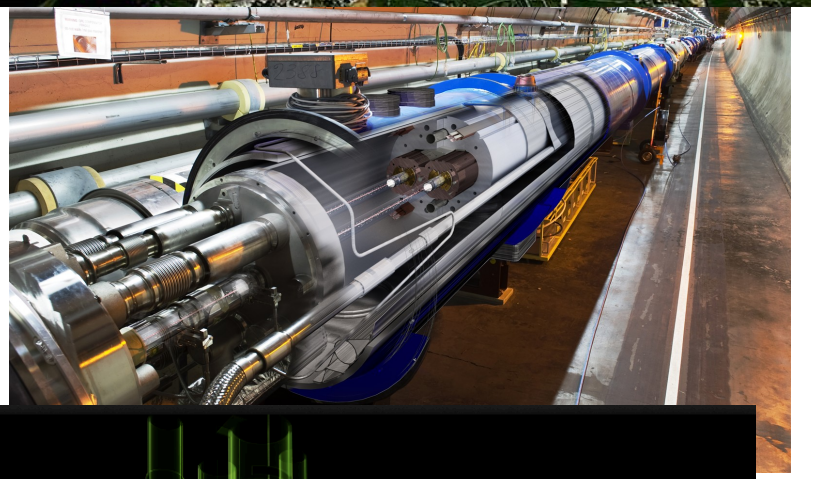
Part 3.
Distributed computing
computing grids
example with Worldwide LHC computing grid (WLCG)

