# Working with ATLAS data

1. the Tier-3 cluster in Geneva
   - description
   - current usage patterns
2. what will change when data arrives
   - volumes and rates
   - how an analysis can be done
   - Data Quality work in Geneva

# The Tier-3 in Geneva

- 61 computers to manage
  - 53 workers, 5 file servers, 3 service nodes
- 188 CPU cores in the workers
- 75 TB of disk storage
- can burn up to 30 kW (power supply specs)

Mostly ready

35 workers have 2 GB/core

18 have 1 GB/core, will be upgraded soon

likely more disk at some stage

# Hardware in Geneva (1)

# Hardware in Geneva (2)

12 worker nodes in 2005
SunFire V20z

21 in 2006
SunFire X2200

and 20 in 2007
SunBlade 6220

# Hardware in Geneva (3)

power and network
cabling of worker nodes

three X4200
for services
(grid, batch,
web, monitoring,
gridftp)

direct line from CERN

# Setup of the Tier-3 in Geneva

- **SLC4.6 on 53 workers 15 of which are login machines**
- **Solaris on four X4500 file servers, zfs**
- **One to two file systems per server, exported by NFS**
- **Torque batch system**
- **NorduGrid middleware, front-end on the batch server (X4200)**
- **Ganglia monitoring and Apache on another server (X4200)**

