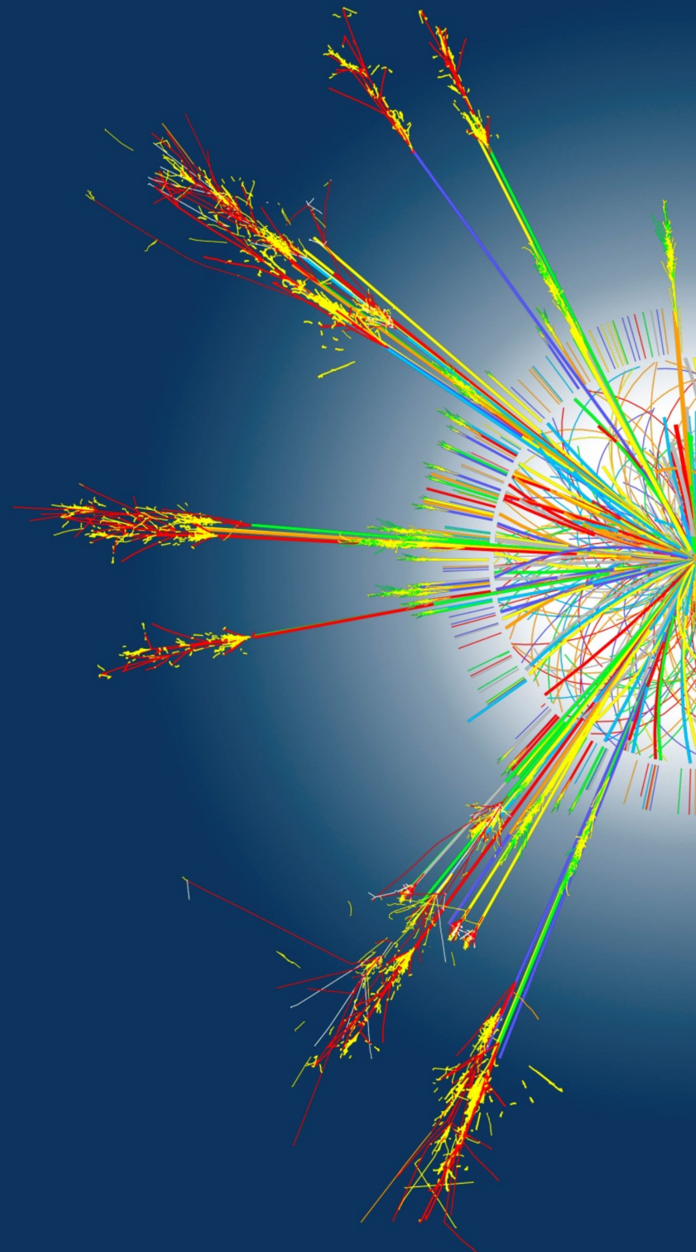
Principles of distributed computing in data-intensive science

Oxana Smirnova
div. of Particle Physics, LU

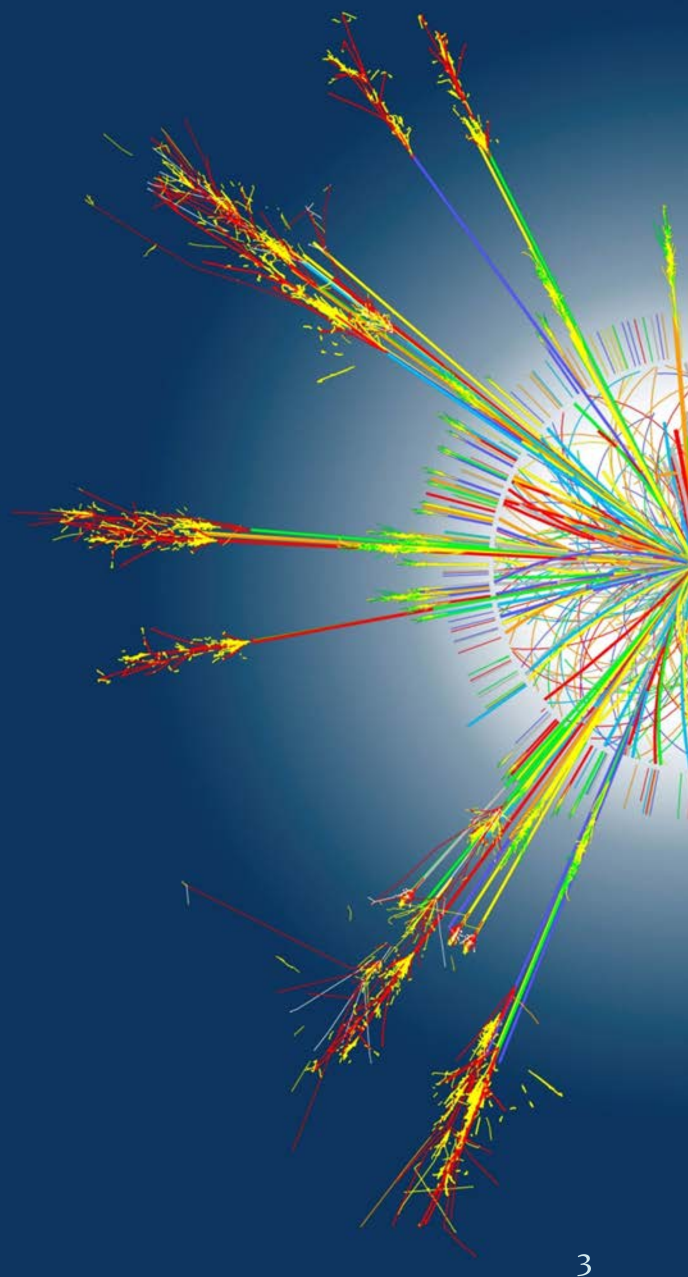
COMPUTE retreat,
20 August 2012, hotel ÅhusStrand



Outline

- E-Science and Big Data
- Context: Large Hadron Collider as a data-intensive research pioneer
- Computing challenges at LHC
- Distributed computing approach and solutions

Big Data



Evolution of science paradigms

Empirical science

- Descriptive

Theoretical science

- Models

Computational science

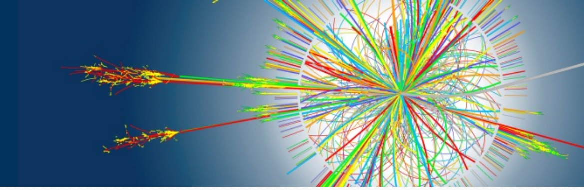
- Simulations

Data exploration (**e-Science**)

- Unifies the rest to explore **Big Data**

after Jim Gray

Data tsunami



All scientific research needs **data** (testing models, finding patterns)

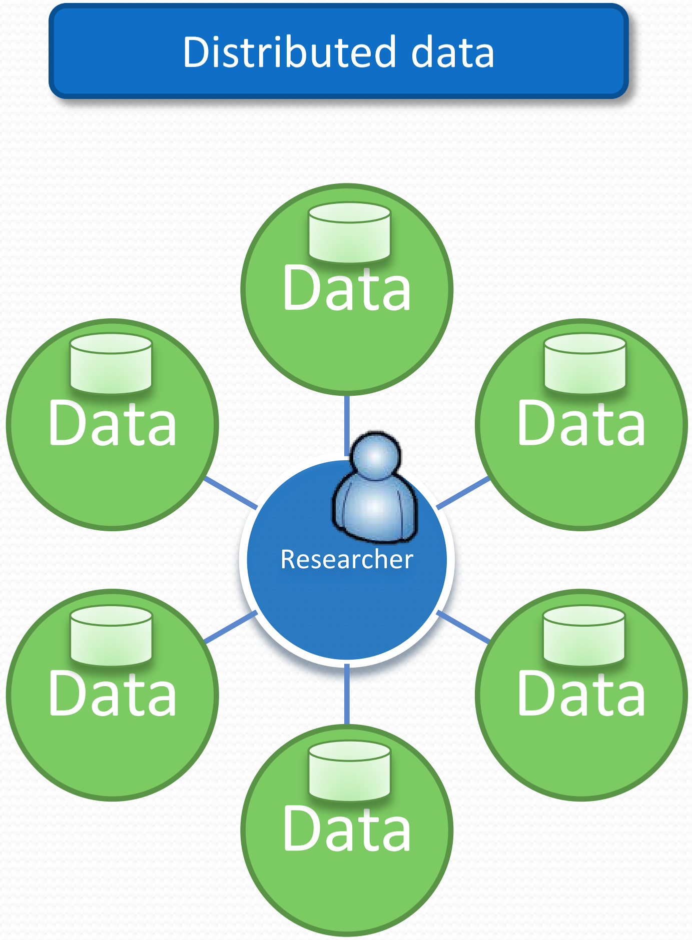
All data and information is **digitised**

