



**UNIVERSITÉ
DE GENÈVE**

ATLAS Tier-3 cluster @ UniGe

Luis March and Yann Meunier (Université de Genève)

**CHIPP GRID: Face To Face meeting
ETH Zurich, March 11th 2016**

Description of ATLAS Tier-3 cluster at UniGe

The ATLAS Tier-3 cluster at UniGe-DPNC is physically located at [Uni Dufour](#) (~ 500 m away from UniGe-DPNC building)

Grid services (Nordugrid):
ARC-CE, BDII, proxy, DPM SE

Batch system (62 nodes):
Worker Nodes = 656 cores
Memory/process = 2.5 – 6 GB

Storage system (DPM):
ATLAS pool = 474.27
Reserved = 466.0 TB
File Servers = 15



ATLAS Space Tokens	Capacity (TB)	Used (TB)	Free (TB)
ATLASGROUPDISK	25.0	9.85	15.15 (~ 60.6%)
ATLASLOCALGROUPDISK	436.0	434.92	1.08 (~ 0.2%)
ATLASSCRATCHDISK	5.0	0.84	4.16 (~ 83.2%)

**Clean-up
needed**

Description of extra Tier-3 cluster at UniGe

Some extra Tier-3 cluster resources at UniGe-DPNC, for different experiments (not only ATLAS), which are also physically located at [Uni Dufour](#)

User Interfaces (login machines for users):

SLC6 (3 nodes) = 48 cores

SLC5 (3 nodes) = 48 cores → Check if still needed by users

In addition to DPM SE, we have NFS disk servers for local storage:

<code>/atlas/users</code>	→ Intended for software development (3 TB)	
<code>/atlas/software</code>	→ Intended for common ATLAS software (local users) (2 TB)	
<code>/cvmfs/*.cern.ch</code>	→ Official software tools for (some) experiments (mounted)	
<code>/atlas/data</code>	→ Data storage for UniGe ATLAS users	108.0 TB
<code>/neutrino/data</code>	→ Data storage for UniGe neutrino users	16.0 TB
<code>/ams/data</code>	→ Data storage for UniGe AMS users	103.0 TB
<code>/icecube/data</code>	→ Data storage for UniGe IceCube users	2.0 TB
<code>/dampe/data</code>	→ Data storage for UniGe DAMPE users	~120.0 TB

Total NFS disk space for local storage = ~5 TB + ~349 TB = ~ 354.0 TB

Operations

Grid services (Nordugrid):

ARC-CE → “nordugrid-arc-ce-5.0.5” (latest one)

DPM SE → “glite-yaim-dpm 4.2.20-1” (we should upgrade it with Puppet)

GGUS ticket/s:

Ticket-ID 117900 → About ATLAS storage (monthly) consistency checks
Data management: **Status = closed**

```
srmls -l srm://grid05.unige.ch:8446/srm/managerv2?SFN=/dpm/unige.ch/home/atlas/atlaslocalgroupdisk/dumps/
```

```
srmls -l srm://grid05.unige.ch:8446/srm/managerv2?SFN=/dpm/unige.ch/home/atlas/atlasscratchdisk/dumps/
```

```
srmls -l srm://grid05.unige.ch:8446/srm/managerv2?SFN=/dpm/unige.ch/home/atlas/atlasgroupdisk/trig-daq/dumps/
```

Scheduled downtime:

22-Feb-2016 → Created at **GOCDB Add Downtime ID = 19935**

Reasons for the downtime:

- Sanity checks for the whole cluster and service machines
- ARC-CE (“nordugrid-arc-ce-5.0.5”) updated
- Update glibc (“Critical” risk glibc remote code execution)

In general, running smoothly:

