INFN-T1 site report

Andrea Chierici

On behalf of INFN-T1 staff

HEPiX fall 2010

Overview

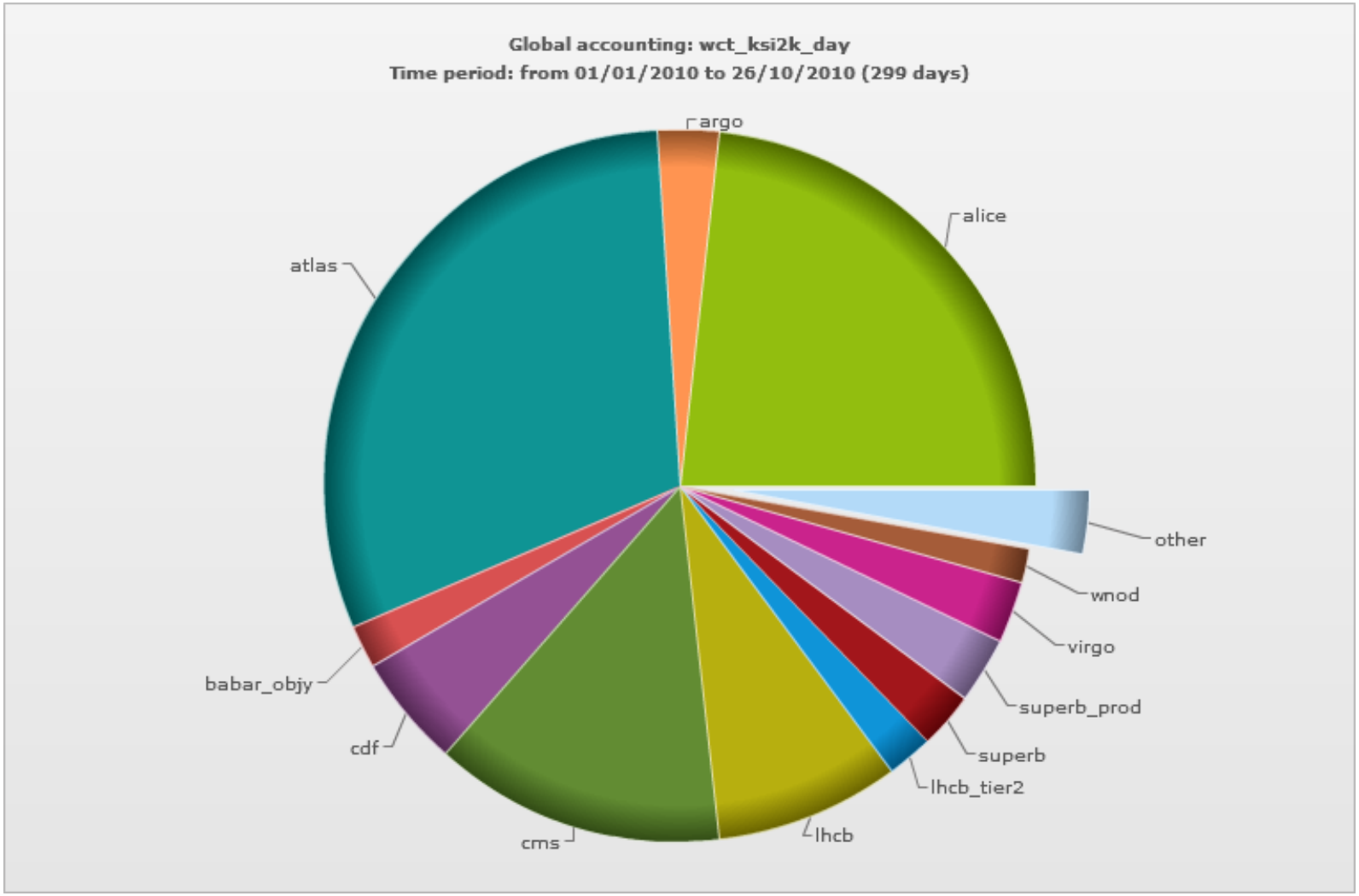
- Farming
- Storage
- User experience

Farming

Computing resources

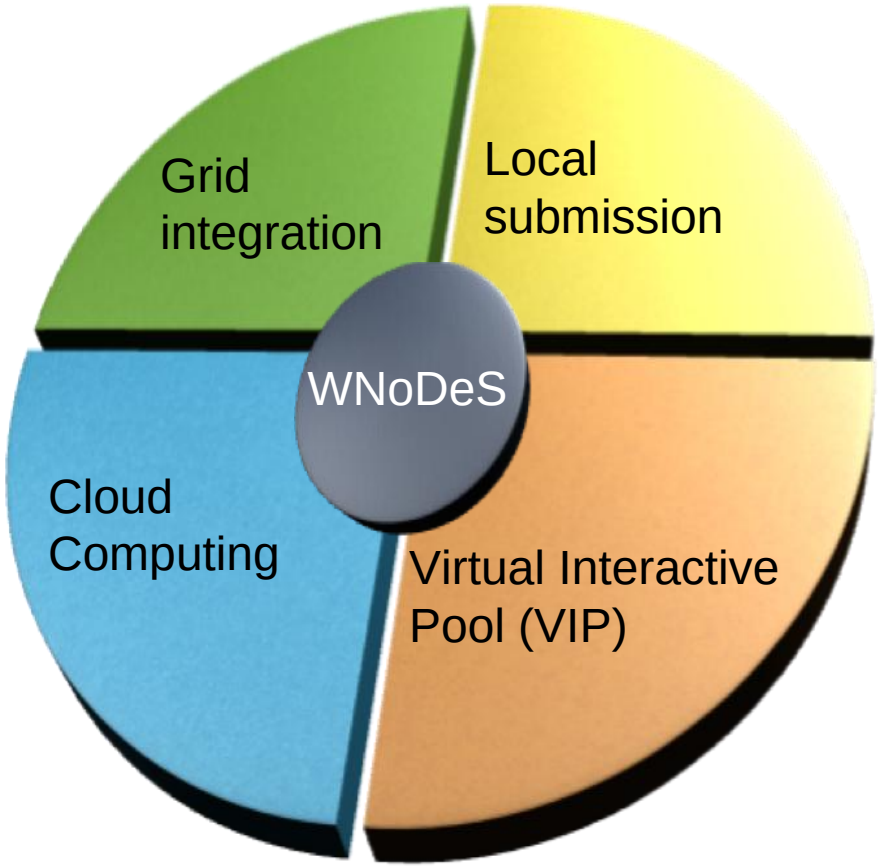
- 80K HS-06
- New tender will add 22.5K HS-06
  - We will reach 100K HS-06 within may 2011
- We host other sites
  - T2 LHCb
  - T3 UniBO (almost ready)