Modern instruments produce digital data in **huge amounts**

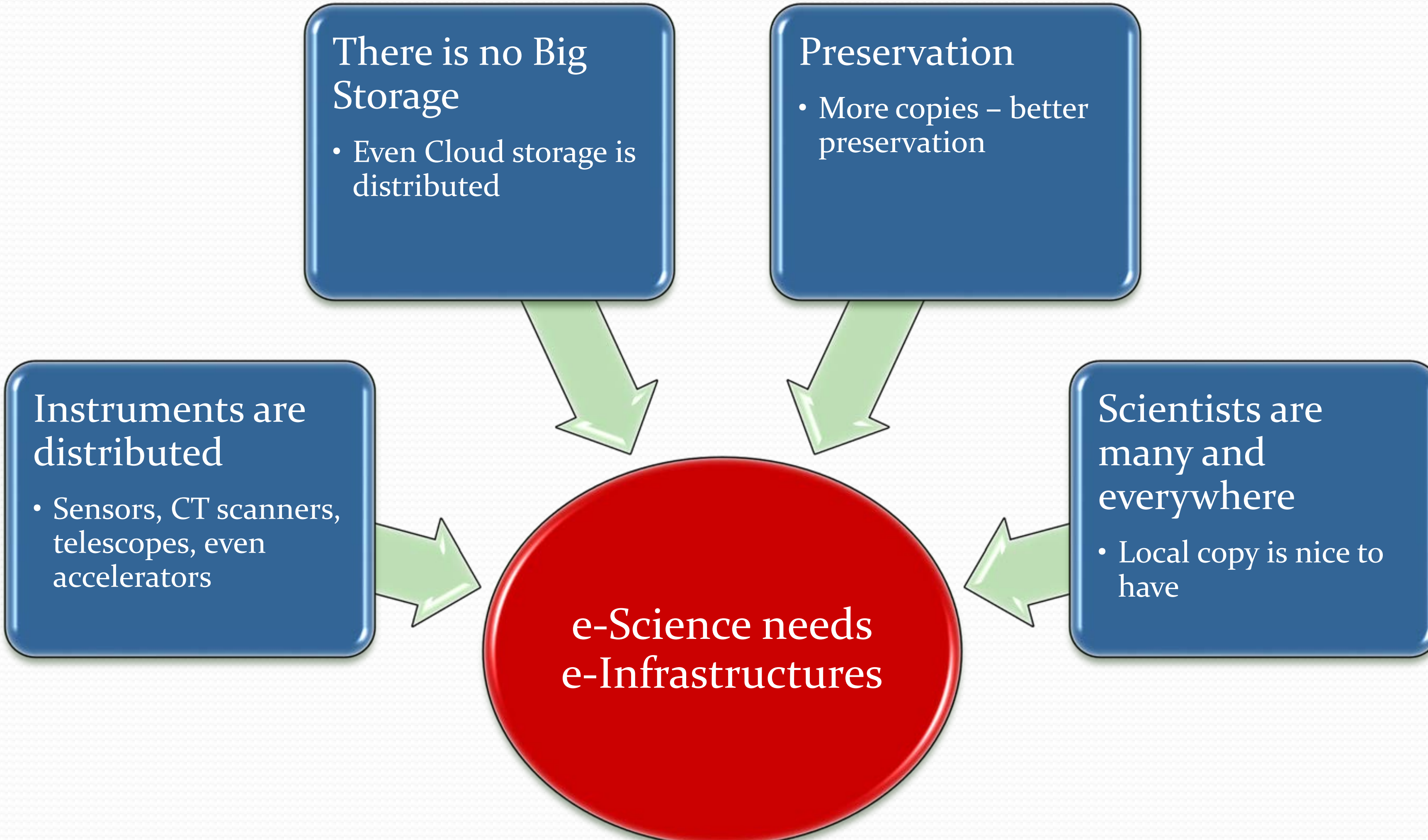
Instruments and data are **accessible by all scientists** on the planet

Modern science is impossible without **global distributed infrastructures** to handle data

Different data access concepts



Why are Big Data distributed?

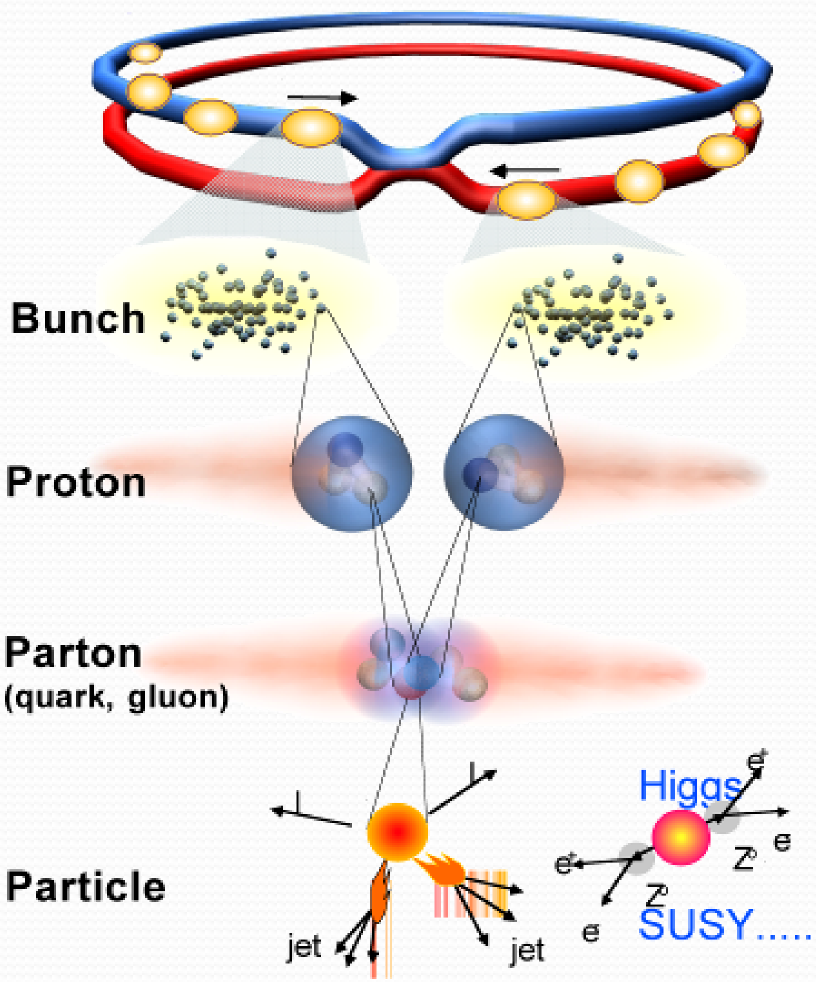


LHC as Big Data pioneer

Short-list of questions for LHC

- Can **gravity** be included in a theory with the other three interactions?
- Why do the particles have the **masses** we observe, and what is the origin of mass?
- How many space-time **dimensions** do we live in?
- Are the known elementary particles fundamental or do they possess **structure**?
- Why is the electrical charge of the electron **equal** and opposite to that on the proton?
- Why are there **three** generations of quarks and leptons?
- Why is there overwhelmingly more **matter** than anti-matter in the Universe?
- Are protons **unstable**?
- What is the nature of the **dark matter and dark energy** that pervade our Galaxy?
- Are there **new states** of matter at exceedingly high density and temperature?
- Do the **neutrinos** have mass, and if so why are they so light?

How LHC produces data



Proton-Proton	2835 bunch/beam
Protons/bunch	10^{11}
Beam energy	7 TeV (7×10^{12} eV)
Luminosity	10^{34} $cm^{-2}s^{-1}$
Crossing rate	40 MHz
Collisions rate ≈	10^{7} - 10^{9} Hz

