mia analisi “da qualche parte”
 - Ad esempio potrei utilizzare la Grid, la Cloud, delle macchine locali, il mio laptop, etc.
 - **Recupero i risultati** del mio lavoro.
- Notate in quanto sopra la interazione molto stretta tra **il mio lavoro e i dati che mi servono**.

Come gestire questo pattern di analisi?

- Dovrei:
 - **Scegliere accuratamente dove mandare i miei job**; in particolare, devono andare in un posto “vicino” al set di dati che mi servono.
 - Utilizzare dei tool per **creare delle repliche locali** dei file di dati che mi servono.
 - Se poi produco qualcosa di interessante, devo poter scrivere i miei risultati nel “repository” da cui ho preso i dati in modo che siano visibili a tutti.
 - Se esagero in quanto sopra, rischio di consumare una quantità notevole di risorse e di tempo.
- Potrei anche cercare di **ottimizzare il più possibile**
 - Ad esempio pre-popolando un sito di dati prima che i job che richiedono quei dati arrivino nel sito.
 - Ogni variazione potrebbe essere sensata e va validata.



Mi fido o no?

- Per implementare quanto sopra, semplificando potrei **scegliere** una di queste 3 opzioni:
 - Lascio che “il sistema distribuito” (essenzialmente la Grid) scopra dove stanno le risorse e decida autonomamente dove mandare i miei job
 - **Conto che il “middleware” Grid sia ottimizzato** per i miei scopi e mi concentro sulla fisica
 - Scelgo autonomamente i centri di calcolo dove mandare i miei job confidando in una selezione semi-automatica e chiedo a questi centri di eseguire i miei job
 - **Mi fido della Grid ma non tanto**; scrivo io del software per rimpiazzare parte del “middleware” Grid.
 - Gestisco tutto io: scelgo i miei resource provider preferiti (magari commerciali), ottengo macchine dedicate (o virtuali, vedi tra poco), controllo direttamente l’esecuzione, pago il conto
 - **Non mi fido della Grid**, sono un sostenitore del Cloud computing puro.
- Attualmente la selezione della ultima opzione (Cloud computing) è ancora marginale nel calcolo scientifico distribuito su larga scala.

Introduzione alla virtualizzazione

- La “**virtualizzazione**”, o *astrazione delle risorse*, **non è particolarmente nuova**
 - L’IBM M44/44X ha introdotto i sistemi a paginazione virtuale e il concetto di macchina virtuale nel 1965.
 - L’emulatore x86 chiamato SoftPC è stato introdotto nel 1988.
 - La prima versione del software open-source XEN è uscita nel 2003.
- Ciò che è cambiato recentemente è il fatto che le tecnologie di virtualizzazione sono ora **pervasive, potenti e (relativamente) facili da usare.**

