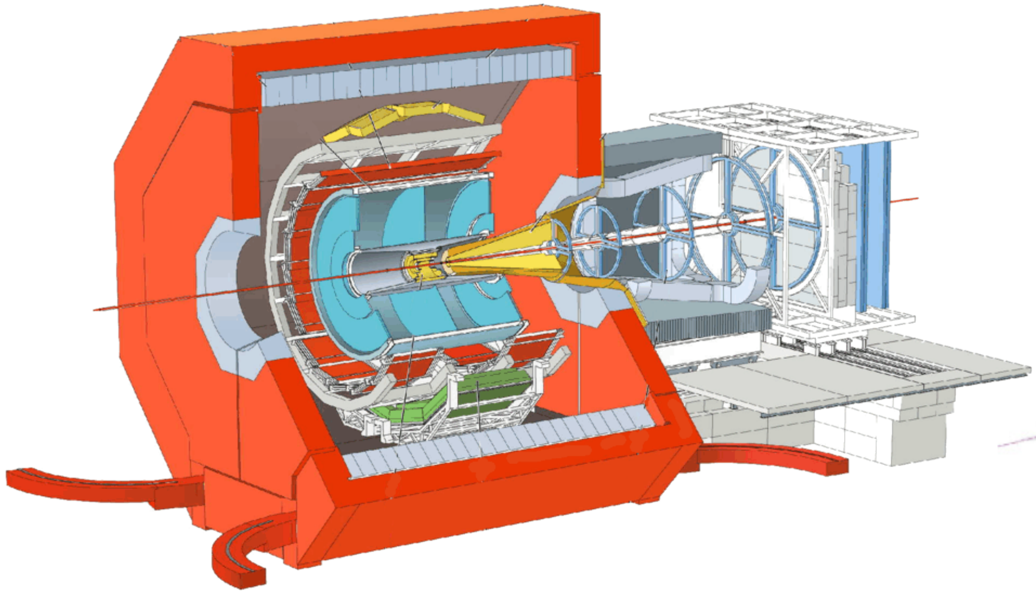
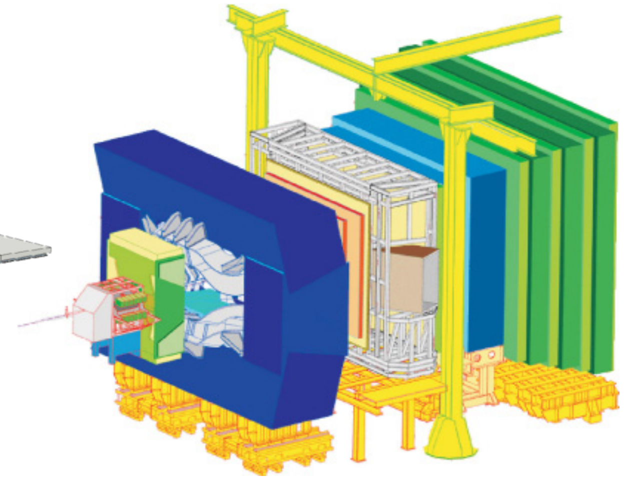
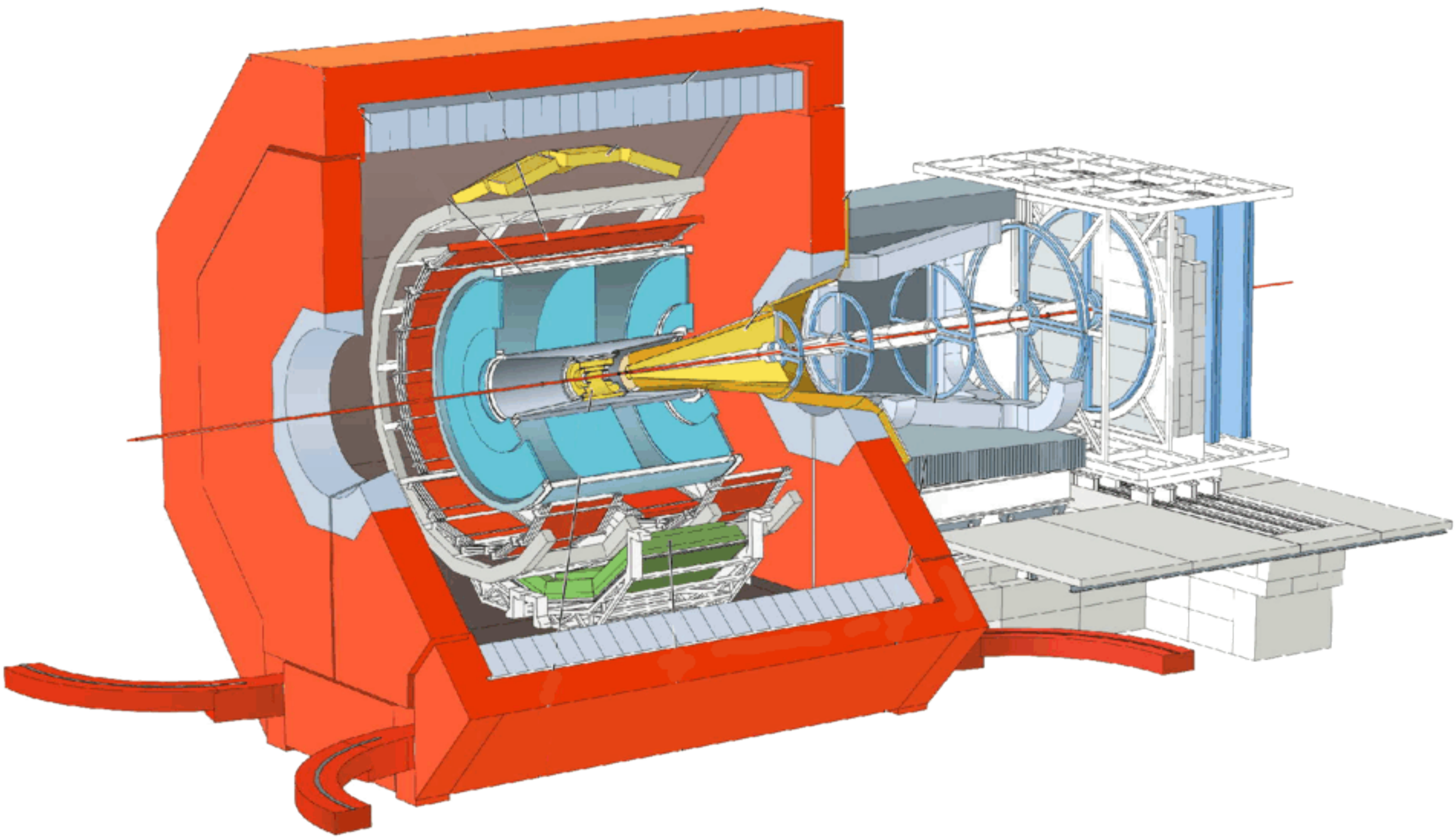



EVOLUZIONE DEL CALCOLO: ALICE E LHCB

Stefano Bagnasco, INFN Torino
Vincenzo Vagnoni, INFN Bologna

ALICE



The original ALICE CM

- For **pp** similar to the other experiments
  - 1.6MB Raw, 0.1MB ESD+AOD
  - 0.11 kHS06 s/event
  - Quasi-online data distribution, calibration and first reconstruction at Tier-0
  - Further reconstructions at Tier-1's

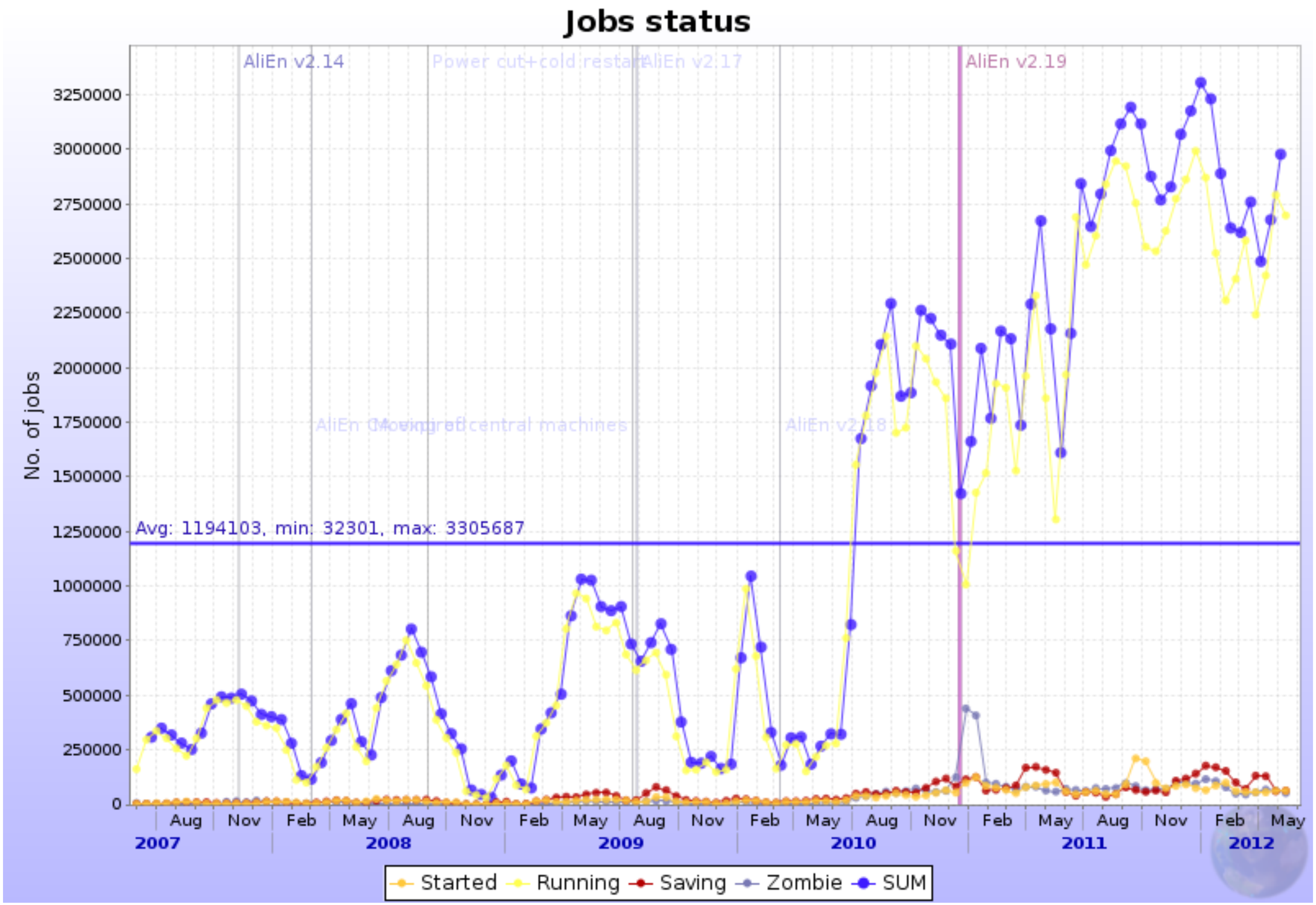
- For **HI** different model
  - 3.5MB Raw, 3.9MB ESD+AOD
  - 2.0 kHS06 s/event
  - Online calibration, alignment, pilot reconstructions and partial data export during data taking
  - Data distribution and Pass1 reconstruction at Tier-0 in the four months after HI run (during shutdown)
  - Further reconstruction passes (one) at Tier-1's

The original ALICE CM

- Three kinds of data analysis
  - **Fast pilot analysis** to tune the first reconstruction at CERN Analysis Facility (CAF)
  - **Scheduled batch analysis** on the Grid (Analysis Trains: ESDs and AODs)
  - **End-user interactive or batch analysis** on AAFs and Grid (AODs and ESDs)

- **T0 (CERN)**
  - Does: first pass reconstruction; calibration and alignment
  - Stores: one copy of RAW, calibration data and first-pass ESDs
- **T1s**
  - Does: reconstructions and scheduled batch analysis
  - Stores: second collective copy of RAW, one copy of all data to be kept, disk replicas of ESDs and AODs
- **T2s**
  - Does: simulation and end-user analysis
  - Stores: disk replicas of AODs and ESDs