Resource usage per VO

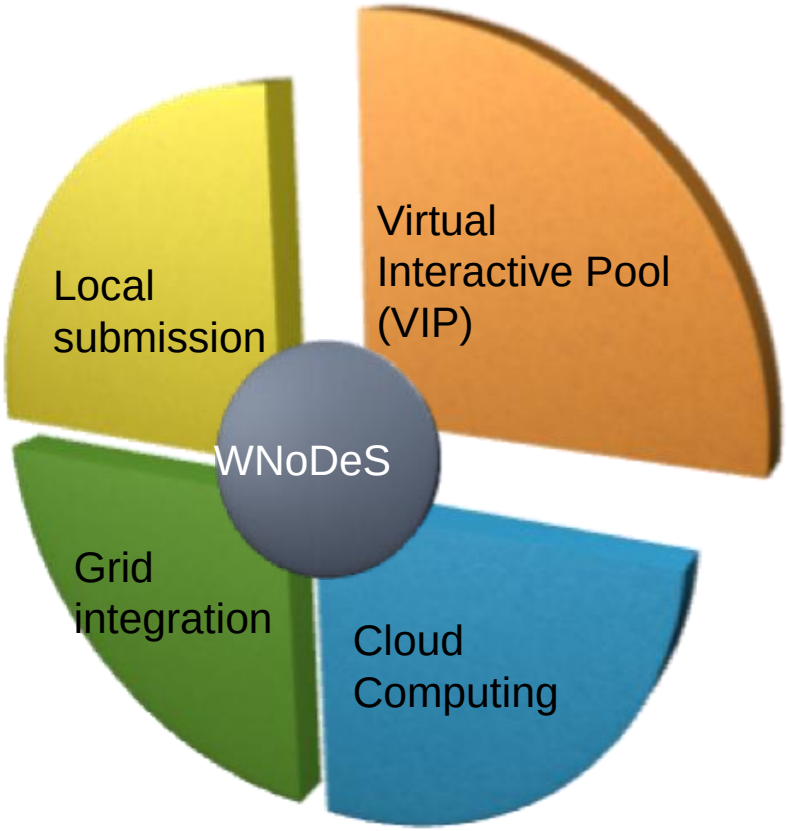


WNoDeS

- The Worker Nodes on Demand Service (WNoDeS) is an INFN-developed architecture that makes it possible to dynamically allocate virtual resources out of a common resource pool.
- WNoDeS is built around a tight integration with LRMS



Virtual Interactive pool (VIP)



VIP a CLI for WNoDeS

VIP interface is more suitable to local user habits

It provides a customized computing resource which is at the same time dedicated.

It can be useful for interactive analysis or software development.

The provisioned virtual resource matches user requirements in terms of RAM, #CPU, Bandwidth and shared file system to be mounted

WNoDeS: specs

- Features
  - Virtual Image selection
  - VLAN support
  - Multi-core
  - Cloud Web-Interface
  - Authentication Gateway
- Technology used
  - virtIO
  - LibguestFS
  - Network throttling

WNoDeS: facts

- Up to 2k VM can run concurrently
  - 20% of our HS-06 computing power
- All VOs supported
  - Some use ONLY WNoDeS
- First deployment outside CNAF done in October
  - Legnaro
- http://web.infn.it/wnodes