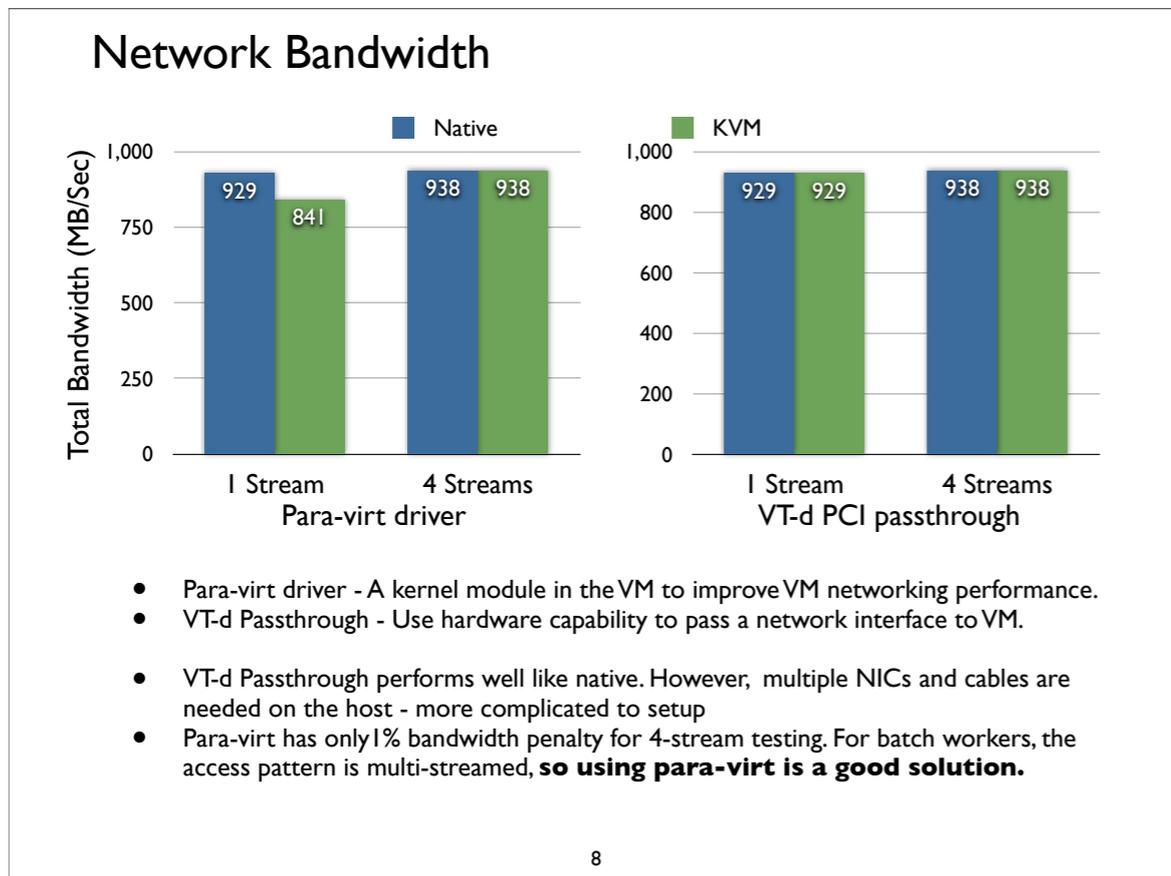
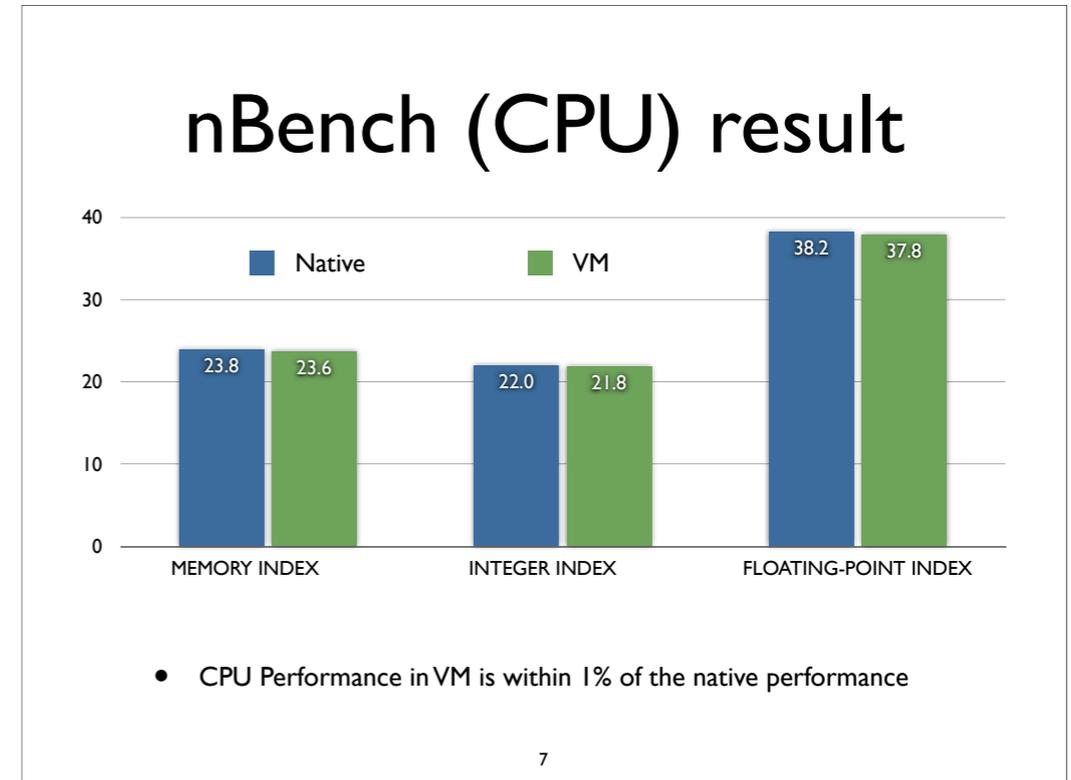