New physics rate ≈ .00001 Hz

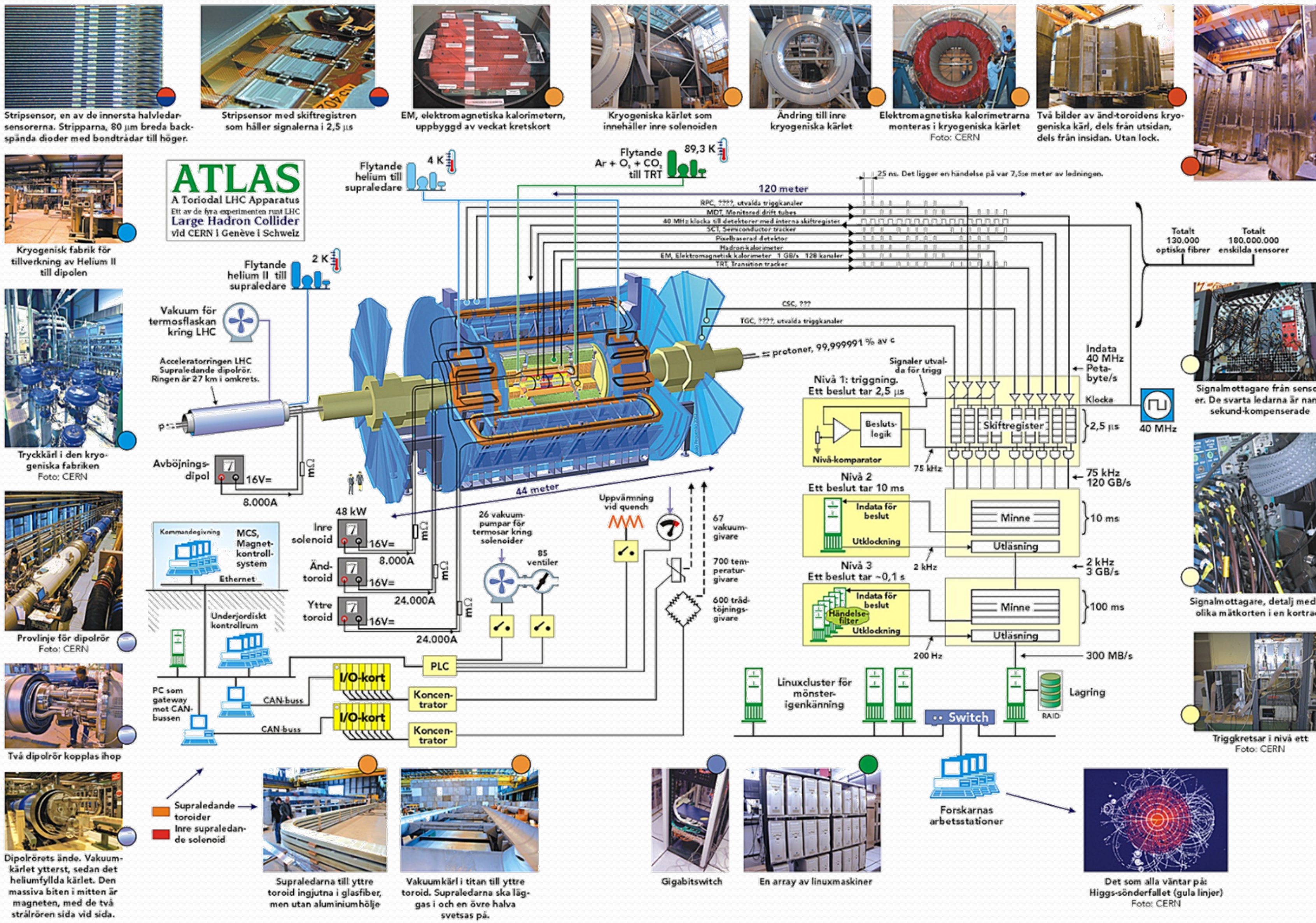
Event selection:
1 in 10,000,000,000,000