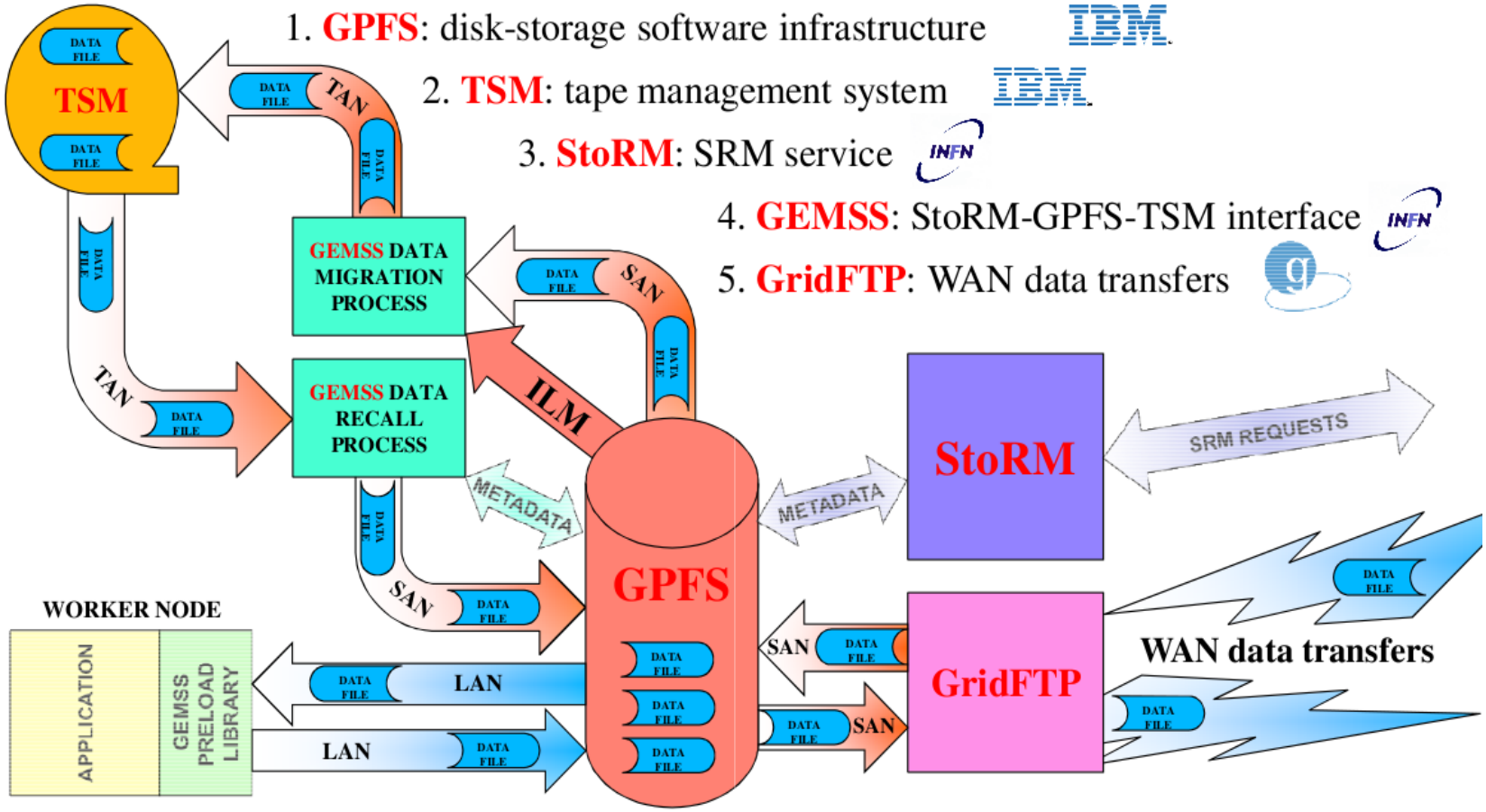
Storage

GEMSS: Grid Enabled Mass Storage System

- Our choice is driven by need **to minimize management effort**:
  - Very positive experience for scalability so far;
  - Large GPFS installation in production at CNAF since 2005 with increasing disk space and number of users;
  - Over 6 PB of net disk space partitioned in several GPFS clusters served by less than 100 disk-servers (NSD + gridFTP);
  - 2 FTE employed to manage the full system;

GEMSS: building blocks

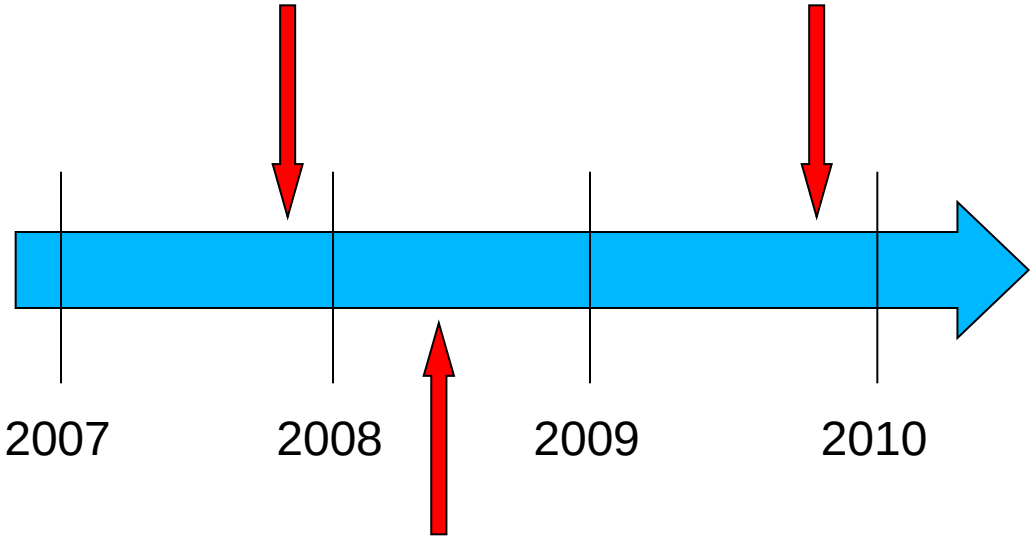
- **Disk-centric system** with five fundamental components
  1. **GPFS**: disk-storage software infrastructure
  2. **TSM**: tape management system
  3. **StoRM**: SRM service
  4. **GEMSS**: StoRM-GPFS-TSM interface
  5. **GridFTP**: WAN data transfers



GEMSS: Timeline

D1T0 Storage Class implemented @Tier1 with StoRM/GPFS since Nov. 07 for LHCb and ATLAS