Virtualizzazione: maneggiare con cautela

- La virtualizzazione apre interessanti scenari. Tuttavia:
 - Spesso comporta ancora **penalizzazioni nelle performance**.
 - Queste penalizzazioni possono **non essere affatto trascurabili nel caso di operazioni di input/output** (lettura/scrittura da disco o da rete).
 - Come **virtualizzare i dischi**? (in quantità significative)
 - Relativamente facile con un modello stile Dropbox. E se uno vuole garanzie di latenza? Di affidabilità? Di throughput?
 - Alcune risorse **non sono molto facili da virtualizzare**: ad esempio, le Graphical Processing Unit (GPU).

Performance della virtualizzazione (1)

- Tipicamente qualche % di penalizzazione per la CPU

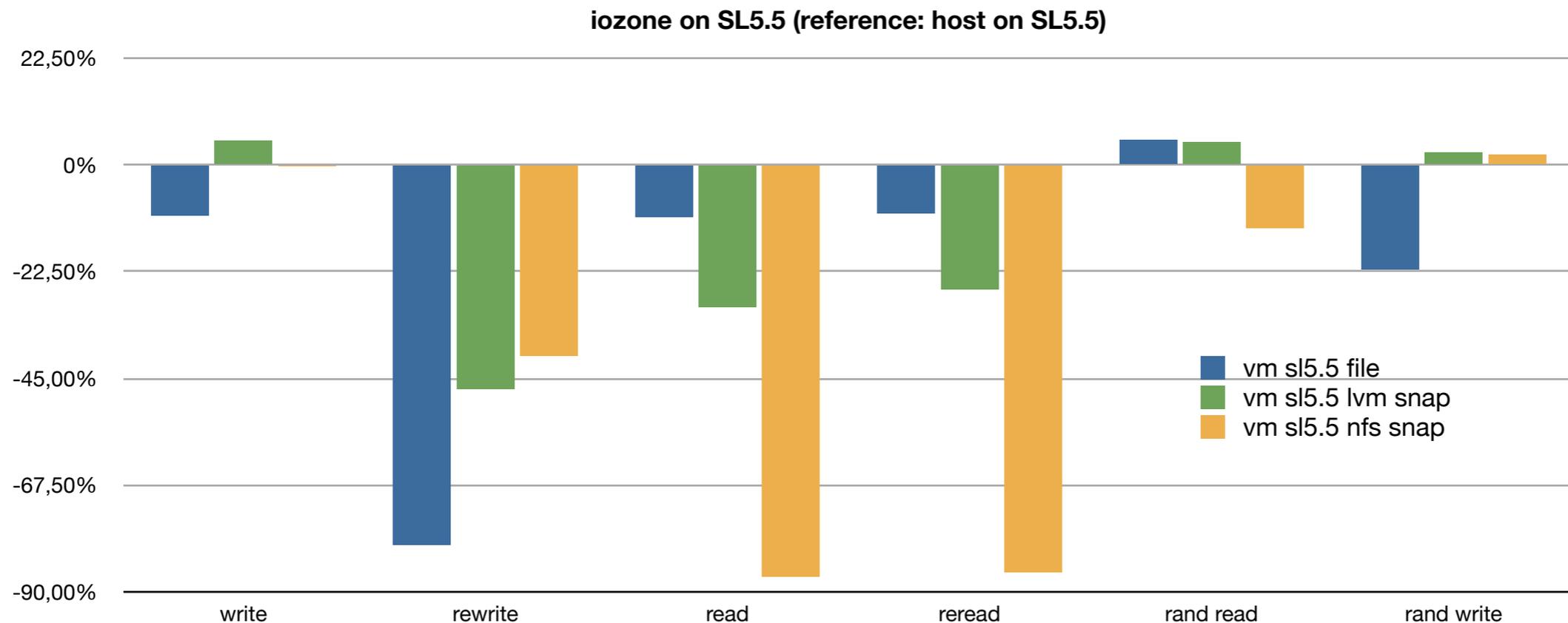


- Anche la rete va abbastanza bene (90% o più della velocità massima teorica)