Five years of running jobs



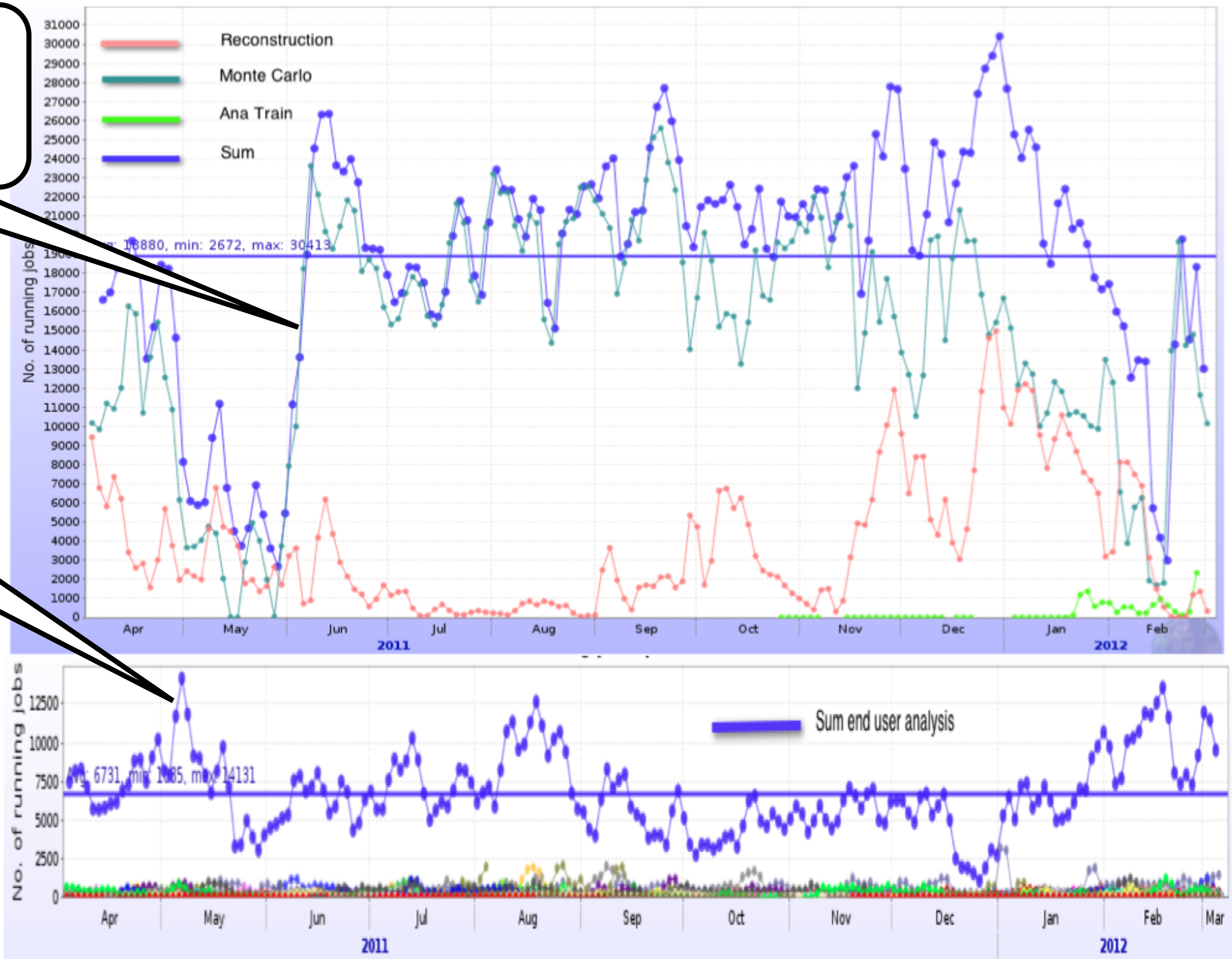
Three job classes

- MC simulation & reco production
  - Low I/O, high CPU efficiency
  - Data export after job completion
  - Managed, scheduled
- Analysis Trains
  - Optimized I/O (read once, do many tasks)
  - Streamlined code (as much as possible…)
  - Managed, scheduled
- User jobs
  - Lowest CPU efficiency
  - Variable job duration, lots of failures, far-from-perfect code
  - Unmanaged, chaotic

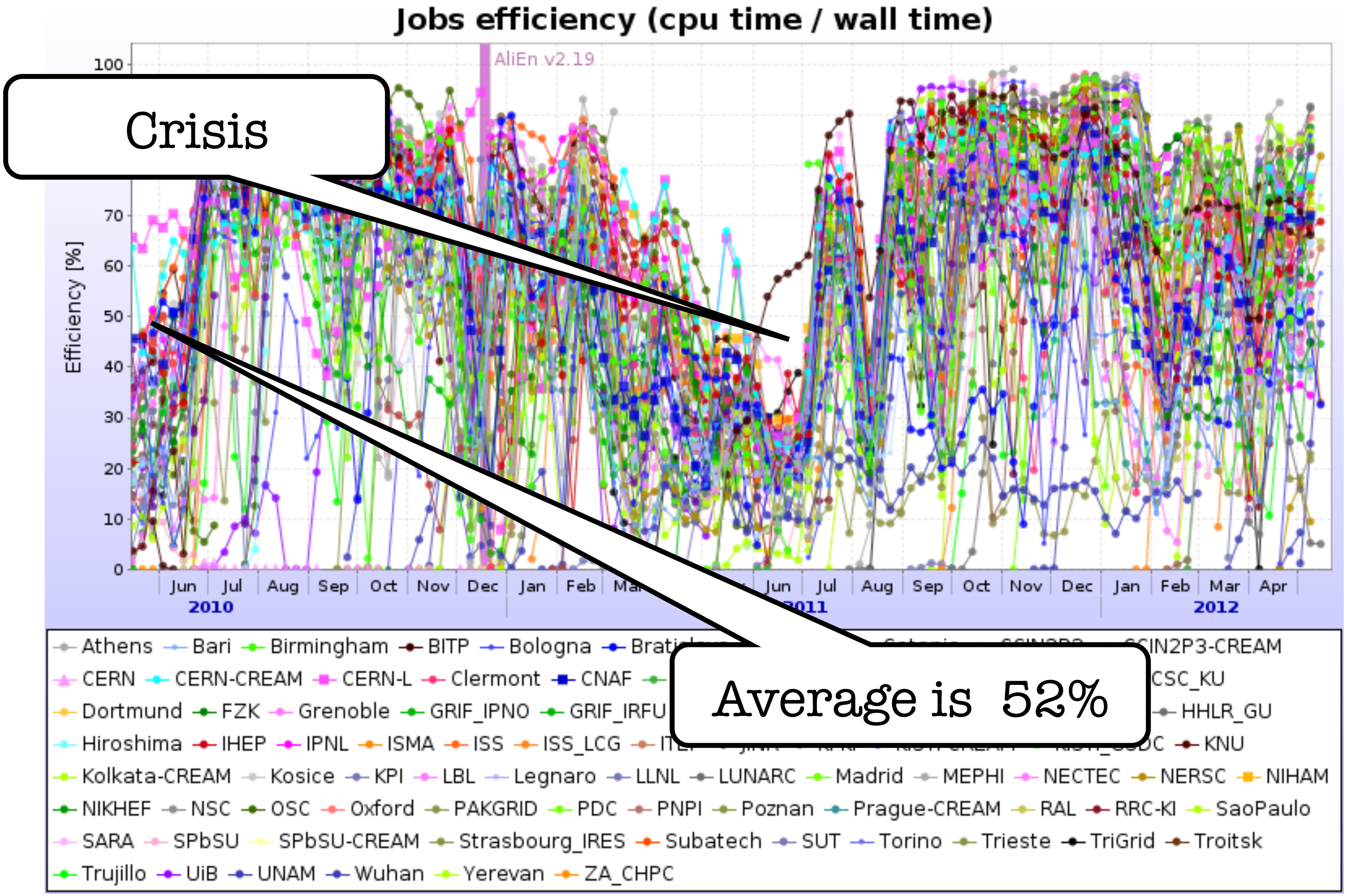
Computing activities

22% Reco &
Analysis Train
56% Montecarlo

22% Chaotic
analysis



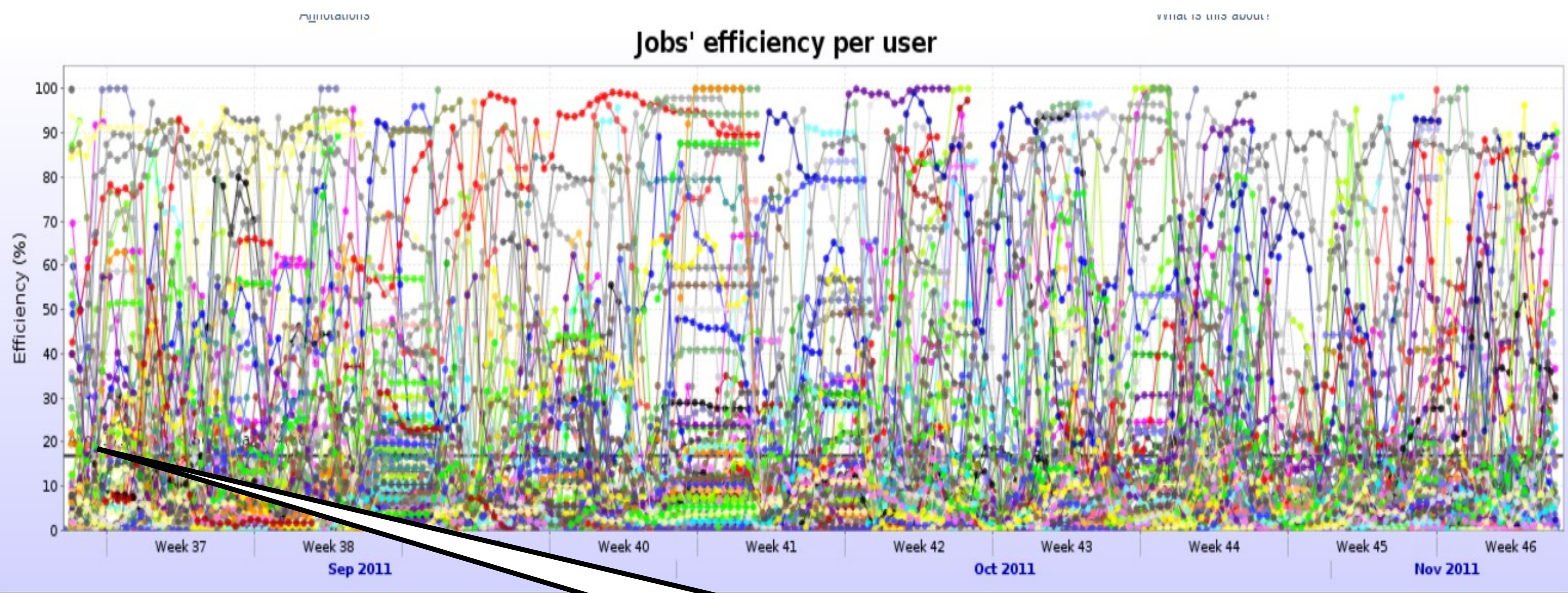
The CPU Efficiency Crisis



What was done

- Modifications of OCDB infrastructure
  - Addition of caches, APIservers
- Improvement of RAW processing
  - Reconstruction algorithm
  - Data access model
- More efficient trains
- Move user analysis from ESD to AOD
- Still…

User jobs



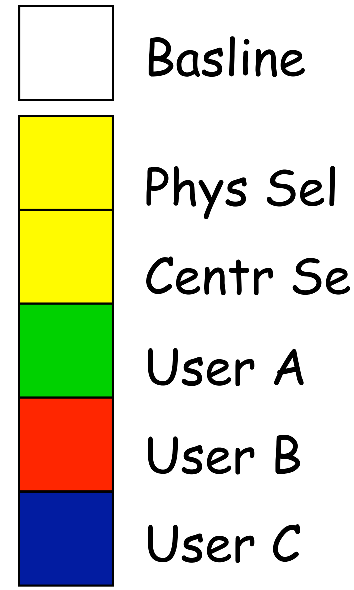
Average is 20%

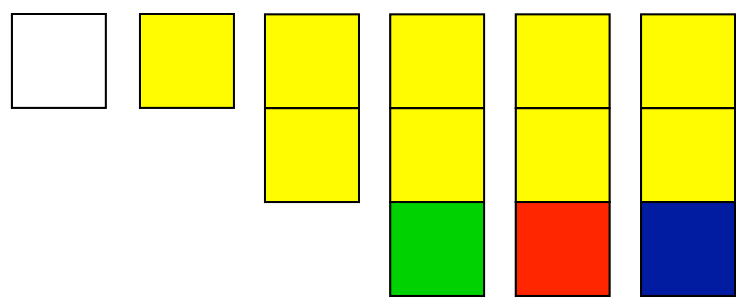
LEGO Framework

- Manage trains using MonALISA
  - Users register wagons
  - Train operators compose trains
- **Automatic testing** per wagon
- Train file generation
- Submission managed by ML
  - Existing LPM infrastructure
- Merging managed by LPM
- **Aim:** allow operators easy running of analysis trains (weekly) getting output on the scale of 1-2 days

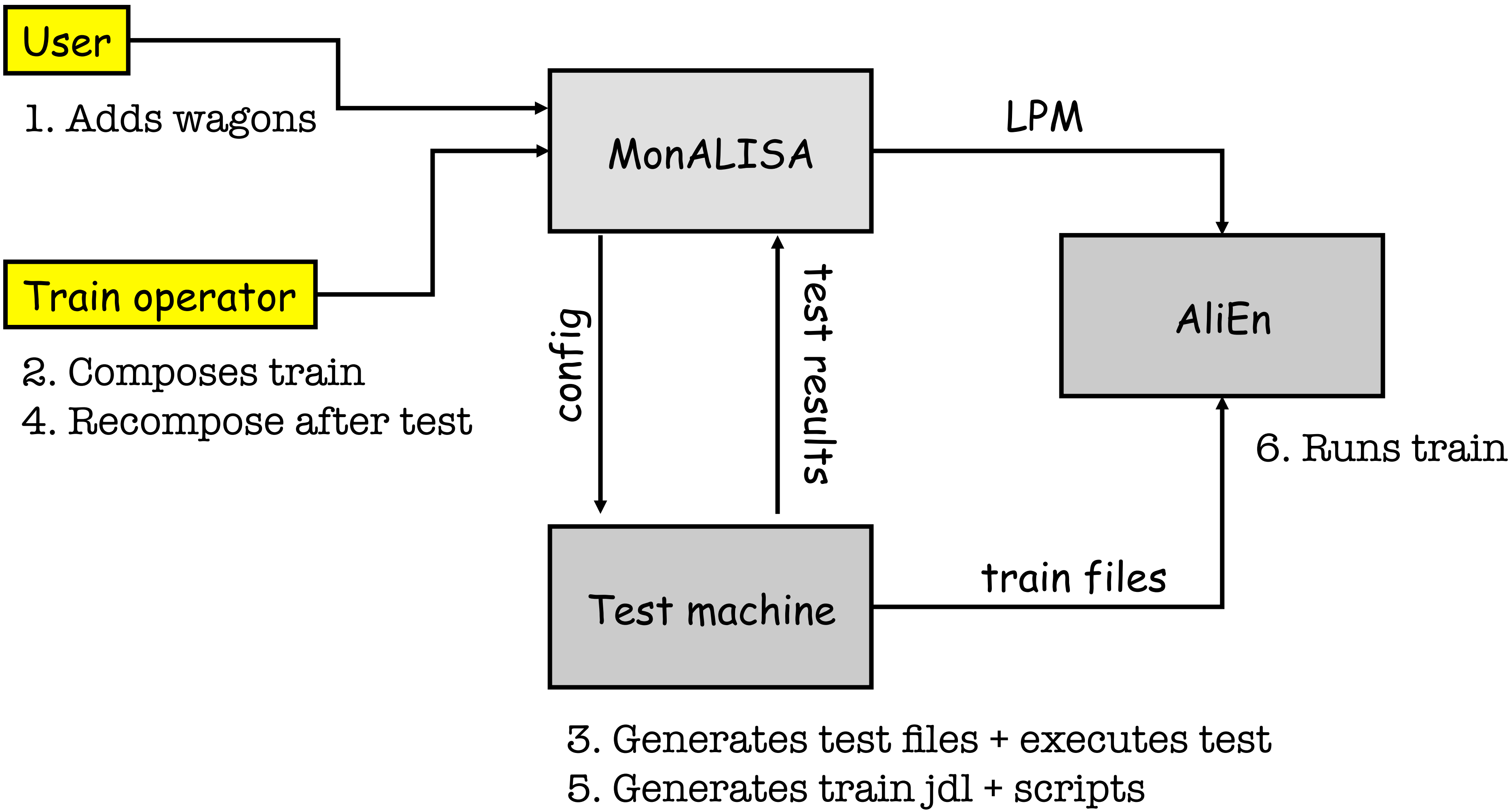
Configuration & Testing

- Train Configuration
  - New class AliAnalysisTaskCfg
  - Contains description of wagons (task macro, libraries, dependencies)
- Testing
  - Runs tests **per wagon**
  - Extracts mem/cpu information
  - Tests also empty "baseline" task