D0T1 Storage Class implemented @Tier1 with StoRM/GPFS/TSM since Oct. 09 for CMS



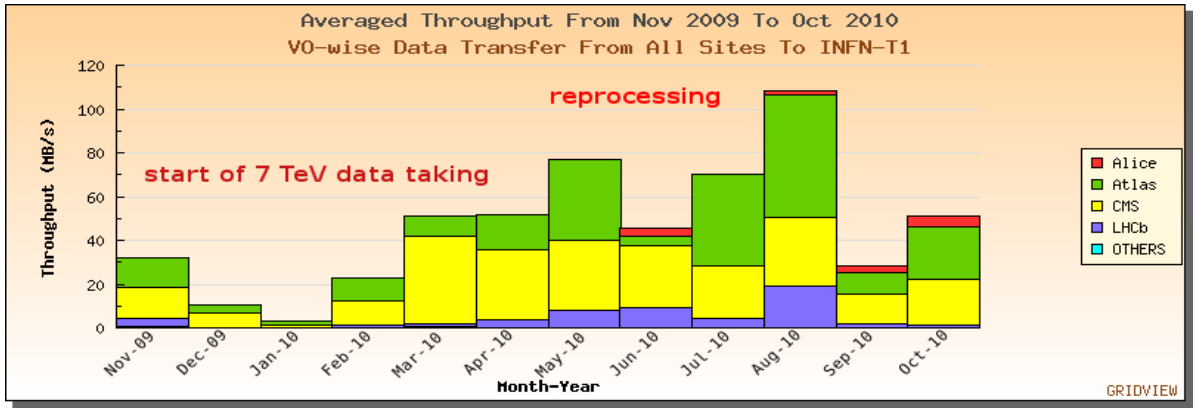
D1T1 Storage Class implemented @Tier1 with StoRM/GPFS/TSM since May 08 for LHCb

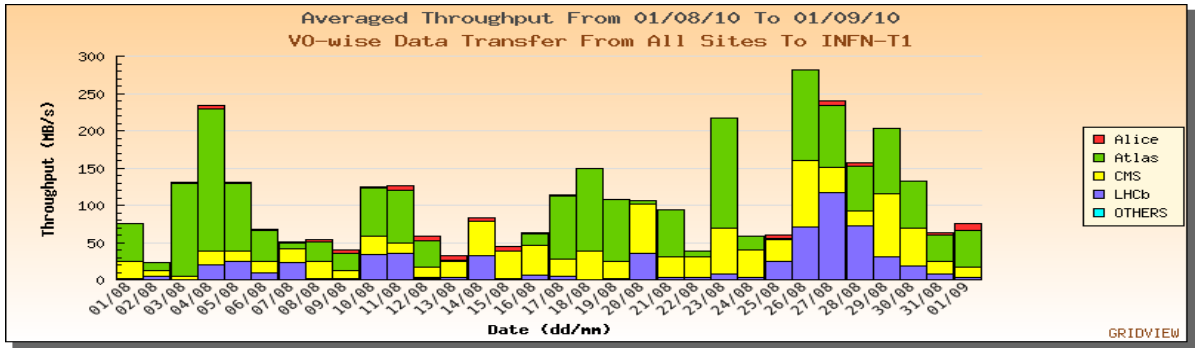
In 2010 ATLAS, ALICE, CMS and LHCb experiments, together with other non-LHC experiments (Argo, Pamela, Virgo, AMS), have agreed to use GEMSS

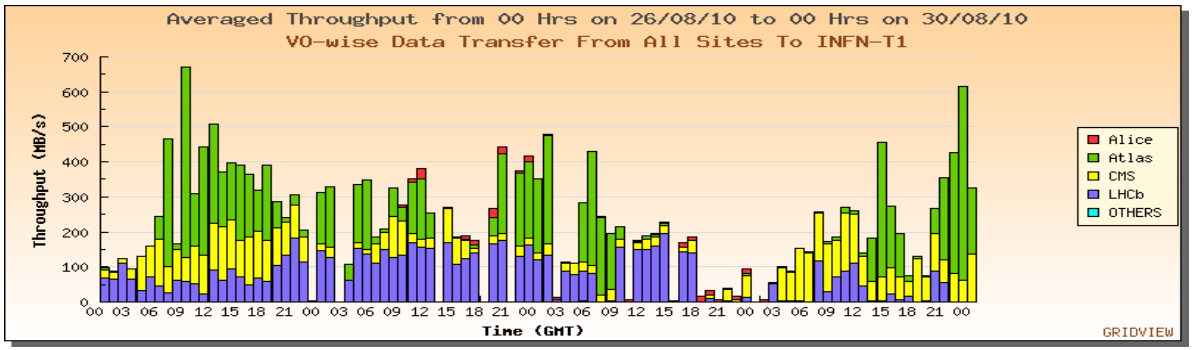
GEMSS is now used by all LHC and non-LHC experiments in production for all Storage Classes:

StoRM 1.5, TSM 6.2, GPFS 3.2.1-23

GEMMS: Throughput from all sites to INFN-T1







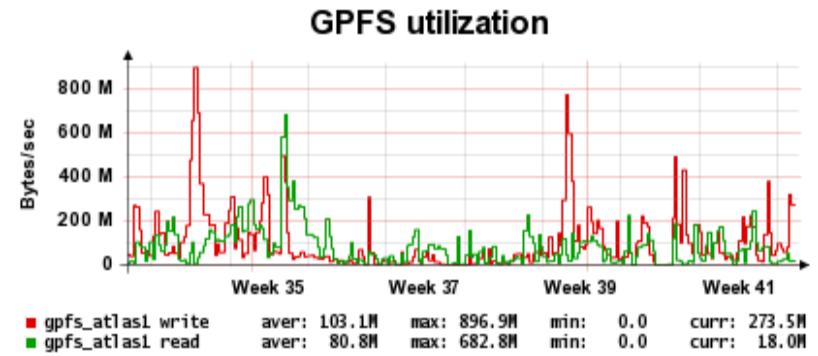
- **Monthly throughput per VO**

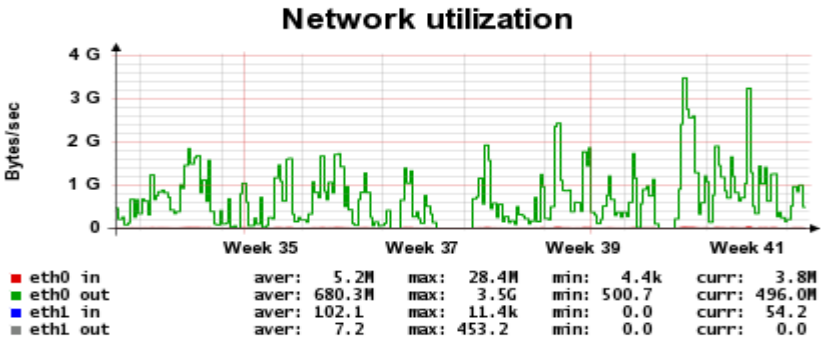
Period	Throughput peak
Monthly	~120 MB/s
Daily	~300MB/s
Hourly	~700 MB/s

- **Daily throughput per VO during August, the most intesive month since data taking**
- **Hourly throughput per VO in the most intensive four days of August**

GEMSS Disk-Only System

GPFS and gridFTP throughput results from the ATLAS experiment, which runs most intensive activities on disk, are:



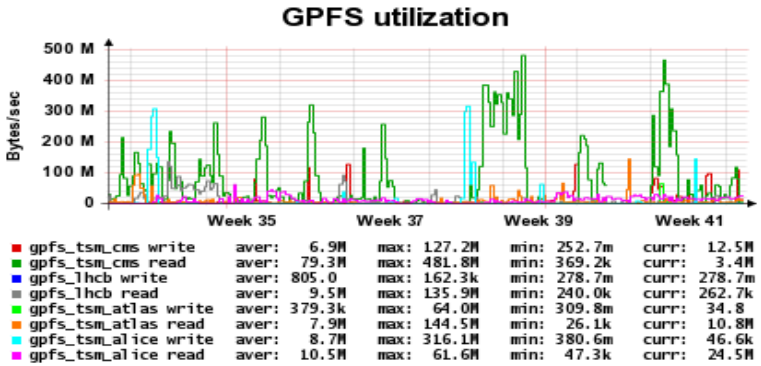


- **ATLAS jobs access from computing farm**
- **ATLAS storage access on WAN**

Service	Throughput peak
GPFS	~900 MB/s
gridFTP	> 3.5GB/s

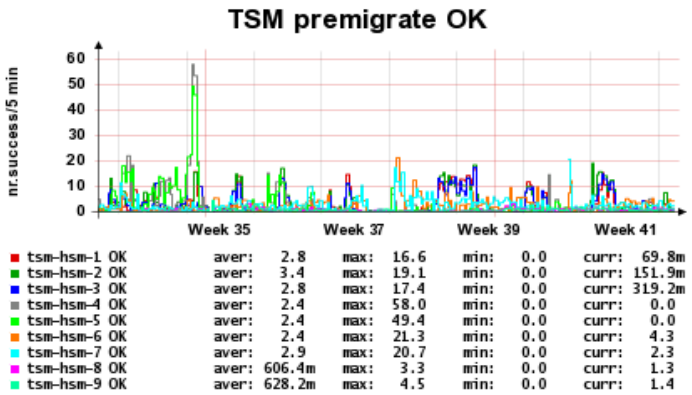
GEMSS TSM HSM System

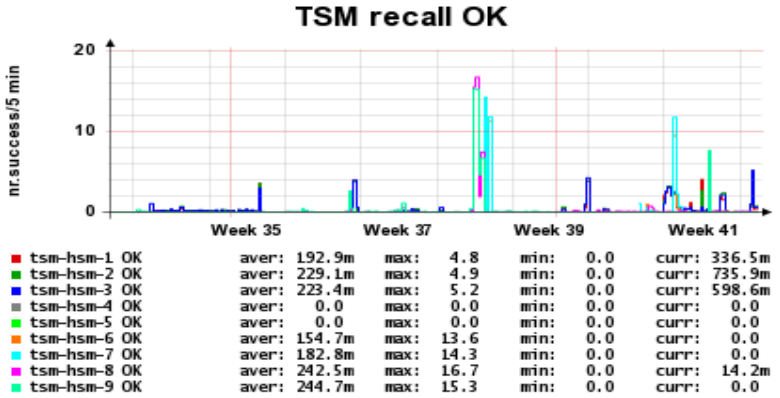
- Example of throughput on TSM HSM tape servers in last two months for all VOs is:



Service	Throughput peak	Efficiency
GPFS	500 MB/s	99%

- Number of succeeded pre-migrate and recalled files for all VOs are:



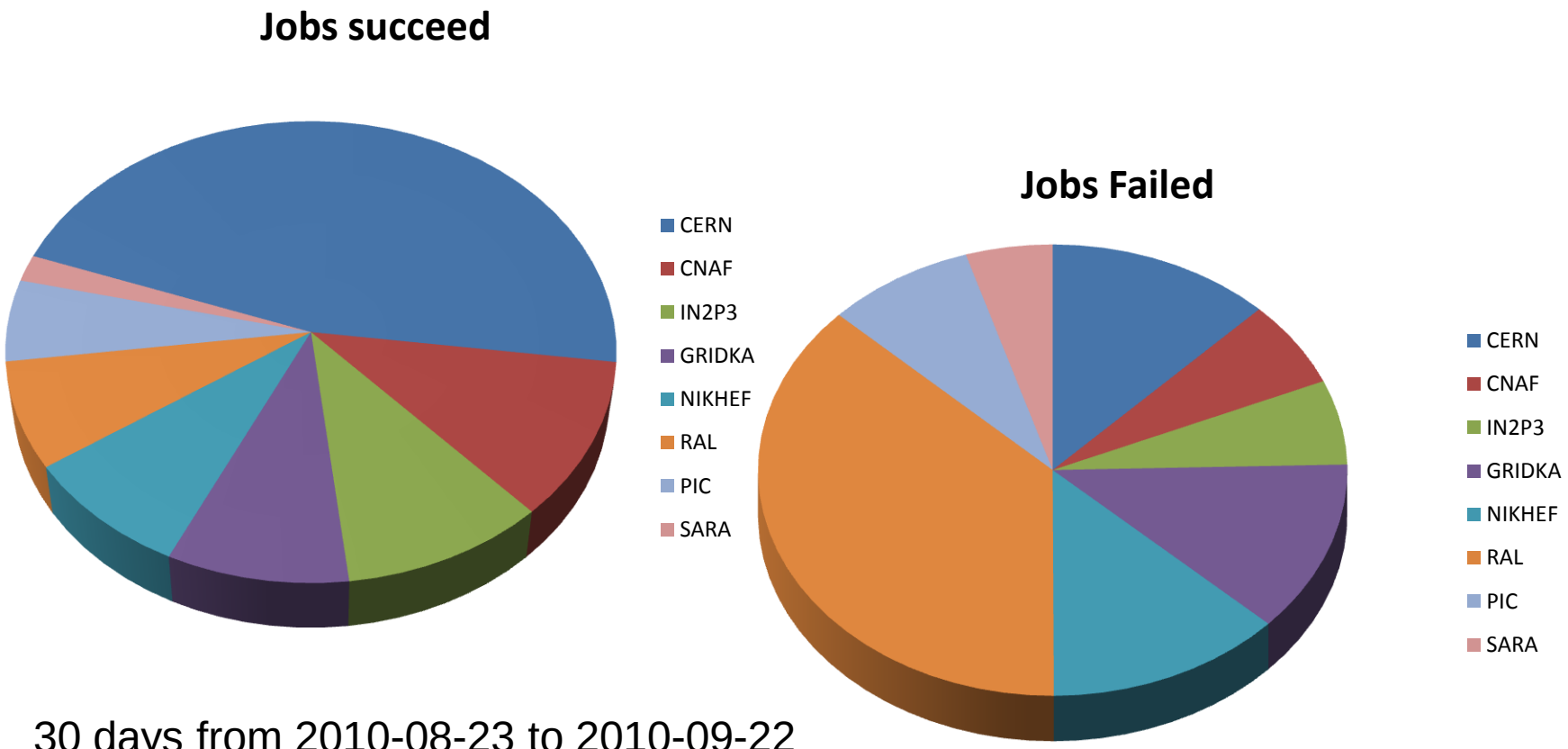


GEMMS Conclusions

- LHC experiments are using GEMSS as MSS at INFN Tier-1 since the starting of data taking.
- The first year of production **has been very promising** in terms of **performances and scalability**.
- **Non-LHC experiments** are now starting using the GEMSS solution.

User experience

Jobs



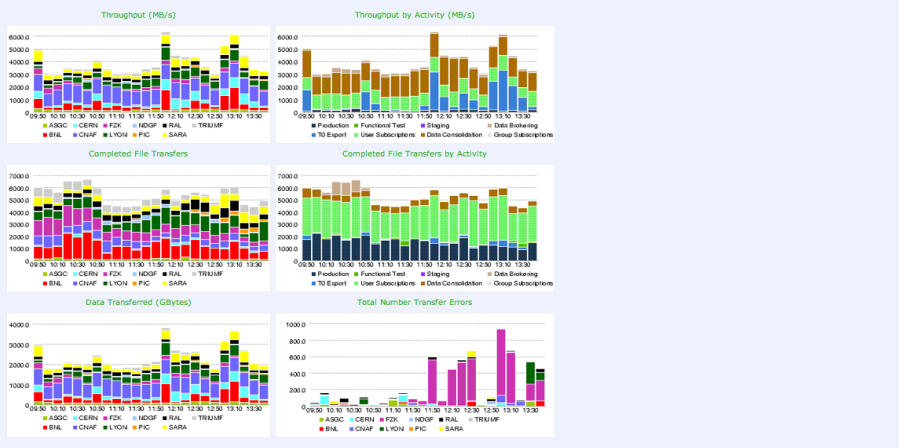
30 days from 2010-08-23 to 2010-09-22

First Tier-1 in terms of succeeded jobs, one of the latest in terms of failed jobs

