GRID COMPUTING TRAINING COURSE

CLOUD COMPUTING AND HPC IN THE ATLAS COLLABORATION

E. FULLANA TORREGROSA

Outline

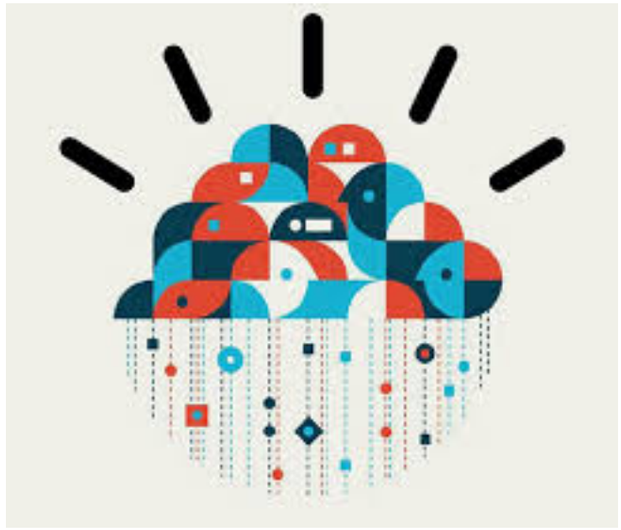

Accept the
PAST,
Manage the
PRESENT,
& Work hard
towards the
FUTURE.
WWW.LIVELIFEHAPPY.COM


TITAN

Outline

Where we are, where we go in high energy physics computing

Outline



Cloud Computing :
-What it is
-Basic jargon
-Why it is important

Outline

High performance computers :

A key player in science (and politics!) not exploited in our community

Outline

Where we are, where we go in high energy physics computing

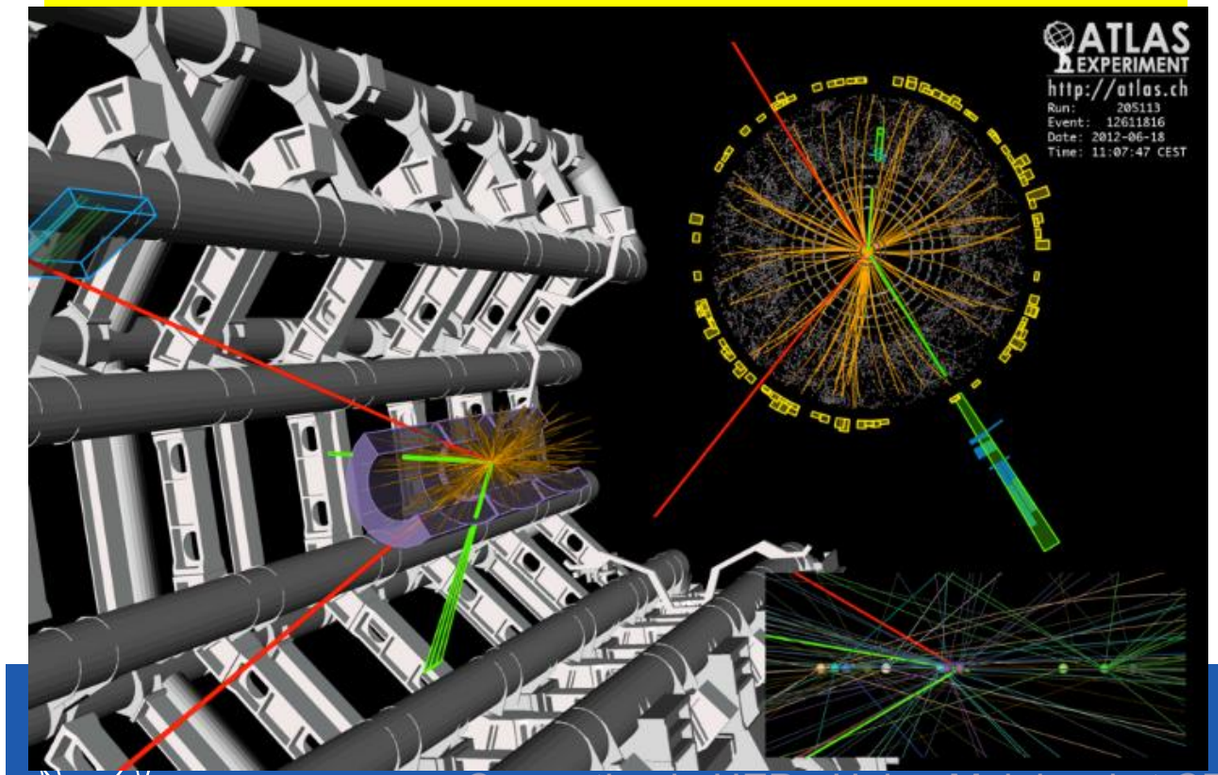
Further information : https://indico.cern.ch/event/630357/contributions/2547039/

The whole Facebook was 100 PB in 2012

https://indico.cern.ch/event/630357/contributions/2547039/

What is the LHC Data?

- 150 million sensors deliver data … 40 million times per second
- Generates ~ 1 PB per second

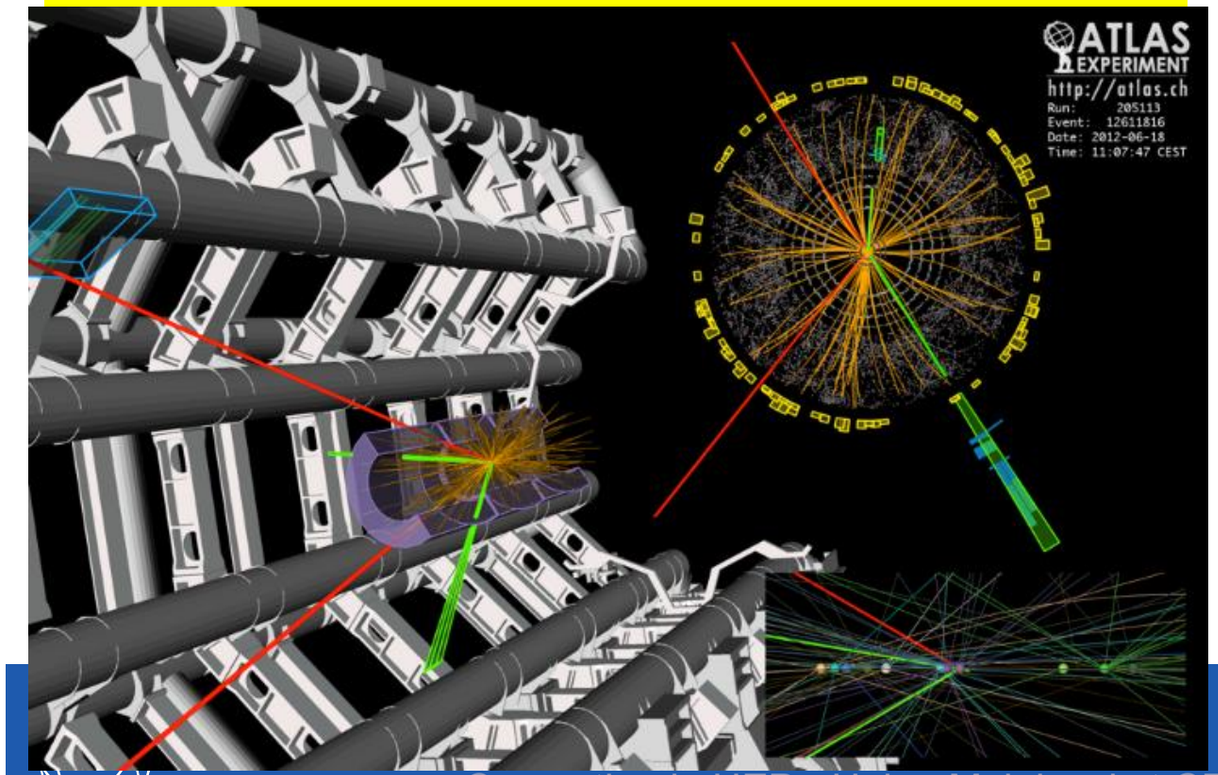


- Raw data:
  - Was a sensor hit?
  - How much energy deposit?
  - What time?
- Reconstructed data:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information
  - …

https://indico.cern.ch/event/630357/contributions/2547039/

What is the LHC Data?