Workflow


User
1. Adds wagons
MonALISA
LPM
Train operator
2. Composes train
4. Recompose after test
config
test results
AliEn
6. Runs train
Test machine
train files
3. Generates test files + executes test
5. Generates train jdl + scripts

The LEGO Framework

Name	Devel_1 (train temporary file dir)
PWG	1
Description	Development testing train #1

Handlers

Name	Macro path (parameters)	Body	Actions
AOD handler	ANALYSIS/macros/train/AddAODHandler.C ( )	handler->Dump();	

Add new handler »

Wagons

Name	Owner	Macro path (parameters)	Libraries	Dependencies	Enabled	Last run	Actions
PhiCorrelations	grigoras	PWG4/macros /AddTaskPhiCorrelations.C ( )	CORRFW,EMCALUtils,JETAN,PWG4JetTasks		✔	8	
PhiCorrelationsQA	grigoras	PWG4/macros /AddTaskPhiCorrelationsQA.C (kTRUE)	libPWG4JetTasks.so,libJETAN.so		✖	2	
test wagon 2	grigoras	something (param)	l1.so	PhiCorrelations,PhiCorrelationsQA	✖		

Add new wagon »

Periods

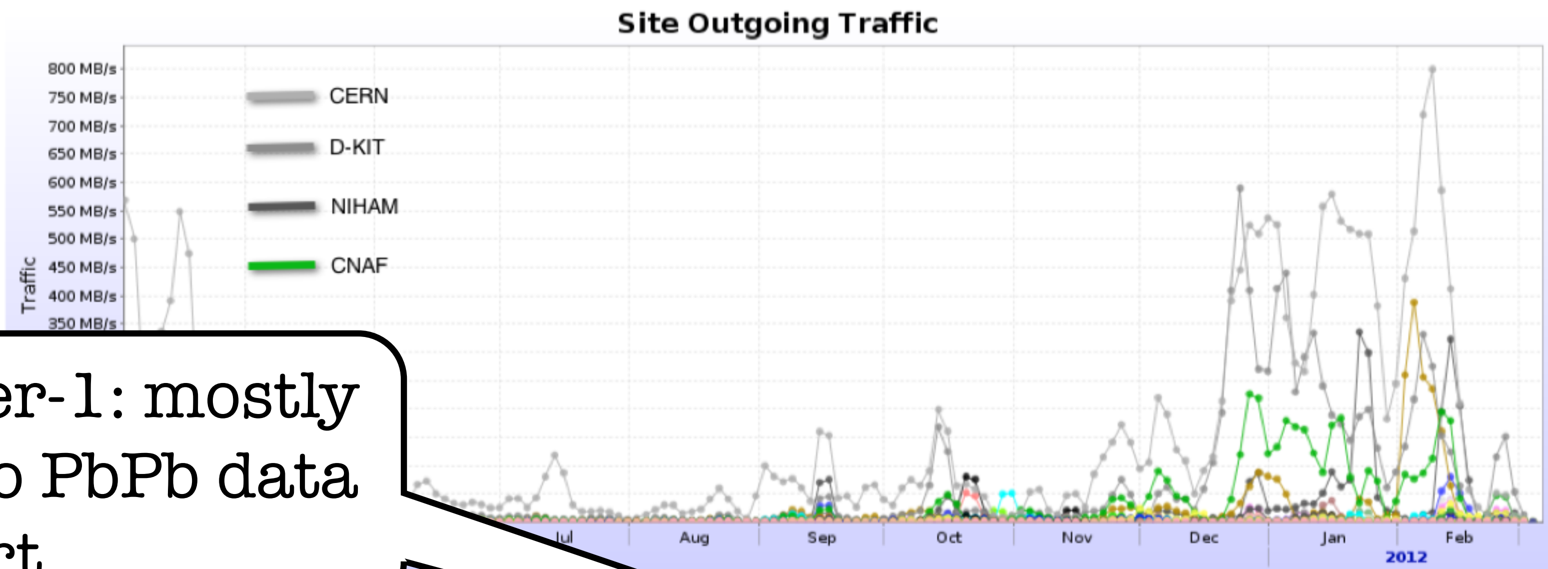
Period name	Reference production	Run list	Description	Last analyzed	Actions
LHC10h(2)	LHC10h(2)	123456:130000,130010,130020	LHC10h - pass2, the ...	2	
AODs_73	FILTER_Pb-Pb_073_LHC10h			8	
AODs	FILTER_Pb-Pb_049_LHC10h_Stage3	136854, 139513, 139514, 139517		3	

Add new period »

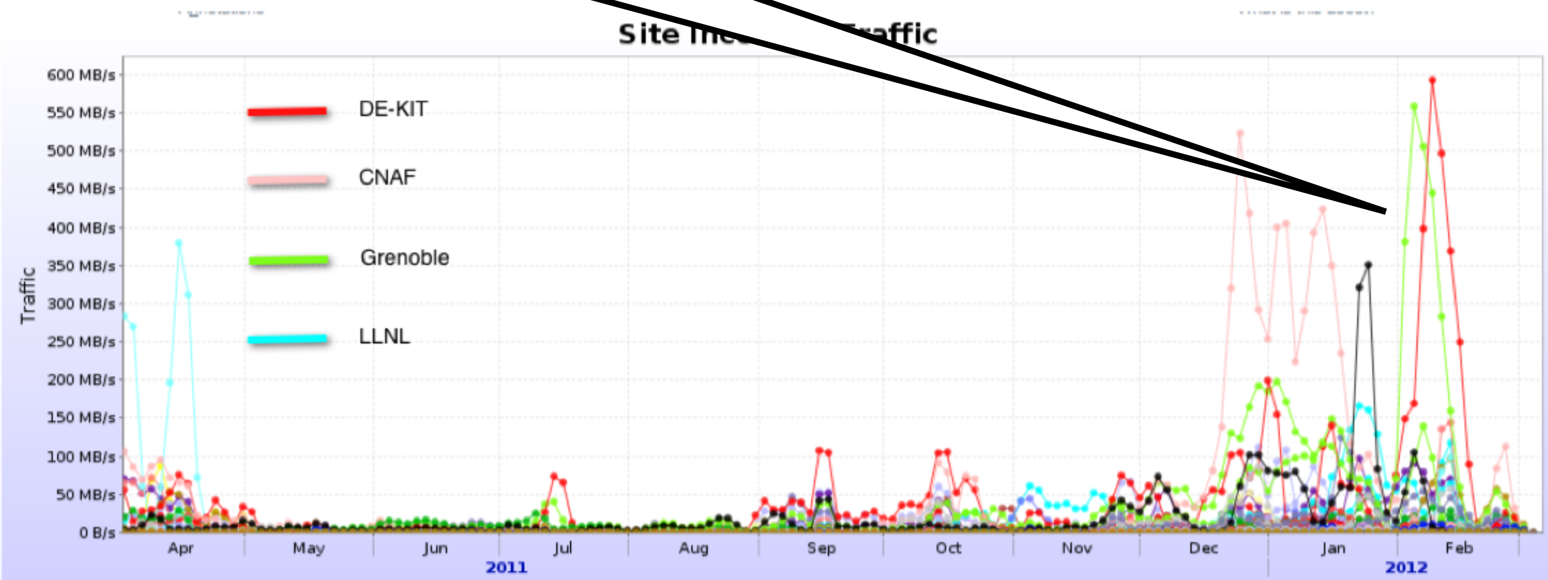
Runs

Run ID	AliRoot version	Testing status	Run status	Actions
8	VO_ALICE@AliRoot::v5-02-05-AN	Finished (1:47 total time)		
7	VO_ALICE@AliRoot::v5-02-04-AN	Started 4d 23:17 ago		
6	VO_ALICE@AliRoot::v5-02-04-AN	Finished (3m 48s total time)		
5	VO_ALICE@AliRoot::v5-02-04-AN	Finished (0m 43s total time)		
4	VO_ALICE@AliRoot::v5-02-04-AN	Finished (1d 19:40 total time)		

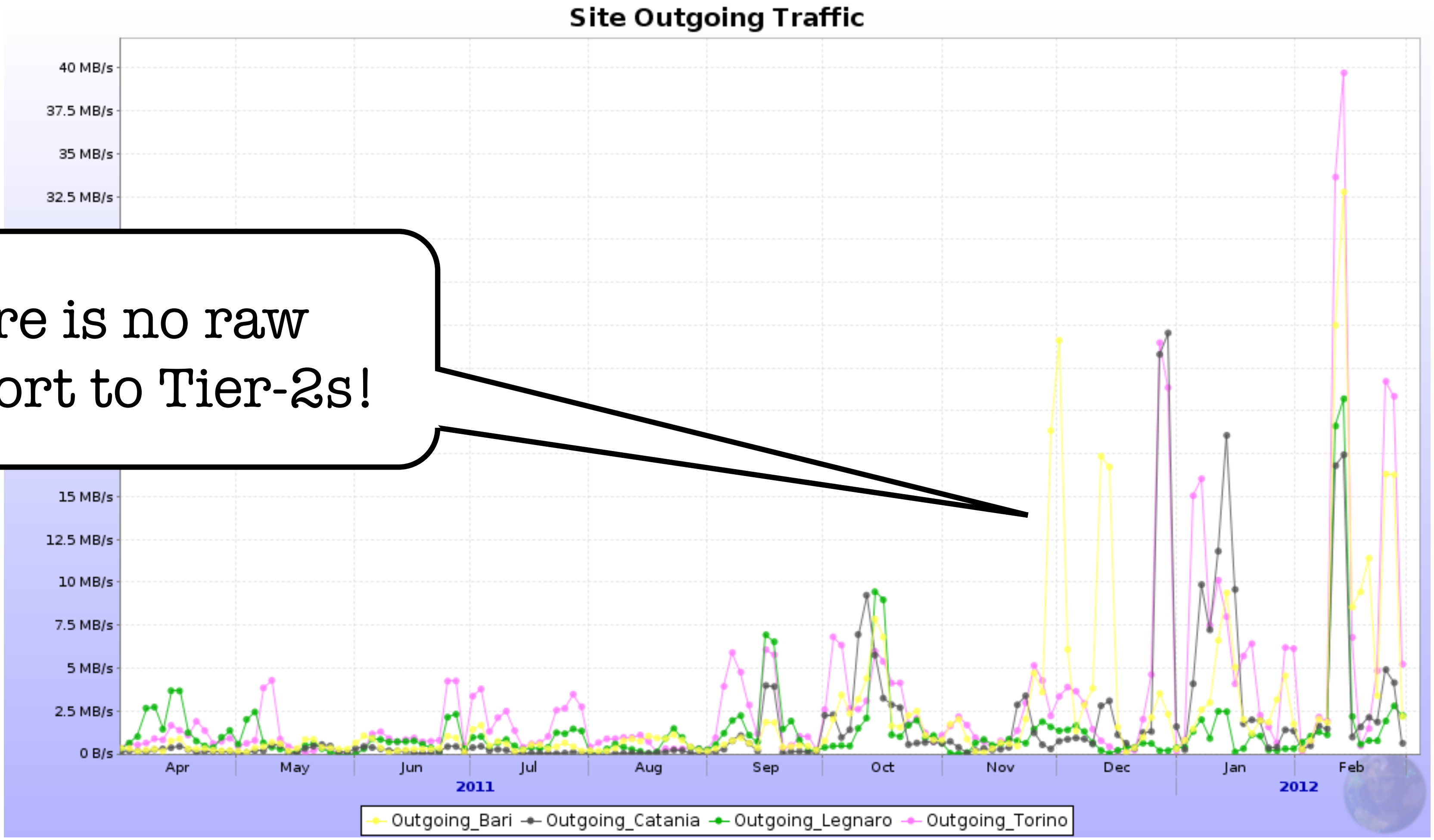
Network traffic: Tier-1



At Tier-1: mostly due to PbPb data export



Network Traffic: Tier-2



There is no raw export to Tier-2s!