Graphics by CERN

A data sample: collision event at LHC

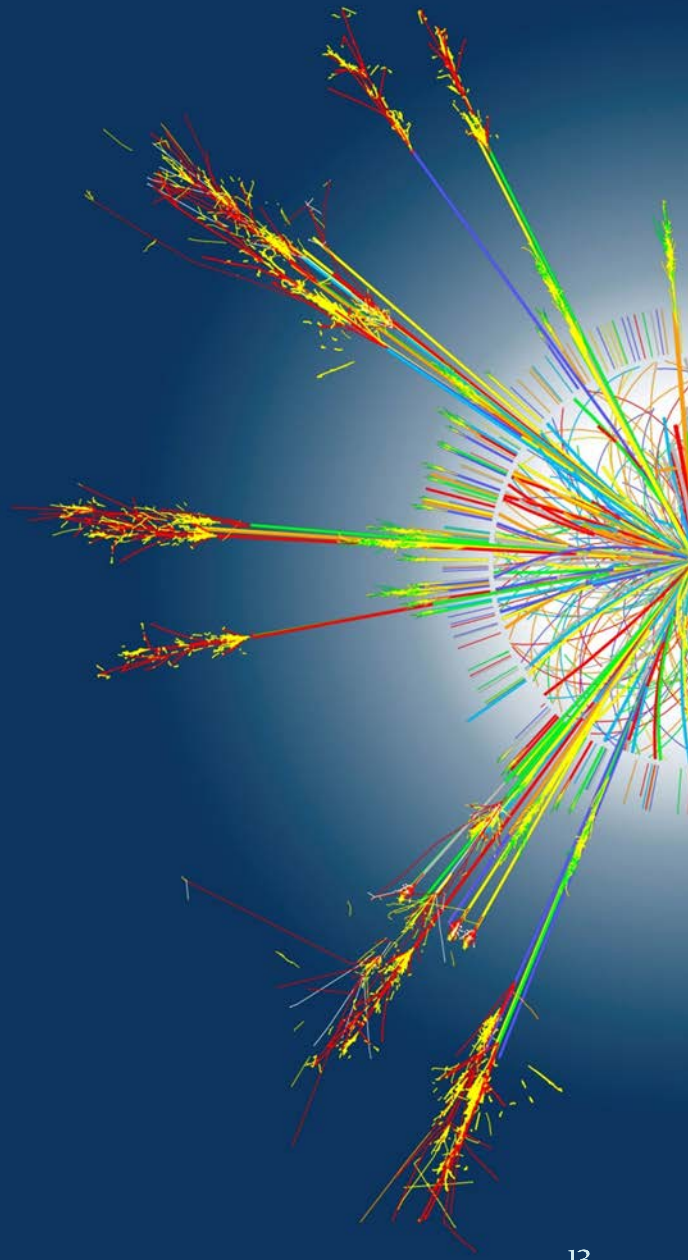
ATLAS video

An instrument at LHC: ATLAS

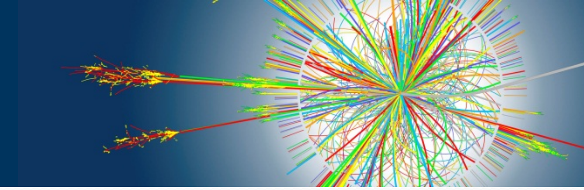


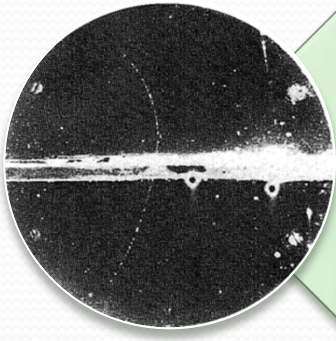
Graphics by QED AB

LHC computing



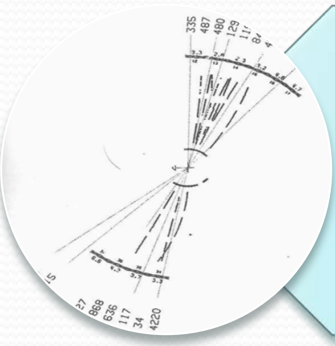
Dealing with data





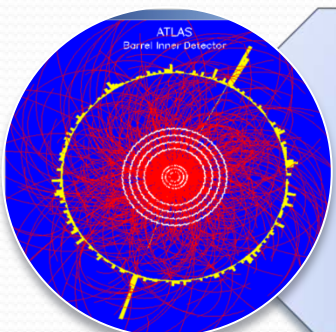
A discovery in 1930-ies

- ~2 scientists in 1country
- pen-and-paper



A discovery in 1970-ies

- ~200 scientists in ~10 countries
- mainframes

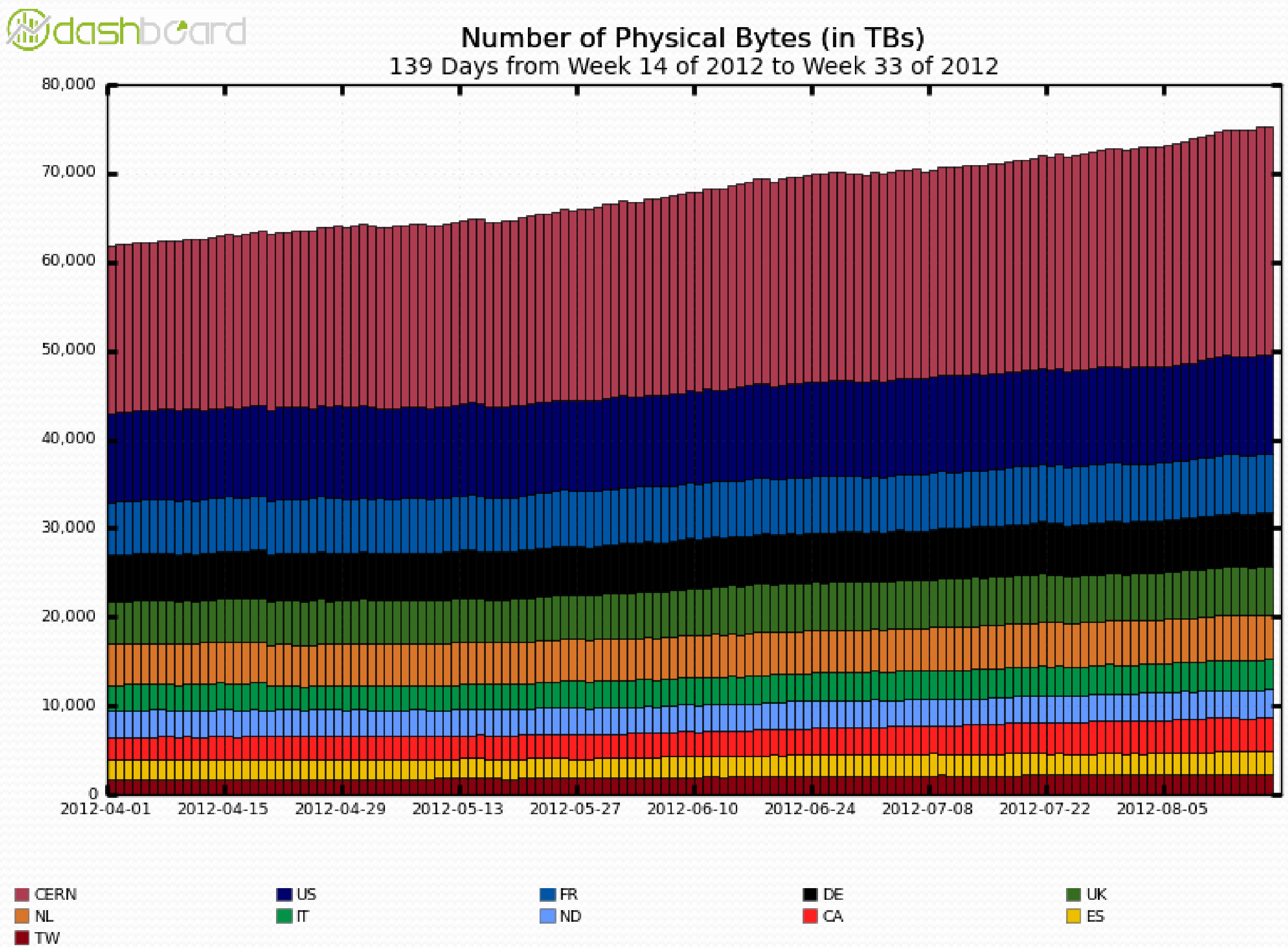


A discovery today

- ~2000 scientists in ~100 countries
- **Grids**

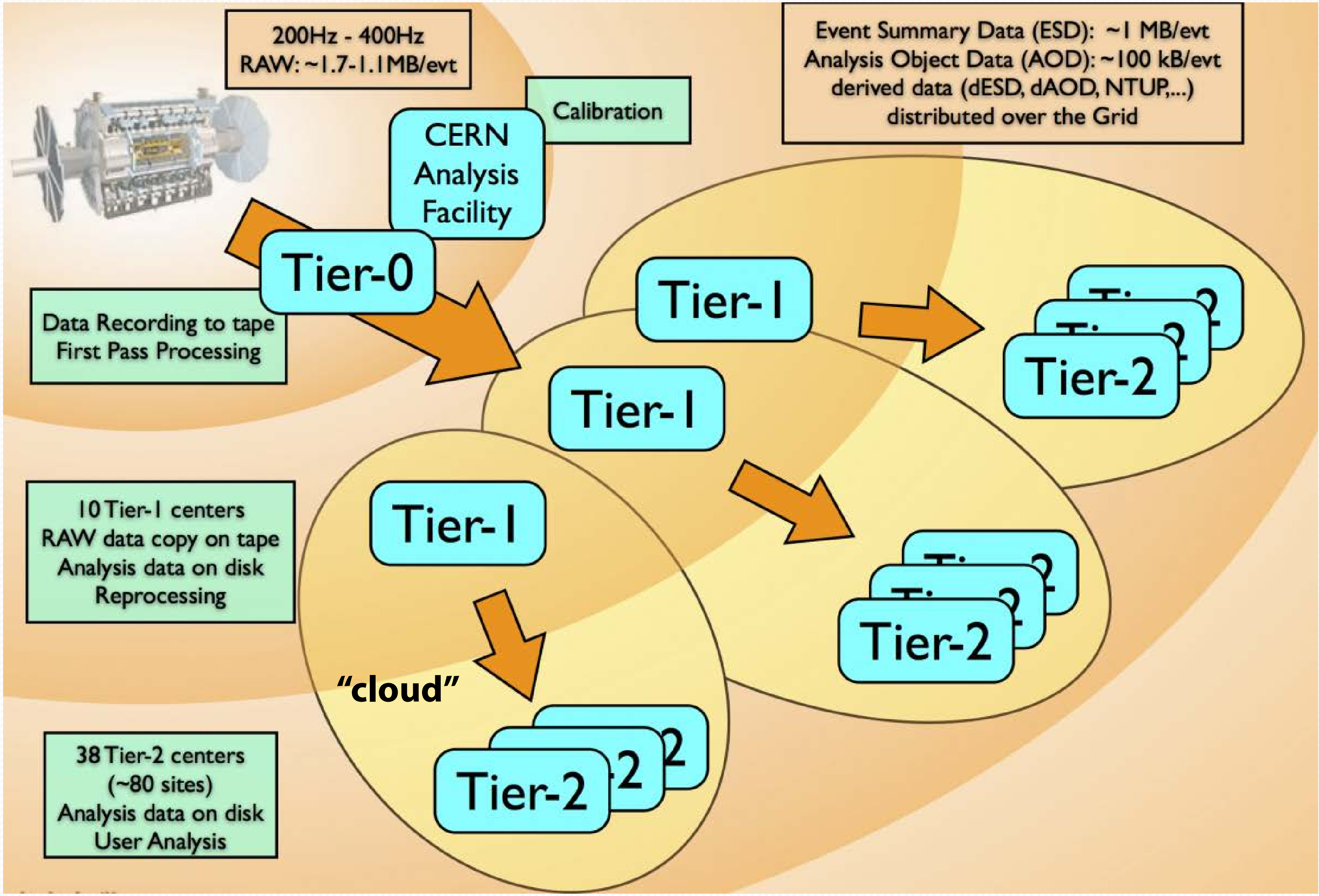
- Relatively simple algorithms
- Analysis is split in very many computing jobs
- Easy to distribute data and jobs
- Distributed computing: Grid

How big is Big?



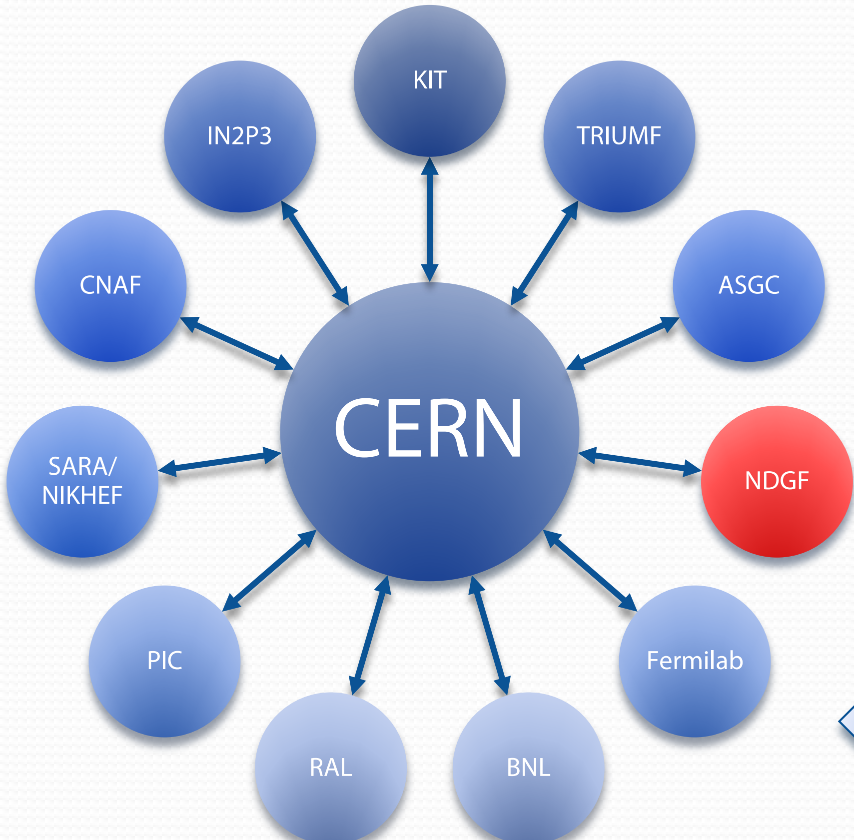
Data collected by ATLAS so far, stored in regional "clouds"

