Towards a data science infrastructure

Challenges and treats to afford the data deluge

Symposium on HPC & Data-Intensive Applications in Earth Science
November 14th, Trieste

Carlo Cavazzoni
Giuseppe Fiameni

http://www.wired.com/wired/issue/16-07

September 2008

The 4 paradigms of Scientific Research

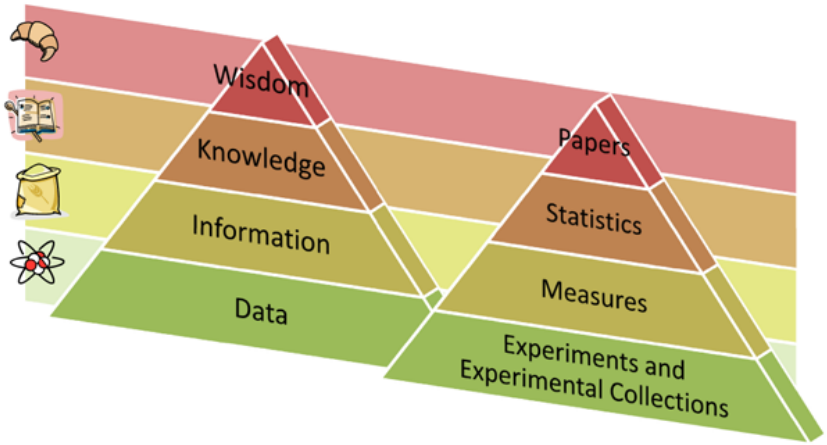
1. Theory
2. Experiment or Observation
3. Simulation of theory or model
   Supercomputers
4. Data-driven (Big Data) or The Fourth Paradigm: Data-Intensive Scientific Discovery (aka Data Science)
   - More data; less models

The Shift Towards a data centric view

- All science is becoming data-dominated
  - Experiment, computation, theory
  - Fourth paradigm
    - http://research.microsoft.com/en-us/collaboration/fourthparadigm/
- Classes of data
  - Collections, observations, experiments, simulations
  - Software
  - Publications
- Totally new methodologies
  - Algorithms, mathematics, culture
- Data become the medium for multidisciplinarity science

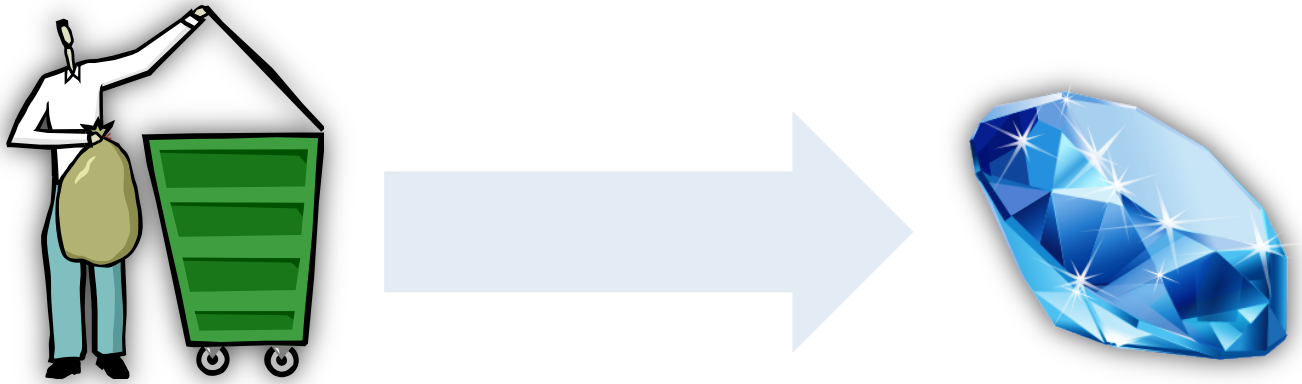
DIKW Process

- ✓ **Data** becomes
- ✓ **Information** becomes
- ✓ **Knowledge** becomes
- ✓ **Wisdom** or **Decisions**
  - Community acceptance of results is important (experiments reproducibility)
  - Volume of bits&bytes decreases as we proceed down DIKW pipeline