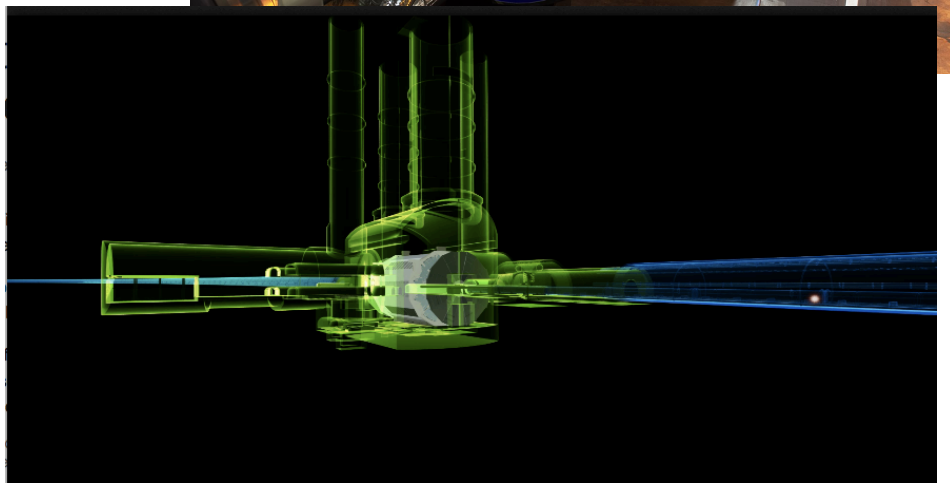
What is LHC

superquick LHC overview for dummies

LHC = Large Hadron Collider

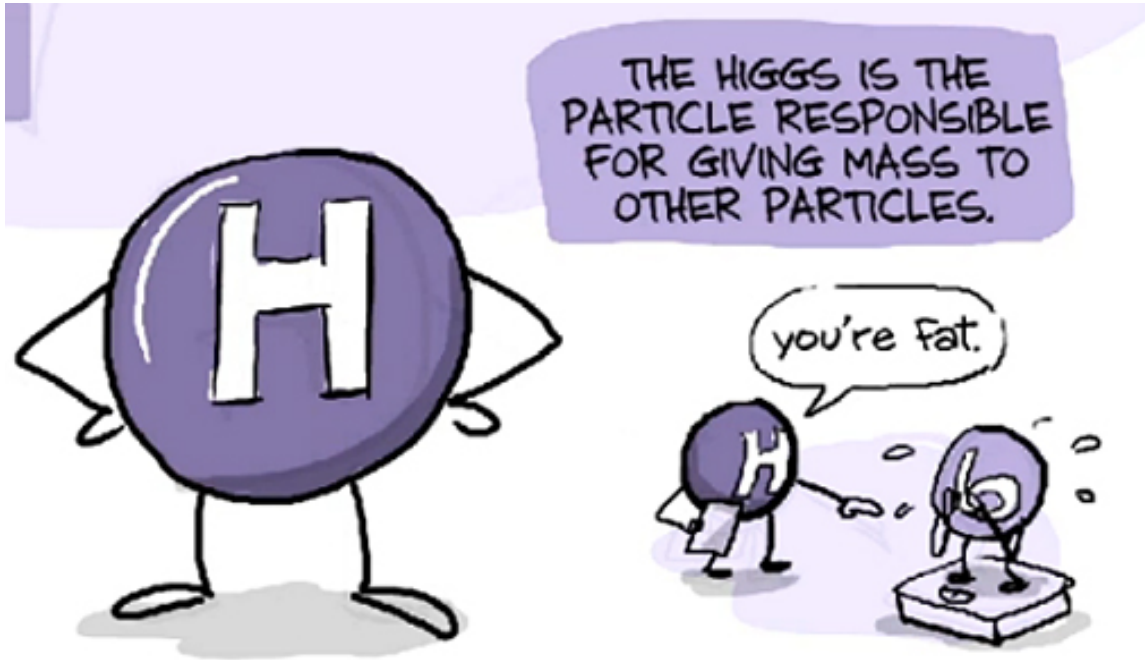
- Large particle accelerator situated underground on the border of France and Switzerland
- LHC collides protons (and also heavy ions (lead) in dedicated runs) – protons and ions are hadrons – therefore the name LHC
- Recreates the conditions just after the Big Bang
  - This is exciting because particles that not normally can be produced, can in these extreme energy densities
- Goal is to understand more of the fundamental physics of nature at higher energy ranges than earlier





LHC

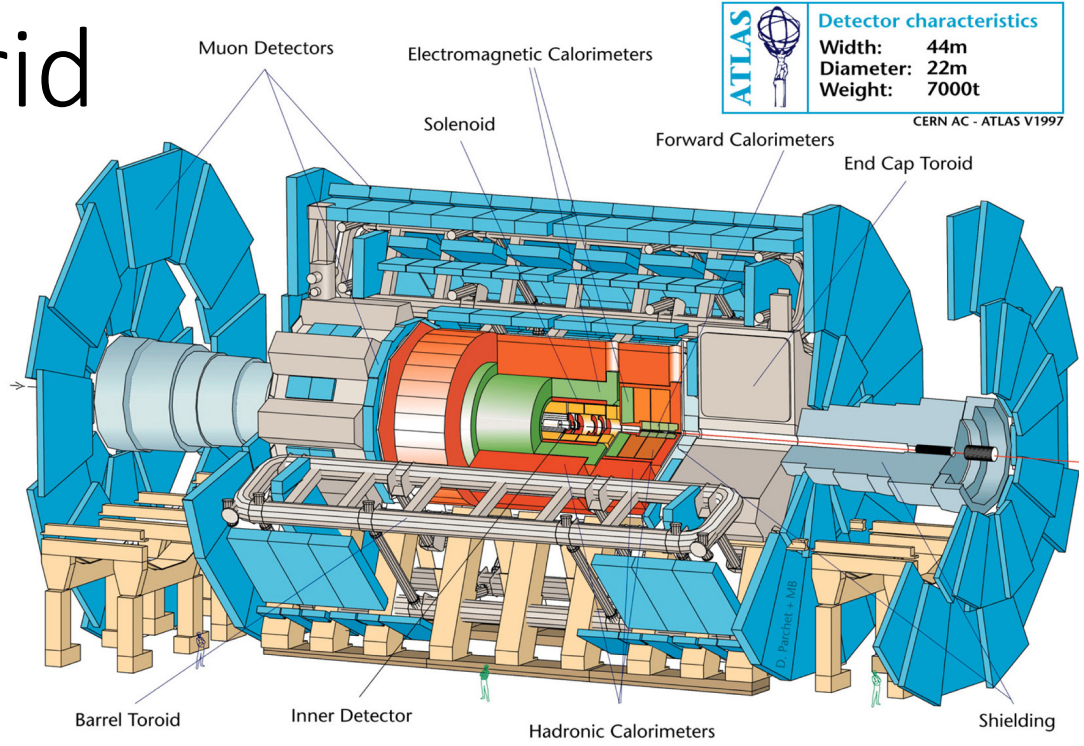
- Started running in 2009 (2008)
- One of the main goals: discover the Higgs particle
  - Higgs particle discovered in 2012 ✓
- 4 main experiments around the LHC ring
  - ATLAS, CMS, ALICE, LHCb
  - These are like digital cameras recording all that happens in the particle collisions
- ~600 million collisions per second → huge amount of data!
  - Dataflow: up to 25 gigabytes-per-second