Cf. Yao et al., J. Phys.Conference Series, (219:032031), 2010

Performance della virtualizzazione (2)

- Molto peggio per quanto riguarda l'input/output di disco
 - 20-30% di penalizzazione rispetto all'hardware non virtualizzato è una misura abbastanza comune



Cf. Salomoni et al., PoS(ISGC 2011 & OGF 31)049, to be published

Come usare la virtualizzazione nel mondo scientifico?

- (nota: di solito il cloud computing si basa fortemente su tecniche di virtualizzazione)
- **Attenzione** ad applicazioni che hanno bisogno di I/O performante
- Possono essere create **macchine virtuali (VM) personalizzate** sulle esigenze dell'utenza
 - Al CNAF abbiamo circa 10 tipi di VM create *ad-hoc*
- Le VM creano un **isolamento tra utenti**
 - Un utente che usa una VM è “protetto” da un altro utente che, sulla stessa macchina fisica, utilizza un'altra VM
 - Le VM **non** migliorano necessariamente la security

Calcolo parallelo su Cloud?

- Esiste una **lacuna significativa** tra l'uso di calcolo parallelo su sistemi dedicati e Amazon EC2
- Cf. E.Walker, Benchmarking Amazon EC2 for high-performance scientific computing, 2008, ;LOGIN: Vol. 33, No. 5
 - [A] performance gap exists between performing HPC computations on a traditional scientific cluster and on an EC2 provisioned scientific cluster. This performance gap is seen not only in the MPI performance of distributed-memory parallel programs but also in the single compute node OpenMP performance for shared-memory parallel programs

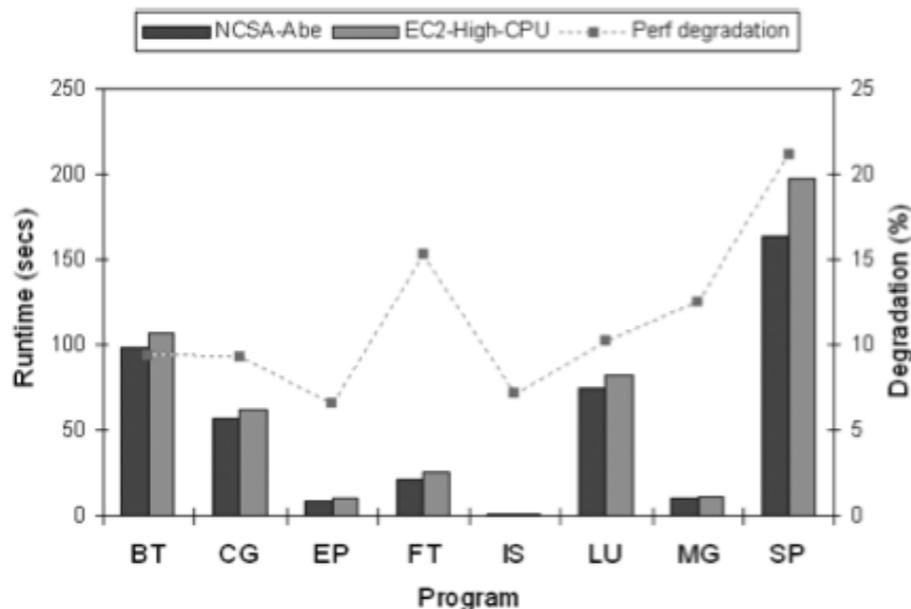


FIGURE 1. NPB-OMP (CLASS B) RUNTIMES ON 8 CPUS ON EC2 AND NCSA CLUSTER COMPUTE NODES. OVERLAID IS THE PERCENTAGE PERFORMANCE DEGRADATION IN THE EC2 RUNS.