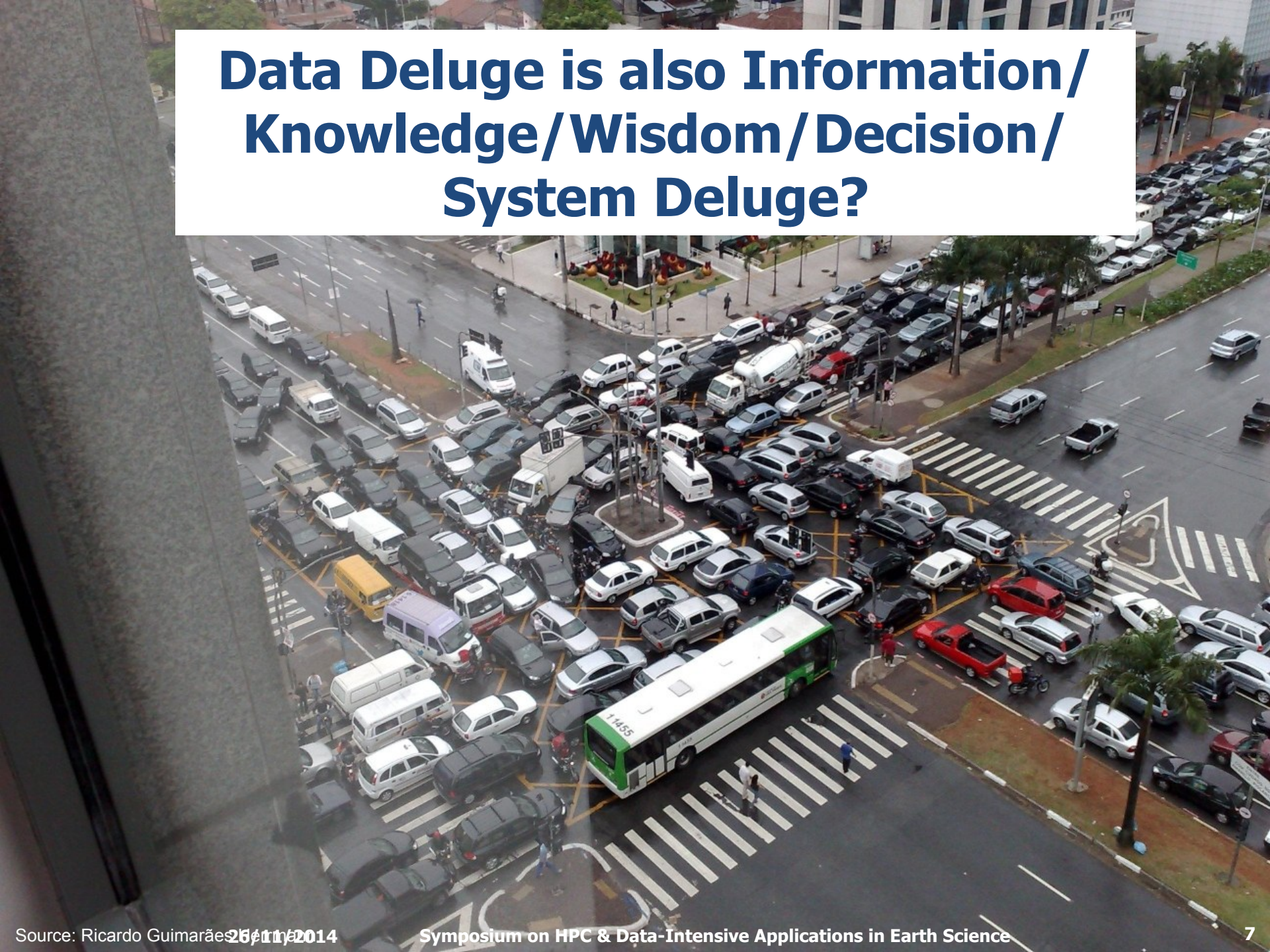


Data as an asset

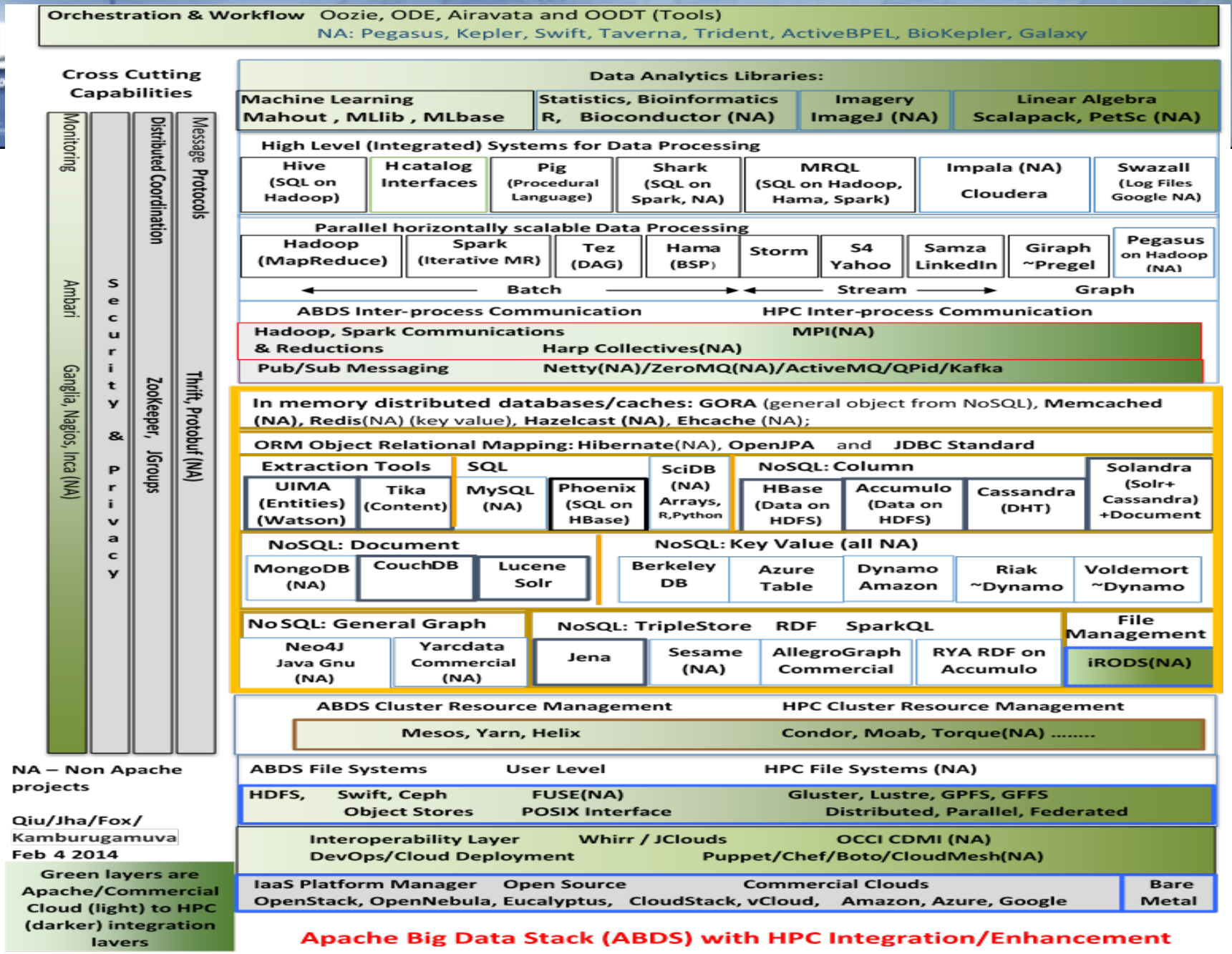
"One of the most significant changes of the past decade has been the widespread recognition of data as an asset rather than the refuse of research."



Data Deluge is also Information/ Knowledge/Wisdom/Decision/ System Deluge?