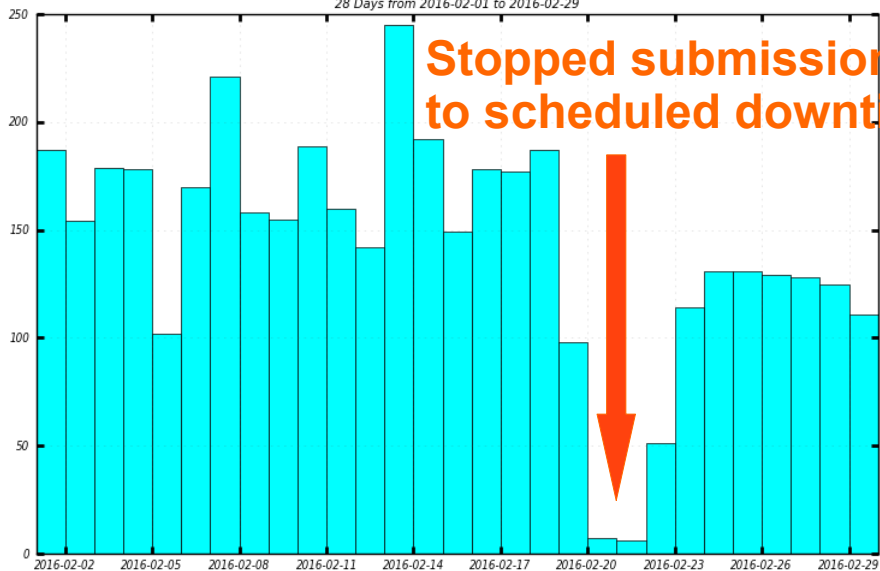
ATLAS Production jobs → UNIGE-DPNC is under testing: Not too many jobs

UniGe local users → Increased activity (job submission) for last month

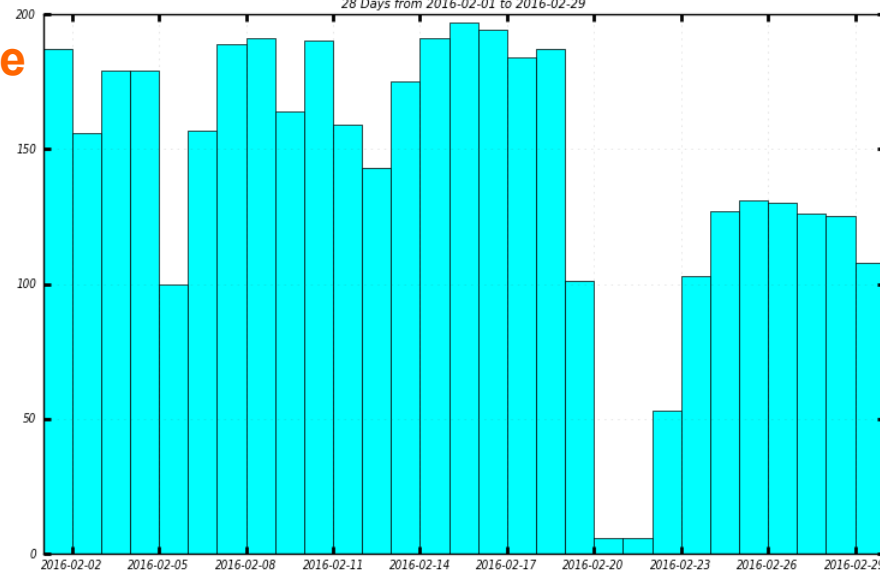
Accounting: ATLAS dashboard (1)



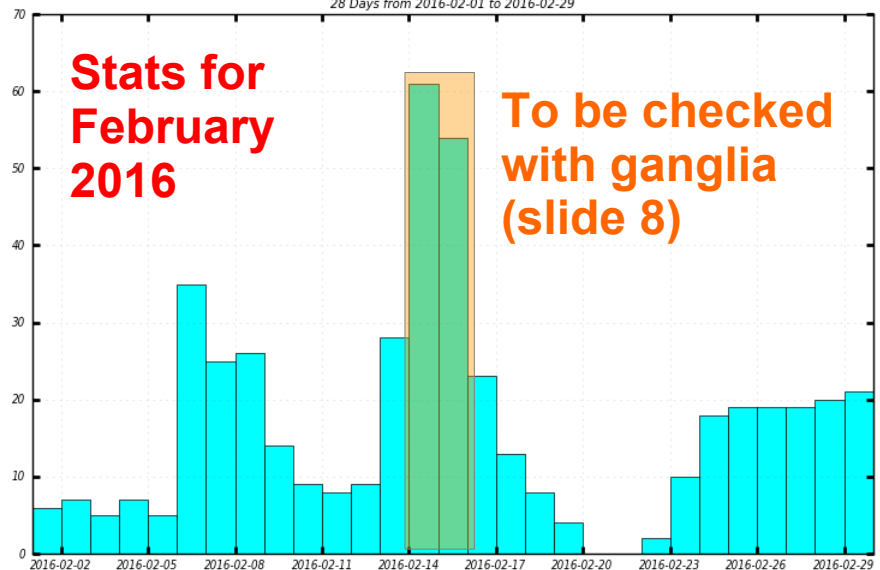
Submitted jobs
28 Days from 2016-02-01 to 2016-02-29