Site availability

2010-01-01 to 2010-09-30

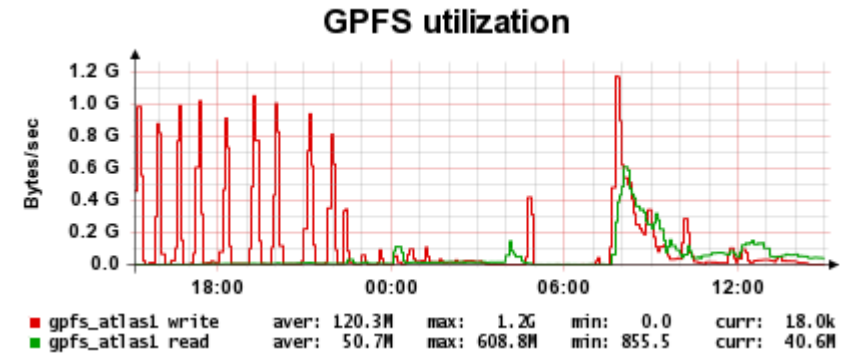
ATLAS activities



Activity Summary ('2010-08-12 09:50' to '2010-08-12 13:50' UTC)

Click on the cloud name to view list of sites

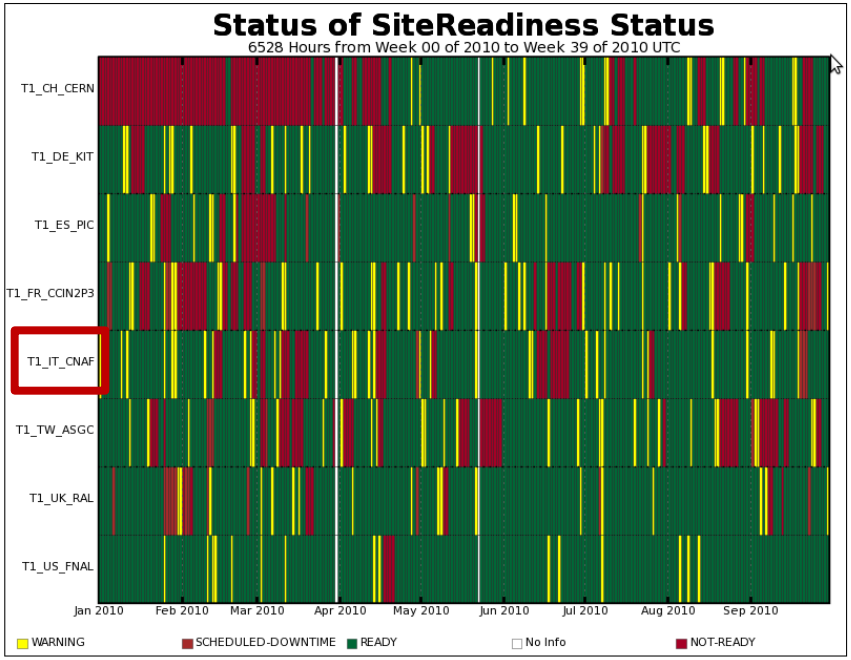
	Transfers			Registrations		Errors			Services
Cloud	Efficiency	Throughput	Successes	Datasets	Files	Transfer	Registration	Services	Grid
ASGC	92%	149 MB/s	2803	55	2802	246	0	0	
BNL	99%	500 MB/s	28166	478	28173	163	0	0	
CERN	87%	373 MB/s	2172	67	2174	332	0	0	
CNAF	98%	1223 MB/s	14001	99	14017	255	0	0	
FZK	83%	204 MB/s	20161	1452	20163	4113	0	0	
LYON	98%	524 MB/s	20971	574	20988	454	0	0	
NDGF	98%	49 MB/s	2681	49	2677	42	0	0	
PIC	100%	34 MB/s	2354	127	2357	6	0	0	
RAL	97%	276 MB/s	12144	458	12158	425	0	0	
SARA	99%	436 MB/s	10607	430	10615	153	0	0	
TRIUMF	100%	132 MB/s	13022	355	13021	0	0	0	

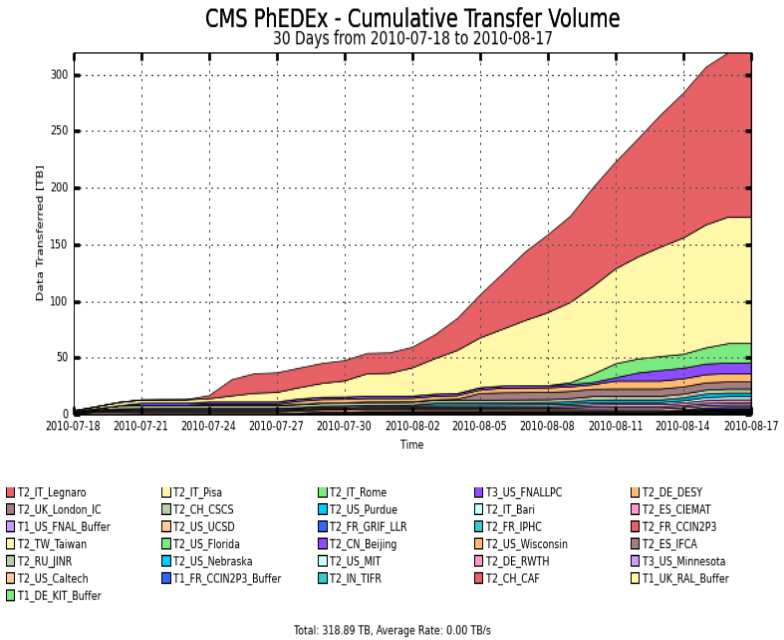


- From ATLAS dashboard: Exceptional throughput (1.2 GB/s) to CNAF due to
  . Tier-0 export;
  . data consolidation;
  . user subscription of data to the site.