# Swiss ATLAS Grid clusters

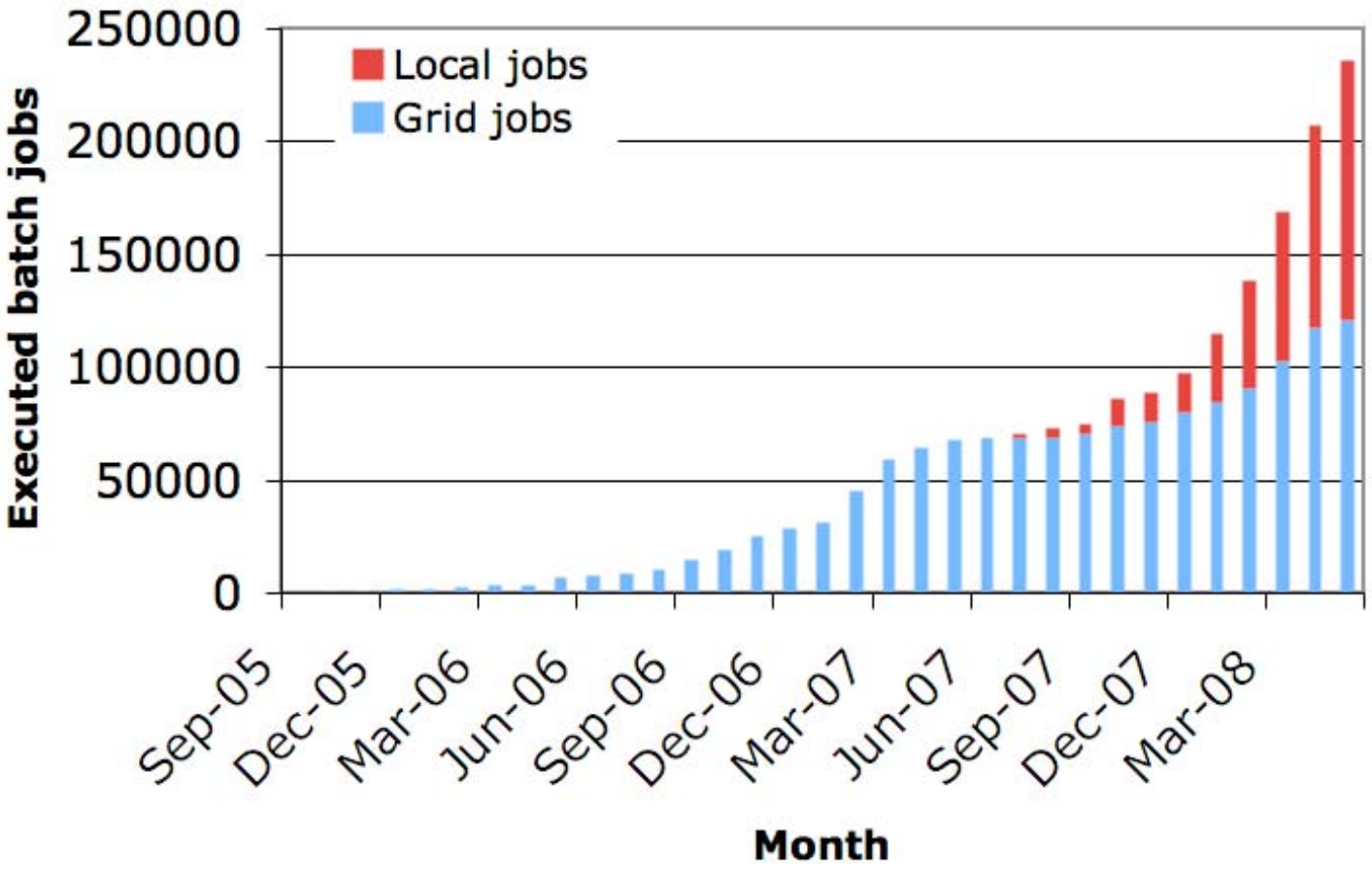
**ATLAS Grid Monitor**

*2008-04-16 CEST 13:56:05*

Processes: Grid Local

| Country | Site | CPUs | Load (processes: Grid+local) | Queueing |
|---|---|---|---|---|
| Denmark | Steno (DCSC/KU) | 1640 | 799+124 | 60+1 |
| Norway | EPF (UiO/FI) | 19 | 14+3 | 7+0 |
| | Hyperion (UiO/USIT) | 275 | 0+188 | 0+59 |
| | Norgrid@NTNU | 38 | 0+17 | 0+0 |
| | Titan A (UiO/USIT) | 2234 | 271+1407 | 0+0 |
| Slovenia | SiGNET | 204 | 204+0 | 43+0 |
| Sweden | Hive (Swegrid, C3SE) | 101 | 72+29 | 0+4 |
| Switzerland | Bern ATLAS T3 Cluster | 28 | 24+0 | 3+0 |
| | Bern UBELIX T3 Cluster | 512 | 73+394 | 46+303 |
| | Geneva ATLAS T3 | 152 | 120+0 | 24+0 |
| | Manno PHOENIX T2 | 443 | 0+139 | 0+0 |
| *TOTAL* | *11 sites* | *5646* | *1577 + 2301* | *183 + 367* |

# Batch jobs in Geneva since 2005



# Current use

1. **local cluster for the Geneva group and friends**
   - log in and have an environment to work with ATLAS software, both offline and trigger
     - develop code, compile,
     - interact with ATLAS software repository at CERN
   - work with nightly releases of ATLAS software, normally not distributed off-site but visible on /afs
   - disk space
   - final analysis tools, in particular ROOT
   - a convenient way to run batch jobs
   - around 40 users, so far rather happy
2. **grid computing**
   - ATLAS production via NorduGrid (most of the batch use)
   - my users do not see a point in it yet, to be honest

# Analysis of simulated data

## A typical way to work so far

1. **First of all, copy the dataset to Geneva.**
2. **Ask questions later.**

## Typical dataset size  O(100 GB)

| | | |
|---|---|---|
| accounting done on | 26-May-08 | |
| datasets | 141 | |
| size in /atlas/dataN/datasets | 11.0 | TB |
| average size | 79.9 | GB |

# ATLAS dataset sizes

| Estimate of data for a final physics analysis | | | |
|---|---|---|---|
| **data taking period** | 3 | years | |
| seconds/year | 1.00E+07 | | |
| raw data rate | 3.20E+08 | B/s | |
| raw data rate | 200 | events/s | |
| raw data volume | 9.60E+15 | B | |
| data in events | 6.00E+09 | events | |
| AOD volume | 1.23E+15 | B | |
| | | | |
| **one stream** | 25% | | |
| total AOD for a stream | 279 | TB | collaboration |
| 25% selection | 69.8 | TB | physics group |
| 5% selection | 14.0 | TB | individual |