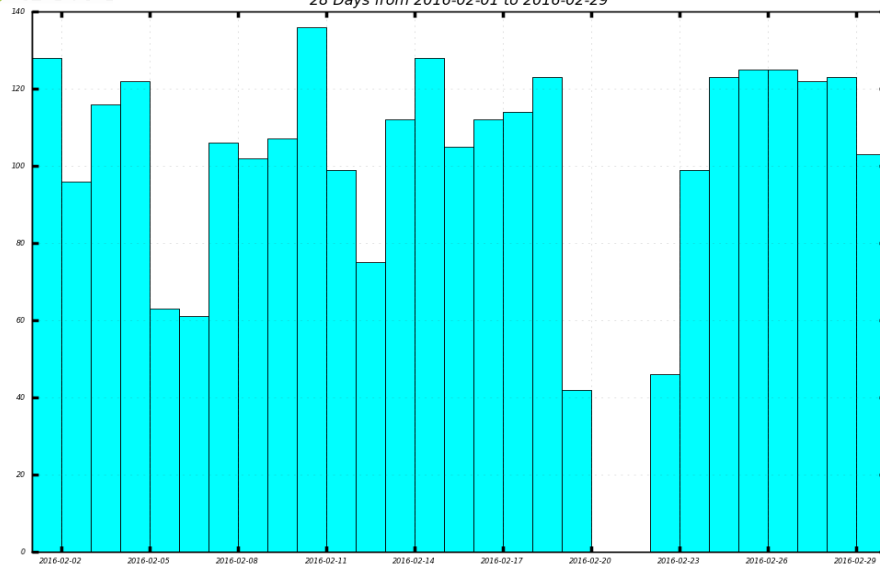
Completed jobs
28 Days from 2016-02-01 to 2016-02-29