- 150 million sensors deliver data … 40 million times per second
- Generates ~ 1 PB per second



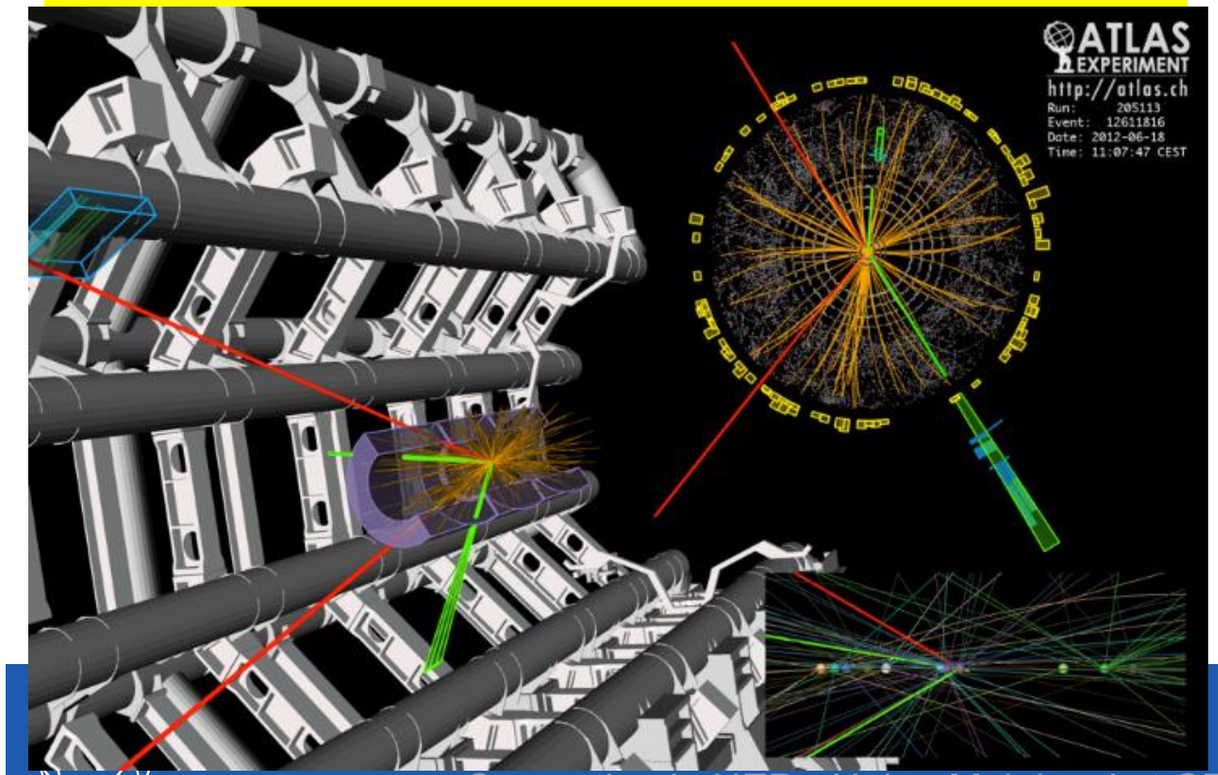
- Raw data:
  - Was a sensor hit?
  - How much energy deposit?
  - What time?

- Reconstructed data:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information
  - …

https://indico.cern.ch/event/630357/contributions/2547039/

What is the LHC Data?

- 150 million sensors deliver data … 40 million times per second
- Generates ~ 1 PB per second



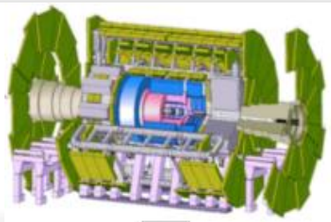
- Raw data:
  - Was a sensor hit?
  - How much energy deposit?
  - What time?

- Reconstructed data:
  - Momentum of tracks (4-vectors)
  - Origin
  - Energy in clusters (jets)
  - Particle type
  - Calibration information
  - …

https://indico.cern.ch/event/630357/contributions/2547039/

2) where this number comes from?

HEP Computing



~40 MHz ~ PB/s

L1 Trigger (HW)

~100 kHz

HL Trigger (SW)

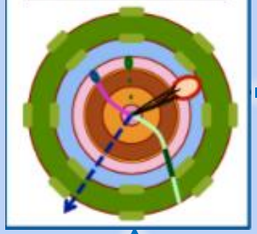
"Online" – Real time
Operated and funded as part of the detector

~1 kHz

WLCG

"Raw Data" ~ 1-10 GB/s

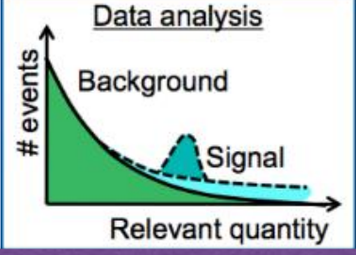
Reconstruction



Calibration

Monte Carlo Simulations

Data analysis



"Offline" - Asynchronous

WLCG: An international collaboration to distribute and analyse LHC data

Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists

https://indico.cern.ch/event/630357/contributions/2547039/

Nature of the Computing Problem

- Enormous numbers of collisions of proton bunches with each other
  - Data from each collision are small (order 1…10 MB)
  - Each collision independent of all others

this is a key point in high energy physics computing

https://indico.cern.ch/event/630357/contributions/2547039/

Nature of the Computing Problem

but it does not mean we cannot use them

- No supercomputers needed
  - Most cost-effective solution is standard PC architecture (x86) servers with 2 sockets, SATA drives (spinning or SSD), Ethernet network
  - Linux (RHEL variants: Scientific Linux, CentOS) used everywhere

https://indico.cern.ch/event/630357/contributions/2547039/

Nature of the Computing Problem

this point drives one the needs for future requirements in computing high energy physics

It does not scale linearly!!

- Calculations are mostly combinatorics – integer (rather than floating-point) intensive

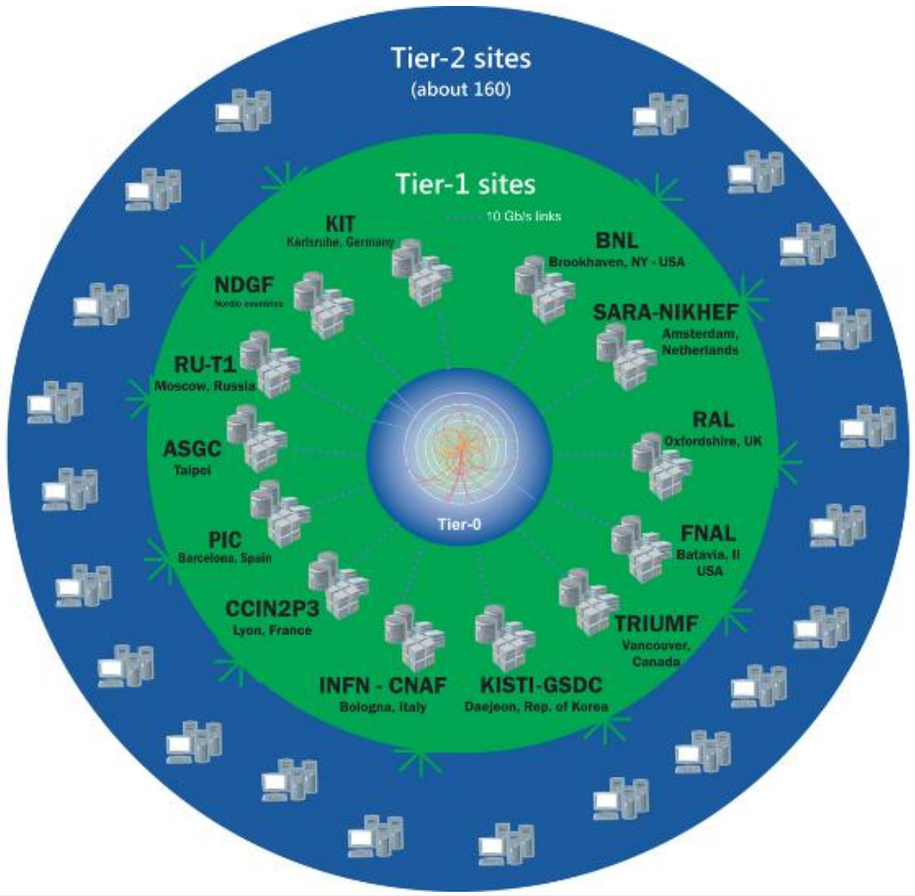
https://indico.cern.ch/event/630357/contributions/2547039/

The Worldwide LHC Computing Grid

Tier-0 (CERN):
data recording, reconstruction and distribution

Tier-1:
permanent storage, re-processing, analysis

Tier-2:
Simulation, end-user analysis



~170 sites, 42 countries

~750’000 cores

~1’000 PB of storage

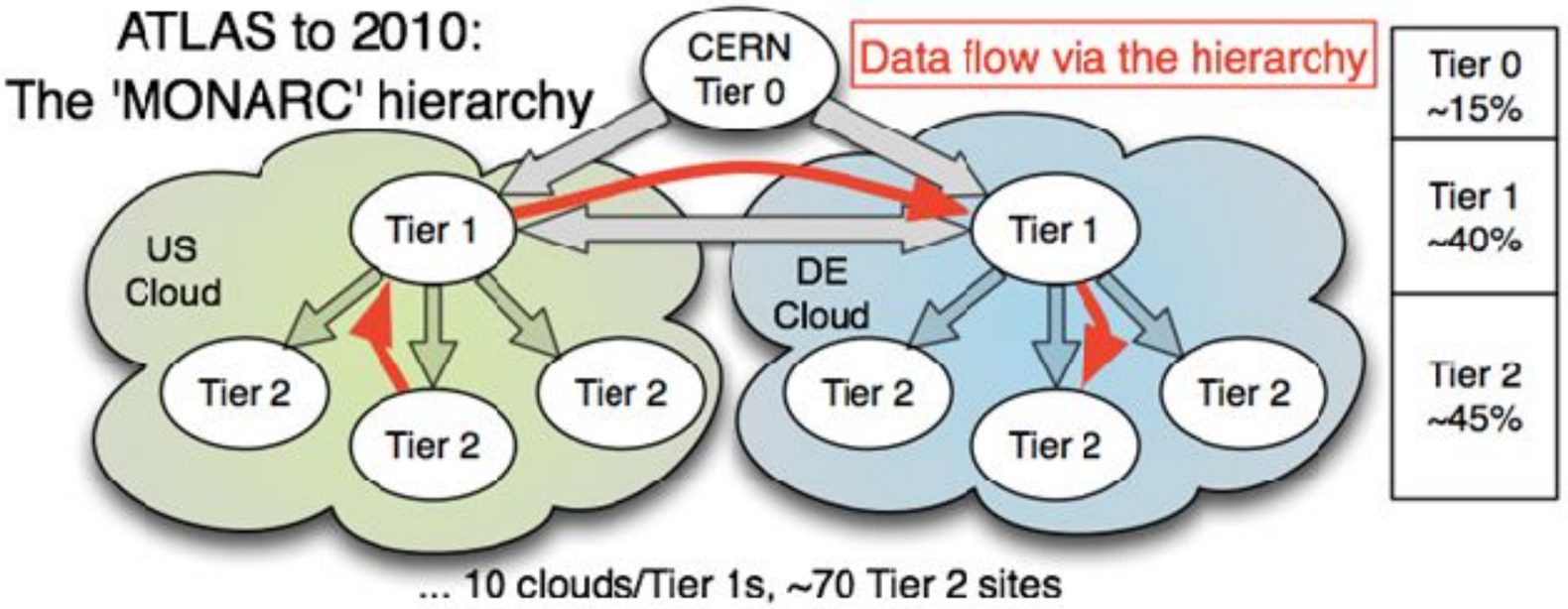
> 2 million jobs/day

10-100 Gb links

WLCG: An international collaboration to distribute and analyse LHC data

Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists

Computing Model Evolution



Original model:

Static strict hierarchy

Multi-hop data flows

Lesser demands on Tier 2 networking

Virtue of simplicity

Designed for <~2.5 Gb/s within the hierarchy

Today:

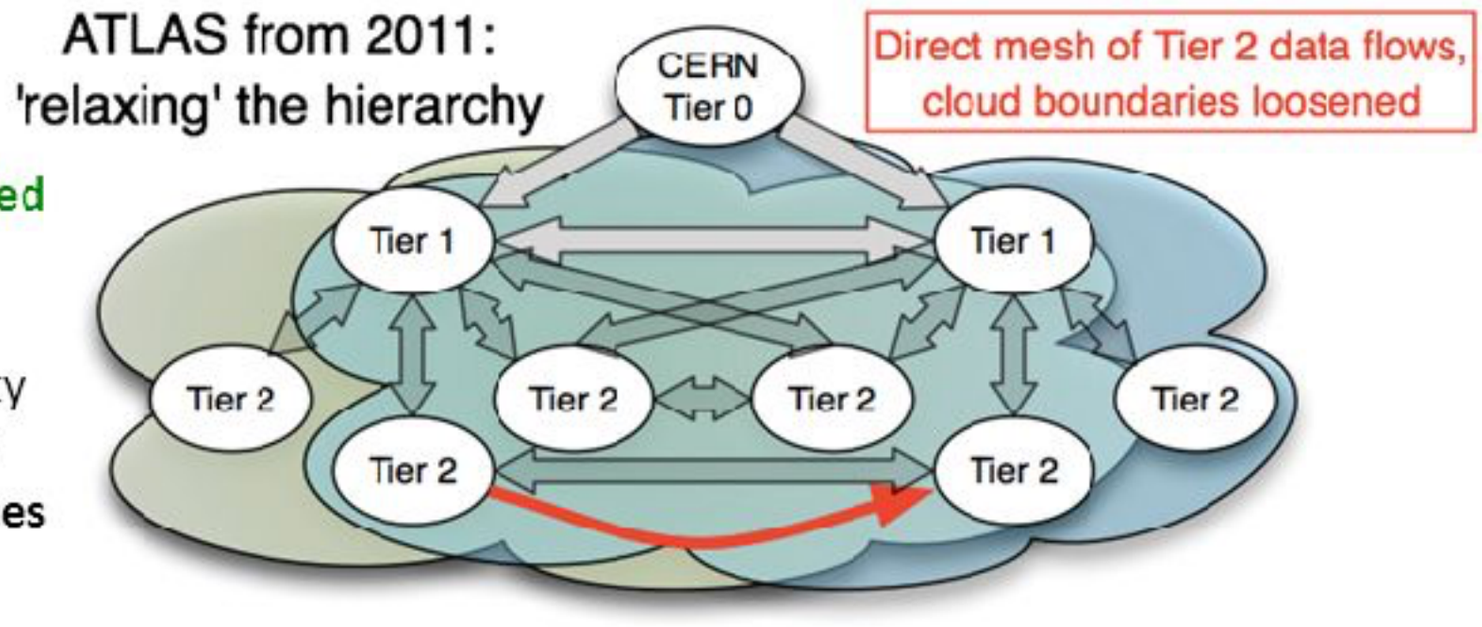
Bandwidths 10-100 Gb/s, not limited to the hierarchy

Flatter, mostly a mesh

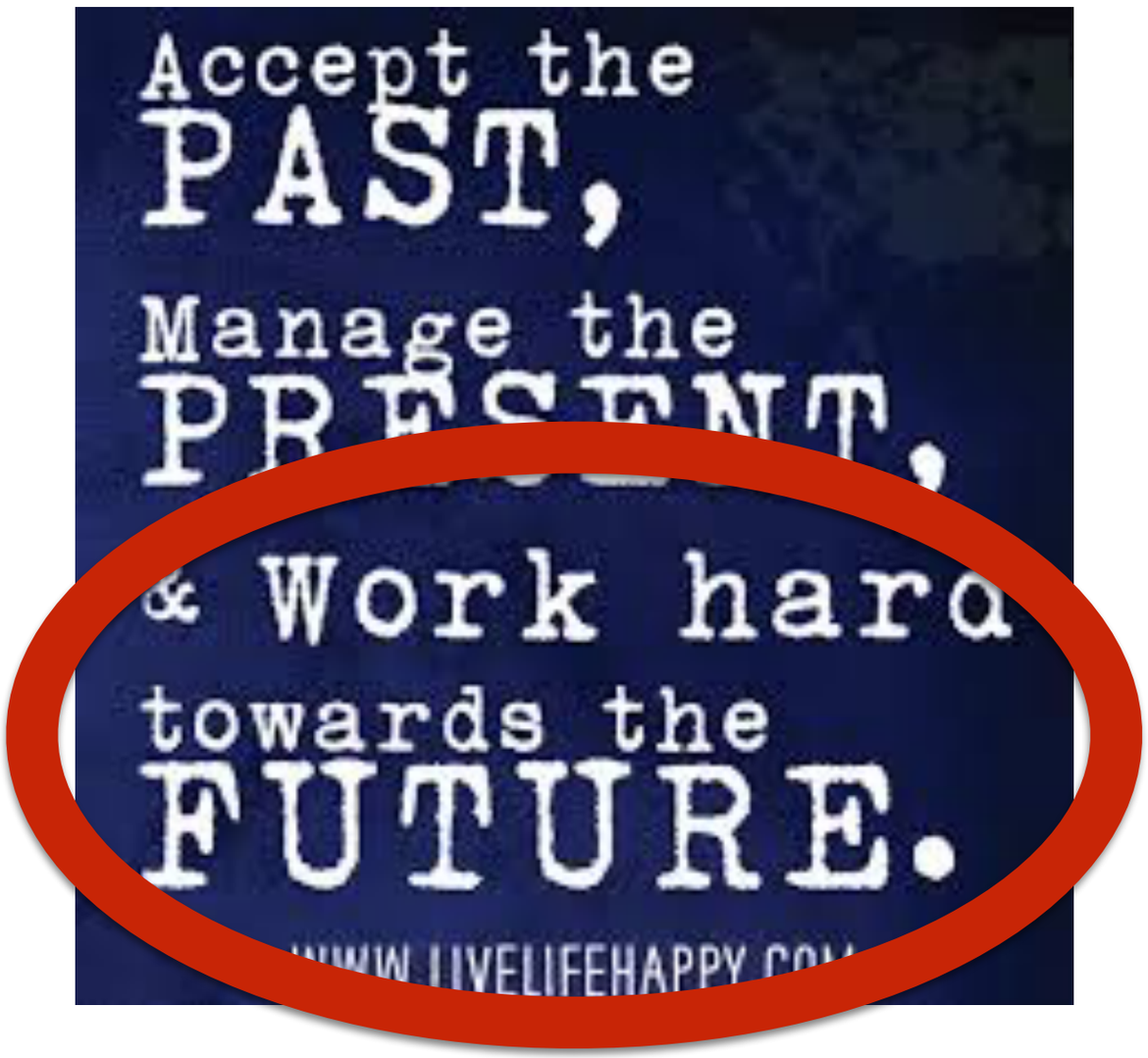
Sites contribute based on capability

Greater flexibility and efficiency

More fully utilize available resources



Outline



Where we are, where we go in high energy physics computing

today

LHC / HL-LHC Plan

High Luminosity LHC

Run I | Run II | Run III | HL- Run IV, V…

Run 1 | Run 2 | Run 3 | Run 4 - 5…

7 TeV | 8 TeV | LS1: splice consolidation, button collimators, R2E project | 13 TeV | EYETS | LS2: injector upgrade, cryo Point 4, Civil Eng. P1-P5 | 14 TeV | cryolimit interaction regions | LS3: HL-LHC installation | 14 TeV energy | 5 to 7 x nominal luminosity

2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2037

75% nominal luminosity

experiment beam pipes

nominal luminosity

experiment upgrade phase 1

radiation damage

2 x nominal luminosity

experiment upgrade phase 2

30 fb^{-1} | 150 fb^{-1} | 300 fb^{-1} | 3000 fb^{-1} Integrated luminosity

HL-LHC desde el 2026 hasta el 2037!!