Source: Ricardo Guimarães Herrmann

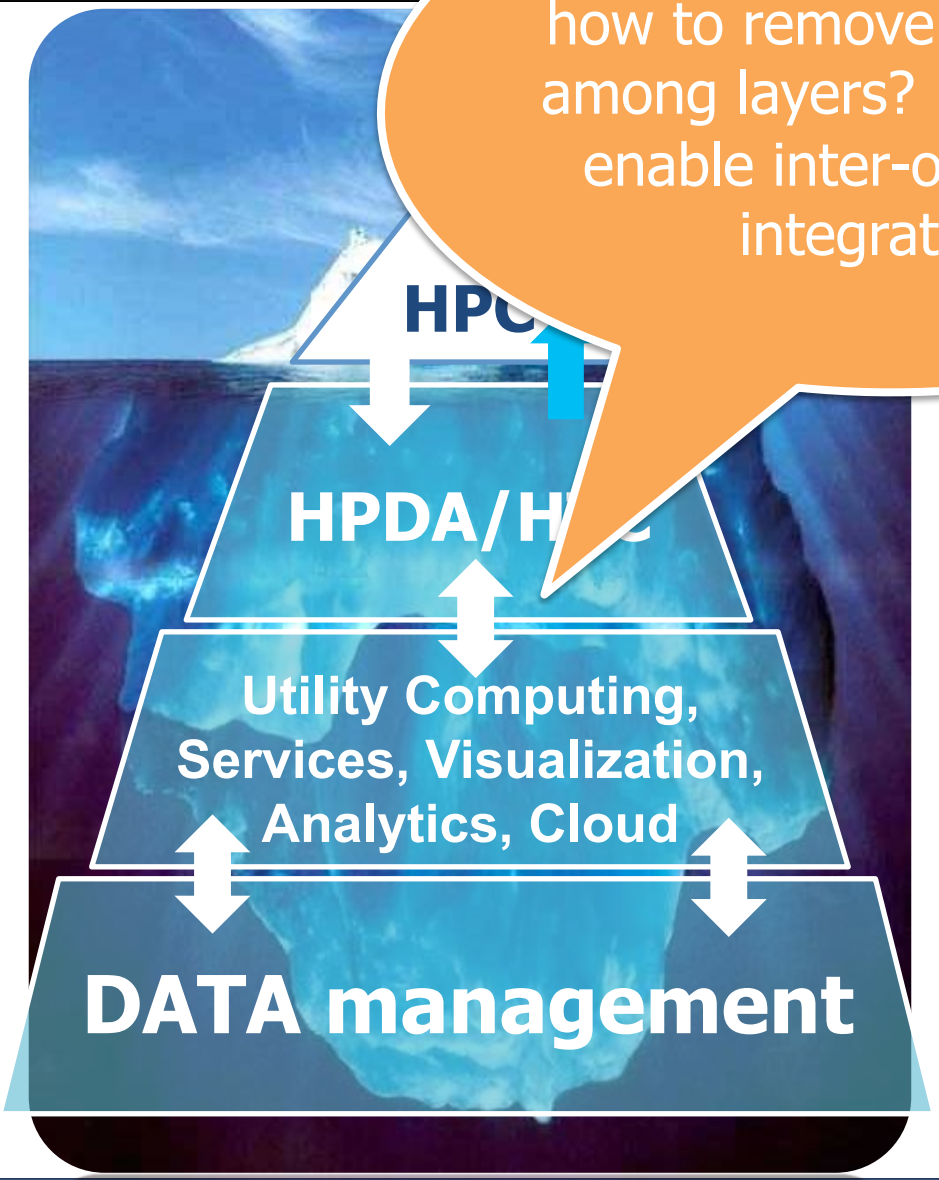


Apache Big Data Stack (ABDS) with HPC Integration/Enhancement

*Courtesy of **Geoffrey Fox***

Pyr...berg

Fundamental question: how to remove boundaries among layers? How do we enable inter-operability/ integration?



Velocity

Variety

Science Computing Environments

- **Large Scale Supercomputers** – multicore nodes linked by high performance low latency network
  - Increasingly with GPU/Accelerator enhancement
  - Suitable for highly parallel simulations
- **High Throughput Systems** – cluster/farm typically aimed at pleasingly parallel jobs
  - Classic example is **LHC data analysis**
- **High Performance Data Analytics -** combination of high computation resources and analytics methods
  - MapReduce
  - Graph
  - Relation and NoSQL database
- **Federation of** resources to enable convenient access to multiple backend systems including supercomputers
- Use **Services** (**SaaS**)
  - **Portals** make access convenient and
  - **Workflow** integrates multiple processes into a single job
  - **Visualization** either remote or in-situ

Classic HPC systems do help

- 3D physics simulations
- Linear Algebra
- FLOPS

HPC

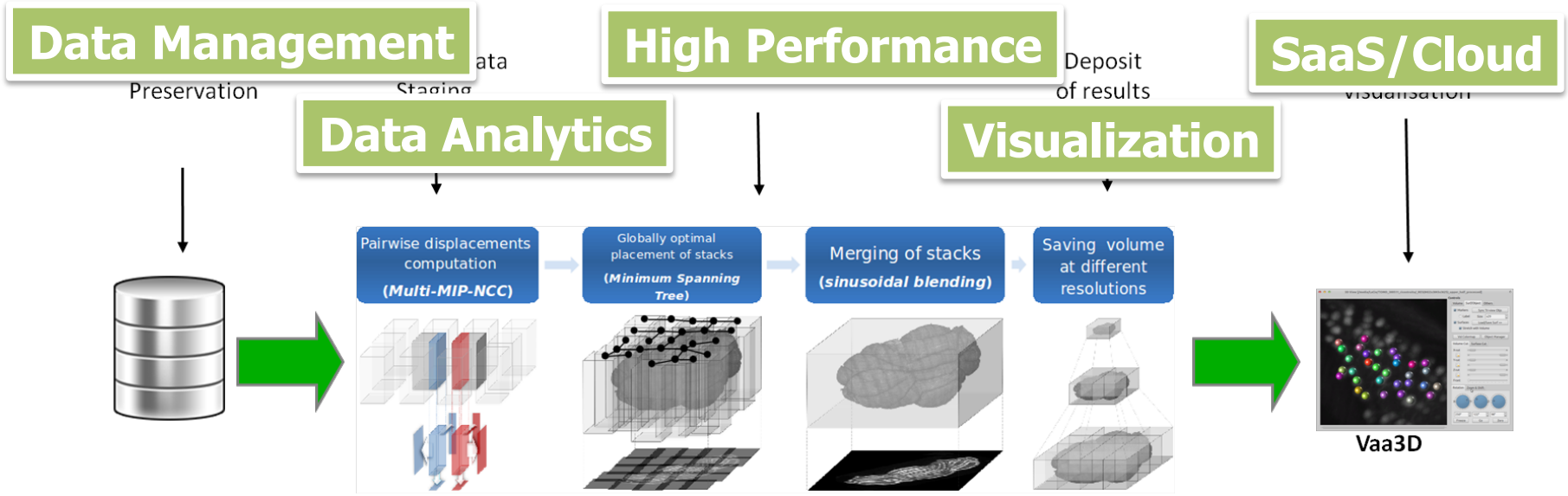
How to overcome this bottleneck? (Non volatile memory, SSD, MapReduce)

- Graph oriented algorithm
- GTEPS/IOPS

Data Analytics

Emerging use case: the HBP case study

- **Human Brain Project: understand the human brain**
  - six ICT platforms to **Neuroinformatics**, Brain Simulation, High Performance Computing, Medical Informatics, Neuromorphic Computing and Neurorobotics