**A different scale. Ramp-up *may or may not* be gradual**

- **groups will agree and make subsets of streams, hopefully via central ATLAS production**
- **individuals will make further derived datasets, submitting jobs to the grid and getting results to a “user space”**

# Transfer and processing times

**Data transfer and processing times**

| | | |
|---|---|---|
| gigabit line speed | 1.00E+09 | bps |
| assumed line utilization | 50% | |
| realistic rate per minute | 3.5 | GB |
| realistic rate per hour | 210 | GB |

| data volume | transfer time | | events | analysis time [h] | h on 120 CPU |
|---|---|---|---|---|---|
| 100 GB | 28.6 | min | 5.2E+03 | 0.089 | |
| 1 TB | 4.9 | h | 5.4E+06 | 91 | 0.76 |
| 15 TB | 3.1 | days | 8.1E+07 | 1365 | 11.38 |
| 75 TB | 15.3 | days | 4.0E+08 | 6827 | 56.89 |

- **Data samples up to 15 TB can be copied to Geneva and analyzed by individuals in a coordinated fashion.**
- **O(1 TB) is not a problem, if you have a cluster...**

# Question of user space on the Grid

- **This is the main role of a Tier-3.**
- **Many Tier-2 and even Tier-1 sites in ATLAS have user accounts and user space.**
  - **In the computing model this is not black or white. Some “preference for regional users” at a T2.**
- **Big discussion in ATLAS about user space on T2 sites.**
  - **No conclusion yet.**
  - **This concerns only the space managed by ATLAS DDM.**
- **My opinion**
  - **Computing will be both “central” and “private”. We can’t predict the shares. Daily decision by users.**
  - **Having space gets physicists more interested and more involved.**
  - **We need to get the tools ready. Quotas and shares will in any case be re-discussed as soon as we have some data.**

# Data Quality work in Geneva

- **Need efficient ways to quickly look at fresh real data.**
  - **Feedback in hours, not days. Small data volumes.**
  - **Particularly important when the experiment starts.**
- **The CERN Analysis Facility (CAF) is under pressure. Trigger has 8 machines for Data Quality.**
- **The T3 in Geneva can be an extension of the CAF**
  - **direct line, we work at CERN, we are involved in detector and trigger, 2/3 of users are guests**
- **The same data can be made available to Grid jobs, e.g. sent from Berne.**

# Data and CPU for DQ work

| **Esitmate for Data Quality work** | | | |
|---|---|---|---|
| run time | 12 h | | |
| raw data volume | 12.6 TB | | |
| data fraction to monitor | 15% | 1% | 0.10% |
| volume [TB] | 1.9 | 0.1 | 0.013 |
| events | 1.3E+06 | 8.6E+04 | 8.6E+03 |
| transfer time [h] | 9.22 | 0.61 | 0.06 |
| reco time in CPU hours | 4235 | 282 | 28 |
| reco hours on 120 CPU | 35 | 2.35 | 0.24 |