Distributing collected data



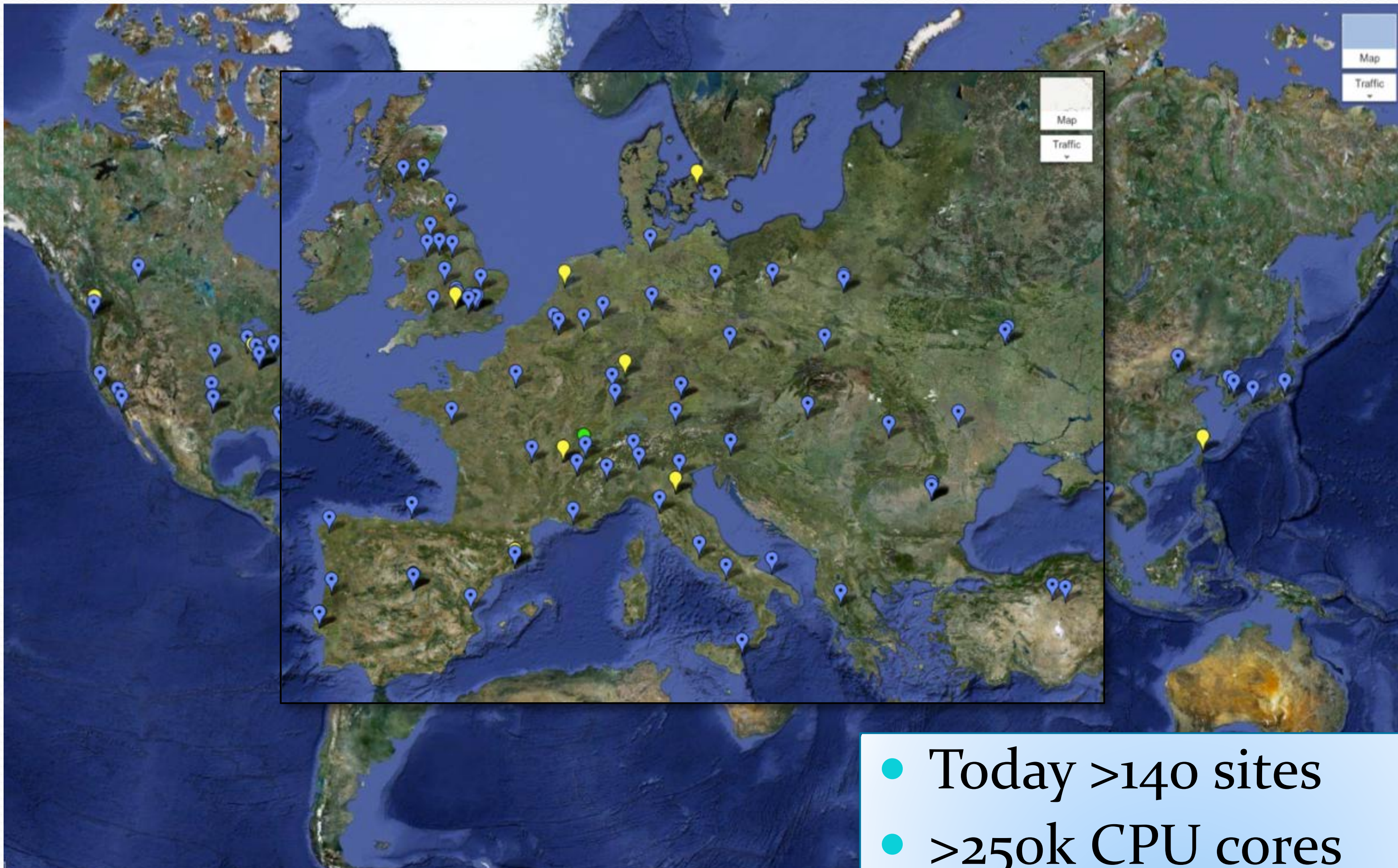
Graphics by iSGTW

Computing infrastructure for LHC



- ***WLCG*** – *Worldwide LHC Computing Grid*
  - Federates worldwide computing and storage resources
  - Started operating in 2005
- Hierarchical structure
  - **Tier0 at CERN**
    - **11 Tier1 sites** (connected by the private network)
      - **138 Tier2 sites** (research networks)

WLCG Grid sites



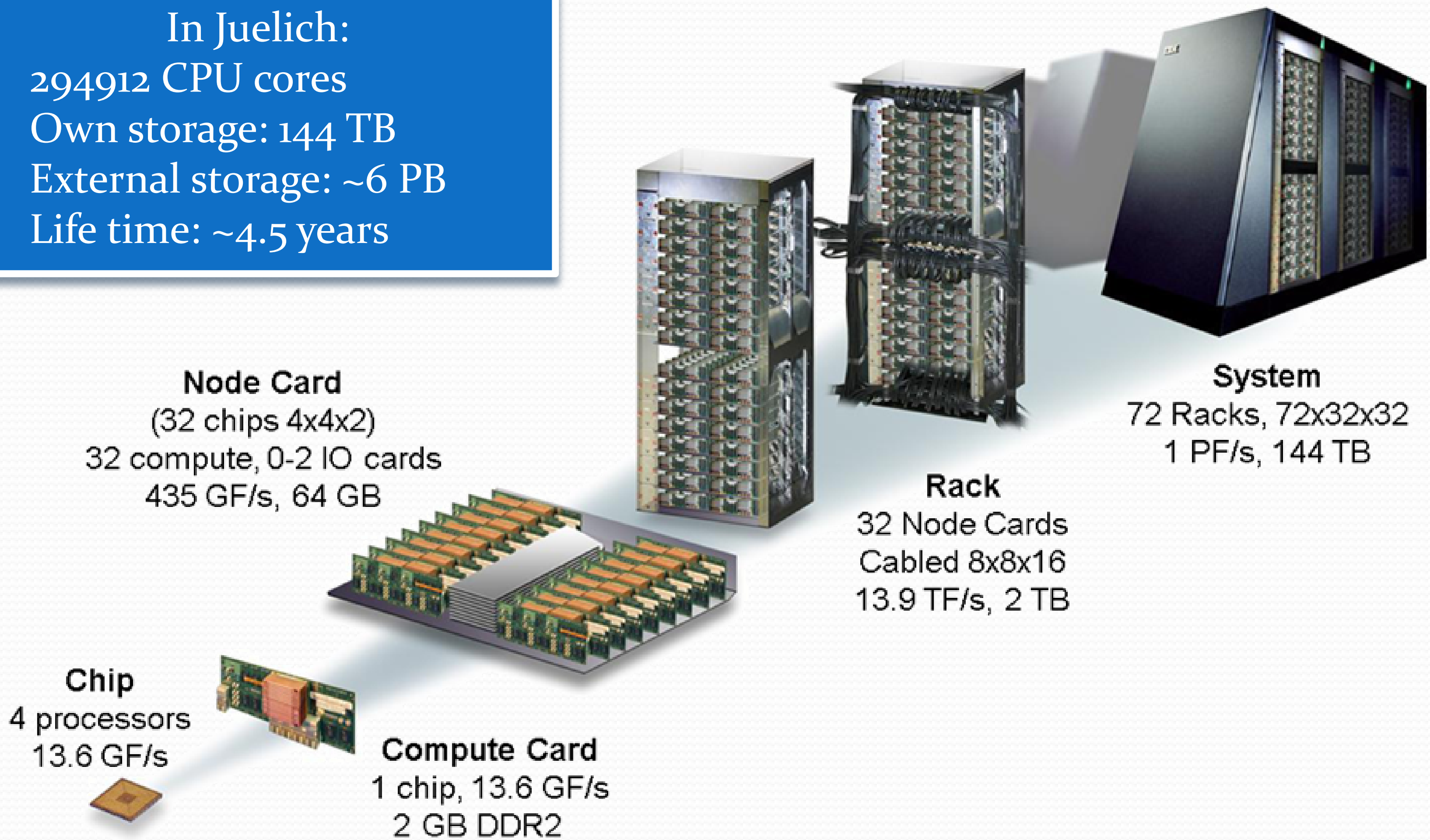
- Today >140 sites
- >250k CPU cores
- >150 PB disk

Tier 0 Tier 1 Tier 2

For comparison: Blue Gene/P

In Juelich:

- 294912 CPU cores
- Own storage: 144 TB
- External storage: ~6 PB
- Life time: ~4.5 years



Graphics by IBM

Grid looks for Higgs

Global Effort → Global Success

Results today only possible due to extraordinary performance of accelerators – experiments – Grid computing

Observation of a new particle consistent with a Higgs Boson (but which one…?)