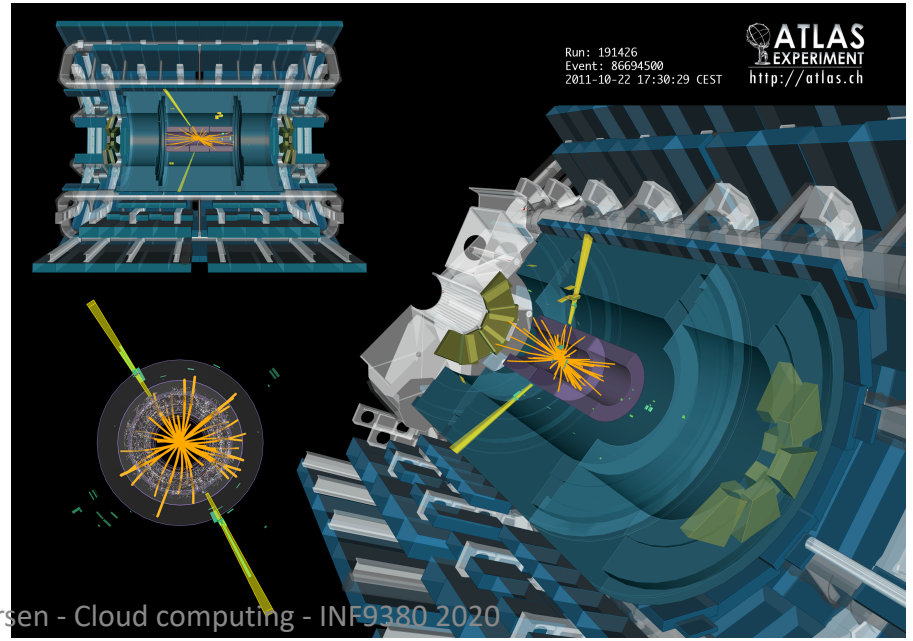


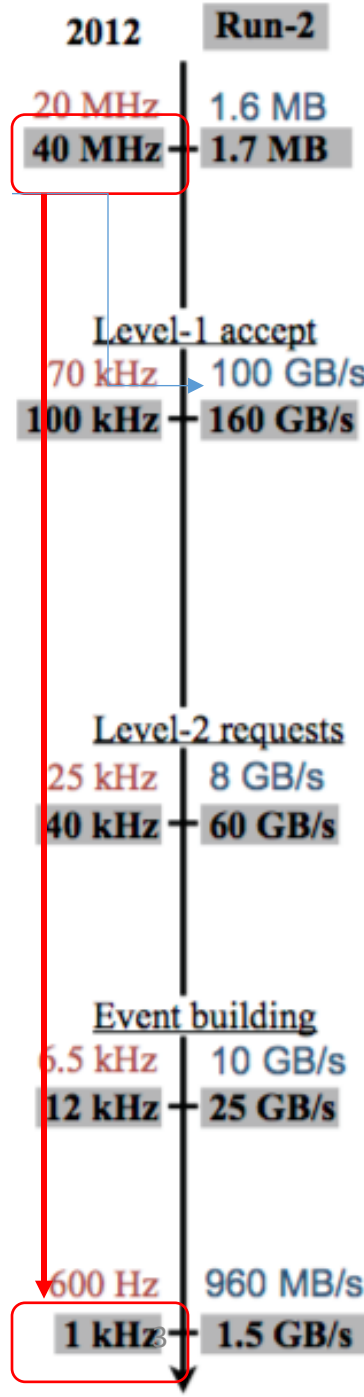
From experiment to grid

ATLAS example

- Detector (here ATLAS) records all the particle hits in the detector
- RAW data processed at CERN and collisions "reconstructed"
  - 1000 events per second (1 kHz) input to CERN (the rest is filtered out)
  - Average streaming throughput of RAW data 800 MB/s – 1000 MB/s