Running jobs
28 Days from 2016-02-01 to 2016-02-29



Successful jobs
28 Days from 2016-02-01 to 2016-02-29



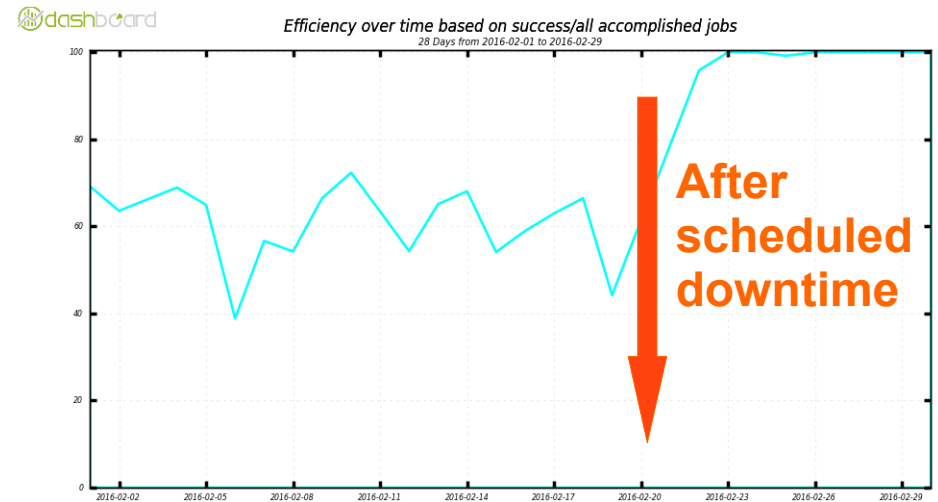
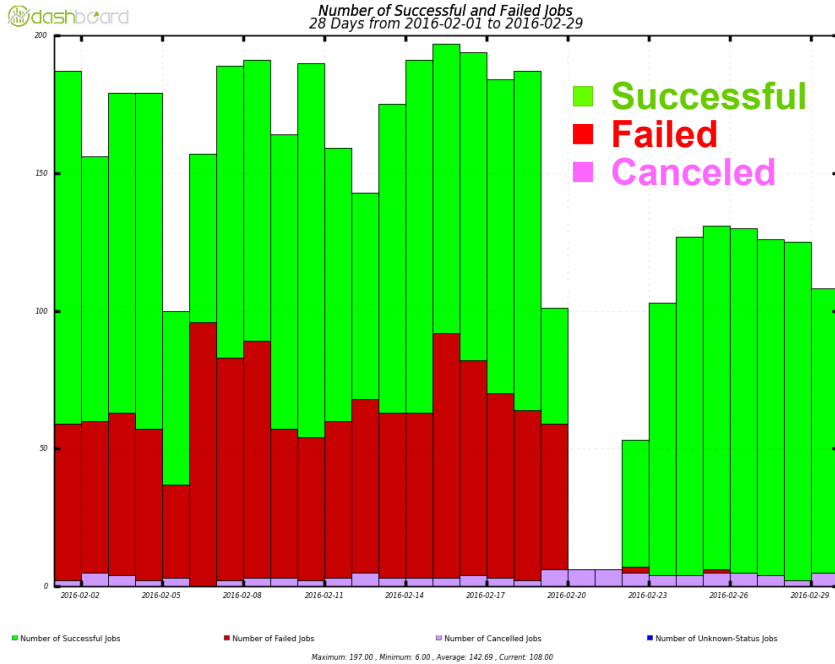
Others

Maximum: 61.00 , Minimum: 0.00 , Average: 16.38 , Current: 21.00

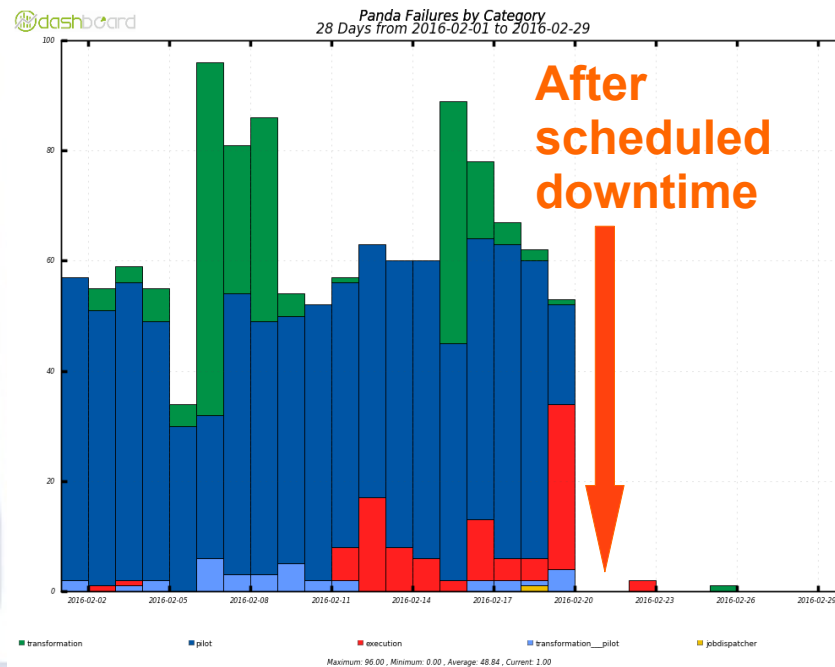
Others

Maximum: 136.00 , Minimum: 0.00 , Average: 97.00 , Current: 103.00

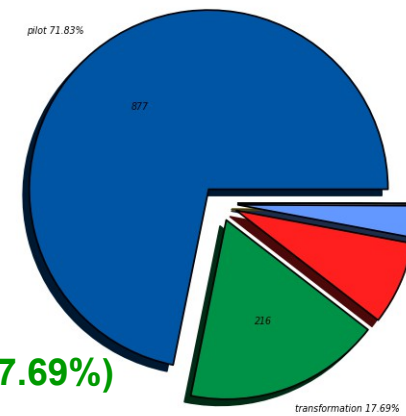
Accounting: ATLAS dashboard (2)