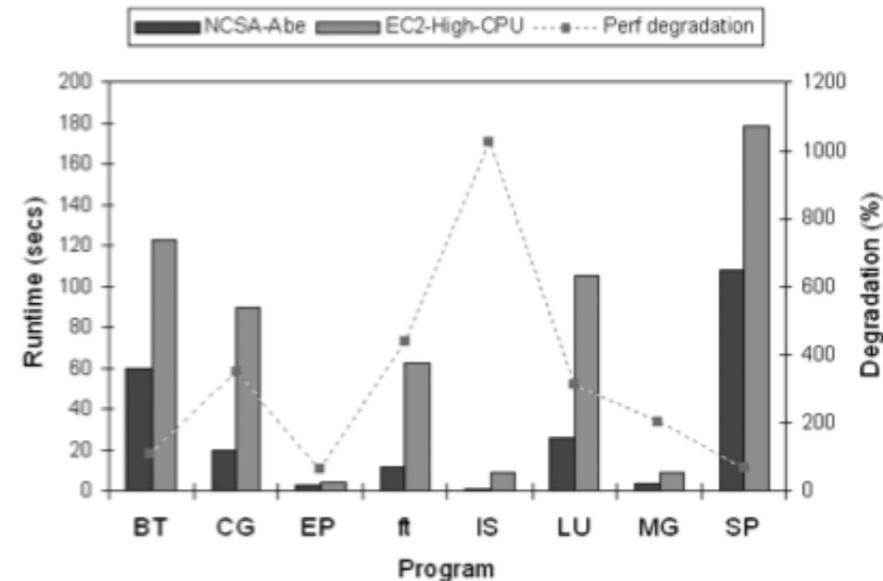


FIGURE 2. NPB-MPI (CLASS B) RUNTIMES ON 32 CPUS ON THE NCSA AND EC2 CLUSTER. BT AND SP WERE RUN WITH 16 CPUS ONLY. OVERLAID IS THE PERCENTAGE DEGRADATION IN THE EC2 RUNS.

Quanto costerebbe?

- Si nota spesso un **decremento esponenziale delle performance su Cloud rispetto ai soldi investiti.**
 - Sostanzialmente occorre cercare che il problema da risolvere non richieda un cluster di macchine troppo grande.
- Cf. J.Napper, P.Bienintesi, Can Cloud Computing Reach the TOP500?, UCHPC-MAW'09

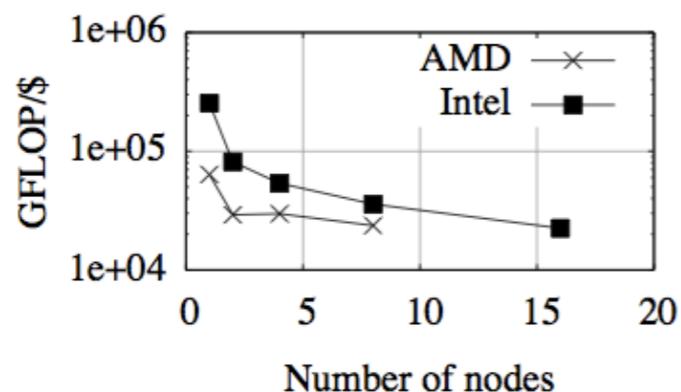


Figure 2: By cluster size. Performance for money spent on EC2 using AMD and Intel instances by cluster and problem sizes. The y-axis is log scale.

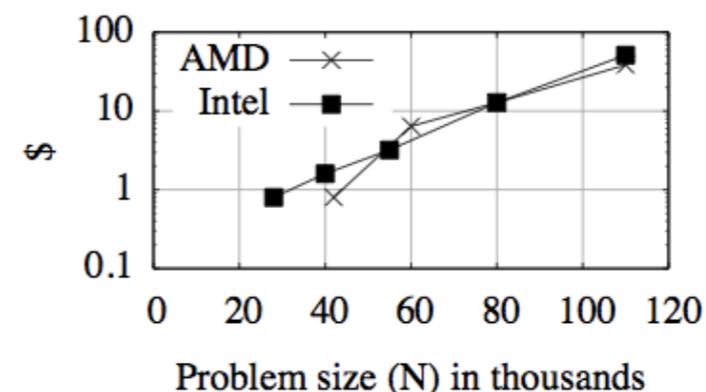


Figure 3: By problem size. Performance for money spent on EC2 using AMD and Intel instances by cluster and problem sizes. The y-axis is log scale.