- Data distributed around the world through the grid for further processing and analysis
- LHC-total ~30 Pb per year







WLCG created to provide necessary storage and processing power to the LHC experiments

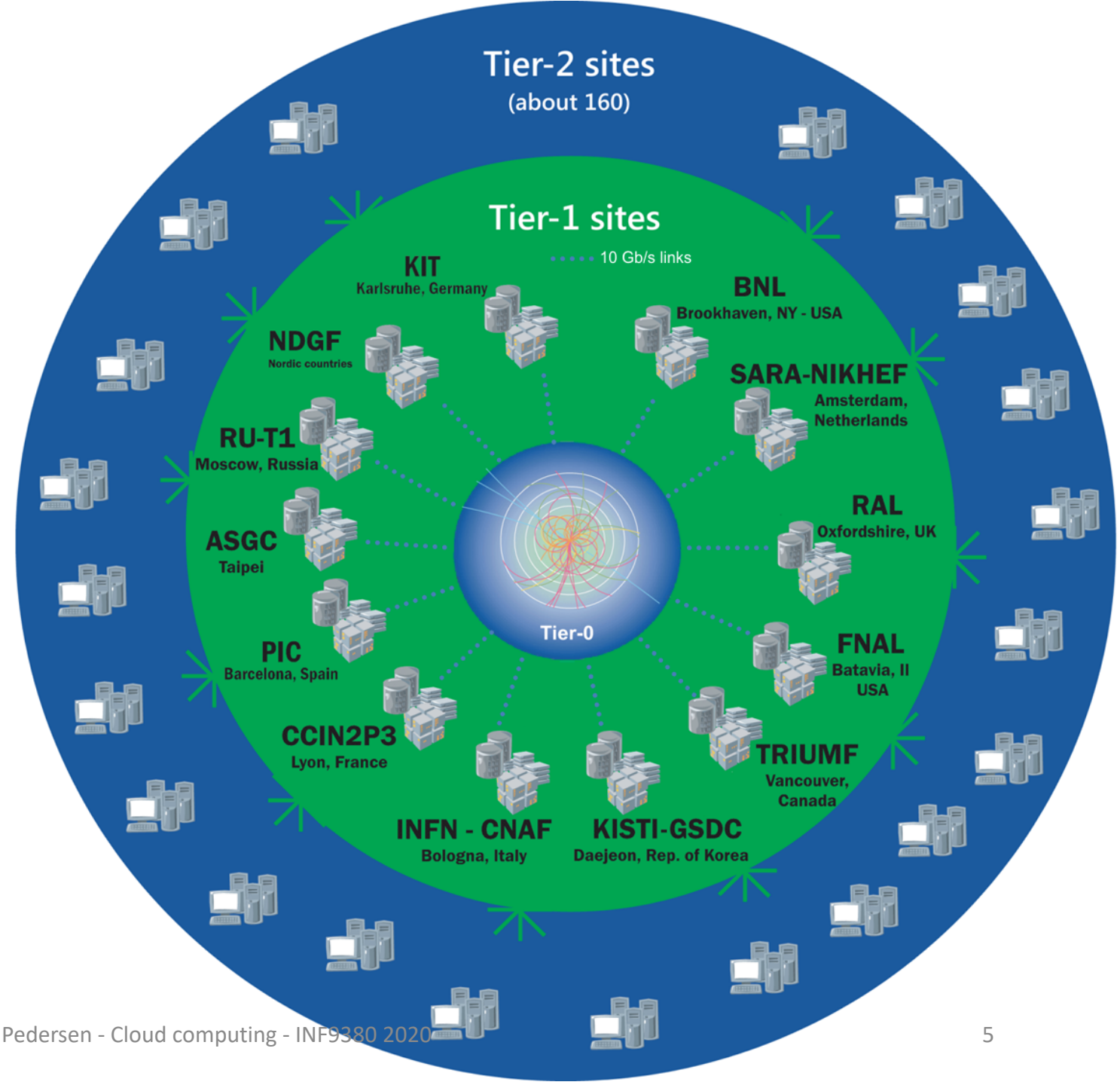
- LHC and its experiments produce an enormous amount of data (~30 petabytes a year).
- Needs storing, distribution, and processing.
- Users need to easily access the data and analyze it.
- The mantra is “Jobs go to data”.