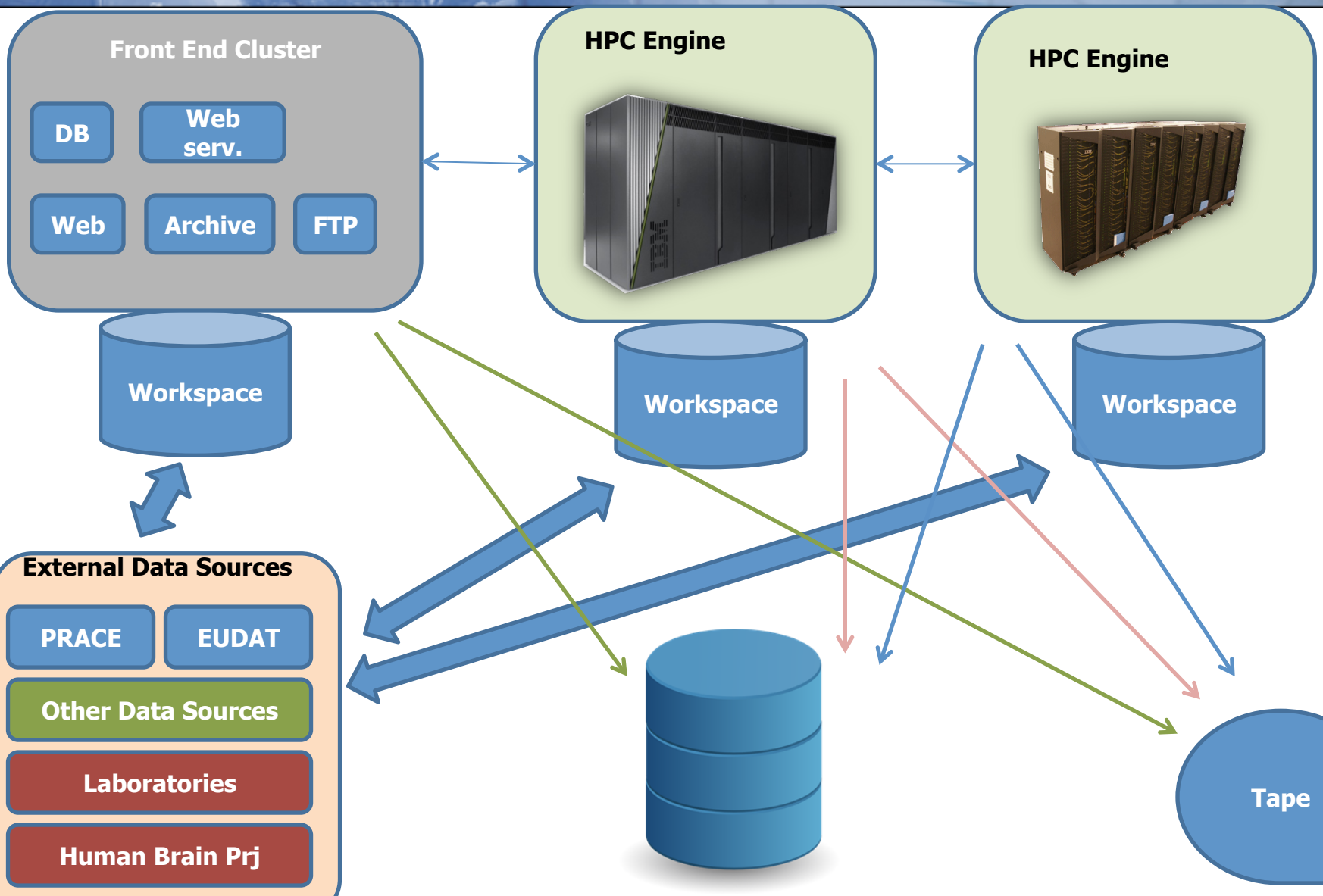


Other emerging use cases

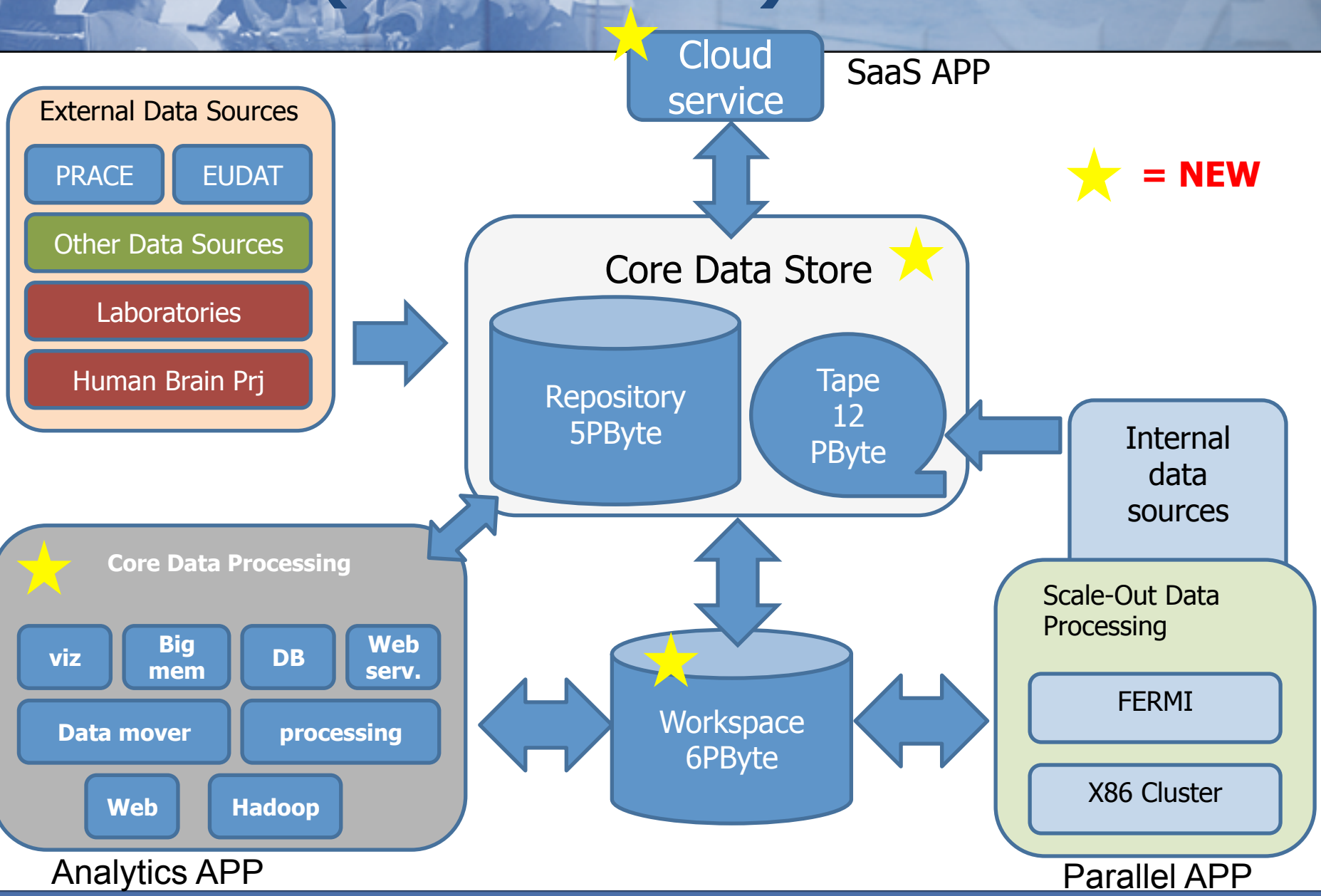
- ✓ Neuroimage processing
- ✓ Genome data processing (NGS)
- ✓ Geo Science
- ✓ High Throughput Material Science
- ✓ Pharmacy, Oil&Gas and Finance
- ✓ Astrophysics

They require more and more data processing & analytics
Together with the typical HPC Number crunching
Work-load