Civil engineering: Lay-out, permits, tenders, preparation | Excavation period | Surface buildings

Components: Prototypes/Spec | FAV

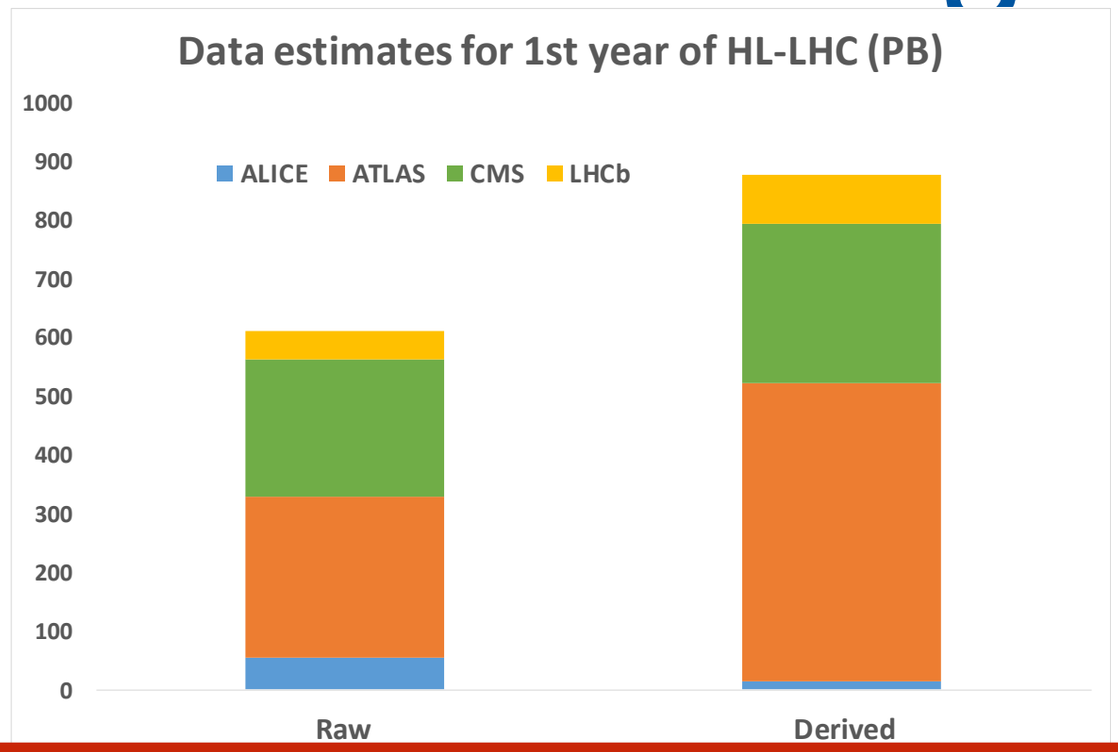
Tech Infrastructure: Specifications | FAV | Installation

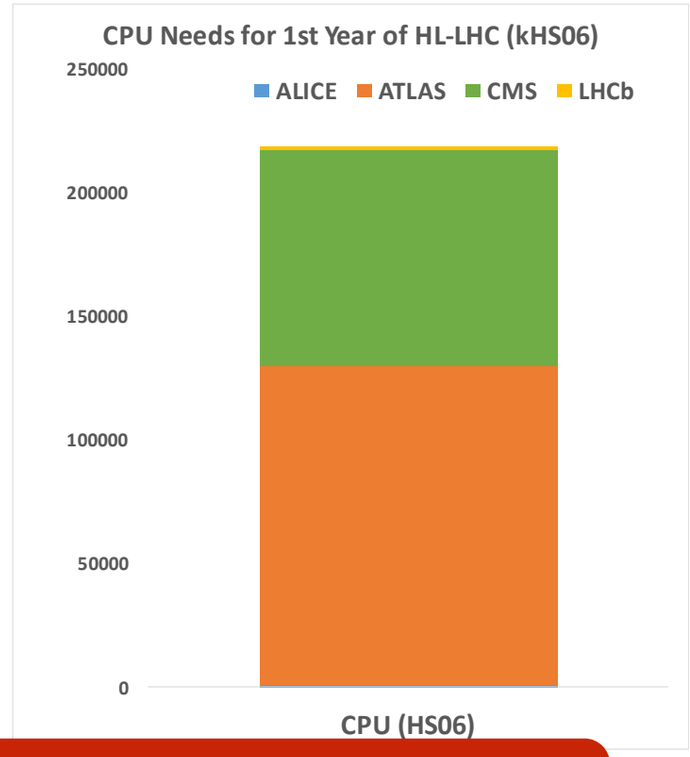
FAV = Fabrication, Assembly and Verification

Frederick Bordry (CERN) @ ATLAS Week 21/06/17

https://indico.cern.ch/event/630357/contributions/2547039/

Future Challenges for LHC





Data:
- Raw 2016: 50 PB → 2027: 600 PB
- Derived (1 copy): 2016: 80 PB → 2027: 900 PB

CPU:
- x60 from 2016

A factor of 10 more storage needed

A factor of 60! more CPU

What can we do?

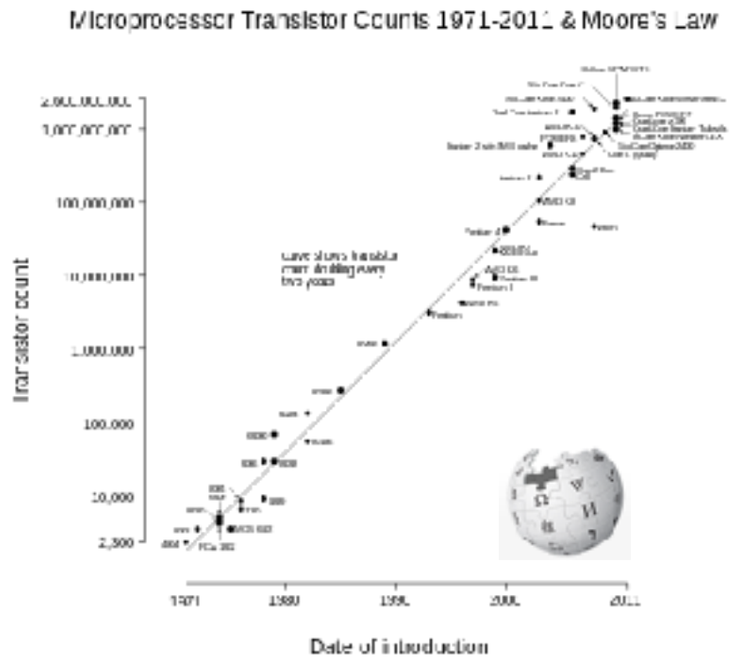
- Assuming 20% per year from technology, still factors missing in terms of cores, storage etc

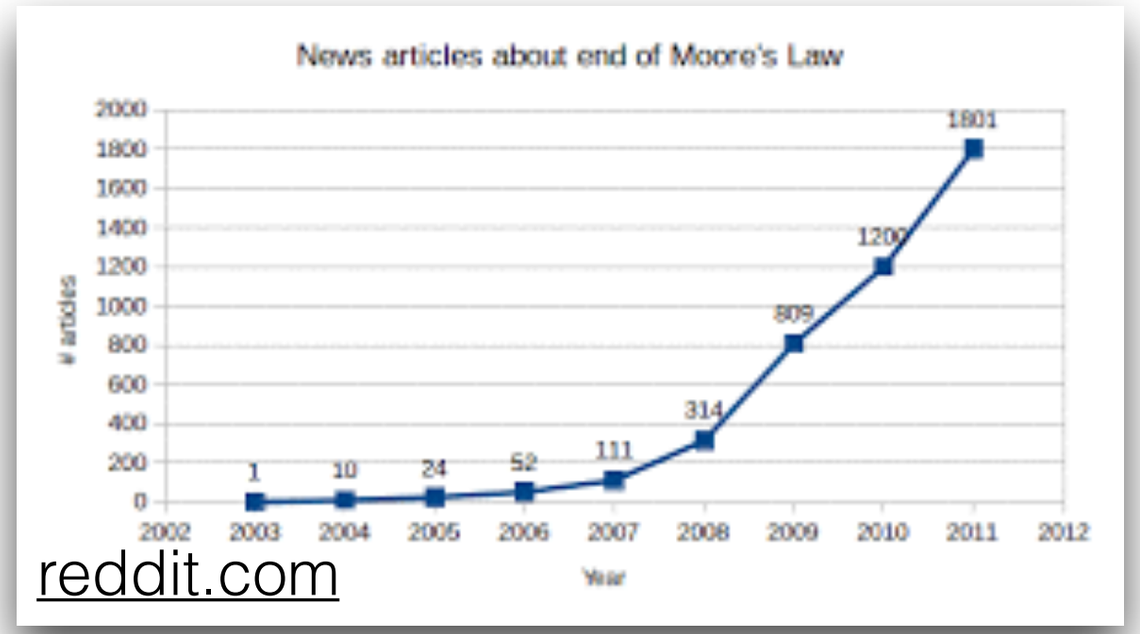
- Moore's law coming to an end for business and financial reasons?