Data distribution
Layered infrastructure:
Tier 0, 1 and 2 (and 3)

Tier 0: Raw data from experiments
Recorded on tape and processed at Tier 0

Tier 1: Data distributed from Tier 0 to Tier
Analysis requiring access to large subsets of raw, processed or simulated data takes place here

Tier 2: Data distributed from Tier 1 to Tier-
End user analysis and Monte Carlo simulation



WLCG: hardware, middleware, analysis software

Hardware

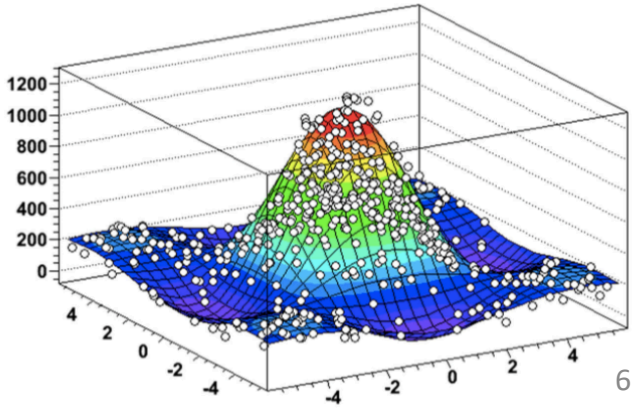
- Diverse hardware at Tier 1-s
  - Hardware varies from site to site
  - Tools must handle this diversity
- Storage and storage tools
  - Tape and disk
  - dCache, CASTOR (**C**ERN **A**dvanced **STOR**age manager), ENSTORE (Fermilabs data storage system)

Middleware

- Placed between the operating systems of the computers and the physics application software used by physicists
- Many different initiatives, and many in use by WLCG
  - ARC, HTCondor-CE (works nicely with HTCondor scheduling system), UNICORE, CREAM …