Ancora sui costi

- Tra gli altri: recente studio di una università americana sull'**uso di Amazon EC2 per simulazioni di fisica** (A.Melo, Vanderbilt U, CHEP 2010 Proceedings)

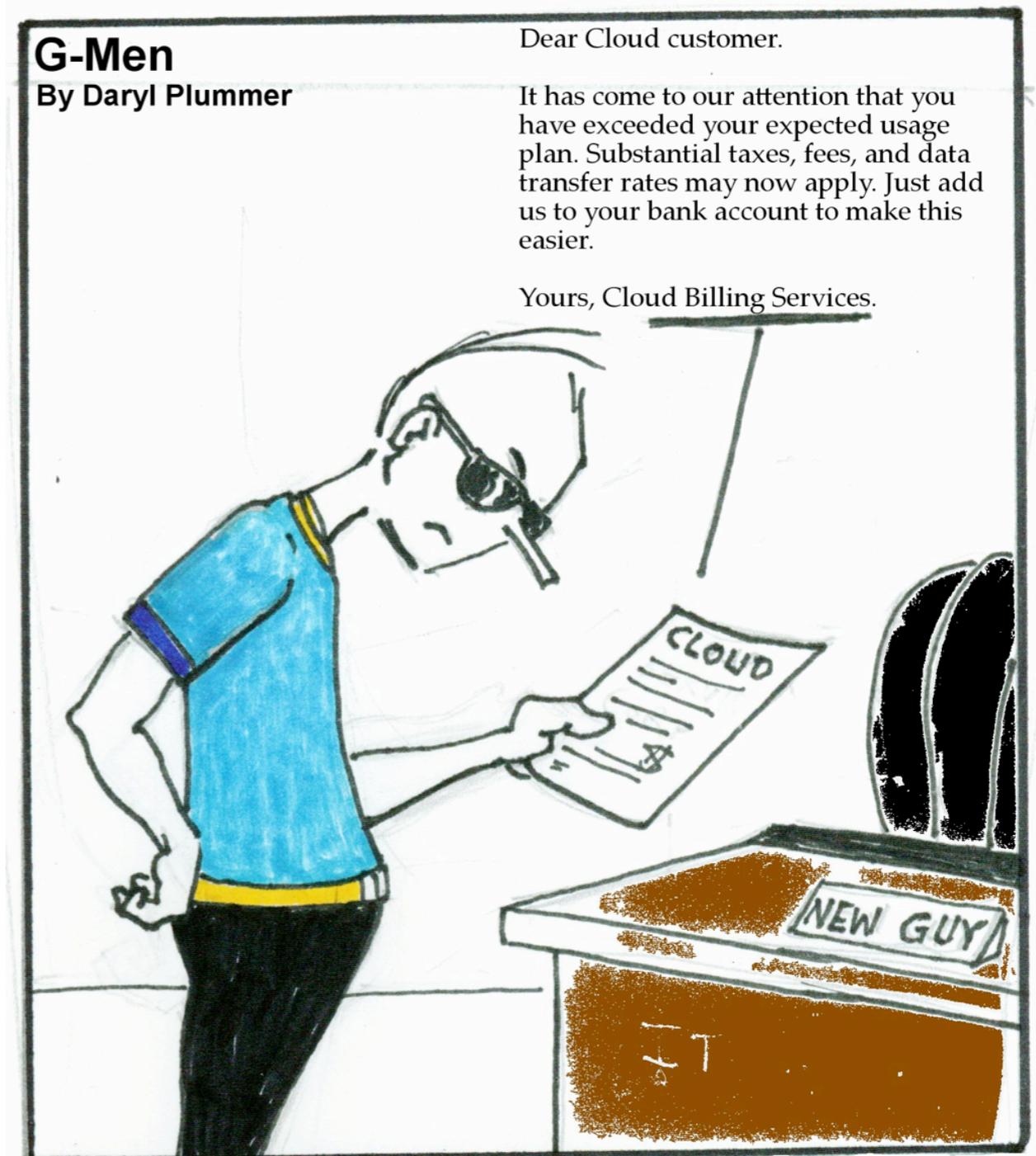
Site	Events/ Core-Hour	USD/Core- Hour	USD/1000 events
Vanderbilt	244	\$0.024	\$0.098
Amazon	208	\$0.17 (varies)	\$0.817

- **EC2 ~8.3x more expensive**
- EC2 probably **only reasonable if CPUs are needed for a short amount of time**

- CPU-bound tests involving EC2 large instances
 - EC2: ~85% performance compared to Vanderbilt systems
 - This is likely an optimal scenario for EC2
- Vanderbilt costs include hardware, power, facilities, bandwidth and support costs

Costo “a consumo”

- La “**elasticità**” del Cloud Computing (“posso darti tutte le risorse che desideri” - *a fronte di una valida carta di credito*) può avere evidenti effetti indesiderati
 - Analogo ai piani tariffari dei telefoni cellulari (“davvero ho ecceduto i miei minuti gratis?”)