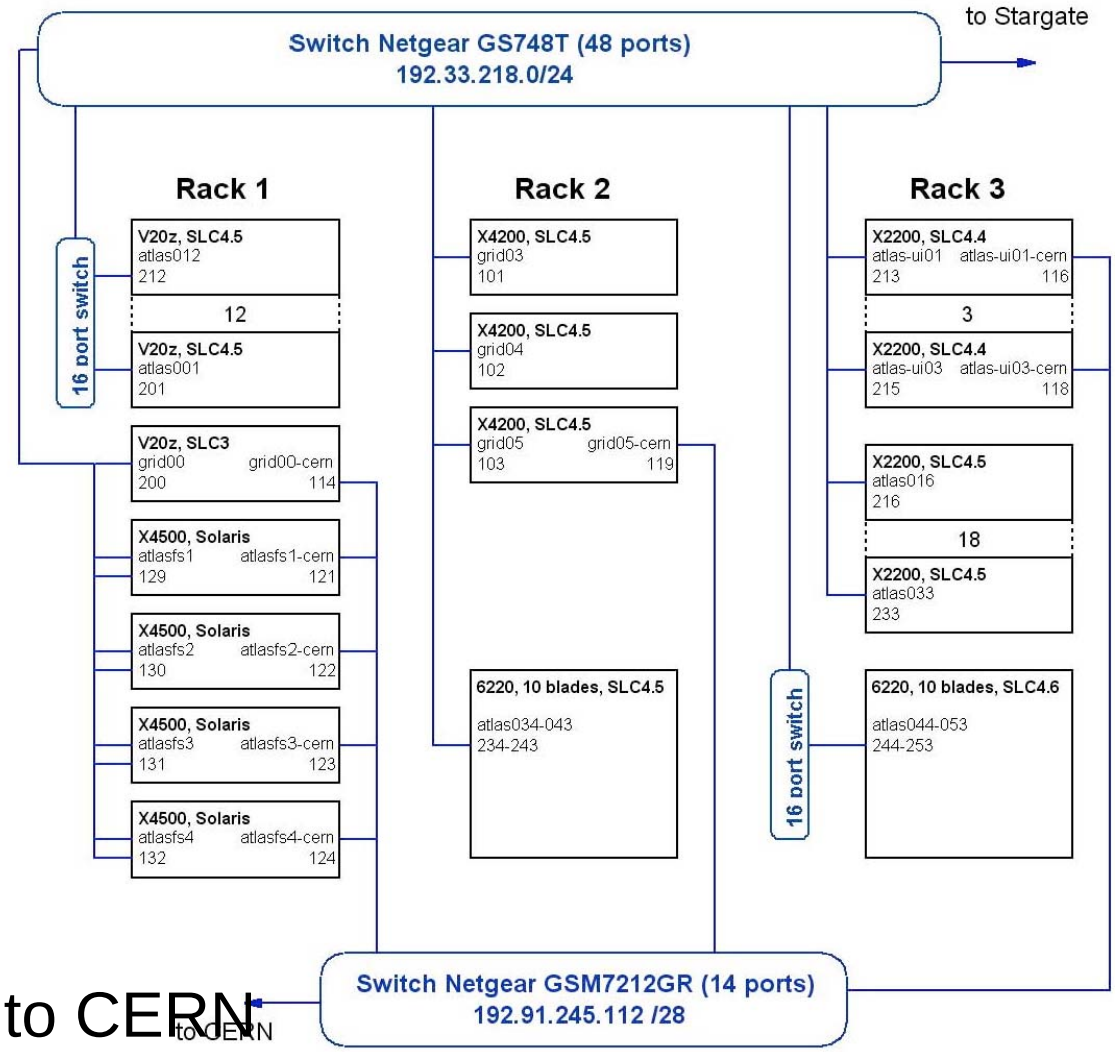
- **entire Express Stream would be too much for GE**
- **monitoring of the order of 1% of data is possible**

# Where we are, what next

- **The Geneva T3**
  - **Setup mostly done. No major changes or additions foreseen before the first data.**
  - **Three roles**
    - **running some production via Grid, like a T2 (~there since 2005)**
    - **a T3 site, for final analysis of data (~there)**
    - **enable a quick check of small amounts of fresh data (to be established)**
- **The CSCS T2 as we see it**
  - **Good to see it used by ATLAS production! Congratulations!**
  - **Usage by Swiss ATLAS physicists to be tried out**
    - **data movement**
    - **submission of jobs**