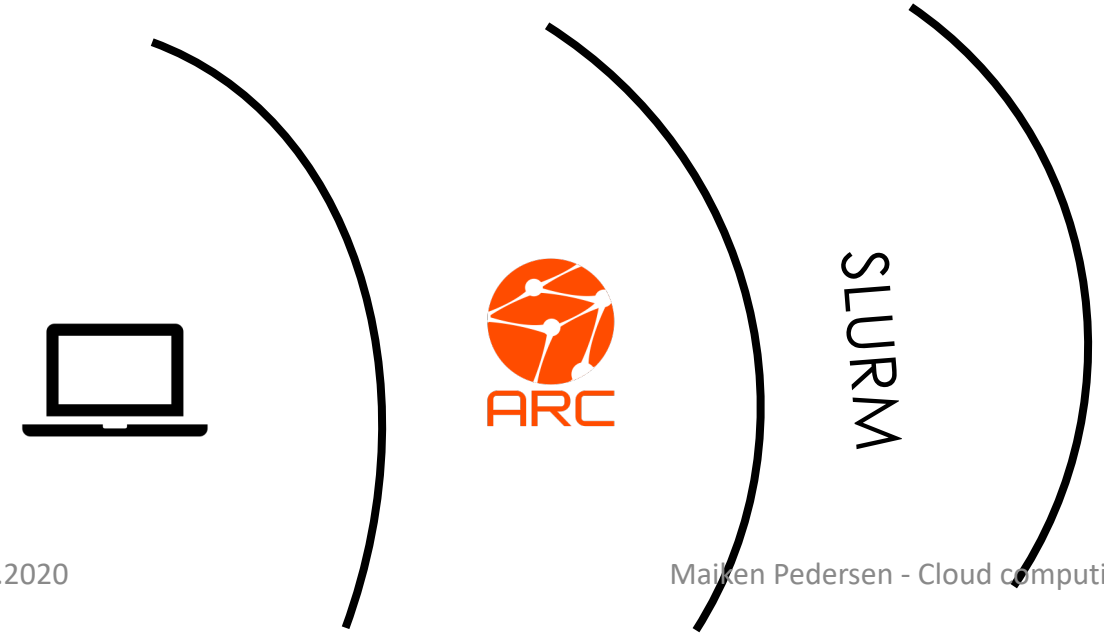
Physics analysis software

- Common software applications through Cern Virtual Machine File System (CVMFS)
  - This includes core software libraries, tools and frameworks for data management and event simulation as well as infrastructure and services for software development, and the support of analysis and database management.



Middleware layer

- An additional layer to enable computing on a distributed and diverse set of compute clusters connected in a computing grid
- The job is submitted from outside – as opposed to inside from the login-node for a normal cluster