Memory set to 2 GB per process (by default):
Some ATLAS jobs require more than 2 GB
Local users can set it up to 6 GB
Declared at AGIS for ATLAS job submission



Panda Failures by Category (Sum: 1,221)



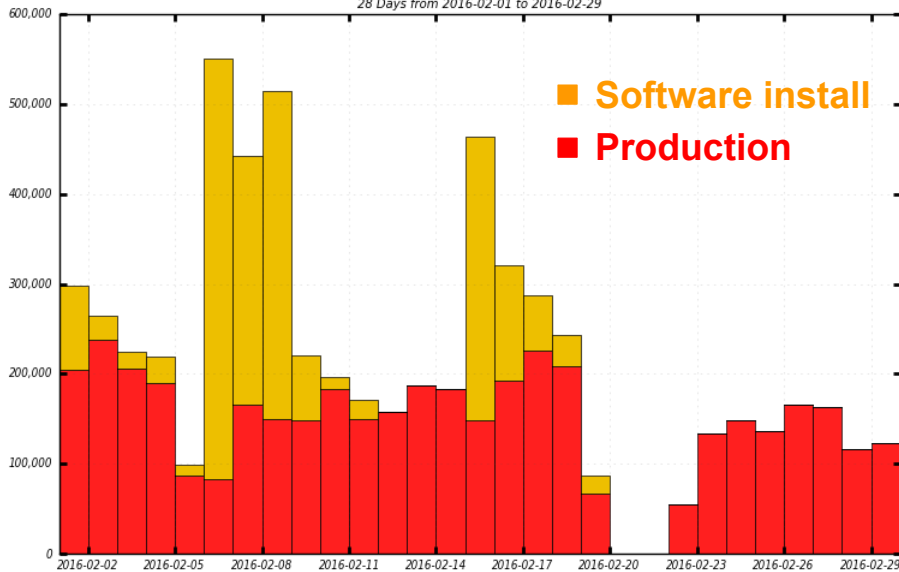
- Pilot (71.83 %)
- Transformation (17.69%)
- Execution (7.53%)
- Transformation_pilot (2.87%)

■ pilot 71.83% (877.00) ■ transformation 17.69% (216.00) ■ execution 7.53% (92.00)
■ transformation_pilot 2.87% (35.00) ■ jobdispatcher 0.08% (1.00)

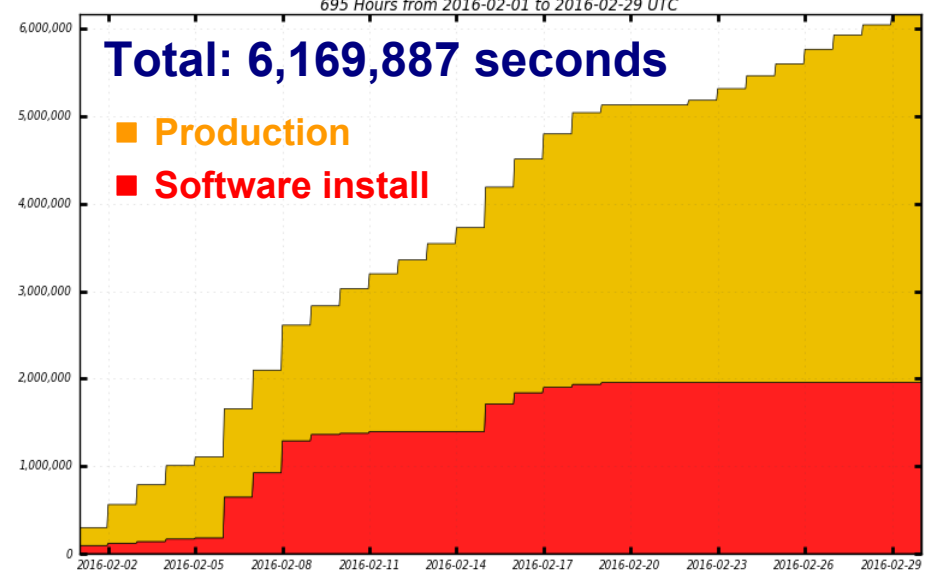
Accounting: ATLAS dashboard (3)