Portabilità

- Quali standard devo usare nel Cloud Computing? Ci sono degli standard?
 - Anche troppi, a mio parere
- Il fatto che le “Cloud” possano non essere affidabili (o meno affidabili delle infrastrutture che abbiamo localmente) significa che devo rivedere le mie applicazioni per renderle robuste e “generiche”
 - Potrebbe non essere facile. Ad esempio, il software di offline di un singolo esperimento LHC (ATLAS) è composto da **più di 7 milioni di linee di codice** (la maggior parte in C++)
- Cf. K.R.Jackson, K.J.Runge, R.C.Thomas (LBNL), Seeking Supernovae in the Clouds: A Performance Study, HPDC 2010

Come “nebulizzare” il mondo scientifico quindi?

- Il Cloud Computing, con la sua facilità d’uso nell’allocazione delle risorse, è di grande interesse anche per il calcolo scientifico.
- Tuttavia **è richiesto un modello evolutivo** che dia la possibilità di confermare i successi ottenuti ad esempio dal Grid Computing per grandi collaborazioni
 - Non è chiaro se sia fattibile (tecnicamente e politicamente) “privatizzare” il mondo del calcolo scientifico distribuito. Un mix pubblico/privato potrebbe avere certamente senso.

Una soluzione tra Grid e Cloud

- Al CNAF di Bologna abbiamo sviluppato un software per la virtualizzazione dinamica delle risorse
 - Nome impronunciabile: **WNoDeS (Worker Nodes on Demand Service)**
 - Le risorse possono essere presentate come risorse “Grid” o come risorse “Cloud”
 - Basato su Linux, pensato per essere scalabile
 - Attualmente gestisce **2000 macchine virtuali all’interno del Centro di Calcolo** ed è in produzione da 18 mesi. Attraverso WNoDeS sono ormai passati svariati milioni di job di fisica.
 - Cf. <http://web.infn.it/wnodes> per maggiori informazioni.

Il portale Cloud di WNoDeS



The screenshot shows the WNoDeS Cloud Portal interface. At the top, there is a navigation bar with 'MY RESOURCES', 'NEW RESOURCE', and 'CONTACT US'. A user profile dropdown menu is open, showing the text '/C=IT/O=INFN/OU=Personal Certificate/L=CNAF/CN=Davide Salomoni' and a user icon. Below the navigation bar, the main heading is 'Create a New Virtual Machine' with a 'Need Support?' button. The current VO is 'cms (Change VO)'. There are four numbered steps: 1. Hardware, 2. Operating System, 3. Keys, and 4. Create. Below the steps, there is a section for selecting a configuration. The 'MEDIUM' option is selected and circled in red. The 'MEDIUM' option details are: 2 cores, 3.5 GB RAM, 100 GB HD, 200 Mb/s throughput. A green arrow points to the right at the bottom of the configuration list.

WNoDeS
Grid Resources via Cloud Interface

MY RESOURCES NEW RESOURCE CONTACT US

/C=IT/O=INFN/OU=Personal Certificate/L=CNAF/CN=Davide Salomoni

Create a New Virtual Machine

Need Support?

Current VO: cms (Change VO)

1 Hardware 2 Operating System 3 Keys 4 Create

Select the preferred configuration between the existing, contact us if you need more customization

<input type="radio"/>	SMALL	1 core, 1.7 GB RAM, 50 GB HD, 100 Mb/s throughput
<input checked="" type="radio"/>	MEDIUM	2 cores, 3.5 GB RAM, 100 GB HD, 200 Mb/s throughput
<input type="radio"/>	LARGE	4 cores, 7 GB RAM, 200 GB HD, 400 Mb/s throughput
<input type="radio"/>	EXTRA-LARGE	8 cores, 14 GB RAM, 400 HD, 800 Mb/s throughput

X.509 Digital Certificate

Resource selection
(note network
bandwidth throttling)

Creazione di risorse virtuali

Need Support?