- ATLAS saturation of the OPN link between CNAF and CERN on July 13th and 14th 2010. 4000 jobs accessing storage from computing farm.

CMS activities



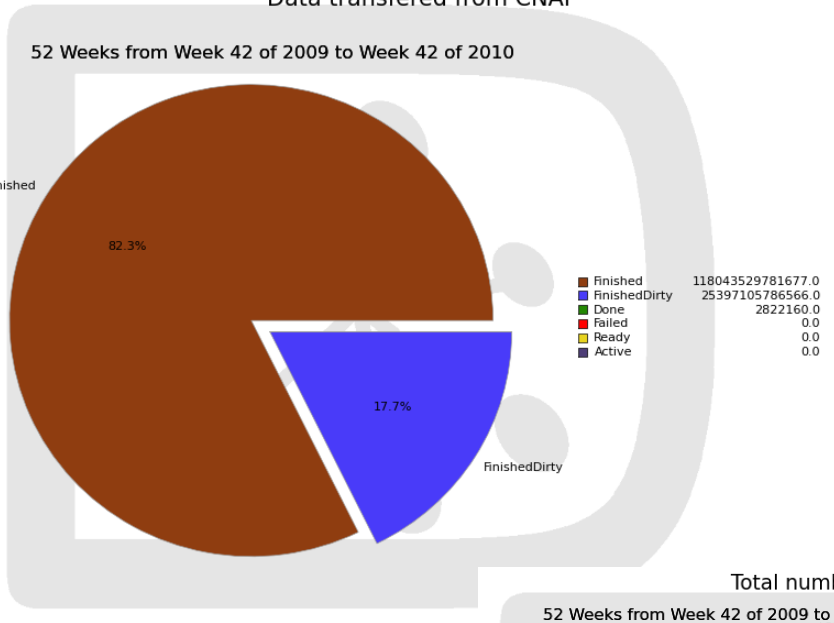


Status of Tier1 readiness in 2010

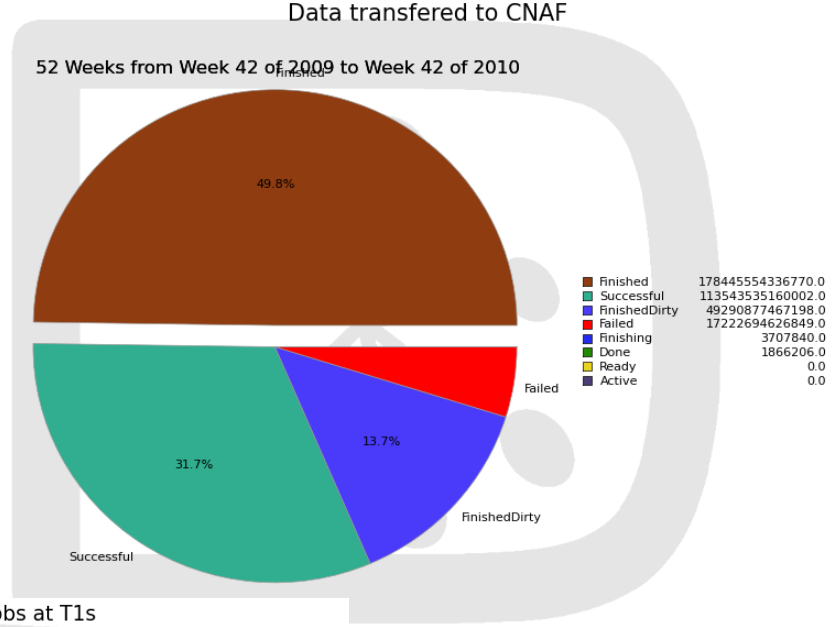
Transfer Volume from CNAF to other sites from mid-July to mid-August

LHCb activities

Data transfered from CNAF

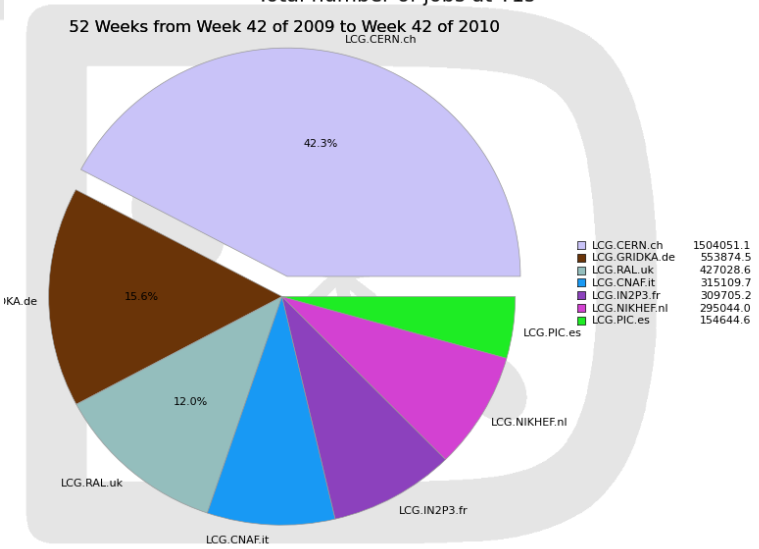


Data transfered to CNAF



Generated on 2010-10-21 13:33:25 UTC

Total number of jobs at T1s



Generated on 2010-10-18 13:56:49 UTC

Questions?

Backup Slides

GEMSS layout for LHC Experiments at INFN Tier-1