- Large effort spent to improve software efficiency

- Exploit multi-threading, new instruction sets,

- but we can still do one more thing





reddit.com

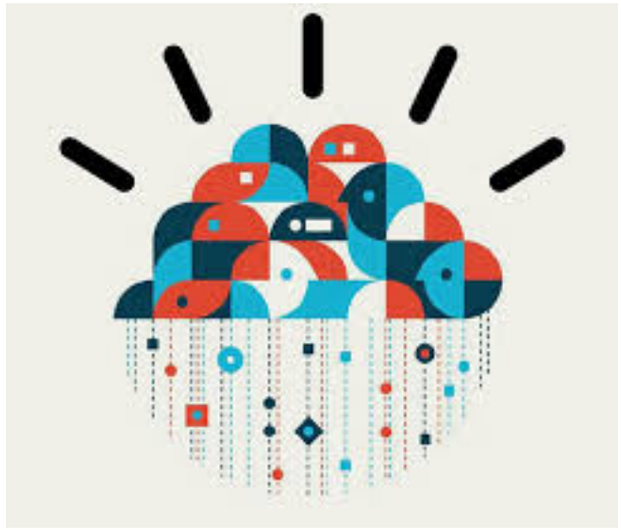
https://indico.cern.ch/event/630357/contributions/2547039/

Opportunistic resources

- Today this has become more important
  - Opportunistic use of:
    - HPC facilities
    - Large cloud providers
    - Other offers for “off-peak” or short periods
  - All at [illegible] cost (for hardware)
- But scale and cost are unpredictable

- Also growing in importance:
  - Volunteer computing (citizen science)
    - BOINC-like (LHC@home, ATLAS/CMS/LHCb@home, etc)
    - Now can be used for many workloads – as well as the outreach opportunities

Outline



Cloud Computing :
-What it is
-Basic jargon
-Why it is important

Further information : https://indico.cern.ch/event/178466/

– Cloud computing is a computing paradigm shift where computing is moved away from personal computers or an individual server to a "cloud" of computers.

Wikipedia - December, 2007

– Cloud computing is a style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet.

Wikipedia - July, 2009

– Cloud computing is a metaphor used by Technology or IT Services companies for the delivery of computing requirements as a service to a heterogeneous community of end-recipients.

Wikipedia - May, 2012

– Cloud Computing is a style of computing where scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies.

Thomas Bittman - Gartner

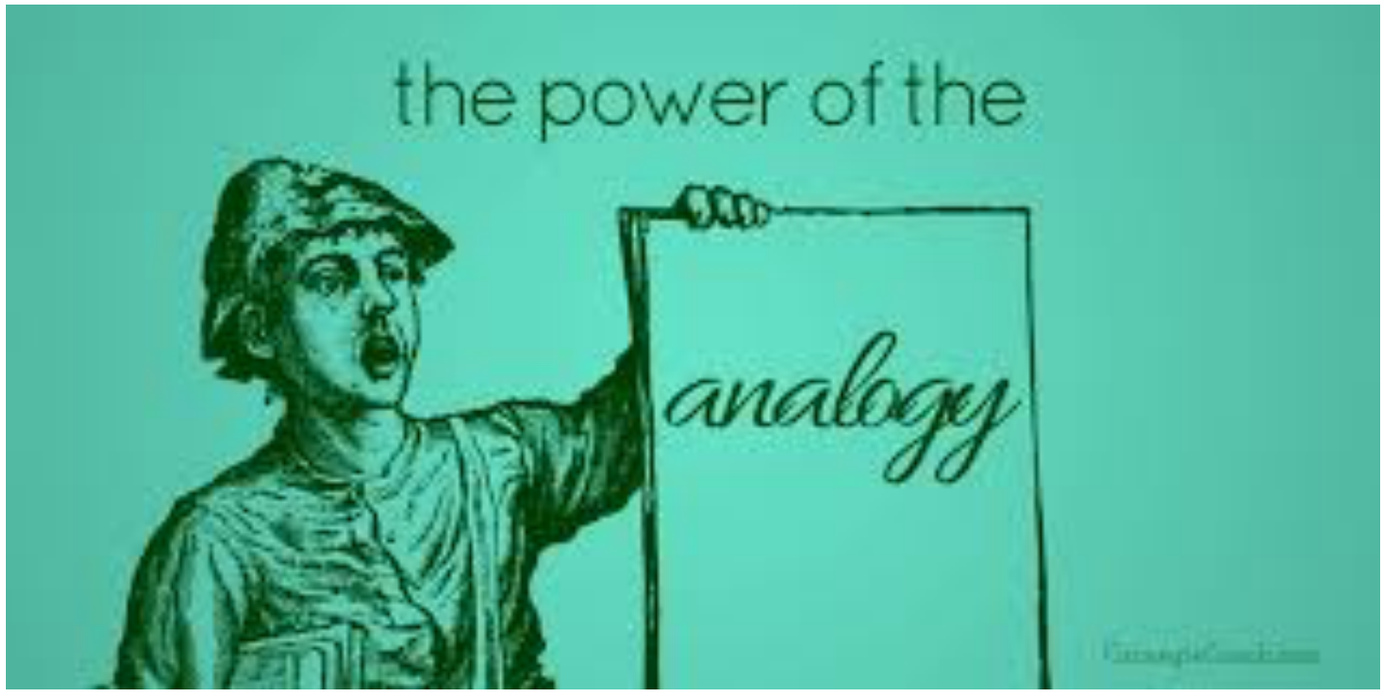
– Cloud Computing is an emerging IT development, deployment and delivery model, enabling real-time delivery of products, services and solutions over the Internet.

Frank Gens - IDC

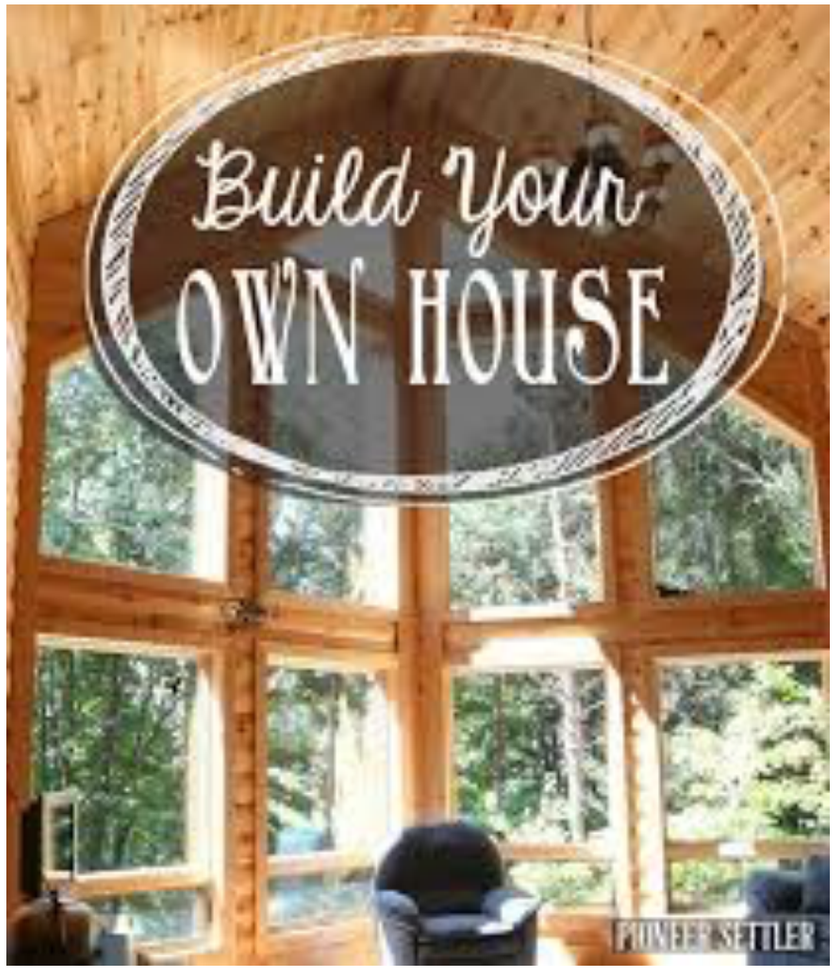
- Cloud computing describes a service model that combines a general organizing principle for IT delivery, infrastructure components, an architectural approach and an economic model – basically, a confluence of grid computing, virtualization, utility computing...

Dan Kusnetzky, Rachel Chalmers - 451 Group

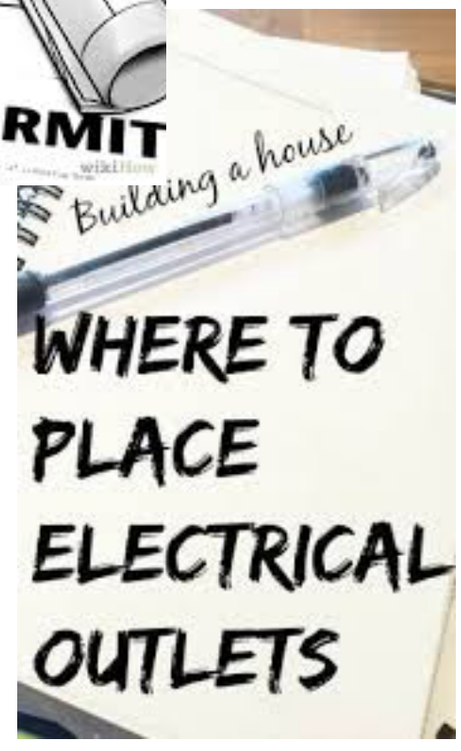
Still not clear, any analogy????



Making your own IT services is like…



The 10 most important things I learned building a house.
TheSunnySideupBlog.com



and you have to take care of it!!!