HPC "island" Infrastructure



(data centric) Infrastructure

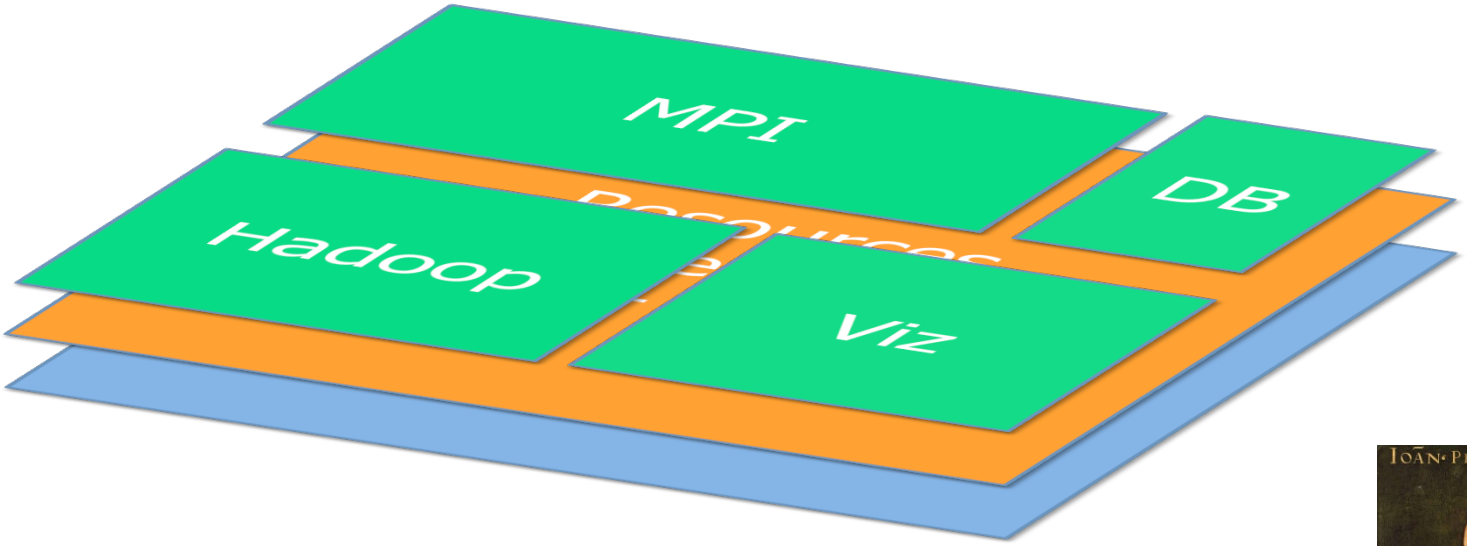


The PICO system

	Total Nodes	CPU	Cores per Nodes	Memory (RAM)	Notes
Compute login node	66	Intel Xeon E5 2670 v2 @2.5Ghz	20	128 GB	
Visualization node	2	Intel Xeon E5 2670 v2 @ 2.5Ghz	20	128 GB	2 GPU Nvidia K40
Big Mem node	2	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16	512 GB	1 GPU Nvidia K20
BigInsight node	4	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16	64 GB	32TB of local disk
SSD Storage					40 TB

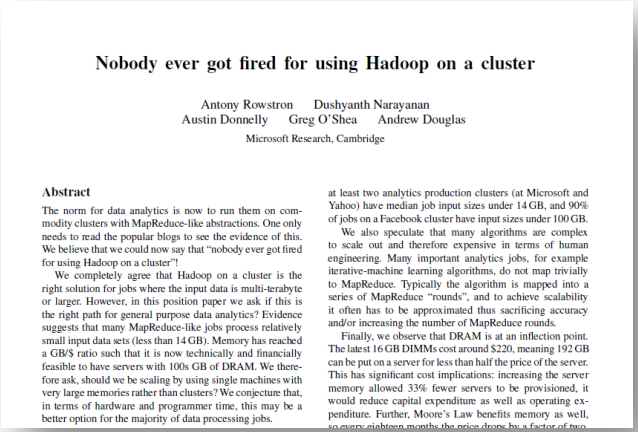
http://www.hpc.cineca.it/hardware/pico

The PICO System



New Services

- **Data management**
  - Preservation
  - Sharing
  - Moving
- **Data analytics ☺**
  - MPI/OpenMP
  - Higher GTEPS/IOPS
  - MapReduce (Hadoop/Spark)
- **Relational and NoSQL database**
- **Visualization**
- **Cloud**
  - Support for RDMA
  - Various O.S. (Microsoft)
- **Higher integration among the systems**
  - Everything is connected through IB
- ??



Nobody ever got fired for using Hadoop on a cluster

Antony Rowstron Dushyanth Narayanan
Austin Donnelly Greg O'Shea Andrew Douglas
Microsoft Research, Cambridge

Abstract

The norm for data analytics is now to run them on commodity clusters with MapReduce-like abstractions. One only needs to read the popular blogs to see the evidence of this. We believe that we could now say that "nobody ever got fired for using Hadoop on a cluster"!

We completely agree that Hadoop on a cluster is the right solution for jobs where the input data is multi-terabyte or larger. However, in this position paper we ask if this is the right path for general purpose data analytics? Evidence suggests that many MapReduce-like jobs process relatively small input data sets (less than 14 GB). Memory has reached a GB/$ ratio such that it is now technically and financially feasible to have servers with 100s GB of DRAM. We therefore ask, should we be scaling by using single machines with very large memories rather than clusters? We conjecture that, in terms of hardware and programmer time, this may be a better option for the majority of data processing jobs.

at least two analytics production clusters (at Microsoft and Yahoo) have median job input sizes under 14 GB, and 90% of jobs on a Facebook cluster have input sizes under 100 GB.

We also speculate that many algorithms are complex to scale out and therefore expensive in terms of human engineering. Many important analytics jobs, for example iterative-machine learning algorithms, do not map trivially to MapReduce. Typically the algorithm is mapped into a series of MapReduce "rounds", and to achieve scalability it often has to be approximated thus sacrificing accuracy and/or increasing the number of MapReduce rounds.