Middleware layer

For this to work, the following main components are needed

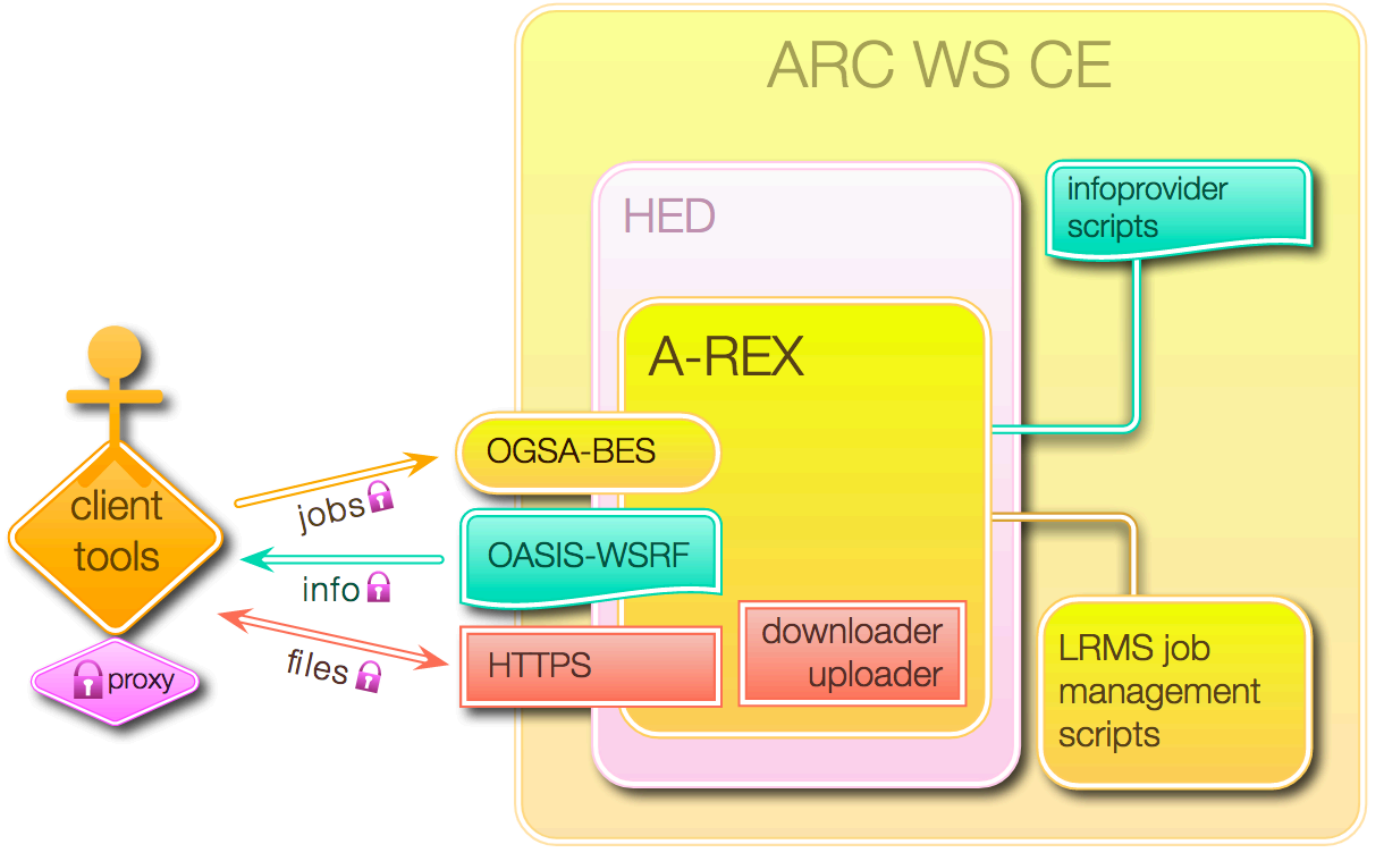
- Agreed job script language to define job (xrls, jdl, etc..)
- Middleware on cluster listening for incoming jobs on given ports
- Middleware translating the job description to the local batch system language (e.g. SLURM)
- Middleware downloading needed input-data
- Middleware making sure to update the status of the job – i.e. pending, running, finished so that the client can collect the job when done