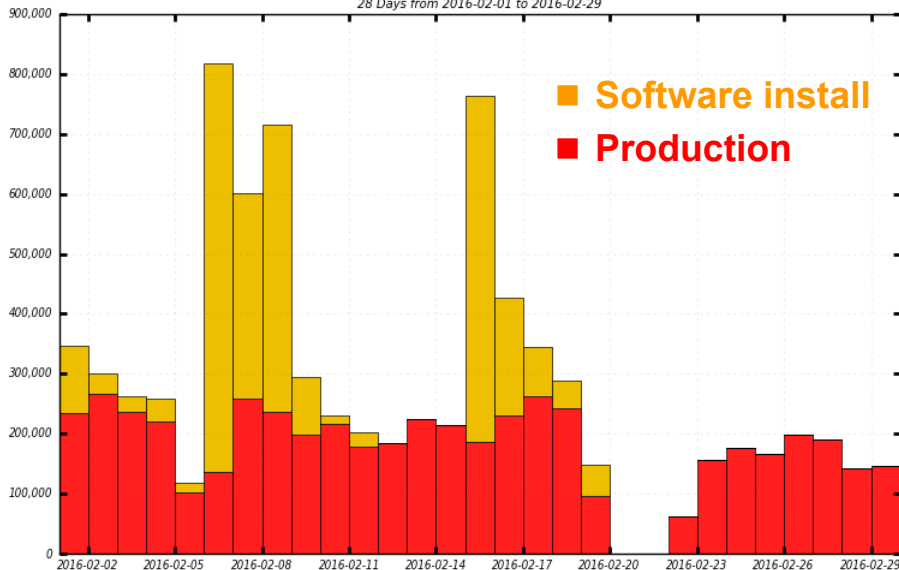
Overall CPU consumptions All Jobs (Time Stacked Bar Graph)
28 Days from 2016-02-01 to 2016-02-29



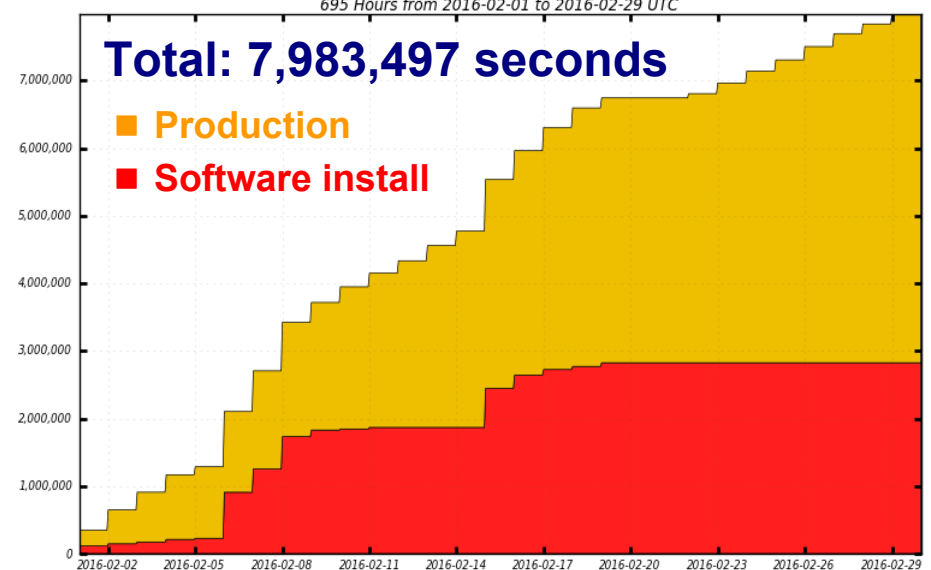
Overall CPU consumptions All Jobs (Cumulative Graph)
695 Hours from 2016-02-01 to 2016-02-29 UTC