The disk crisis

- Most Storage Elements nearly full
  - Aggressive purge campaign gained some time
  - But we were writing faster than we could delete
  - (And eventually we ran out of garbage)

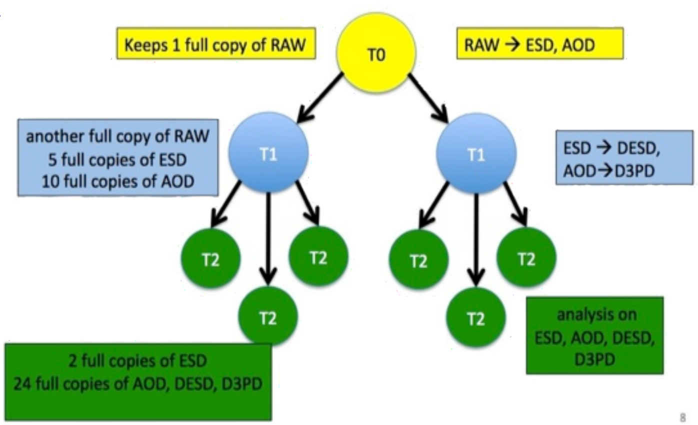
AliEn SE	Statistics						Xrootd info			
AliEn name	Size ▲	Used	Free	Usage	No. of files	Type	Size	Used	Free	Usage
ALICE::CERN::ALICEDISK	2.179 PB	1.745 PB	444.2 TB	80.09%	23,510,045	CASTOR	2.251 PB	2.034 PB	221.6 TB	90.39%
ALICE::FZK::SE	1.786 PB	1.017 PB	787.4 TB	56.95%	11,712,018	FILE	1.829 PB	1.33 PB	511.5 TB	72.69%
ALICE::CNAF::SE	989.7 TB	835.8 TB	153.9 TB	84.45%	8,211,844	FILE	989.7 TB	788.6 TB	201.1 TB	79.68%
ALICE::KISTI_GSDC::SE2	966.8 TB	30.64 TB	936.2 TB	3.169%	878,124	FILE	966.8 TB	50.95 TB	915.8 TB	5.27%
ALICE::NIHAM::FILE	895 TB	470 TB	425 TB	52.52%	12,489,240	FILE	894.9 TB	465.4 TB	429.5 TB	52.01%
ALICE::LLNL::SE	688 TB	166.9 TB	521.1 TB	24.26%	3,789,593	FILE	687.8 TB	339.8 TB	348 TB	49.4%
ALICE::LBL::SE	644.5 TB	92.21 TB	552.3 TB	14.31%	2,304,594	FILE	572.9 TB	86.46 TB	486.4 TB	15.09%
ALICE::CCIN2P3::SE	546 TB	519.7 TB	26.29 TB	95.18%	6,166,502	FILE	545.6 TB	476.5 TB	69.11 TB	87.33%
ALICE::Prague::SE	538.7 TB	285.7 TB	253 TB	53.03%	2,851,047	FILE	538.7 TB	292.4 TB	246.3 TB	54.28%
ALICE::CERN::EOS	500 TB	10.93 TB	489.1 TB	2.186%	569,501	FILE	505 TB	172.5 TB	332.5 TB	34.16%
ALICE::Torino::SE	469.5 TB	195.8 TB	273.7 TB	41.69%	3,059,567	FILE	469.5 TB	391.6 TB	77.89 TB	83.41%

ALICE Storage: some figures

- 52 disk SEs, 8 tape SEs (T0 and T1s)
  - 43× xrootd (some with underlying distributed FS), 2× DPM, 4× CASTOR, 3× dCache
- 20PB in 200M files (replicas included)
- Default 2 replicas for any file, usually 3, 4 for production jobs
- 2 copies of the raw data on MSS:
  - Full copy at CERN T0
  - One distributed copy at T1s (full runs)

Destitution of the Monarc

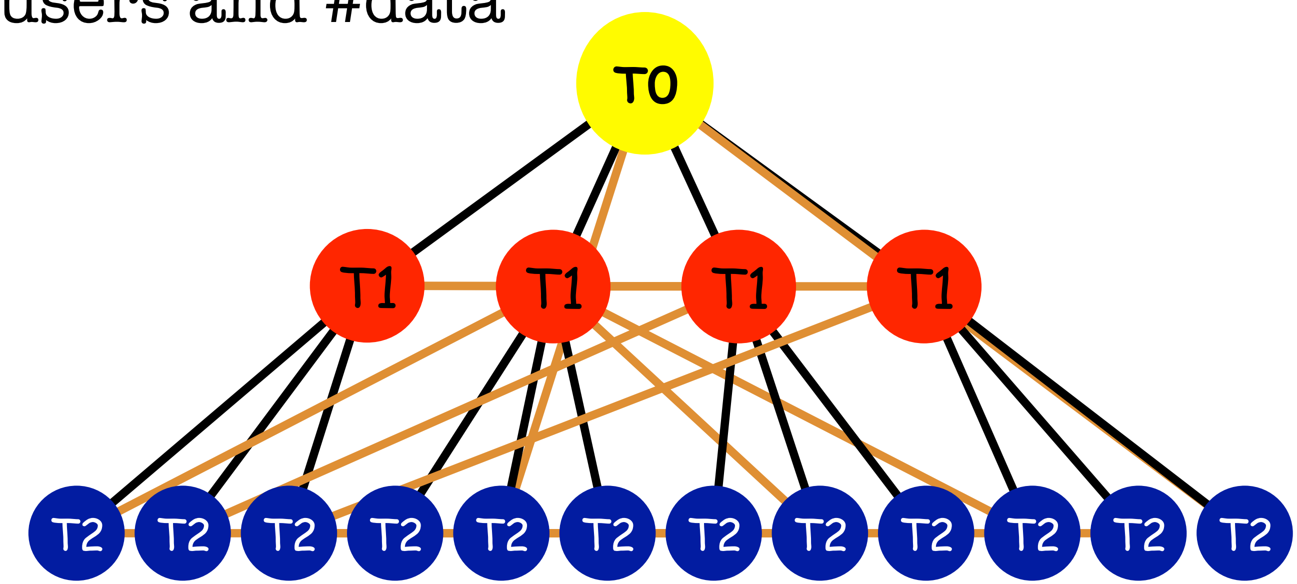
- The Monarc model was based on a “rigid” distribution of tasks between centres of different size and role
- The Grid is becoming less and less structured and tiered
- The difference between T1s and T2s is disappearing
  - Size doesn’t matter: US Tier-2s are LARGE
  - Custodial storage and better network
  - But the latter is about to change



Federico.Carminati@cern.ch

A "cloud" over the Grid

- T0-1-2(-3) hierarchy tends to be softened by user-driven data placement and transfer
- T1 and T2 are becoming equivalent in the network (LHCONE)
- The network is still the least undersubscribed resource we have
- No longer disk space but **network bandwidth** will scale with #users and #data



Federico.Carminati@cern.ch

Data is still the problem

- Data placement is the main problem, particularly for analysis
  - "Predictive" data placement (ATLAS & CMS) or "opportunistic" or "adaptive" (ALICE – need single catalogue)
  - Data distribution "per se" works very well
  - With "infinite" disk space the two are equivalent
  - Increasing the disk more difficult than increasing CPU
  - Quotas & monitoring more difficult for data than CPU

Federico.Carminati@cern.ch

Data access methods in ALICE

- Central catalogue of logical file names
  - With owner:group and unix-style permissions
  - Size, MD5 of files
  - Metadata on subtrees

- Each LFN is associated a GUID

- Any number of physical file names can be associated to an LFN
  - Like root://<redirector>//<HH>/<hhhhh>/<GUID>
    - HH and hhhhh are hashes of the GUID

Data access methods in ALICE

- Exclusive use of xrootd protocol
- Jobs are (usually) only downloading configuration files
- Data files are accessed remotely
  - The closest replica to the job, local replica first
- At the end of the job N replicas are uploaded from the job itself (3× ESDs, 4× AODs, etc...)
- Scheduled transfers only for raw data exports

SE Monitoring

Host Parameters

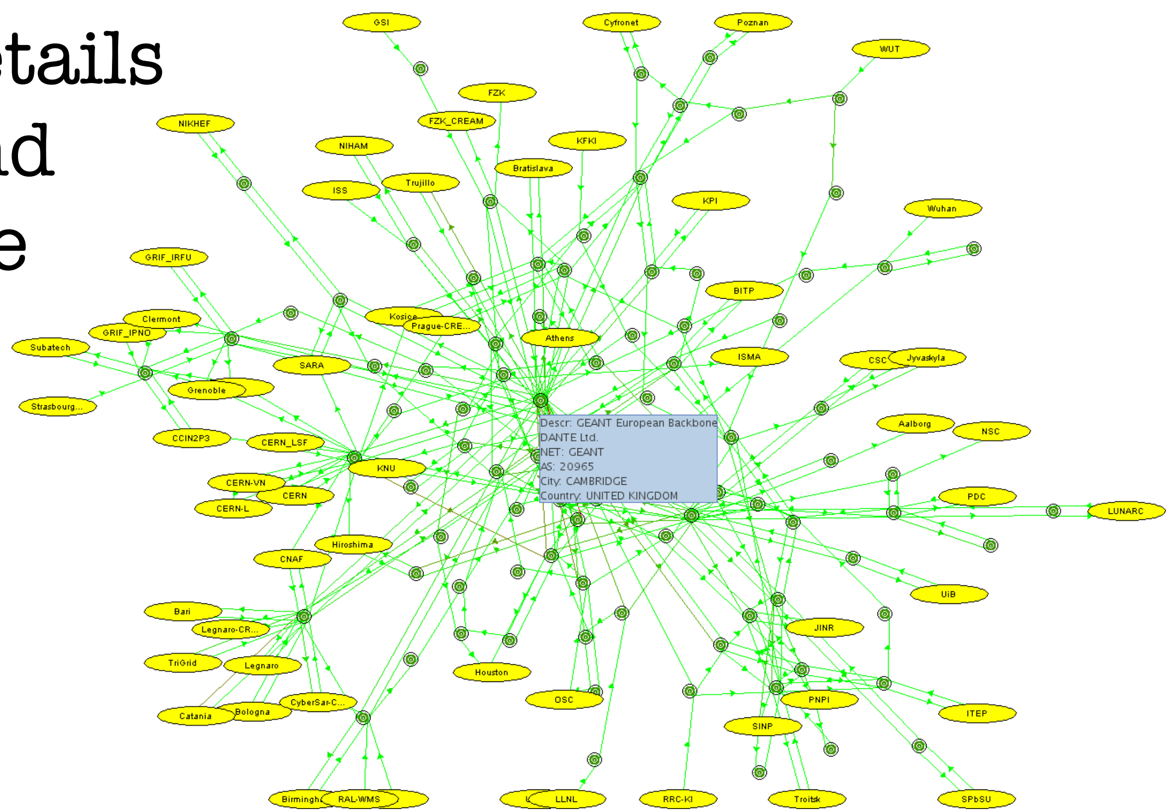
- Integrated in the overall monitoring of ALICE
  - xrootd plugin package includes an ApMon-based host and xrootd monitoring daemon
- Collected by the central repository and aggregated per cluster

Functional tests

- Add/get/delete performed every 2h
  - From a central location
  - Using the full AliEn suite (like any user or job)
- Results archived for a “reliability” metric
  - Last week × 25% + last day × 75%

Network topology discovery

- Site MonALISA instances perform between every pair of them
  - Traceroute / tracepath
  - Bandwidth estimation
- Recording all details we get a good and complete picture of the network topology



A dynamic "distance" metric from an IP address to a SE

- Starting from the network topology
  - Same site, same AS, same country, continent…
- Last functional test results excludes non-working SEs
- Altered by
  - Reliability
  - Remaining free space
  - A small random factor to assure "democratic" data distribution

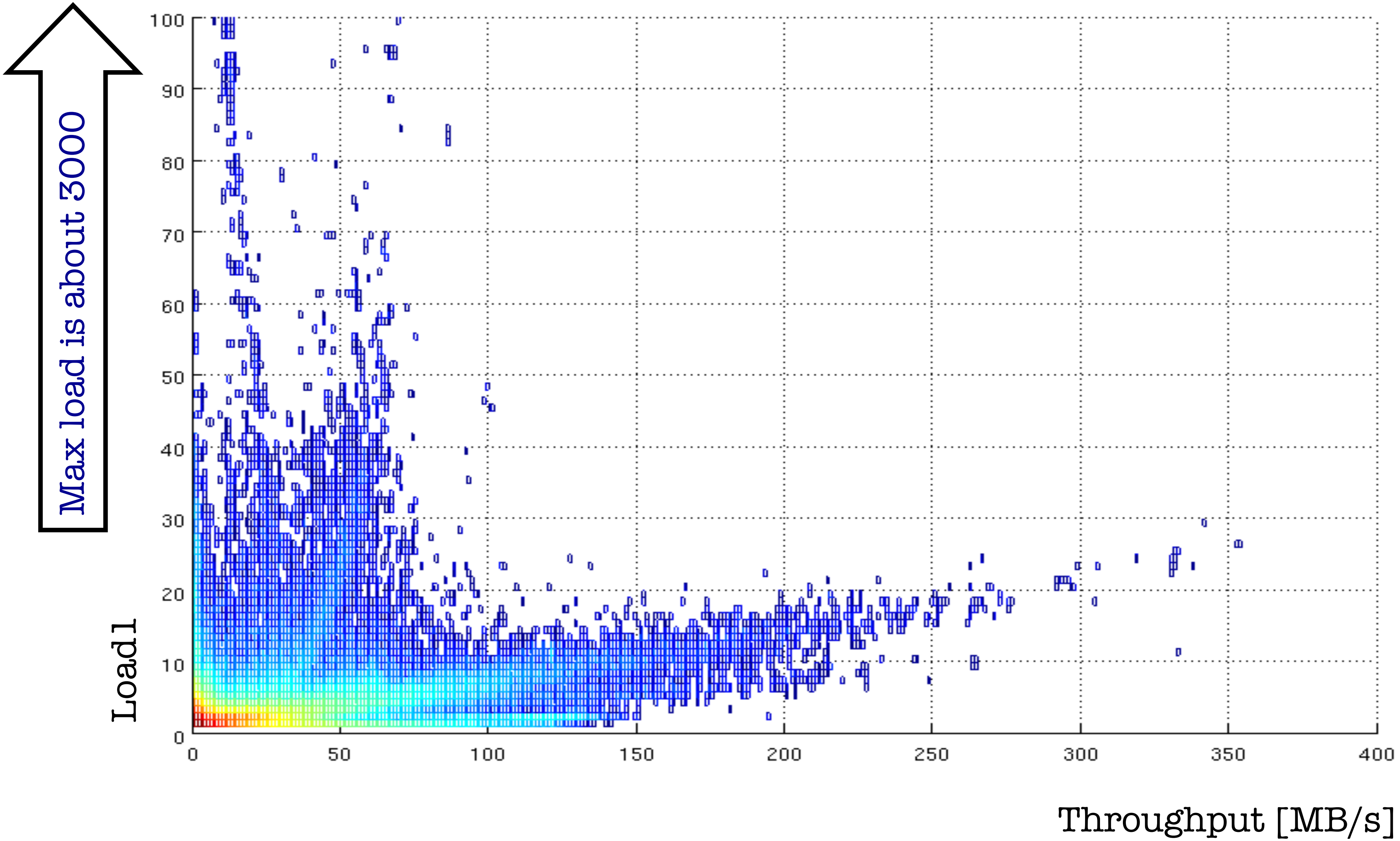
SE discovery

- Reading from the closest working replica
  - Simply sorting by the metric, including the non-working SEs, as last resort

