Group Space Management
&
Group Job Monitoring

Ueda I.
ATLAS S&C workshop
2011.04.05.

Group Space

Global quota vs Specific sites

Questions asked:

- A) which groups would like to utilize the global quota and automatic choice of destination
- B) which groups would like to keep spaces at specific sites, and if so, which sites
- C) which groups would like the mixture of the two

Tentative Results:

- A: 4, B: 2, C: 7

The DDM-dev has been consulted to implement C

- to satisfy all the groups (A + B + C)
- it is a big change in the system, and will take several months for implementation

Group space quota in 2011

We have asked the CREM to confirm the quota in 2011

- hopefully will be increased ~ next week

While waiting for the implementaion of the “global quota”

- Should we move the group spaces from T2s to T1s for some groups who goes A and C?
  - specified T2 spaces will be kept
  - the total volume per group will be within the 2011 quota

Panda Dynamic Data Placement (PD2P)

- There is a plan to activate PD2P for group data
  - only for ‘Group Production’ data? or anything in the Group Spaces?
- the replicas will be created in T2_DATADISK
  - will not eat the group space quota

Group Job Monitoring

Definition of “groups” — Task Request

Task (data10_7TeV.00167607.physics_Egamma.merge.r1774_p327_p333_p491) Request Parameters

Input Attributes

Project	data10_7TeV
Input dataset	data10_7TeV.00167607.physics_Egamma.merge.AOD.r1774_p327_p333
Transformation	Reco_trf.py
Transformation Version	16.0.3.8.3
Transfromation Cache	TopPhys
Transformation Parameters	--athenaopts,--ignoreerrors,--omitvalidation,DBRelease,autoConfiguratio
Values	NONE,True,ALL,15.1.2,everything,NONE,NONE,NONE,NONE,NONE,NONE,NON
Total Number Of Output Events	1684000
First File Number in Input Dataset	1
Total Input Files	1684
Number Of Events for Output File	1000
CPU per Event	10
Memory Usage	3500
Grid Flavour	panda@it
E-mail	Junji.Tojo@cern.ch

Output Atributes

Output Task Name :	data10_7TeV.00167607.physics_Egamma.merge.r1774_p327_p333_p491
Output Datasets	DAOD_ELLOOSE18
Priority	650
SW Release :	16.0.3
State	running
Request Time :	Mon Apr 4 17:30:32 2011
Comments	group production SM WZ
Project Mode	spacetoken=ATLASDATADISK;reprocessing=yes;lumiblock=yes;diskcount=6
Modification Time :	Never
Modified by	None
Request Id :	292776
Total_req_jobs :	170
Total_done_jobs :	6
Total_avail_jobs :	1684
Physics group :	gr_SM
Queue :	default
Ctag :	p491
Bug report :	N/A

Total_avail_jobs : 1684
Physics group : gr_SM
Queue : default
Ctag : p491

The group names are defined in the Task Requests

Definition of “groups” — pathena

How to make group defined outDS

```
$ pathena --official --voms atlas:/atlas/groupName/Role=production --outDS group.groupName.[otherFields].dataType.Version ...
```

where groupName for SUSY is phys-susy, for instance. See the document ATL-GEN-INT-2007-001 for dataset naming convention. Note that you need to have the production role for the group to produce group-defined datasets. If not, please request it in the ATLAS VO registration page. For example, to produce datasets for Higgs working group,

The group names are defined in VOMS and specified in the pathena command line

atlas VO Registration

/atlas/det-indet	production	Approved
/atlas/det-larg	production	Approved
/atlas/det-muon	production	Approved
/atlas/det-tile	production	Approved
/atlas/perf-egamma	production	Approved
/atlas/perf-flavtag	production	Approved
/atlas/perf-idtracking	production	Approved
/atlas/perf-jets	production	Approved
/atlas/perf-muons	production	Approved
/atlas/perf-tau	production	Approved
/atlas/phys-beauty	production	Approved
/atlas/phys-exotics	production	Approved
/atlas/phys-gener	production	Approved
/atlas/phys-hi	production	Approved
/atlas/phys-higgs	production	Approved
/atlas/phys-lumin	production	Approved
/atlas/phys-sm	production	Approved
/atlas/phys-susy	production	Approved
/atlas/phys-top	production	Approved
/atlas/trig-daq	production	Approved