Overall Wall Clock consumptions All Jobs (Time Stacked Bar Graph)
28 Days from 2016-02-01 to 2016-02-29



Overall Wall Clock consumptions All Jobs (Cumulative Graph)
695 Hours from 2016-02-01 to 2016-02-29 UTC



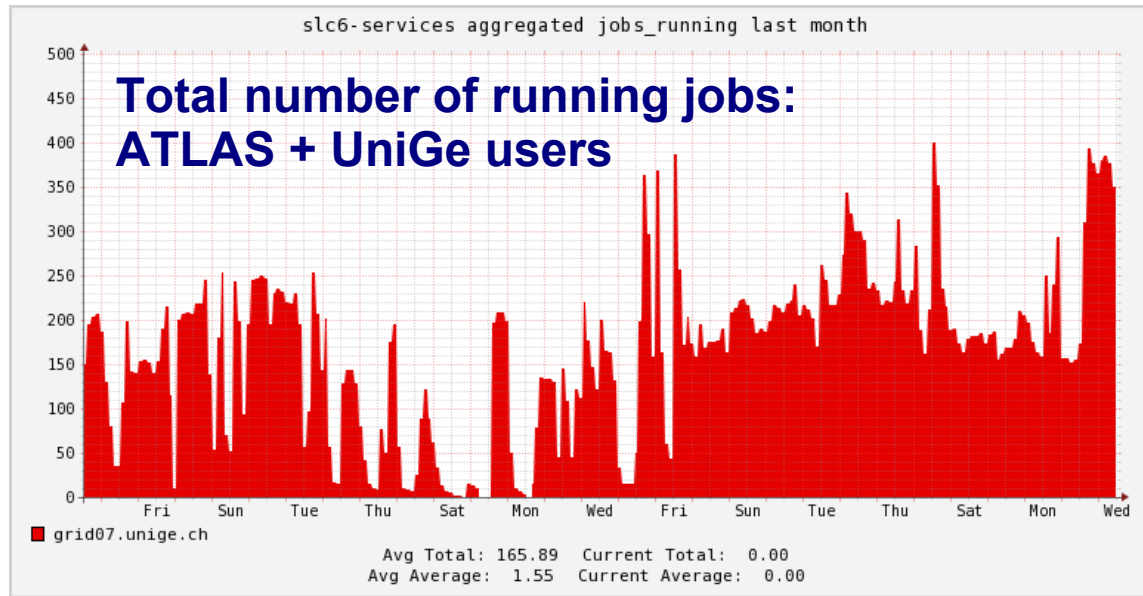
■ install ■ gangarobot-pft

Maximum: 817,500 , Minimum: 0.00 , Average: 275,293 , Current: 146,700

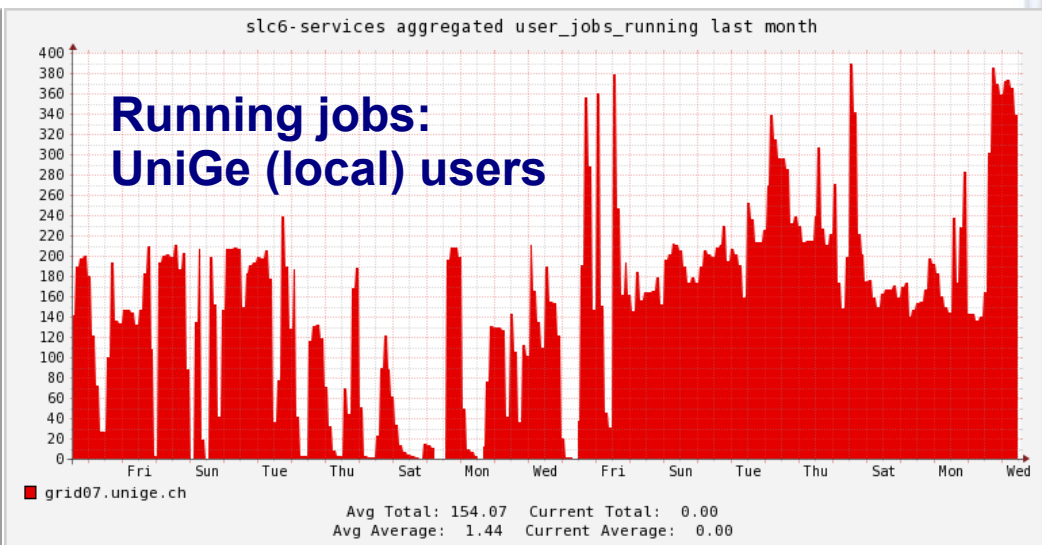
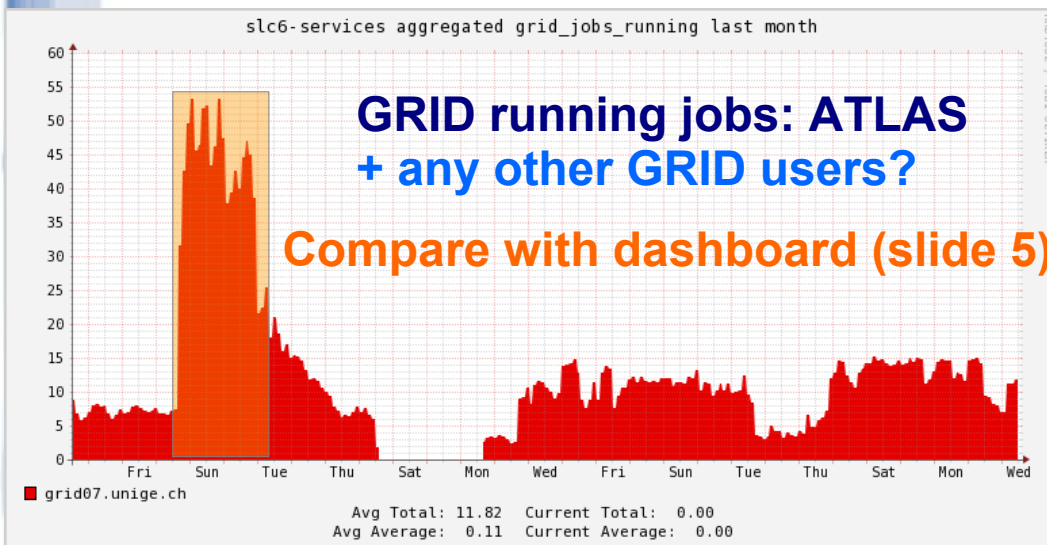
■ gangarobot-pft (5,158,647) ■ install (2,824,850)

Total: 7,983,497 , Average Rate: 3.19 /s

Accounting: Ganglia at the UniGe cluster



Number of running jobs for last month (March 9th 2016 → last month)



Outlook

Operations:

- CPU/cores** → We have up to 240 cores to be added at the cluster
- Batch system** → We would like to move to **SLURM** (currently Torque/PBS)
- Testbed** → We could use some of these CPUs to be tested with SLURM
- GPU machines** → **Funding request submitted to UniGe: Waiting for a reply**
Finally, GPUs would be added into **Baobab HPC cluster**
- ATLAS Production** → We we would like to receive ATLAS production jobs (testing)
We should cross-check/review our accounting

Storage:

- Disk servers** → We would like to add some other disk servers to DPM
11 File Disk Servers with SCL5 (upgrade to SLC6 only if necessary)
- DPM SE** → We would like to move to **Puppet** (currently YAIM)
- Testbed** → We would like to make a small testbed:
 - 1 service machine: Puppet**
 - 1 Head Node: DPM (newer version than current one)**
 - 1 File Disk Server: Data to be managed by DPM**

Network:

- Upgrade to 10 Gb/s** → **Funding request submitted to UniGe: Waiting for reply**

Basically, data transfers from/to NFS disk servers would be faster by a factor ~ 8-9