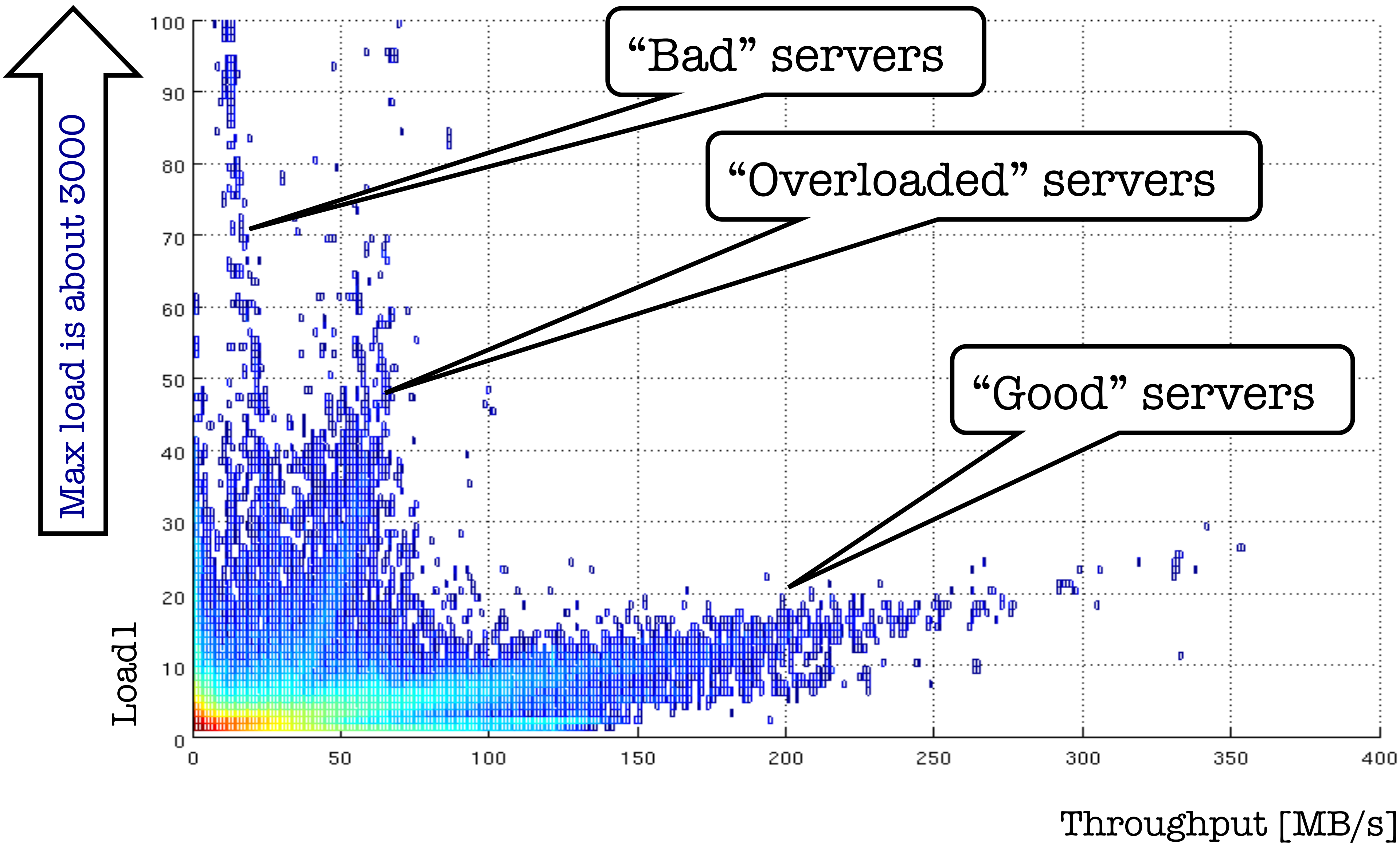
- Writing to the closest working SEs
  - Each SE is associated a tag (“disk”, “tape”,…)
  - Users indicate the number of replicas of each type
    - Default is “disk=2”
  - Not excluding (but not encouraging) the option of specific target SEs
  - Keep asking until the requirements are met or no more SEs left to try

CPU efficiency vs load

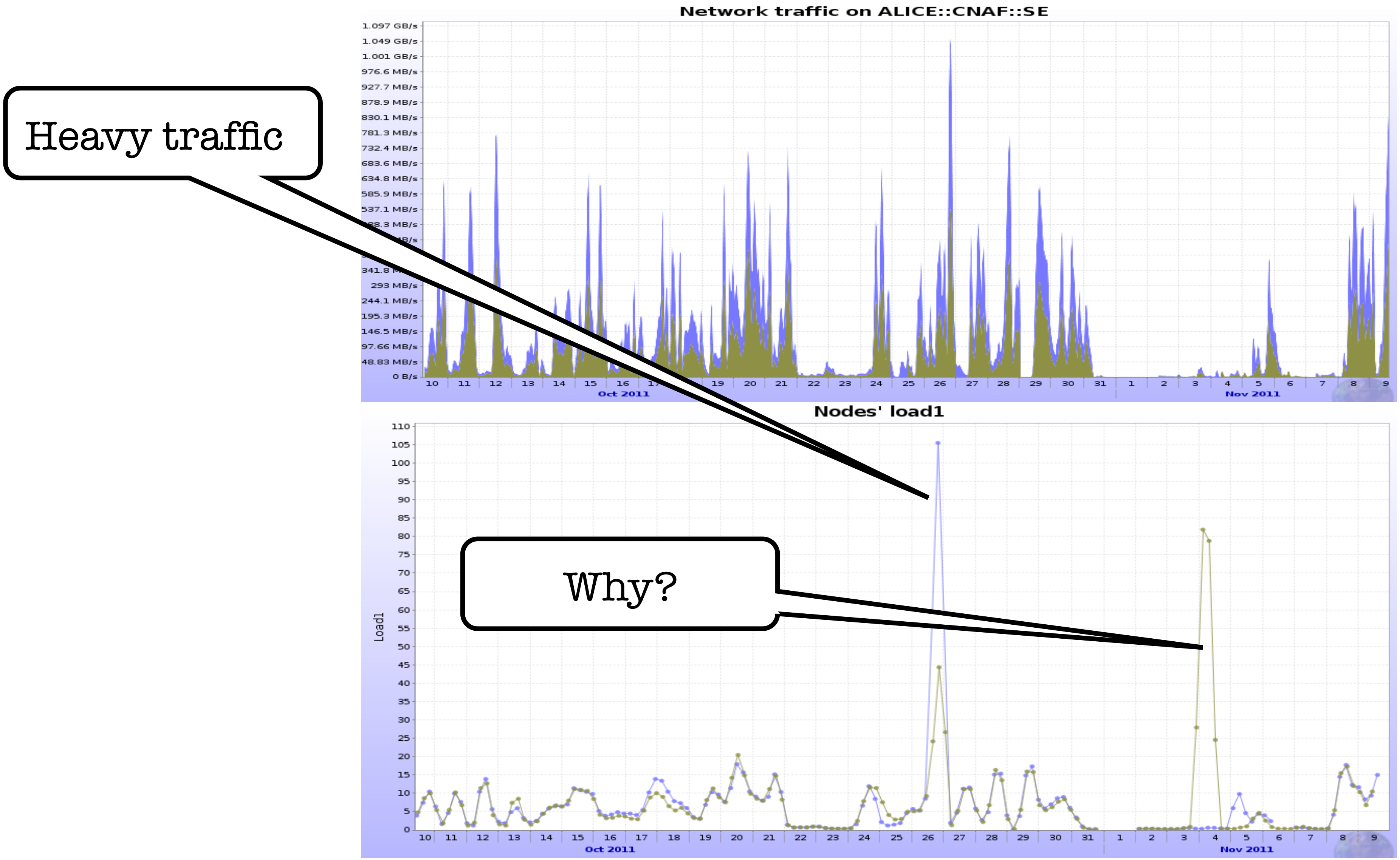
Throughput vs Load



Throughput vs Load

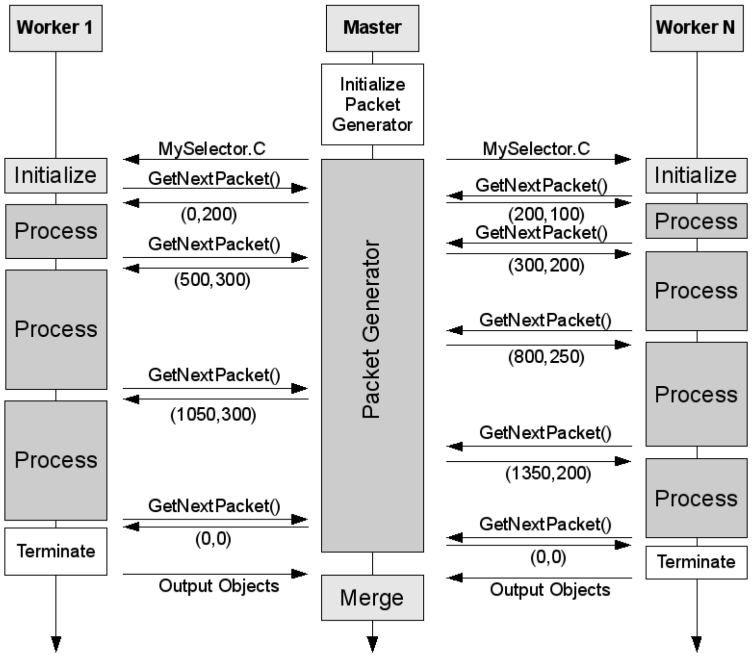


Even good servers have problems



proof

- **Parallel ROOT Facility**: ROOT framework to build a parallel and interactive analysis facility
- **Event-based parallelism**: process single physics events in parallel, merge final results
- **Interactive**: no queue, user has the resources immediately



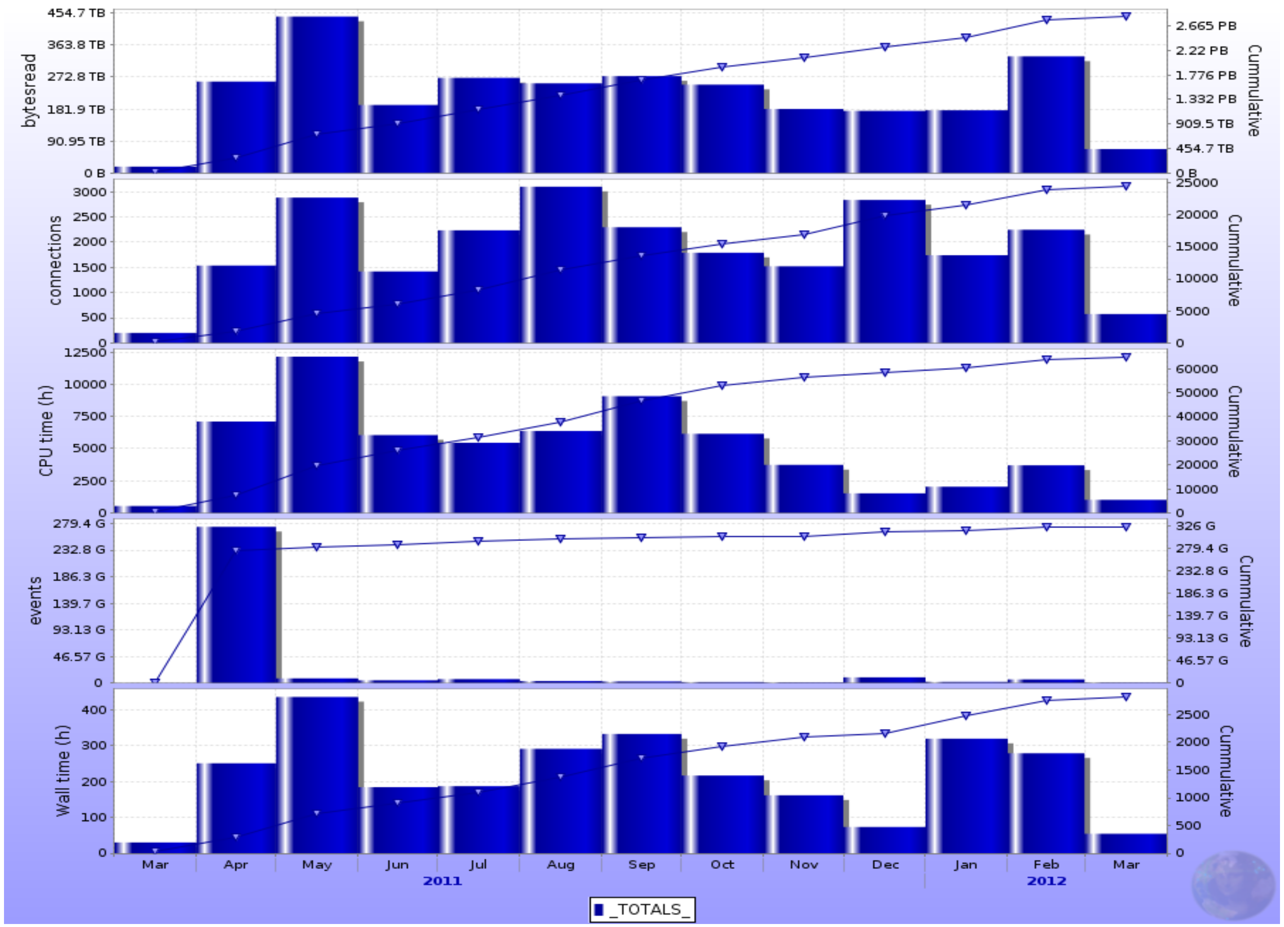
Resources used uncontinuously and in small bursts during certain periods of the day (i.e., working hours)

Dedicated resources often underused

How to optimize resources exploitation and absorb peak loads:
good use case for virtualization and the cloud

AAFs are popular facilities

«**CAF** averages 15-20 users per week, usually during the normal working hours, and processes about 100 TB of data, equivalent to 600 – 800 millions events. During the last year the total amount of data that has been processed on CAF is 3.109 PB for 0.28×10^{12} events»



AAF Monitoring

ALICE PROOF Clusters – ALICE Grid Monitoring with MonALISA

alimonitor.cern.ch/stats?filter_0=&filter_0_9=&filter_0_8=&filter_0_7=&filter_0_6=&filter_0_5=&filter_0_4=&filter_0_3=&

MonALISA Repository for ALICE

MonALISA – MONitoring Agents using a Large Integrated Services Architecture

My jobs | My home dir | Catalogue browser | Repository Home | Administration Section | ALICE Reports | Events XML Feed | Firefox Toolbar | MonaLisa GUI

ALICE Repository

- ALICE Repository
  - Google Map
  - Shifter's dashboard
  - Run Condition Table
  - Production Overview
  - Production info
  - Job Information
  - SE Information
    - Status
    - Files
    - xrootd
    - CERN Castor2x
    - AFs
      - Overview
      - Details
        - Traffic
        - Load
        - Sockets
  - Services
  - Network Traffic
  - FTD Transfers
  - CAF Monitoring
  - SHUTTLE
  - Build system
  - HepSpec
  - Dynamic charts

ALICE PROOF Clusters

What is this about?