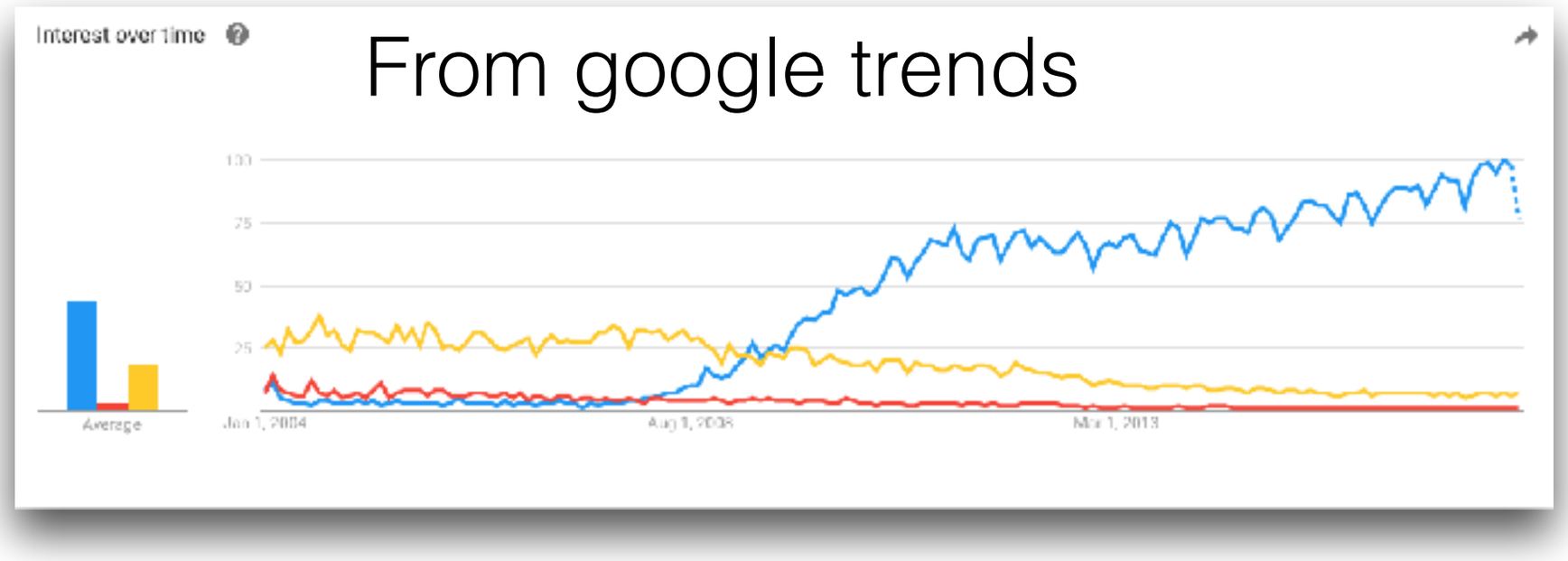
While getting them on the Cloud is more like

but how interesting it is?

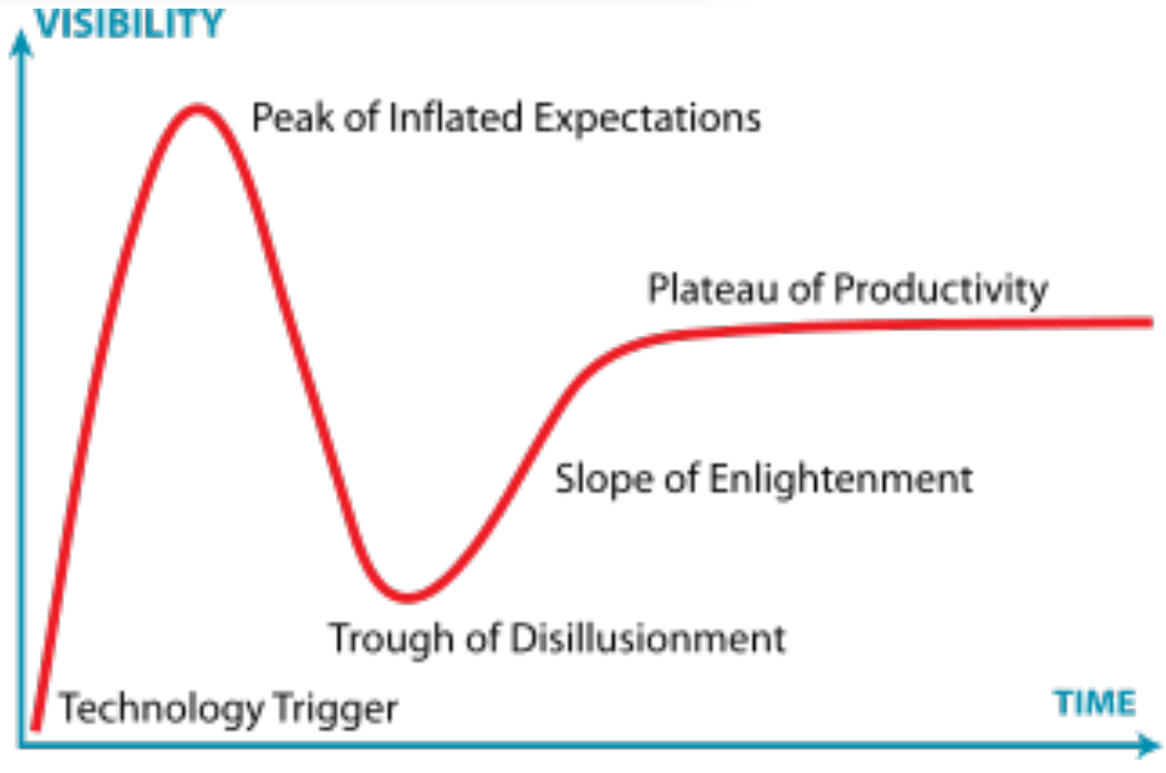
Cloud Computing

Grid Computing

HPC



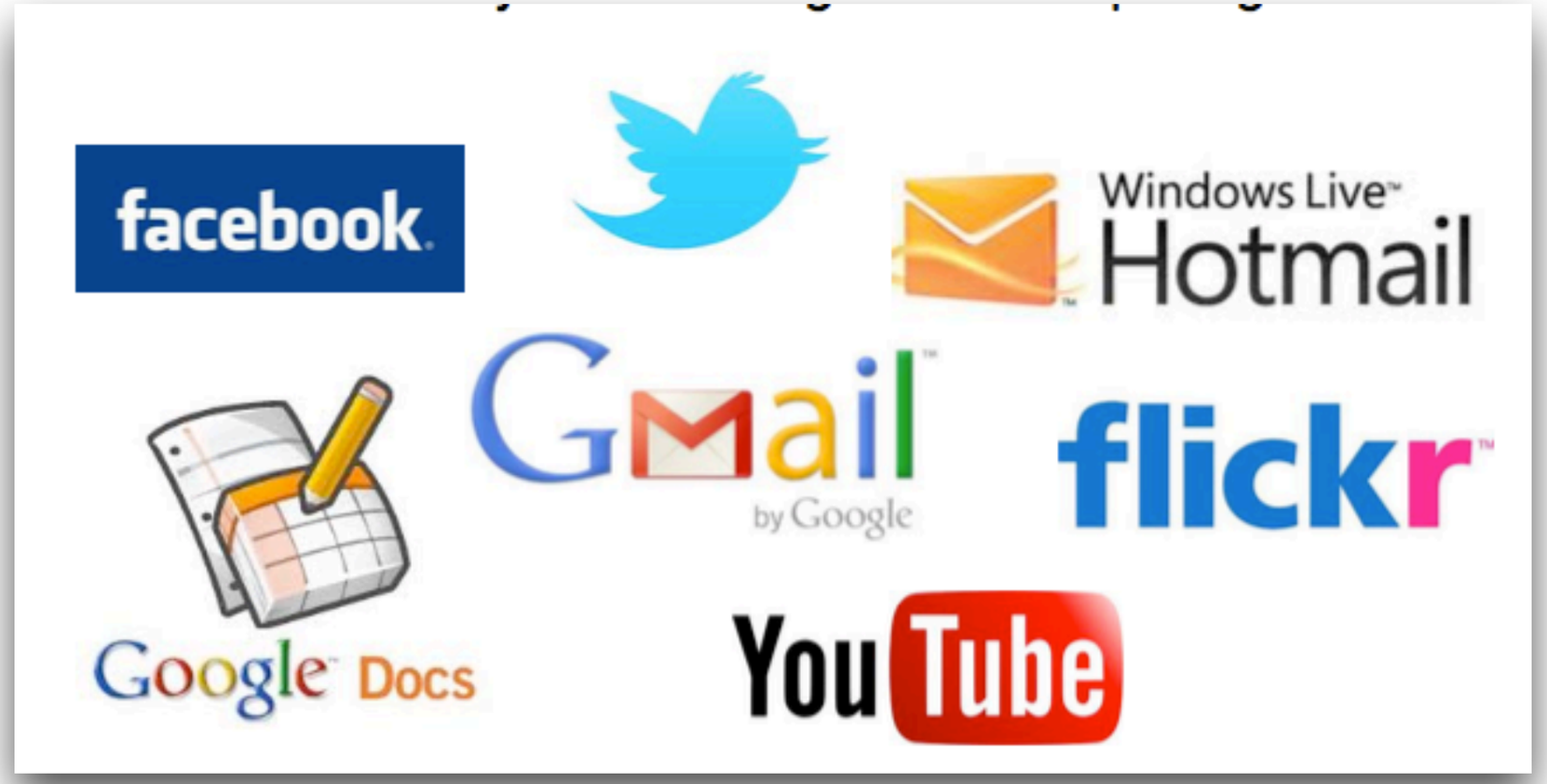
caution is always advisable : we could be on the hype :



https://en.wikipedia.org/wiki/Hype_cycle

in reality is everywhere!! :

in reality is everywhere!! :