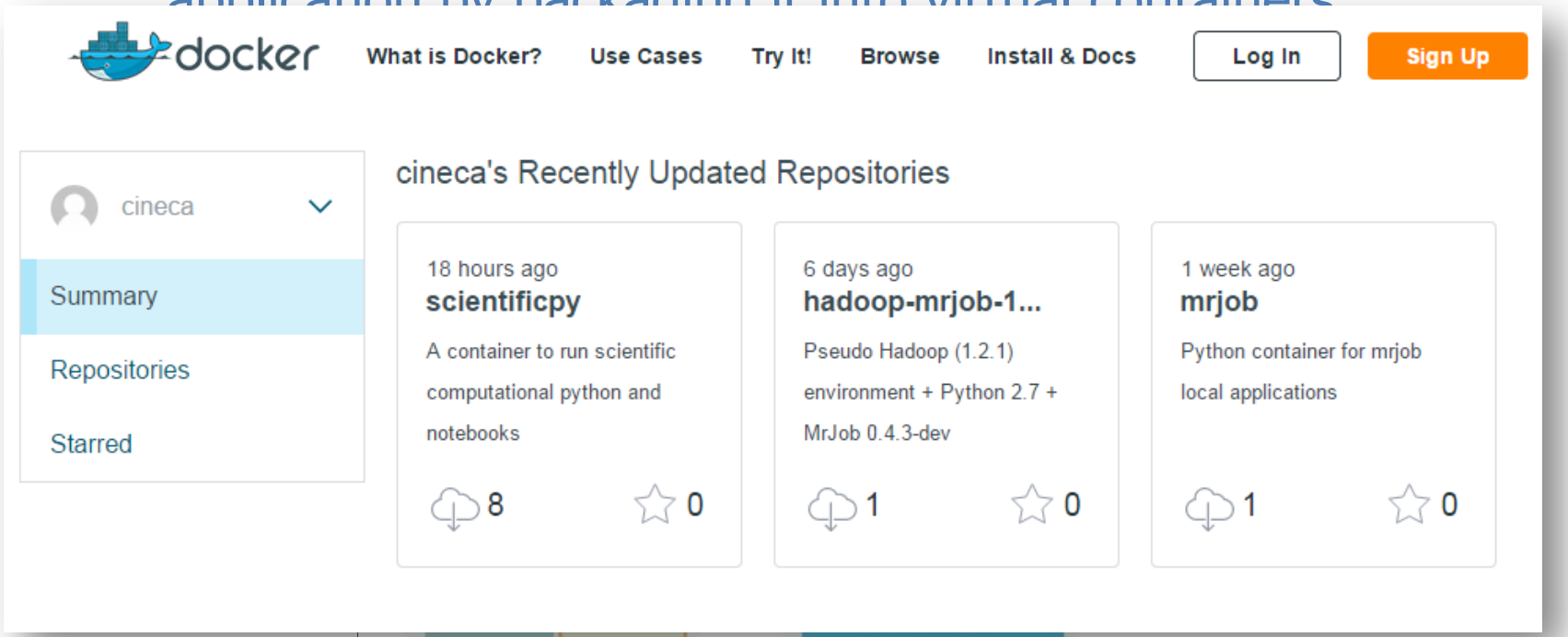
Finally, we observe that DRAM is at an inflection point. The latest 16 GB DIMMs cost around $220, meaning 192 GB can be put on a server for less than half the price of the server. This has significant cost implications: increasing the server memory allowed 33% fewer servers to be provisioned, it would reduce capital expenditure as well as operating expenditure. Further, Moore's Law benefits memory as well,

What Applications work in Clouds

- **Embarrassing parallel** applications with roughly independent data or spawning independent simulations
  - **Long tail** of science and integration of distributed sensors
- **Commercial and Science Data analytics**
  - that can use MapReduce and need an **interactive** environment
  - that use legacy applications
- **Industries** that require high isolation (tenancy)

Virtualization is dead, long live to virtualization

- ✓ HPC in a container
  - light weight virtualization mechanism which isolates each application by packaging it into virtual containers



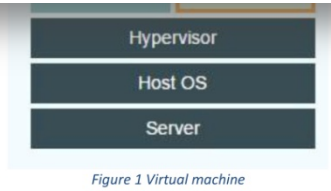


Figure 1 Virtual machine

Figure 2 Docker Containers

http://www.docker.com/whatisdocker

Conclusion

"Large-scale science and engineering problems **require collaborative use of many compute and data resources**, including supercomputers and large-scale data storage systems, all of which **must be integrated with applications and data that are developed by different teams of researchers** or that are obtained from different instruments and all of which are at different geographic locations."

Initiatives

- Tools and techniques for massive data analysis , Cineca Bologna (PATC Course), Dec 15th-16th 2014 (complete ☹)
- PICO: the CINECA solutions for Big Data Science , Cineca-Bologna, Dec 5th 2014
- Call for interest at this link . Deadline is Monday, November 17th at 9:00 am **http://goo.gl/Wtv8oM**

The EUDAT Project

- European Data Infrastructure for addressing data management challenges at EU level
- 5 core services
  - Long Term Preservation
  - Staging towards computational facility
  - Sharing
  - Search through meta-data
  - Deposit (like DropBox)
- Many scientific communities involved
  - EPOS
  - ENES
  - LifeWatch
  - .....

EUDAT

Thanks for your attention!

www.hpc.cineca.it