Cluster list

	Cluster					ROOT	Aggregated disk space			AF xrootd		xrootd
Name	**Online**	**Status**	**Proof master**	**Workers**	**Users**	**Version**	**Total**	**Free**	**Used**	**Running**	**Latest**	**Version**
1. CAF		Stable	alice-caf.cern.ch	114	1	v5-33-02b	159.7 TB	7.772 TB	151.9 TB	1.0.50	1.0.50	20100510-1509_dbg
2. CAF_TEST				-	-		-	-	-			
3. JRAF		Maintenance sin...	jraf.jinr.ru	8	0	v5-33-02b	2.014 TB	1.91 TB	106.4 GB	1.0.50	1.0.50	20100510-1509_dbg
4. KIAF		Stable	kiaf.sdfarm.kr	96	0	v5-30-06-1	171.9 TB	108.8 TB	63.1 TB	1.0.50	1.0.50	20100510-1509_dbg
5. LAF				-	-		9.41 TB	7.624 TB	1.786 TB			20100510-1509_dbg
6. SAF		Maintenance sin...	nansafmaster.in2p3.fr	48	1	v5-30-06-1	12.07 TB	3.48 TB	8.592 TB	1.0.50	1.0.50	20100510-1509_dbg
7. SKAF		Stable	skaf.saske.sk	60	1	v5-33-02b	53.72 TB	1.433 TB	52.29 TB	1.0.50	1.0.50	20100510-1509_dbg
8. SKAF_TEST				-	-		-	-	-			
9. TAF		Warming up	pmaster.to.infn.it	102	0	v5-33-02b	49.1 TB	26.79 TB	22.31 TB			
Total				**428**	**3**		**457.9 TB**	**157.8 TB**	**300.1 TB**			

PoD Vs. Cloud

- Proof-on-demand
- Plugins for most LRMS
- ...including gLite
- A test deployment on WNODeS is planned
- But works well only in large sites


Consolidating and extending Tier-2 resources to serve a larger user community

S. Bagnasco[1], D. Berzano[1,2], R. Brunetti[1], S. Lusso[1]

[1] Istituto Nazionale di Fisica Nucleare - Sez. di Torino / [2] Università di Torino

By using mainstream and open source tools such as OpenNebula, GlusterFS and OpenWRT it is possible to rearrange and consolidate the resources of a computing facility counting a few hundreds of physical nodes, the average size of a EGI/WLCG Tier-2. The underlying infrastructure, originally designed around the main "customer", can now be considered as the core of a more agnostic service capable of meeting the demands of smaller groups and even spot users, aiming to gradually eliminate proprietary farms. The experience made at the INFN Computing Center in Torino leverages the IaaS approach to integrate very different use cases: a conventional Tier-2 EGI/WLGC ALICE farm, an interactive PROOF-based ALICE analysis facility, a small medical farm for CT scanning for the INFN M5L project as well as several single instances for diverse spot users.

IaaS Cloud with OpenNebula

- Mainstream tool for Clouds → **Infrastructure as a Service**
- Handles several hypervisors → **KVM** used
- **Modular** → components based on **Ruby** and **Bash** scripts
- **Network isolation** → ebtables, 802.1Q, Open vSwitch
- Virtual Machine **contextualization**
- Cold and live **migration** between hypervisors
- **Sunstone** Web Interface
- **Multiuser** with **auth** (also x509) and **quotas**
- VM display from browser → **noVNC**

OpenNebula.org

Network isolation with OpenWRT

- Each Virtual Farm has its own **Class-C private network**
- MAC addr **isolation** → OpenNebula supports **ebtables**
- **Virtual Router** with **public IP address** for each Virtual Farm

Virtual Router based on OpenWRT

- Linux distribution conceived for embedded devices → **low resources** needed (1 CPU, < 200 MB RAM)
- **NAT**, **DNS** forwarder, **DHCP** server
- **Firewall** and **port forwarding** → **iptables**
- **Web** control and monitoring

OpenWrt

Classes of Hypervisors

*Existing **IGI/WLCG worker nodes** in Torino's Tier-2 were progressively **converted to hypervisors**, while every newly acquired node is immediately configured likewise. The coexistence of different generations of machines suggested their **consolidation into two classes***

Service Hypervisors
aimed to provide high-availability for critical services

- Ideal for critical services and **head nodes** of **Virtual Farms**
- Virtual machines run from a **GlusterFS shared volume**
- Both a **private** and **public** network interface for VMs
- Robust **Live Migration** protects against hardware failures

GLUSTER

Worker Hypervisors
aimed to provide maximum throughput for data crunching

- Ideal for **farm workers** (*i.e., WLCG WN, PROOF node*)
- Virtual machines are run from **LVM partitions** → reduces typical virtual I/O performance loss
- Only **private** network interface available for VMs
- Live Migration is not feasible → run non-critical VMs

Our Customers

IGI/WLCG Worker Nodes

- No differences from physical nodes as seen from the outside
- New virtual nodes are **automatically added** to the local batch system
- One virtual image **cloned** several times → **easy software updates**

WLCG

PROOF (ROOT-based) Analysis Facility for ALICE

- Part of the **ALICE Analysis Facilities** federation
- Number of nodes can be varied **dynamically**
- Difficult to **estimate** the amount of needed resources → they can be assigned to PROOF only when needed and **never get wasted**

ALICE

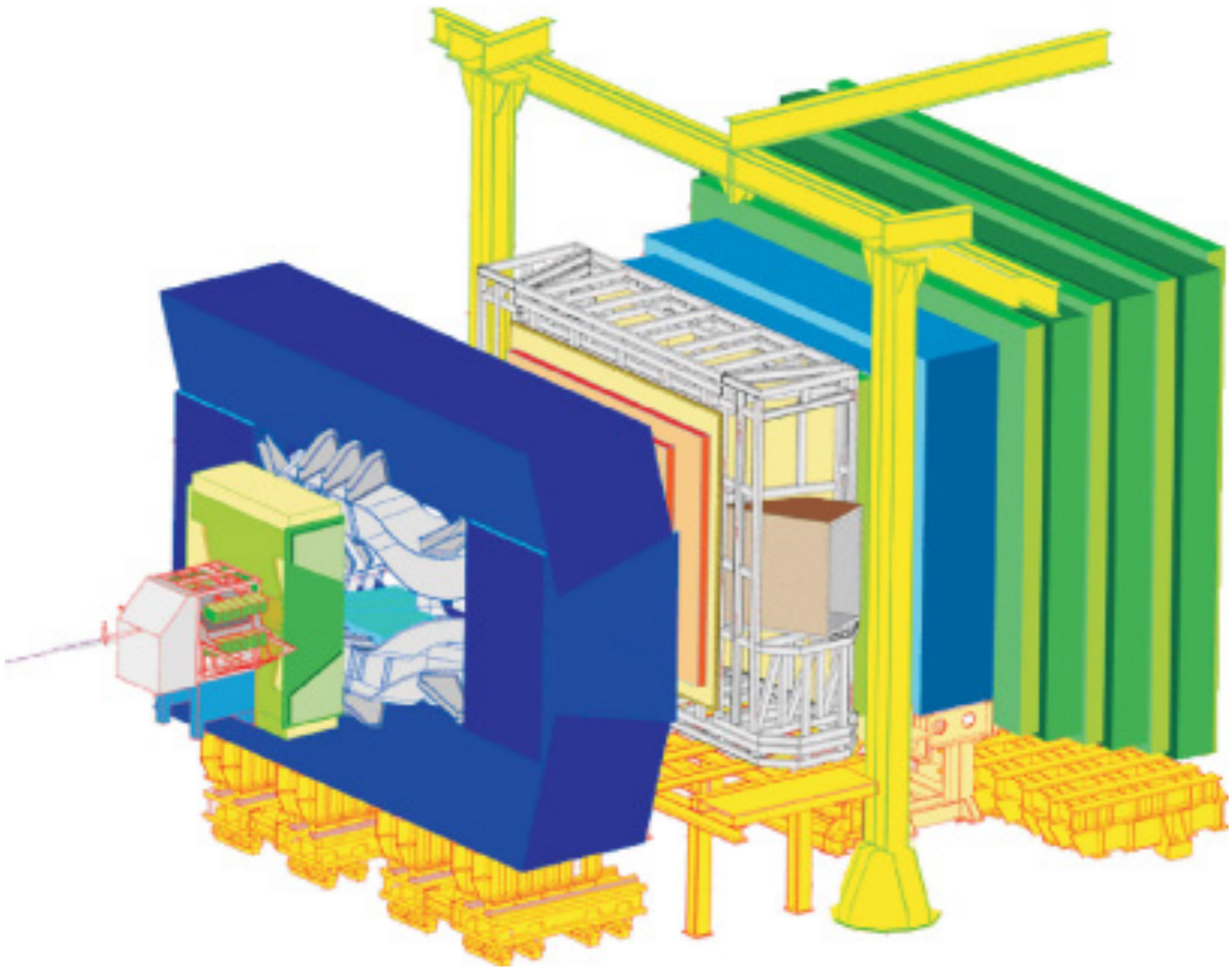
M5L-CAD medical imaging (INFN and diXit spin-off)

- **SaaS** → medical doctor uploads CTs via a Web interface (WIDEN) and is notified of the results after a while
- Torino's cloud provides the IaaS under the hood
- Virtual computing nodes to process CTs are **added dynamically** to **meet the demand**

diXit

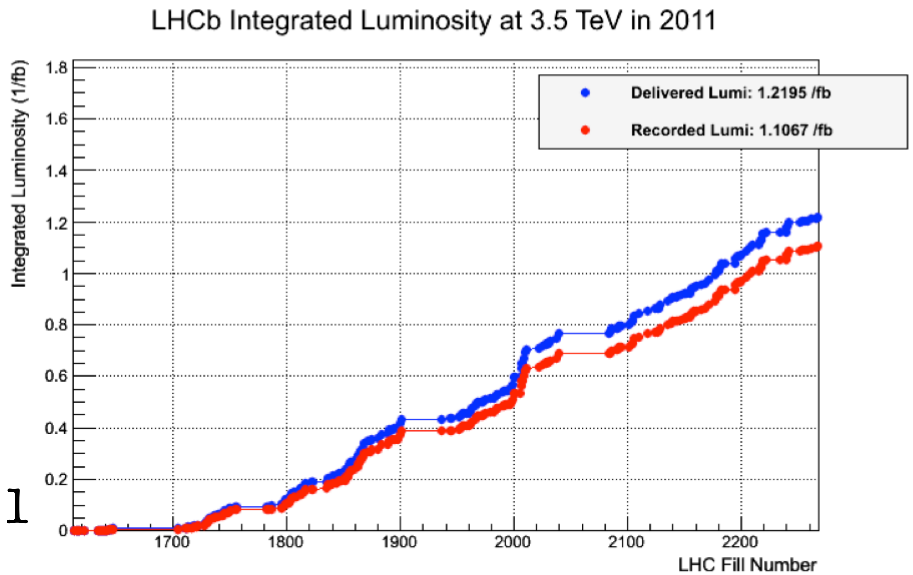
*Immediate creation of **isolated virtual farms** → reduces the hassle of configuring dedicated hardware for each **spot** or **large** use case*

LHCb



LHCb data taking 2011

- Stable data-taking conditions for LHCb
  - ~90% efficiency and increasing instantaneous luminosity



- Accumulated over 1 fb^{-1} of data for physics
  - instantaneous luminosity kept constant throughout the fills