Group activities on the Panda Monitor

Servers: CERN:OK Logger:OK

Active tasks: CA:6 CERN:2 DE:25 ES:9 FR:67 IT:17 ND:7 NL:18 TW:3 UK:14 US:30
Bamboo task brokerage, job submissions, status over last 12 hours

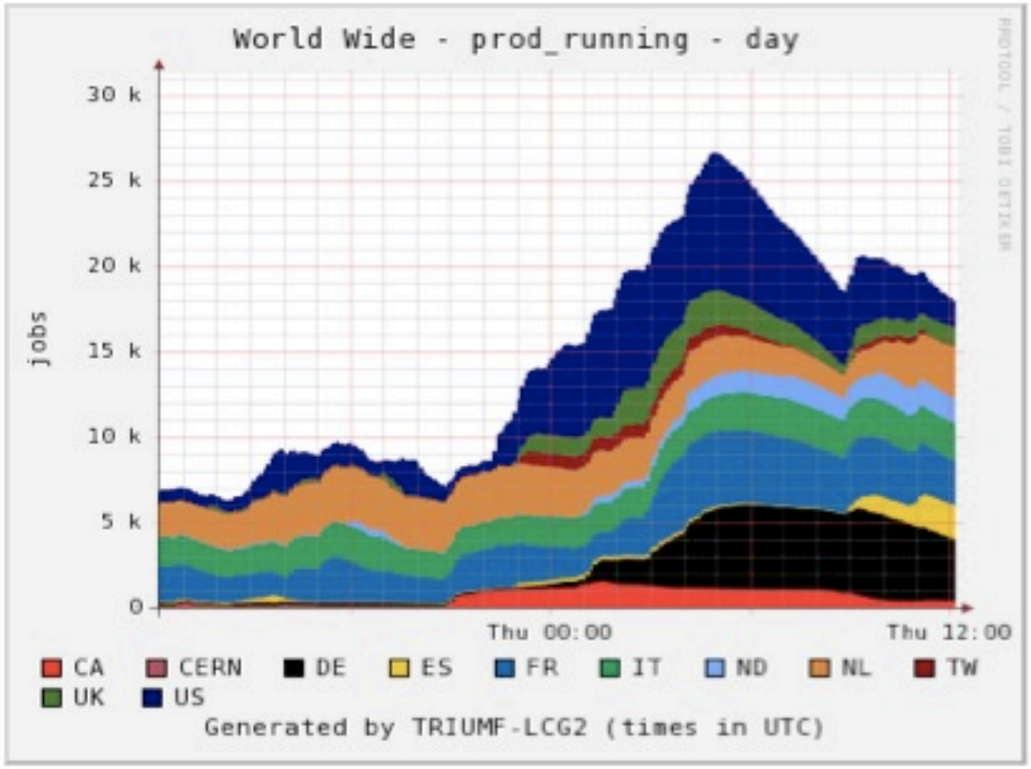
Jobs updated >12 hrs ago: **activated:** 1193 **running:** none
Jobs updated >36 hrs ago: **transferring:** none

Summary plots by cloud

Cloud efficiency history

Ganglia World Wide Summary

More plots: ALL jobs in CA last hour Go



Panda Monitor shows the on-going activities

- per processing type
- per group
- per user

Production job summary at 03–31 12:13, last 12 hours (Details: errors, nodes)
Processing types: evgen(82) merge(2155) pile(263) prod_test(10) reco(1419) reprocessing(6835) simul(48679) validation(2345)
Working groups: BPhysics(2) Exotics(52) Generators(80) Heavylon(176) Higgs(21337) Physics(8970) Reprocessing(5421) SM(2) Susy(1) Top(20449) Validation(2206) gr_Detector(16) gr_SM(2800) gr_Susy(22) gr_Top(5) heavyion(20) higgs(52) validation(177)
Users: Junji.Tojo(28) areinsch(1327) borut.kersevan(51117) c.gwenlan1(80) douglas(10) katarzp(22) leonardo.carminati(10) menke(16) r.henderson(2327) strandbe(5411) yhj(1440)