# Backup slides

# Networking




direct line to CERN IT

# The functionality we need

- our local cluster computing
  - log in and have an environment to work with ATLAS software, both offline and trigger
    - develop code, compile,
    - interact with ATLAS software repository at CERN
  - work with nightly releases of ATLAS software, normally not distributed off-site but visible on /afs
  - disk space
  - use of final analysis tools, in particular ROOT
  - a convenient way to run batch jobs
- grid computing
  - tools to transfer data from CERN as well as from and to other Grid sites worldwide
  - ways to submit our jobs to other grid sites
  - a way for ATLAS colleagues to submit jobs to us

# Things to do (and to research)

a) **configuration of worker nodes**
   - configuration of the SLC, 32 bit compatibility libs and etc
   - torque batch client
   - other added software, as requested by the users

b) **storage management**
   - try going for Solaris on X4500 + zfs + NFS and nothing else
   - if not enough, try a solution for storage management that makes an abstraction from physical servers (e.g. dCache or DPM)
   - the network connections of the file servers can easily be saturated, try “trunking” of lines to the switch

c) **grid services and batch system**
   - SLC on those nodes
   - torque batch system, a different scheduler
   - choice of gLite not excluded

d) **multiple ways to do data transfers, redundancy is needed**

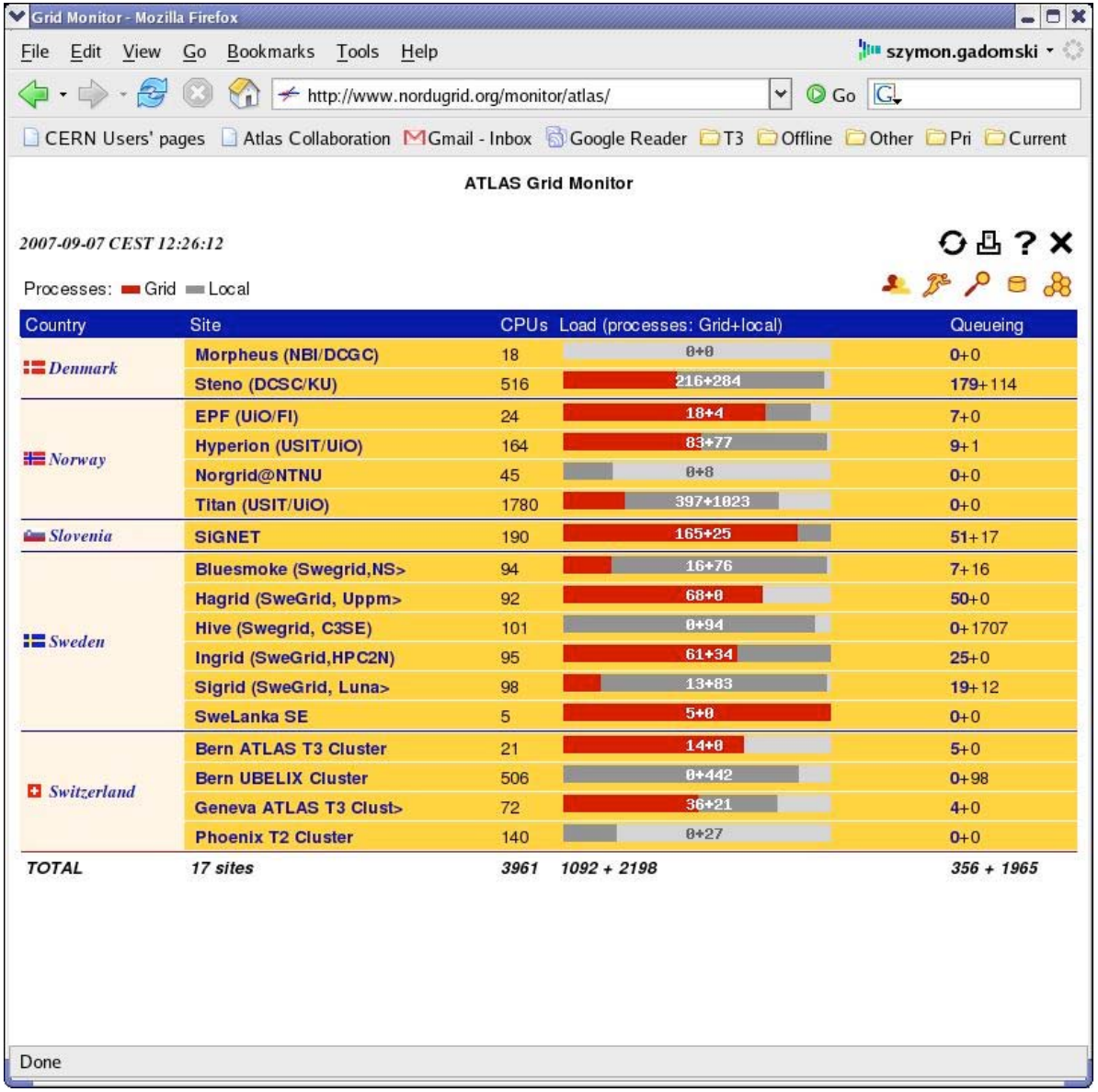
e) **tools for interactive use of multiple machines (e.g. PROOF, Ganga)**