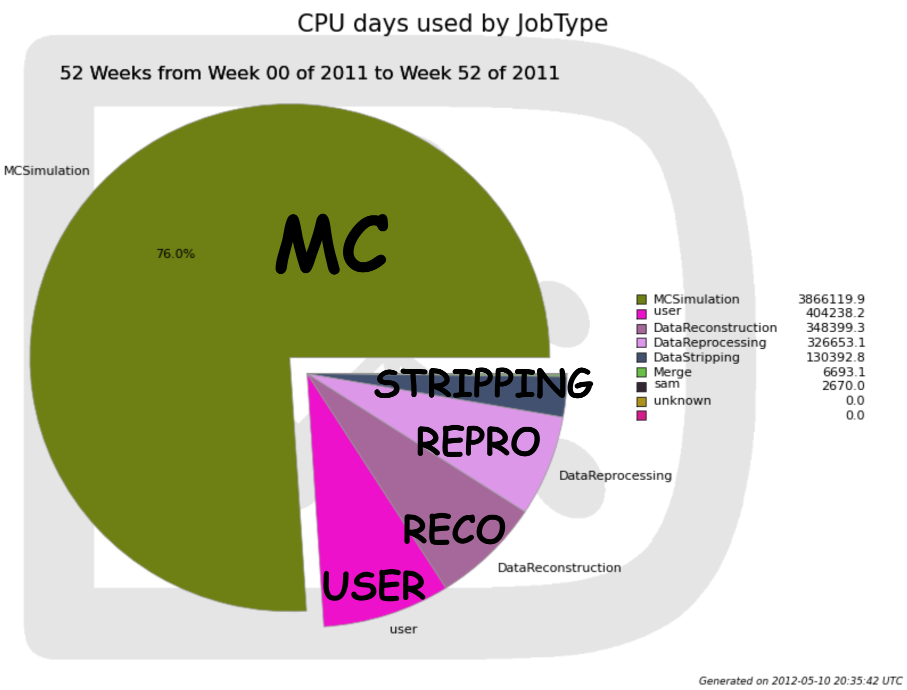
- Average visible collision multiplicity of 1.5
  - below the 2.5 value from 2010
  - average HLT trigger rate slightly below 3 kHz

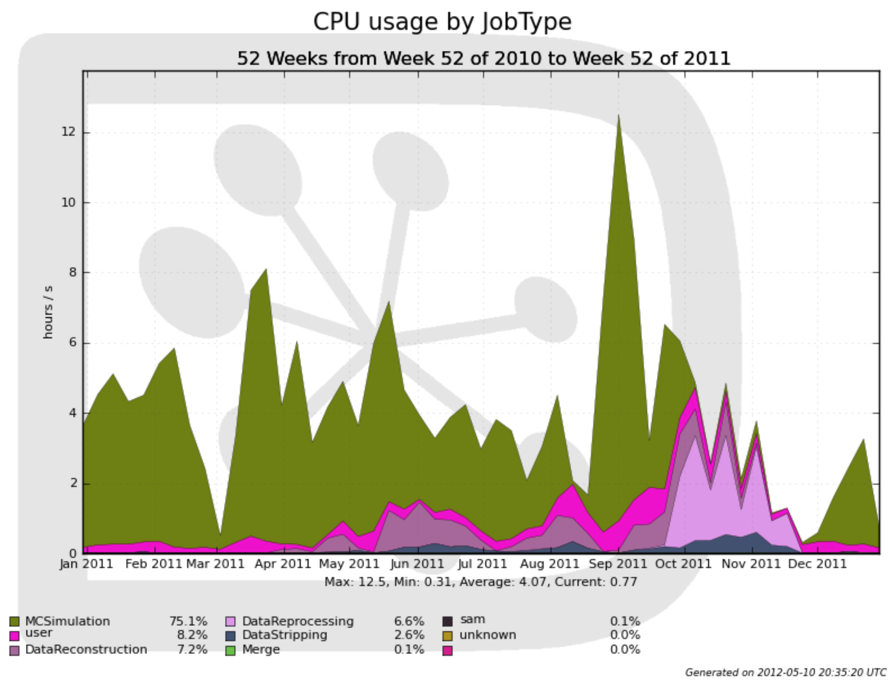
Summary of computing activities

- Simulation
  - Mainly used for identifying background and evaluating acceptances and efficiencies
  - Simulates an ideal detector, however with realistic geometry
  - Event generation and detector response tuned to real data
- Real data handling and processing
  - Distribution to Tier1s (RAW)
  - Reconstruction (SDST)
  - Stripping and streaming (DST)
  - Group-level production (μDST)
- User analysis
  - MC and real data processing
  - Detector and efficiency calibration
  - End-user analysis (usually off-Grid: Tier3 or desktop)

Use of computing resources in 2011

- Mainly MC simulation (76%)
  - Then: user analysis, reconstruction, reprocessing and stripping





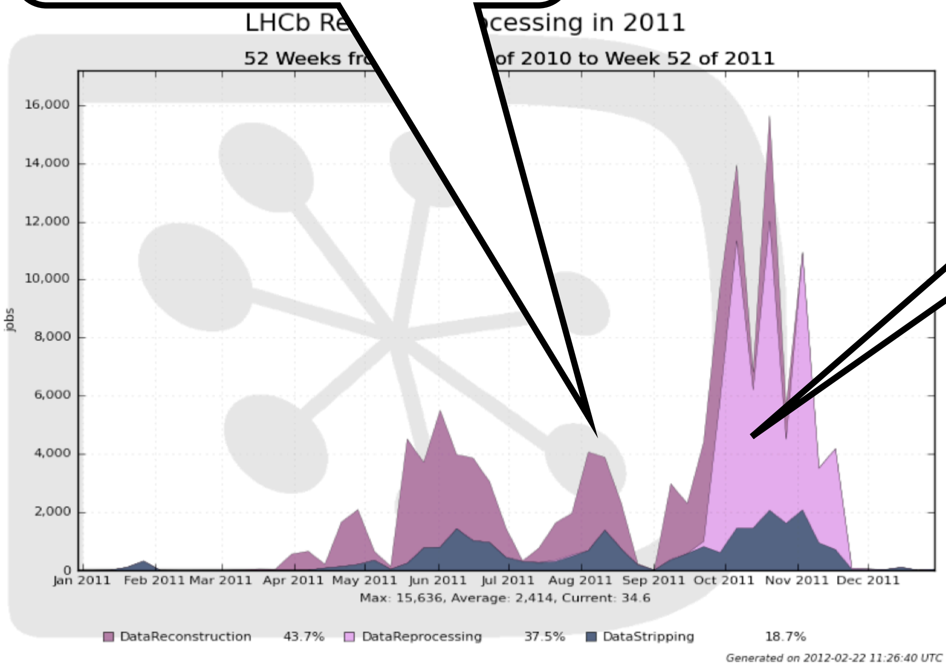
Reconstruction and reprocessing

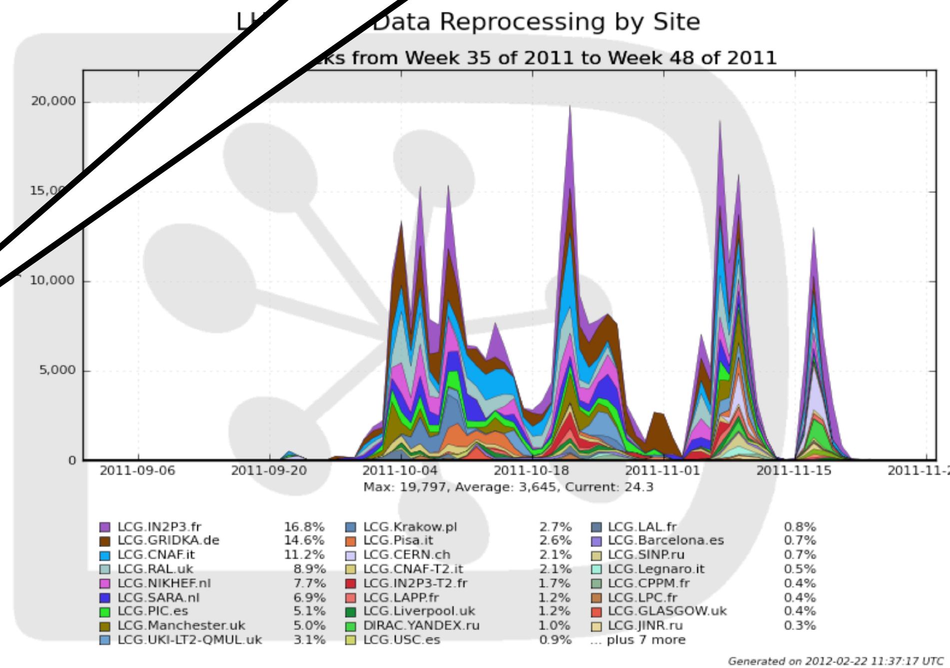
- RAW data is uploaded to CERN and distributed among the Tier1s
  - It is then processed quasi-online (Reconstruction), and data samples for physics are preselected (Stripping)
- News in 2011
  - Tier2 CPU resources have been successfully added to reduce the execution time of the full reprocessing
  - Avoids need for an extra peak of CPU power on Tier0/Tier1s
- Selected Tier2 sites were associated to a "close" Tier1 site, and allowed to execute re-processing jobs
  - the RAW file is downloaded to the worker node at the start of the job from the Tier1 Storage Element, there it is processed for about 1 day, and, at the end, the resulting SDST file is uploaded to the Tier1

Recostruction and reprocessing

Normal reco and stripping throughout the year

Full reprocessing campaign



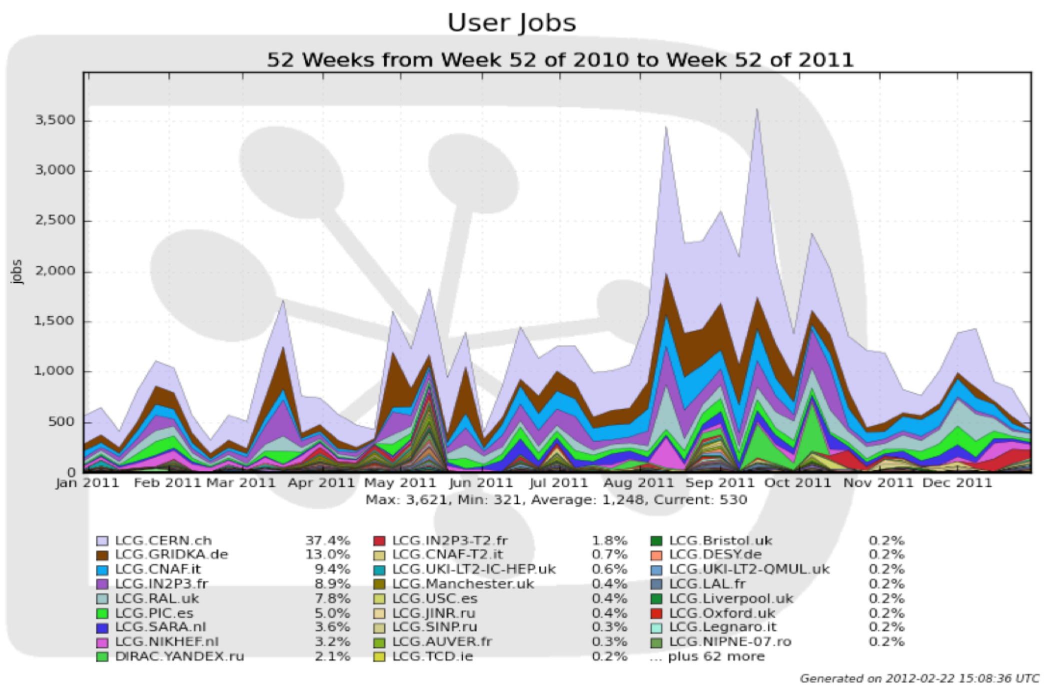


CNAF Tier-1 third site in 2011 in terms of CPU usage during full reprocessing campaign, but many sites involved

User analysis

- CPU activity of LHCb physics users has been about 60% of the sum of all real data processing activities
  - this fraction is steadily increasing

- As expected most of the activity concentrates on the Tier0/1s, with some small contributions from other sites

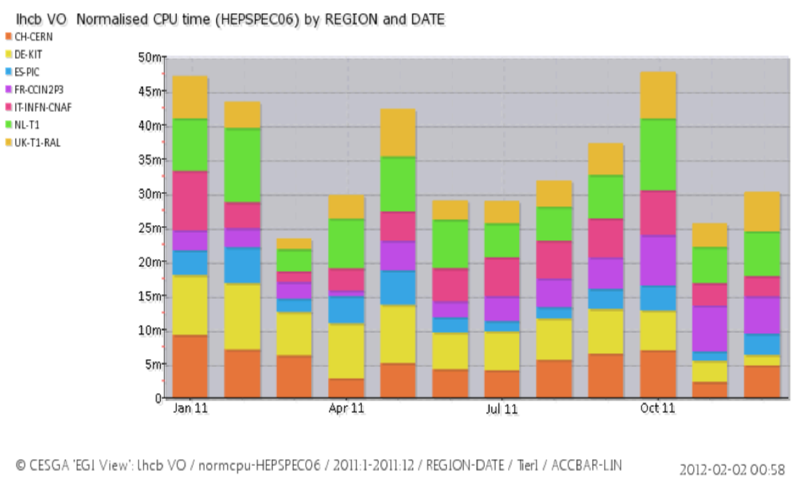


- CNAF Tier-1 third site of the experiment during 2011, after CERN and GRIDKA

Summary of CPU usage

Discrepancies in some sites between DIRAC and WLCG → DIRAC more reliable "by construction"