Historic Milestone but only the beginning

Global Implications for the future

R-D Heuer

Slide by Rolf Heuer, 4 July 2012

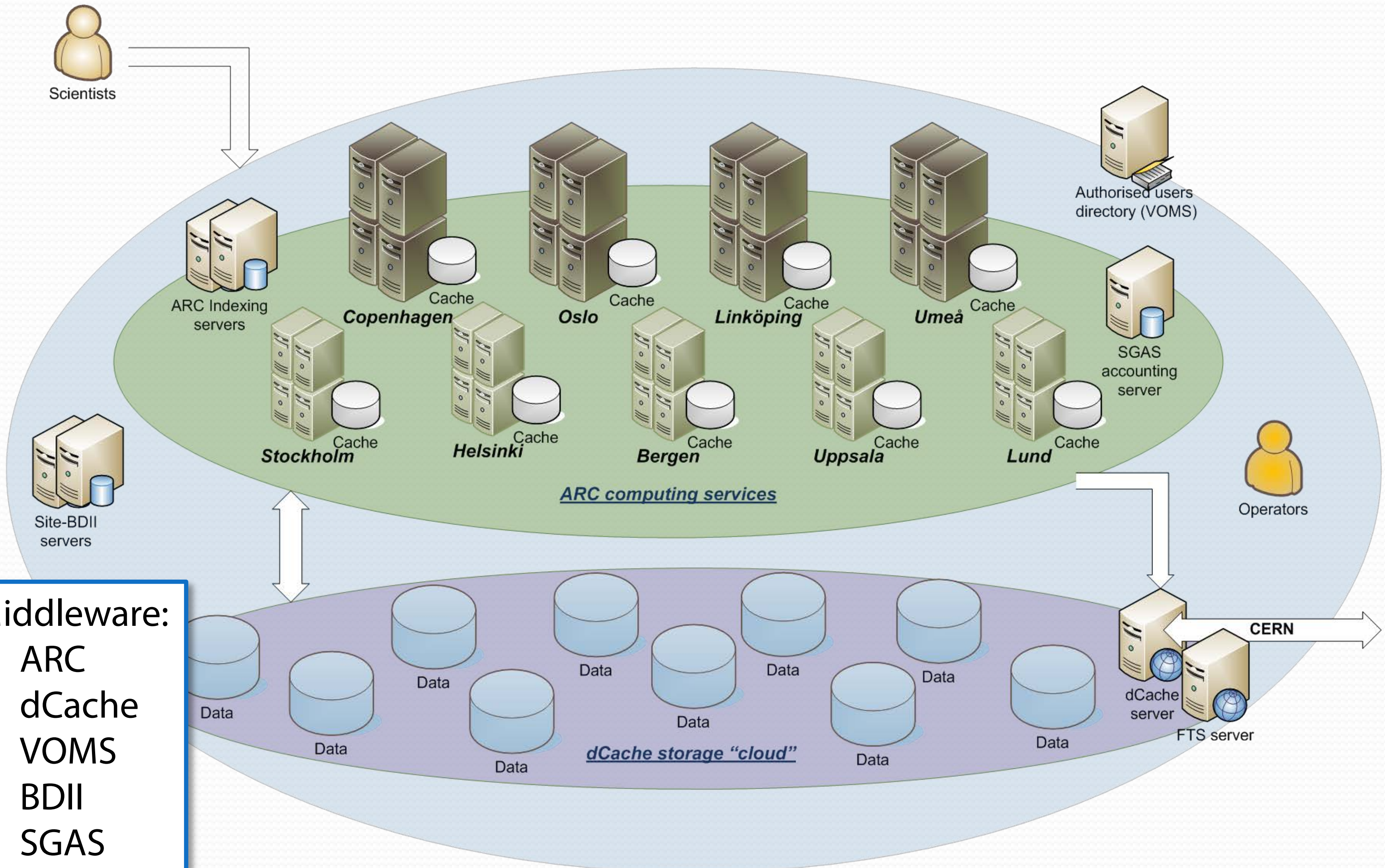
So, what is Grid?

- Many computing and storage systems?
- Network connections?

Grid is a federation of heterogeneous conventional systems, enabled by fast networks and a middleware layer that provides single sign-on and delegation of access rights through common interfaces for basic services

- *Middleware* is software that federates services and creates e-infrastructures
- There are many middlewares and many e-infrastructures

Nordic Grid infrastructure: NDGF



Middleware:

- ARC
- dCache
- VOMS
- BDII
- SGAS
- FTS

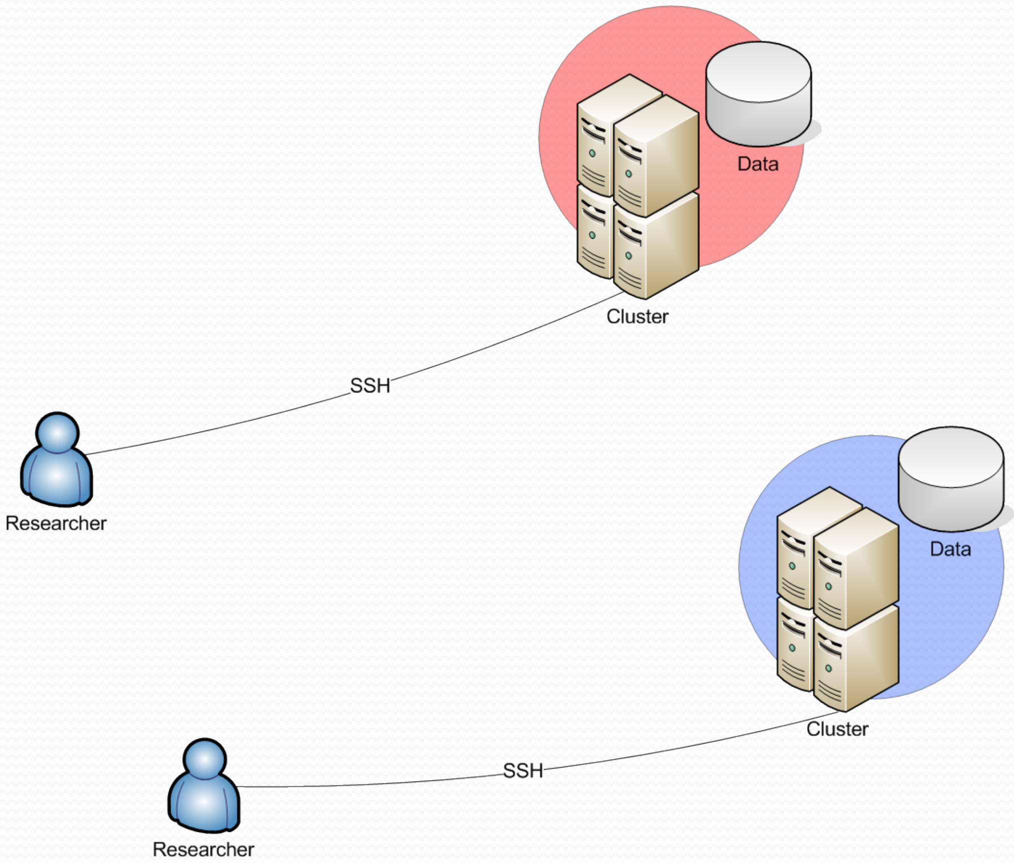
Middleware solutions

What problems does middleware solve?

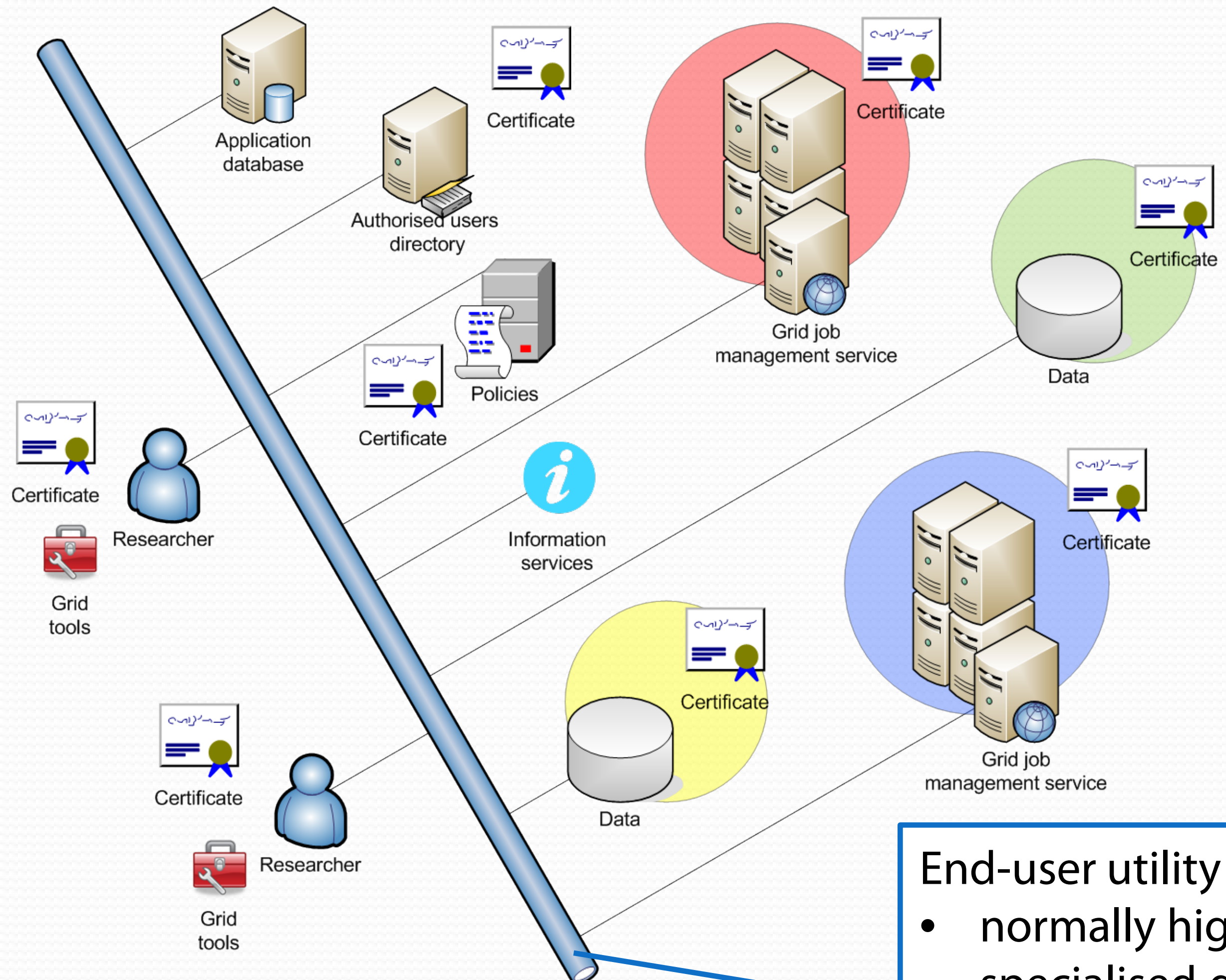
Problem	Solution
Where are my data?	Common catalogue using logical names and unique identifiers for data sets
Every cluster has different batch system (PBS, GE, LSF, SLURM…), what do I do?	Common execution interface is provided by Grid services; each job has a unique ID
How do I describe my task?	Common job description language
I don't even know what is available!	Common information schema and services
I can't remember 200+ passwords!	Single sign-on infrastructure based on certificates; delegation by proxy
Where am I authorised?	Common authorisation and policy decision services
How do I know about my resource usage?	Common accounting service