Current VO: cms (Change VO)

1
Hardware

2
Operating System

3
Keys

4
Create

Give a name to your virtual machine, (max 13 characters)

How many virtual machines you want to create?

NOTE: If you choose to create more than one instance, they will all have the same hw configuration. The names will include an incremental suffix (e.g. "Instancename1*", "Instancename2"...)

CREATE

←



[MY RESOURCES](#)
[NEW RESOURCE](#)
[CONTACT US](#)

/C=IT/O=INFNOU=Personal Certificate/L=Ferrara/CN=Guido Potena

My Virtual Machines List

Manage your virtualized resources.

Refresh List

test	
🕒 15/03/2011-12:20:13	📄 Scientific Linux 5.3 x86_64
👤 8 Core	💾 14 GB
💾 400 GB	🌐 800 Mbs





La pagina sul server https://wnodes-cloud-test-1.c

? Are you sure you want to kill this virtual resource?

✖ Annulla
✔ OK

Copyright 2011, INFN-CNAF Bologna
Viale Bertini Pichat 6/2 40127 Bologna, Italy

Idealmente...

La “Grid” elettrica	La “Grid” (o “Cloud”) computazionale
Non ci si preoccupa (molto) da dove viene l’elettricità , se dall’eolico in Germania o dal nucleare in Francia.	Non ci si preoccupa da dove viene la potenza di calcolo , se da un supercomputer in Giappone o da un laptop in Canada.
L’infrastruttura è chiamata la “rete elettrica”. Connette generatori di corrente con la nostra casa attraverso linee trasmissive, trasformatori, etc.	L’infrastruttura è chiamata “Grid” (o “Cloud”). Connette risorse di calcolo con la nostra casa e fornisce i meccanismi per accedervi.
La rete elettrica è pervasiva: è disponibile dappertutto e vi si accede tramite una presa di corrente.	La “Grid” (o “Cloud”) è pervasiva: risorse remote sono accessibili da laptop, tablet, magari attraverso un semplice browser .
La rete elettrica è una utility : chiedi l’elettricità e la ottieni, pagando quello che ti viene dato.	La “Grid” (o “Cloud”) è una utility : chiedi risorse di calcolo o storage e le ottieni, pagando quello che ti viene dato.

Cf. <http://www.gridcafe.org/>

Purtroppo non siamo ancora a questo punto

Conclusioni

- Il calcolo scientifico distribuito usufruisce a livello mondiale di significative infrastrutture spesso realizzate con finanziamenti pubblici
 - L'ultimo decennio ha mostrato che sono stati ottenuti ottimi risultati nell'utilizzo di tali infrastrutture, ad esempio con il **Grid Computing**.
 - L'Italia in particolare ha spesso giocato e gioca un ruolo di primo piano.
- Il **Cloud Computing** promette di semplificare sostanzialmente l'allocazione e l'uso delle risorse, (forse) di ridurre i costi e di incentivare l'iniziativa privata.
 - Ci sono tuttavia diversi punti di opportunità, politici e tecnici che sono ancora da chiarire.
- Un modello sensato pare essere quello di continuare a usare infrastrutture esistenti e **investire nella ricerca tecnologica per nuovi modelli di calcolo**.
 - Tenendo conto che quello che va bene per la fornitura di 100 o 1000 server "generici" per una azienda o per una Pubblica Amministrazione non va necessariamente bene per una collaborazione scientifica distribuita.

E' tutto

1. **Il Cloud Computing non è altro che un termine modaiolo per il noto concetto di computer collegati via rete. Una Cloud è vapore acqueo.** (Larry Ellison, co-fondatore e CEO di Oracle Corporation, Settembre 2009)
2. **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 per dare scelta al cliente.** (Oracle Cloud Computing FAQ, Ottobre 2010)
3. **La verità è raramente pura, e mai semplice.** (Oscar Wilde, L'importanza di chiamarsi Ernesto, 1895)

Grazie! e-mail: Davide.Salomoni@cnaif.infn.it