Ok, we have an idea of what it is and how important it is; let’s go for a bit of terminology

IaaS

Infrastructure as a service

Provide access to collections of virtualised computer hardware resources

You do :

- Applications
- Runtimes
- Security
- Databases
- Servers

You don’t do :

- Virtualisation
- Server HW
- Storage
- Networking

PaaS

Platform as a service

Facilitate the deployment of applications without the cost and complexity of buying and managing the underlying hardware

You do :

Applications

You don't do :

Virtualisation

Server HW

Storage

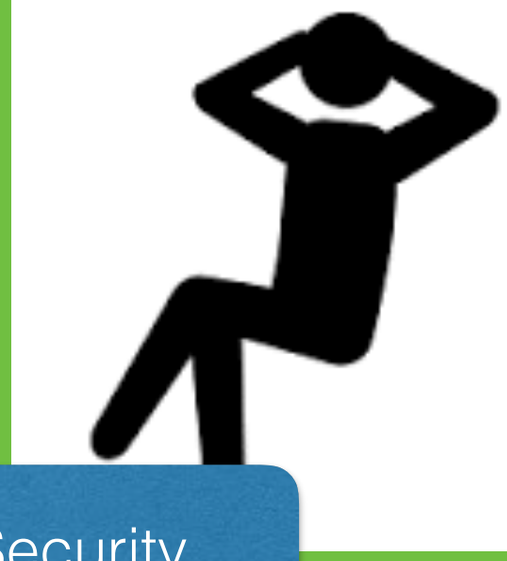
Networking

Runtimes

Security

Servers

Databases



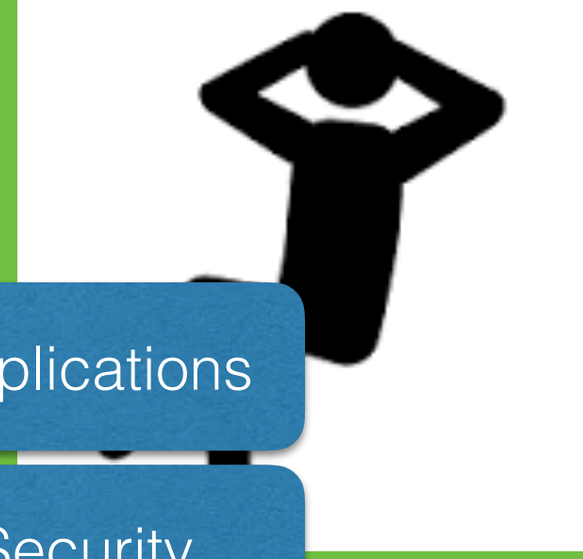
SaaS

Software as a service

Delivery model in which software and associated data are centrally hosted

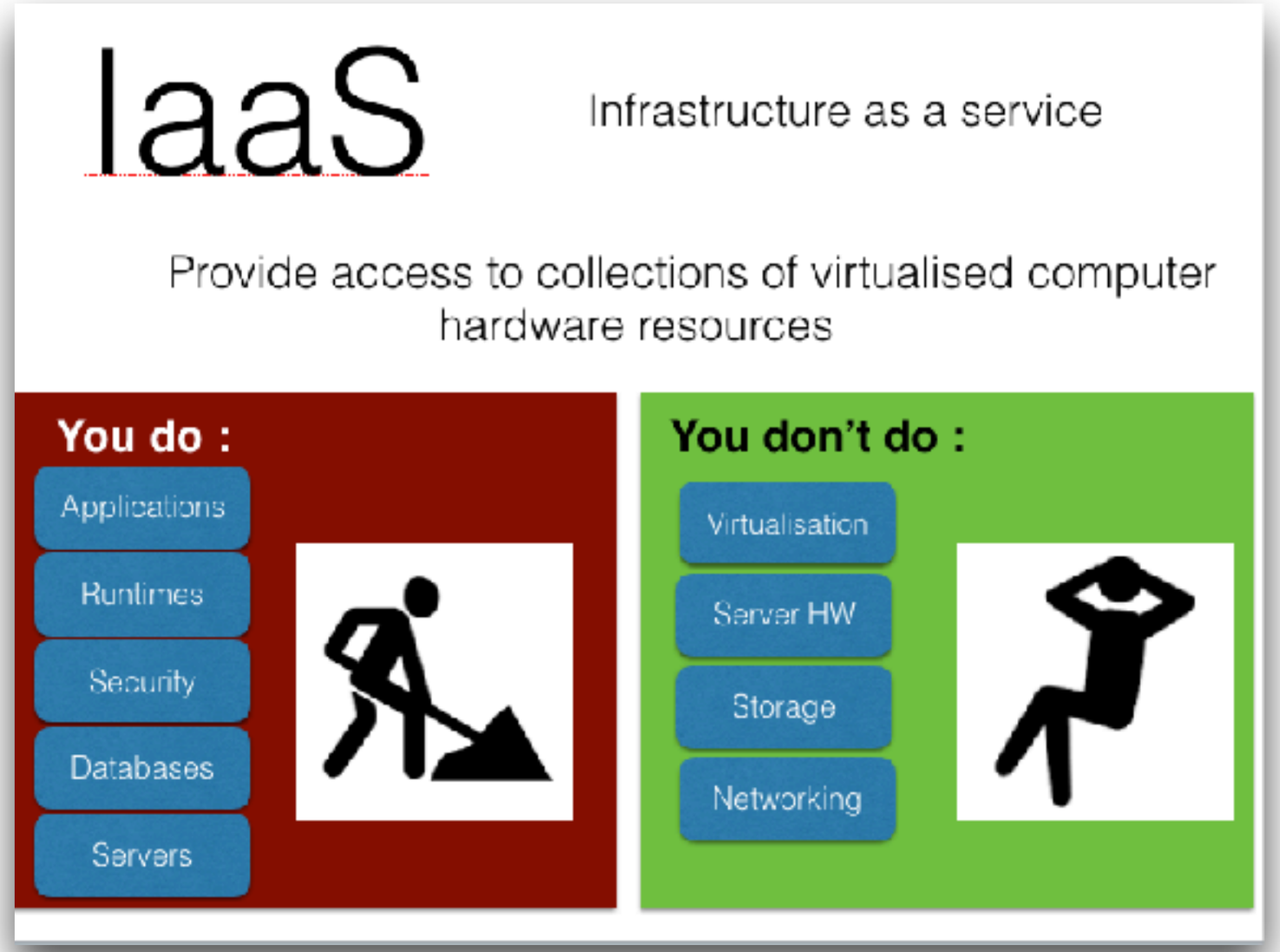
You do :

You don't do :



- Virtualisation
- Server HW
- Applications
- Storage
- Security
- Networking
- Servers
- Runtimes
- Databases

but, can we use it for high energy physics computing?



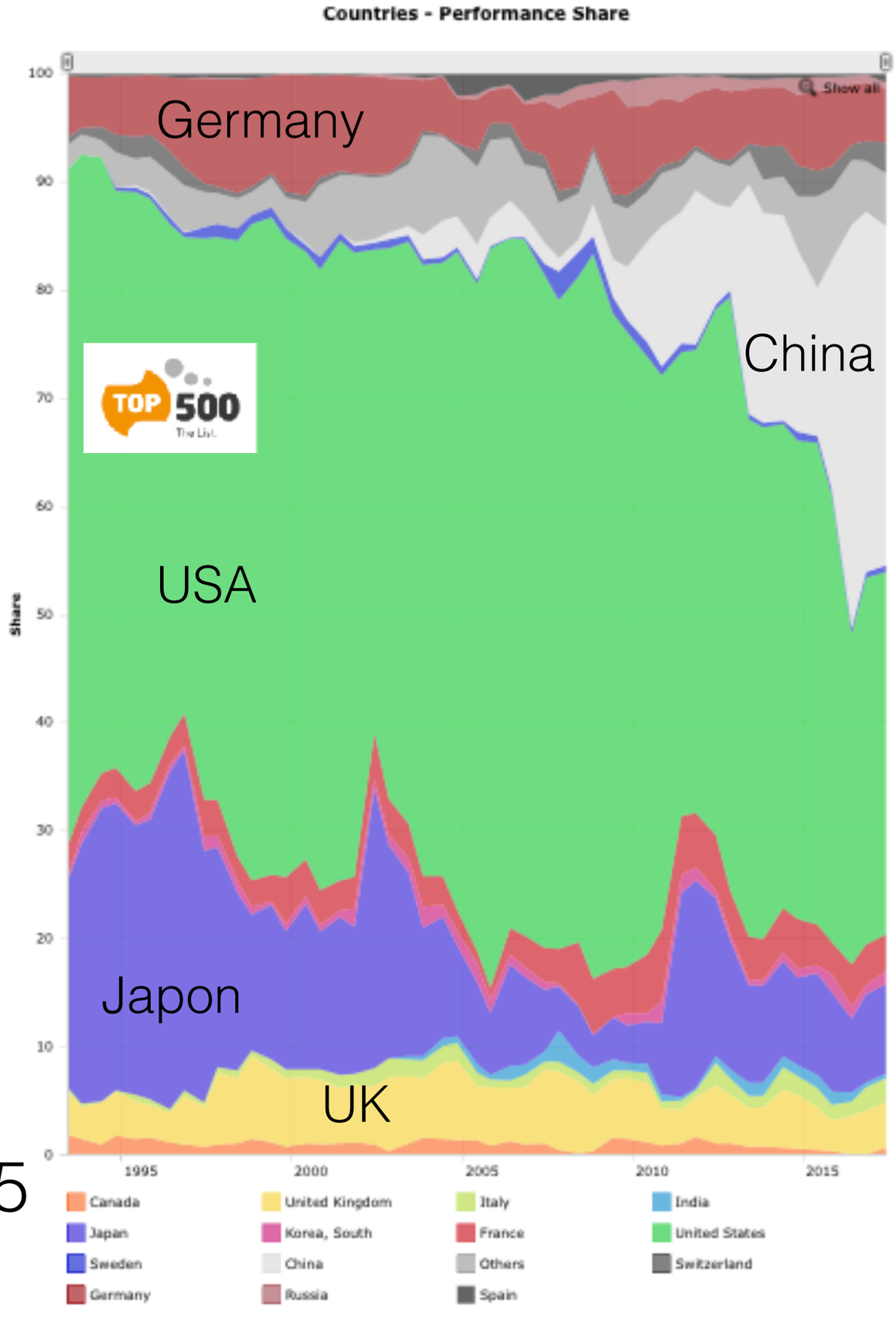
We need to adapt the current architecture but why not?