… and many other services and tools

From the trivial Grid-like system…



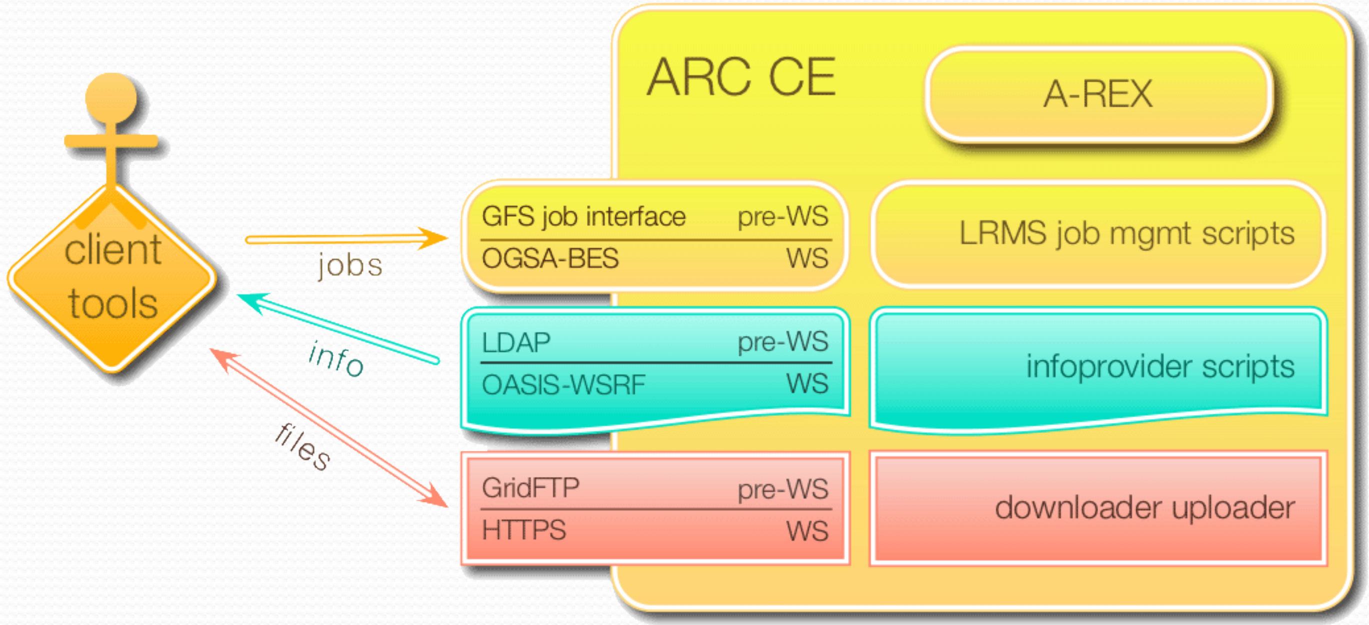
To the full-scale Grid



End-user utility

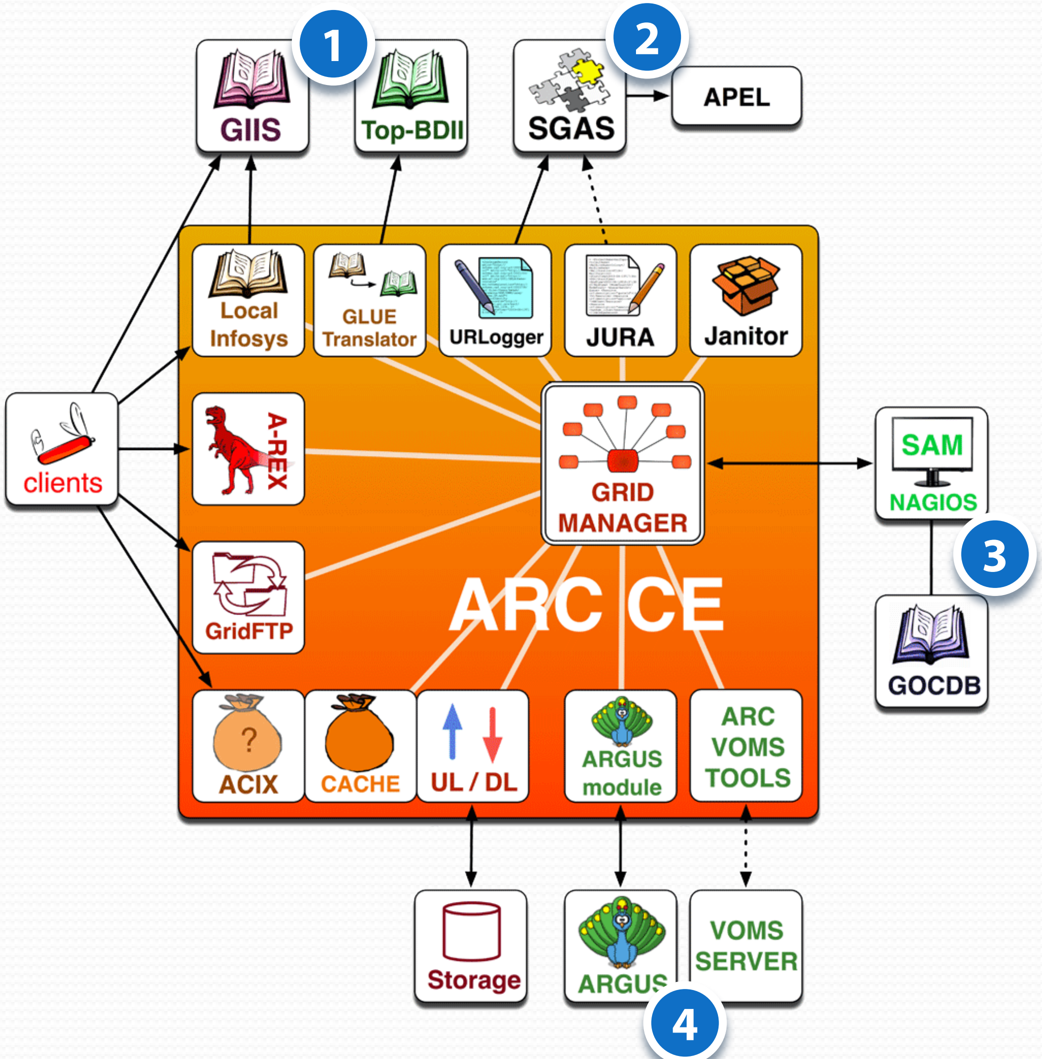
- normally highly specialised depending on research task

Example of a Grid computing service



- Key component of Grid infrastructure
- Universal front-end for different batch systems
- Standard and custom interfaces
- File handling on behalf of the user
- Status information publishing

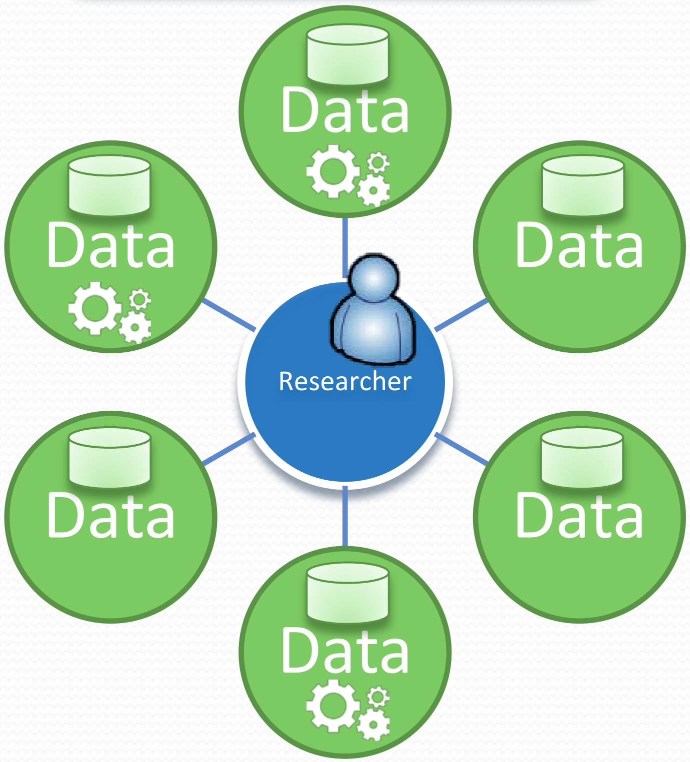
Plugging ARC CE into WLCG infrastructure