f) **grid job submission interfaces (e.g. Ganga, GridPilot)**

g) **general cluster management issues**
   - security issues
   - a way to install the system on multiple machines
   - our own UPS and an automatic shutdown(?)
   - monitoring of temperature, CPU use, network use

# Our system in the NorduGrid

- Geneva is in NorduGrid since 2005
- In company of Berne and Manno (out Tier 2)




| Country | Site | CPUs | Load (processes: Grid+local) | Queueing |
|---|---|---|---|---|
| Denmark | Morpheus (NBI/DCGC) | 18 | 0+0 | 0+0 |
| | Steno (DCSC/KU) | 516 | 216+284 | 179+114 |
| Norway | EPF (UiO/FI) | 24 | 18+4 | 7+0 |
| | Hyperion (USIT/UiO) | 164 | 83+77 | 9+1 |
| | Norgrid@NTNU | 45 | 0+8 | 0+0 |
| | Titan (USIT/UiO) | 1780 | 397+1023 | 0+0 |
| Slovenia | SiGNET | 190 | 165+25 | 51+17 |
| Sweden | Bluesmoke (Swegrid,NS> | 94 | 16+76 | 7+16 |
| | Hagrid (SweGrid, Uppm> | 92 | 68+0 | 50+0 |
| | Hive (Swegrid, C3SE) | 101 | 0+94 | 0+1707 |
| | Ingrid (SweGrid,HPC2N) | 95 | 61+34 | 25+0 |
| | Sigrid (SweGrid, Luna> | 98 | 13+83 | 19+12 |
| | SweLanka SE | 5 | 5+0 | 0+0 |
| Switzerland | Bern ATLAS T3 Cluster | 21 | 14+0 | 5+0 |
| | Bern UBELIX Cluster | 506 | 0+442 | 0+98 |
| | Geneva ATLAS T3 Clust> | 72 | 36+21 | 4+0 |
| | Phoenix T2 Cluster | 140 | 0+27 | 0+0 |
| *TOTAL* | *17 sites* | *3961* | *1092 + 2198* | *356 + 1965* |

# Geneva is a special case

- we have many users who are not with Geneva, but are at CERN and work together with us in physics groups or on the trigger
- we run ATLAS production jobs since several years (Bern does that too)
  - in this respect Switzerland has a distributed Tier 2
- the official data path is a complete nonsense for us
  - logical path is: CERN > FZK > CSCS > Geneva
  - physical path is: CERN > FZK > CERN > CSCS > Geneva
  - rationalizing all this is not a technical problem, but a political one
  - At the moment we are penalized for not installing our computers at CERN, (like many other ATLAS Universities do).
- What can we do?
  - Geneva can become a DDM service location.
    - the DDM are positive
    - need to figure out where and how to have a file catalogue
  - ATLAS will benefit if CERN allows rfcp to Geneva over a direct line. Perhaps ATLAS can help me to negotiate this with CERN IT.