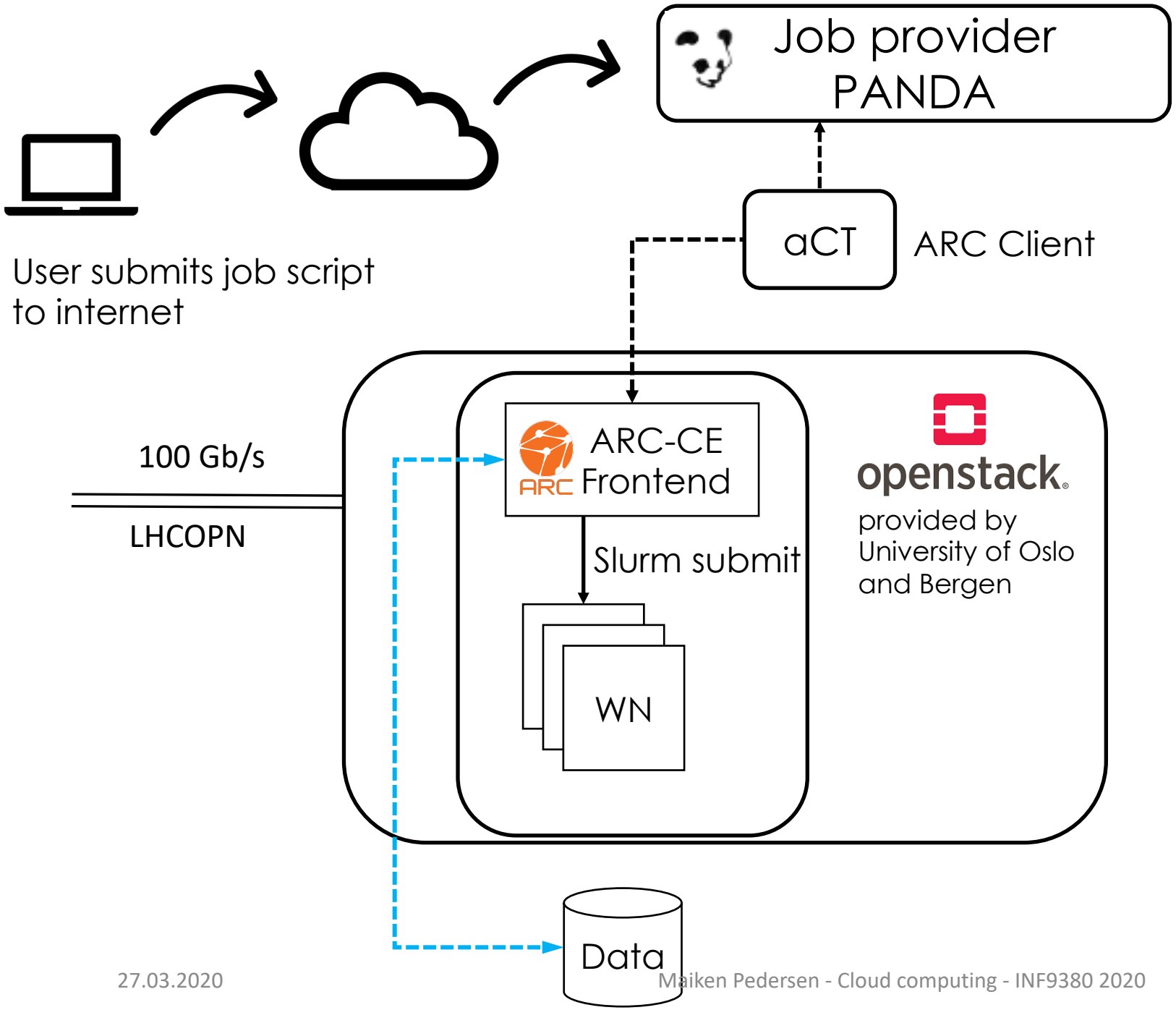
Norway – ARC middleware

- There are various different middleware solutions
- Norway is part of a Nordic Distributed Tier 1 data center using the Nordugrid ARC middleware
  - ARC: Advanced Resource Connector
- Developed by the Nordics
- ~200 sites in the WLCG use ARC

ARC overview

- ARC: middleware to enable computing grids
- Used since 2002, and currently ~200 sites worldwide use ARC
- Job submission
- Information exchange
- File access
- Web-services provided by A-REX
- Once job is submitted to a site, and all preparations done by ARC ((file downloads etc) the job is sent to the local batch system to run.





CERN ATLAS example

- Central job server PaNDA
  - The *PanDA* Production ANd Distributed Analysis system ARC Client submits the job to the grid
- Middleware-layer ARC on the frontend/login grid node

Grid for non-WLCG experiments

- Up until now mostly WLCG experiments have been using ARC
- However, many international collaborations are dealing with the same issues as LHC-experiments:
  - Need for de-centralized storage and compute
  - ➔ Need for middleware layer
- So, maybe in some years your bio-jobs will be sent out to the grid, rather than directly to a specific data-center

Hands-on: submit a job to the grid (or rather, to a small test-site with ARC installed)