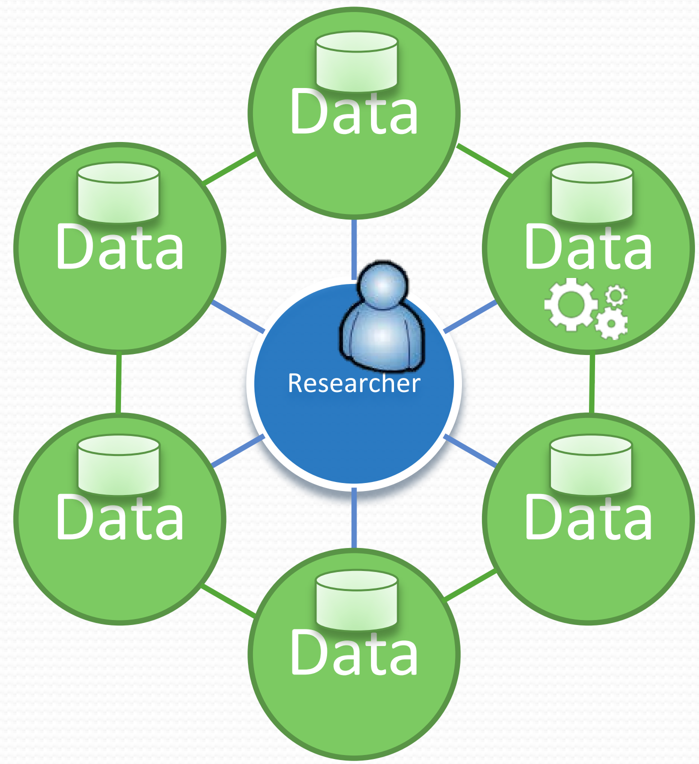
1. Information indexing
2. Accounting records
3. Service monitoring
4. Authorisation services

Move data or move jobs?

Jobs to data:
- Slow network
- Large local storage
- Trivial Grid

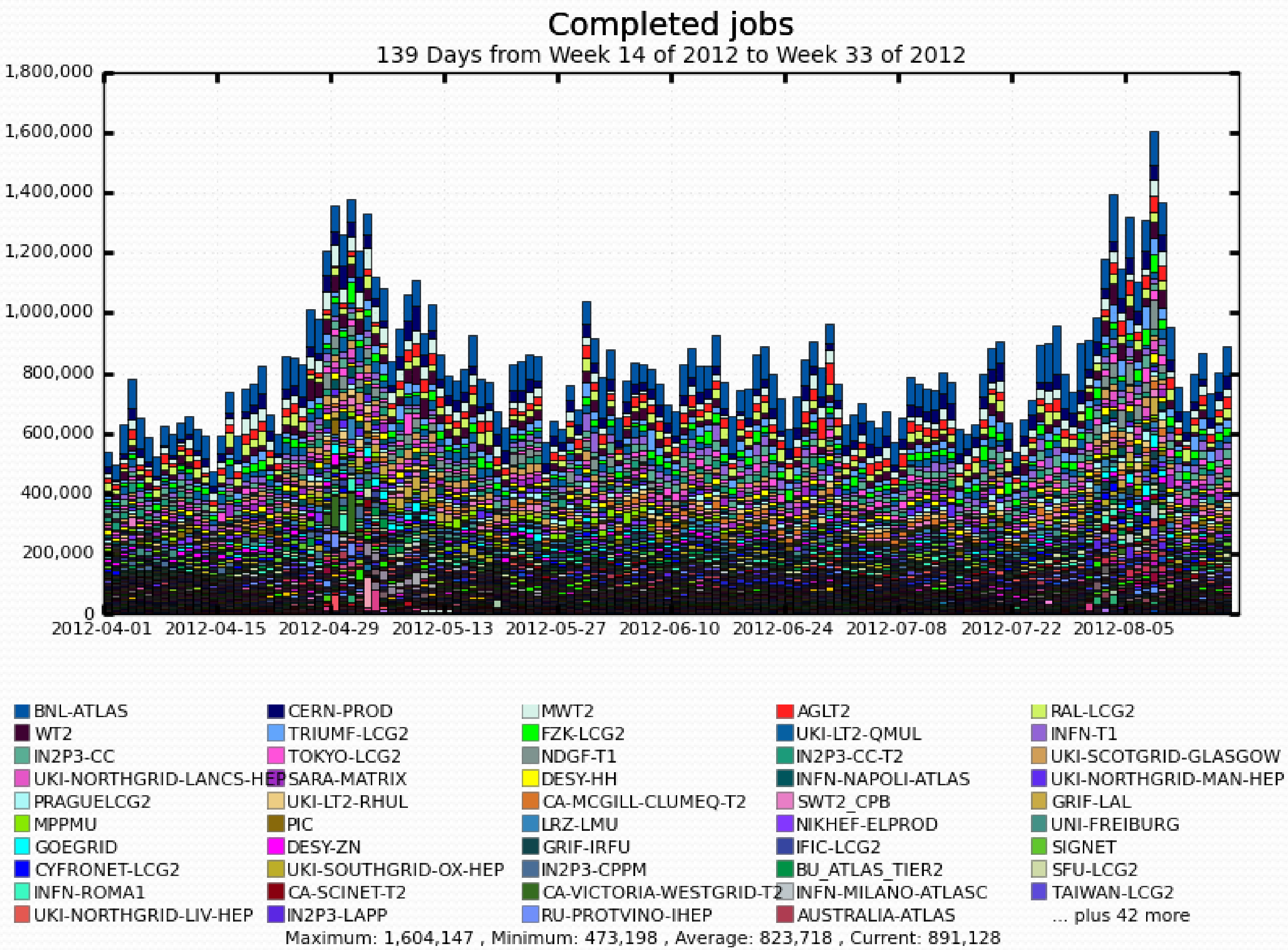


Data to jobs:
- Fast network
- Small local storage
- Complex Grid

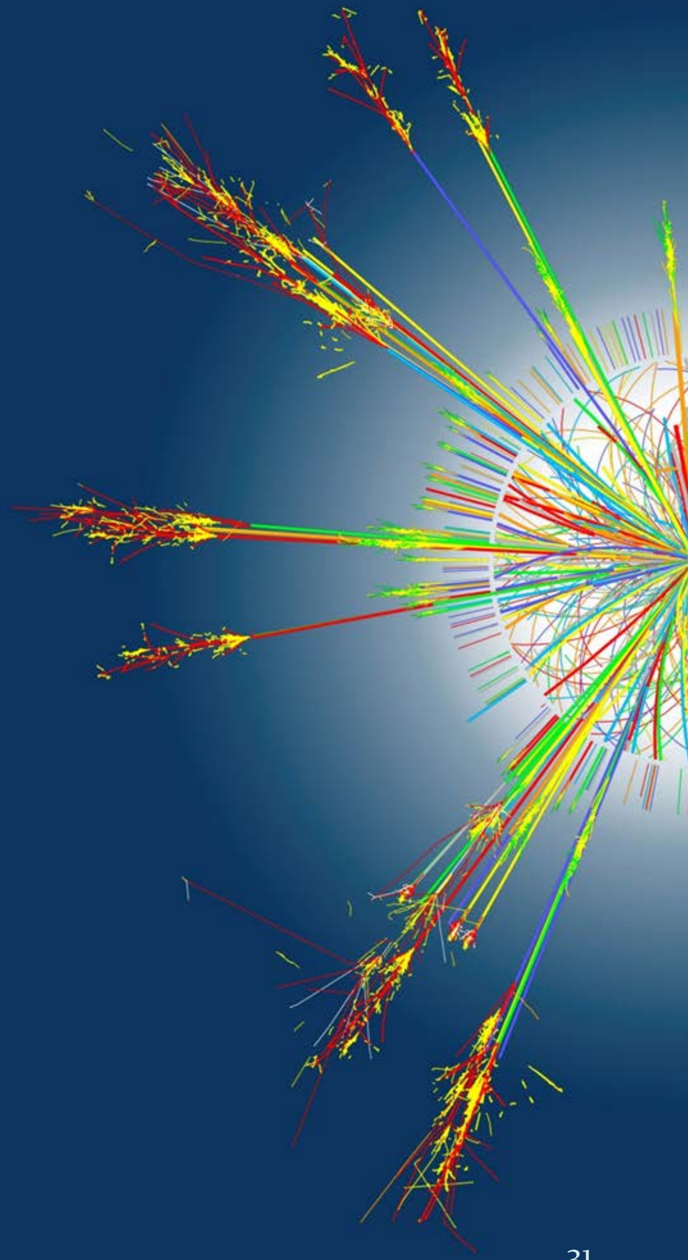


Modern trend is to use hybrid approaches

ATLAS jobs worldwide



Conclusions



Conclusions

- All sciences face rapid increase in digital data volumes
- LHC developed a working solution for Big Data
  - Allows LHC to achieve scientific results almost instantaneously
  - Essentially, it is a Cloud without business model
- There is no alternative to Grid for large scientific efforts
  - Scientific data will always be distributed
  - Global science is a <u>collaborative</u> effort, and so is Grid