Outline

High performance computers :

A key player in science (and politics!) not exploited in our community

100%

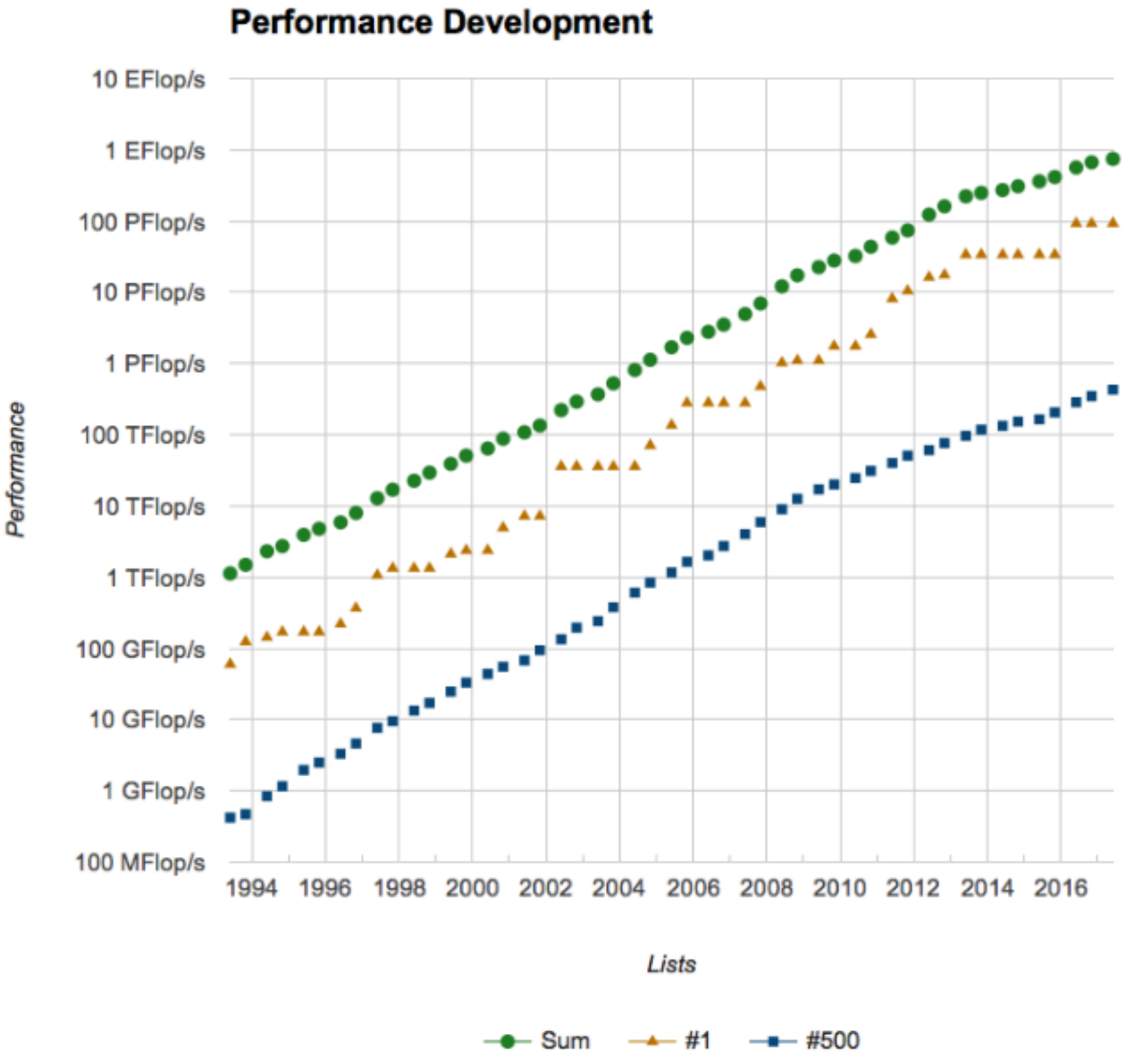


1995

2017

Geopolitical importance of supercomputers

The fall of the US and the raise of China


Performance Development
10 EFlop/s
1 EFlop/s
100 PFlop/s
10 PFlop/s
1 PFlop/s
100 TFlop/s
10 TFlop/s
1 TFlop/s
100 GFlop/s
10 GFlop/s
1 GFlop/s
100 MFlop/s
Performance
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
Lists
Sum
#1
#500

Geopolitically important????

NEWS

US blocks Intel from selling Xeon chips to Chinese supercomputer projects

By Michael Kan

U.S. Correspondent, IDG News Service | APR 10, 2015 5:25 AM PT

Technology

Supercomputers: Obama orders world's fastest computer

By Chris Baraniuk
Technology reporter

30 July 2015 | Technology

Share

In response to the US sanction, China introduced the Sunway TaihuLight supercomputer in 2016, which substantially outperforms the Tianhe-2, and now holds the title as the fastest supercomputer in the world while using completely domestic technology including the Sunway manycore microprocessor.[9]

China aims to build world's first exascale supercomputer prototype by end of 2017

Call it a super-supercomputer

by James Vincent | @jjvincent | Jan 19, 2017, 5:08am EST

So, with the developments of Cloud Computing and the political race for HPC, it seems like the perfect environment to review our computing architecture and profit from these resources

21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015) IOP Publishing
Journal of Physics: Conference Series **664** (2015) 092025 doi:10.1088/1742-6596/664/9/092025

Fine grained event processing on HPCs with the ATLAS Yoda system

Paolo Calafiura[1], Kaushik De[2], Wen Guan[3], Tadashi Maeno[4], Paul Nilsson[4], Danila Oleynik[2], Sergey Panitkin[4], Vakhtang Tsulaia[1], Peter Van Gemmeren[5] and Torre Wenaus[4] on behalf of the ATLAS Collaboration

[1]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA
[2]University of Texas at Arlington, 701 South Nedderman Drive, Arlington, TX 76019, USA
[3]University of Wisconsin, 1150 University Avenue, Madison, WI 53706, USA
[4]Brookhaven National Laboratory, PO Box 5000, Upton, NY 11973, USA
[5]Argonne National Laboratory, 9700 S. Cass Ave, Argonne, IL 60439, USA

E-mail: VTsulaia@lbl.gov

http://iopscience.iop.org/article/10.1088/1742-6596/664/9/092025/pdf

21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015) IOP Publishing
Journal of Physics: Conference Series **664** (2015) 062065 doi:10.1088/1742-6596/664/6/062065

The ATLAS Event Service: A new approach to event processing

P Calafiura[1], K De[2], W Guan[3], T Maeno[4], P Nilsson[4], D Oleynik[2], S Panitkin[4], V Tsulaia[1], P Van Gemmeren[5], and T Wenaus[4] on behalf of the ATLAS Collaboration

[1]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA
[2]University of Texas at Arlington, 701 S Nedderman Dr, Arlington, TX 76019, USA
[3]University of Wisconsin – Madison, Madison, WI 53706, USA
[4]Brookhaven National Laboratory, Upton, NY 11973, USA
[5]Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

E-mail: wenaus@gmail.com

http://iopscience.iop.org/article/10.1088/1742-6596/664/6/062065/pdf

ATLAS Computing NOW

TRADITIONAL WAY

ATLAS Physics Coordination requests n (millions) MC events

ATLAS Computer Farm

Grid Storage Element

event service

NEW PATHS IN HEP COMPUTING

ATLAS Physics Coordination requests n (millions) MC events

Event Service Splits de job in small ranges (even 1 event)

request

input

Grid Storage Element

amazon web services

ATLAS@HOME

Opportunistic resources

output

Object Store
Merging of events

Event Service

- Event processing granularity makes possible the use of opportunistic resources.
- The *flow of events :* one event in, processed and one event out optimises storage
- Now is being commissioned and validated still not fully deployed

Is this jar full?

Summary

- Computing is changing very fast : we must as well evolve
- New computing models requires new ways of thinking to tackle the problems we have to face
- The traditional approach : ATLAS own PC farms, won't become obsolete soon (probably never) but we must prepare ourselves now to be ready for the computing models of the future