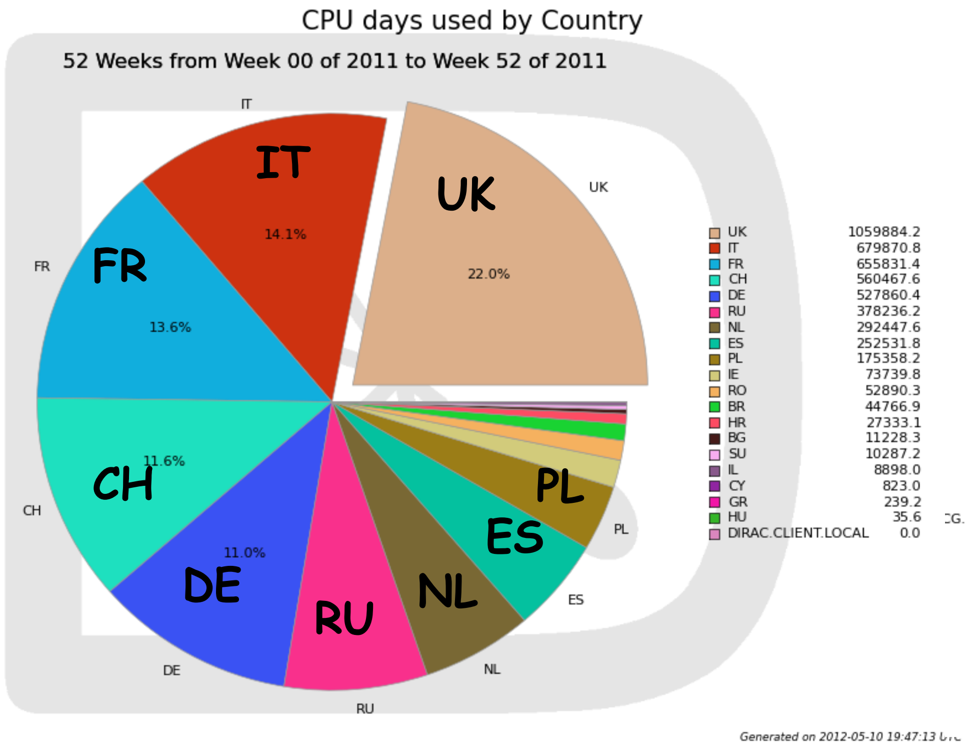
kHS06·y	WLCG	DIRAC	Pledge
CERN	7.2	10.5	21
CNAF	6.4	6.5	8
GRIDKA	8.6	9.5	11
IN2P3	5.4	5.2	18.7
NL-T1	9.5	6.7	15
PIC	4.1	3.9	4.5
RAL	5.9	5.2	12
Total	**47.2**	**47.5**	**90.2**

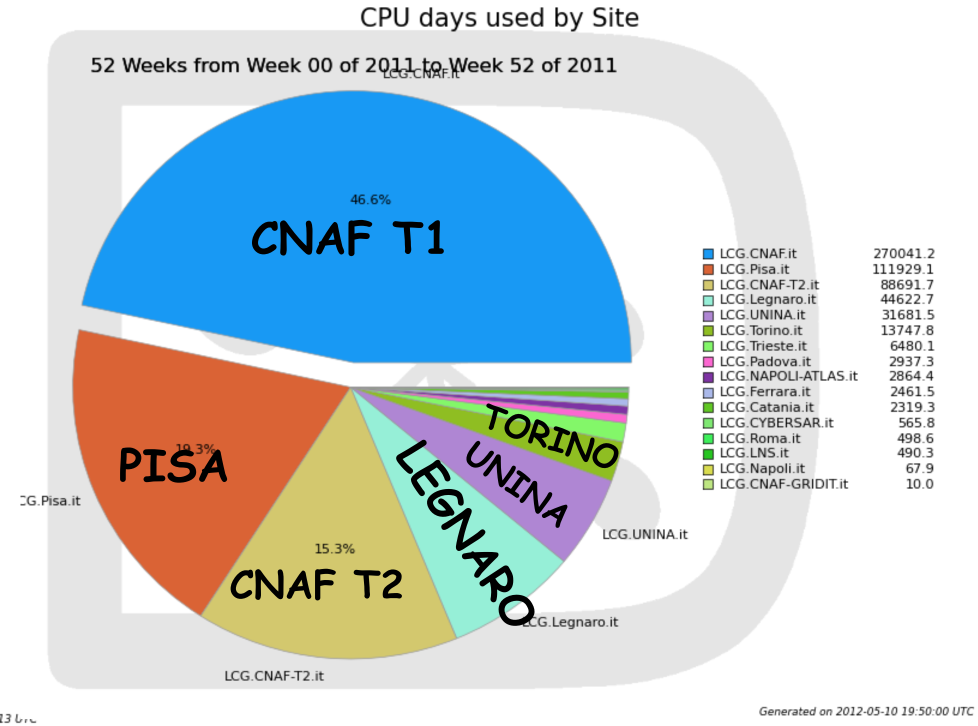


kHS06·y	WLCG	Pledge
FR-T2	7.0	13.2
DE-T2	1.8	3.2
IT-T2	6.4	6
PL-T2	4.3	2.2
RO-T2	0.5	1.6
RU-T2	4.1	2.7
ES-T2	3.4	2.3
CH-T2	1.3	2.7
UK-T2	18.2	6.6
Total	**47.0**	**40.5**
Others	26.2	

CNAF Tier1 amongst the sites with smaller pledges, but larger CPU usage than IN2P3 and RAL!

CPU use by country



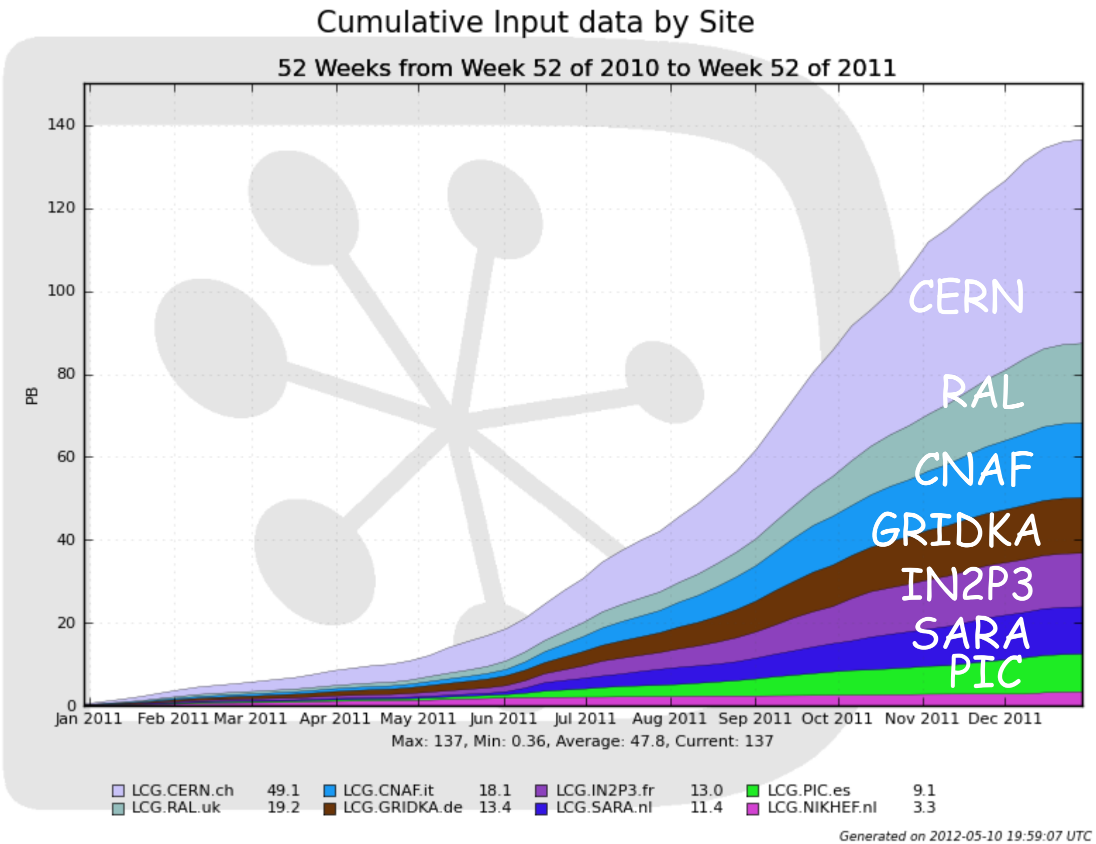


Italy second country in terms of CPU resources made available to LHCb → 14.1% close to fraction of INFN physicists in the experiment

Almost one half of CPU usage from CNAF Tier-1, then PISA (!!!) with 20%, CNAF Tier-2 with 15% and important contributions from Leganro, Napoli, Torino

Data read at T1/TO and disk use

- LHCb has processed almost 140 PB of data at all levels in 2011
  - Mostly by user analysis jobs
  - 18.1 PB at CNAF



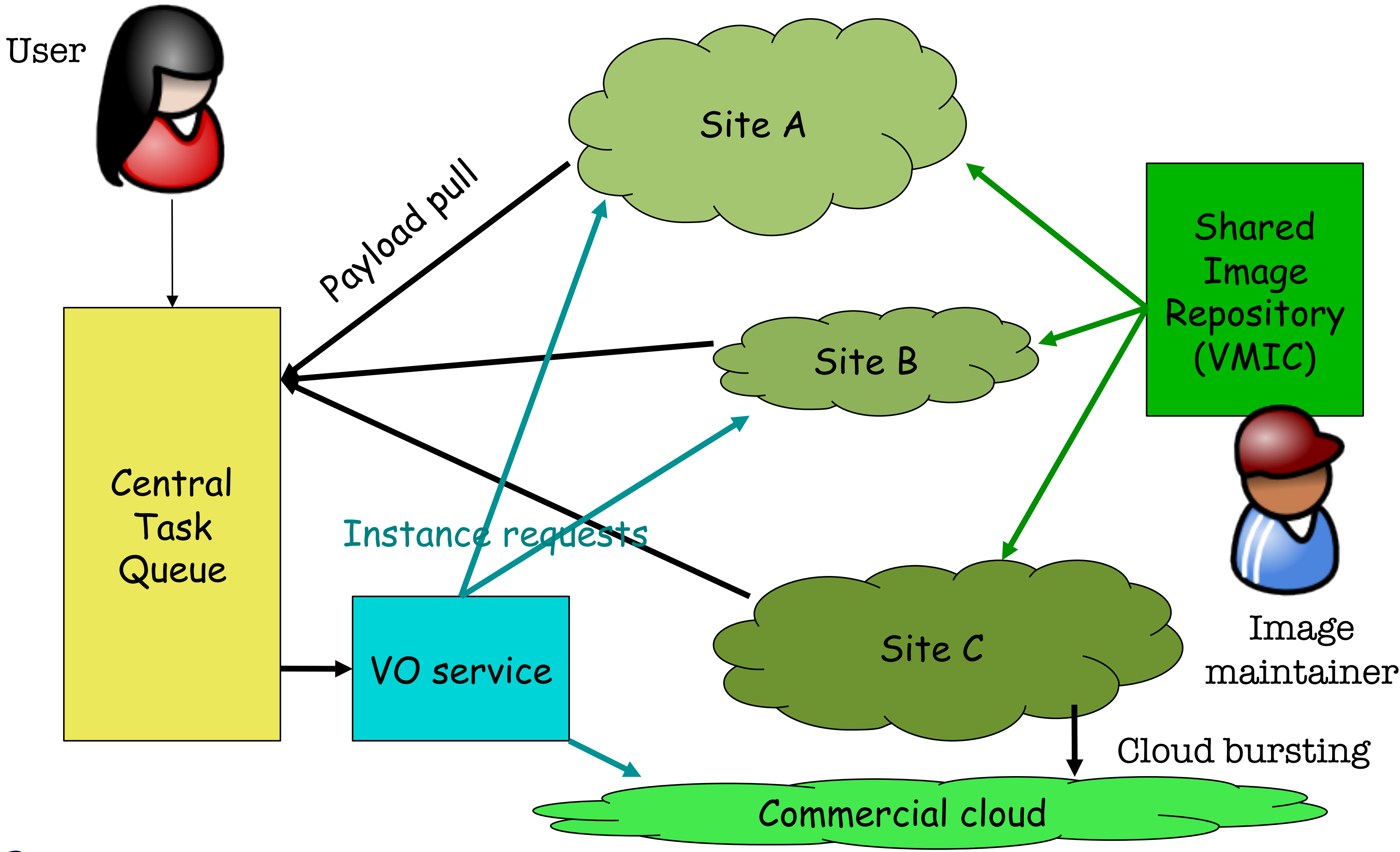
760 TB at 95% occupancy presently at CNAF → large increase with 2012 pledges awaited

Virtualization of LHCb computing

- Objective:
  - Running LHCb jobs that are currently running on Grid batch systems inside LHCb-specific VMs
  - Using cloud infrastructures provided by sites in parallel or in replacement of their current Grid batch systems

- Baseline VM
  - CERNVM is already much used within LHCb (on various platforms)
  - CVMFS is now the baseline software distribution mechanism for LHCb applications, including LHCbDirac

What We Want


User
Site A
Payload pull
Site B
Shared Image Repository (VMIC)
Central Task Queue
Instance requests
Site C
Image maintainer
VO service
Cloud bursting
Commercial cloud

Strawman example

- Start a CERNVM machine on the cloud that:
  - Initializes itself as an LHCbDirac VM
    - Set up the LHCbDirac environment
    - Set up a local configuration
  - Runs a Dirac job agent
    - Matches a job or jobs (depending on the configuration)
    - Execute these jobs (including uploading output data)
    - Loops for matching another job
- Requirements
  - LHCbDirac installed: it is on CVMFS
  - Applications installed: they are on CVMFS
  - Permission to match LHCb jobs

LHCb-DIRAC is working on it

- DIRAC has proven to easily integrate grids and clouds
  - "Belle-DIRAC Setup for Using Amazon Elastic Compute Cloud", J Grid Computing (2011) 9:65–79, R.Graciani et al.
  - "Integration of cloud, grid and local cluster resources with DIRAC", J. Phys.: Conf. Ser. (2011) 331 062009, T.Fifield et el.
- LHCb is currently testing with the extra advantage of CERNVM and CVMFS:
  - "The Integration of CloudStack and OpenNebula with DIRAC" (see contribution at CHEP 12)

What LHCb-DIRAC provides today

- Controlled Submission of VMs
  - Tested EC2 interface,
  - Working on OCCI (missing python client)
  - Dedicated Director per Cloud Manager since it requires access to private info
- Detailed Monitoring of VM usage
  - Already integrated in DIRAC portal
- Close link with DIRAC WMS
  - Similar approach as pilot submission, based on pending load on TQs
- Optimization possibilities via parallel upload of outputs
  - Via local Requests
- Remote Control of Instances
  - Halt (now), Stop (halt after job completion), Pause (stop matching), Resume (restart matching),…

Thanks!