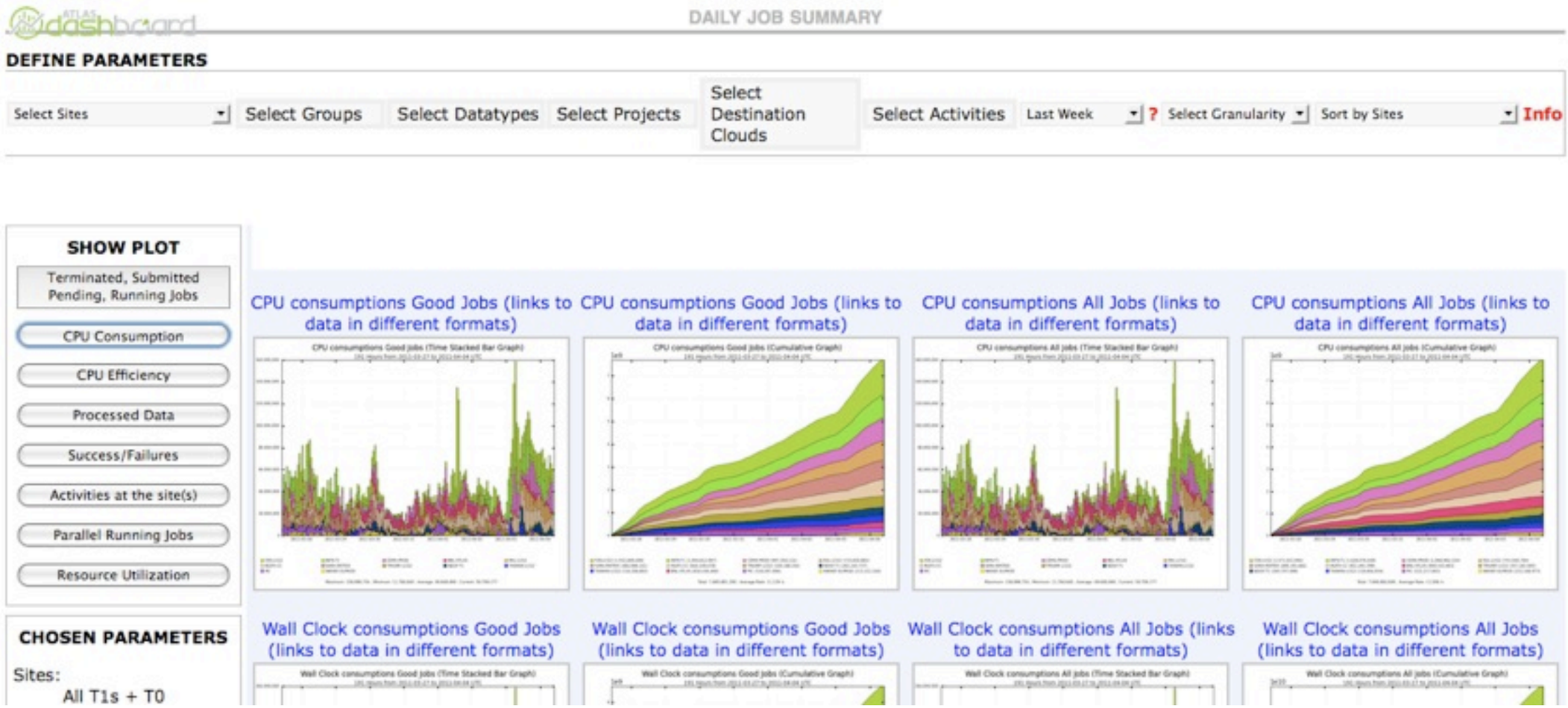
Panda production monitoring page shows group production activities as “gr_Group” as defined in the task request

Panda analysis monitoring page shows pathena jobs for groups with VOMS group names

Analysis job summary at 03–31 12:10, last 12 hours (Details: errors, nodes) pathena analysis queue status
Processing types: ITB_Integration(320) ganga(3042) gangarobot(3189) hammercloud(5230) pathena(32839) prun(65590) usermerge(1)
Working groups: calib-muon(580) det-indet(451) det-muon(1002) perf-jets(84) perf-tau(1452) phys-higgs(3) phys-sm(5740)
Special Handling: localpool(301)

New Monitoring available

http://dashb-atlas-job-prototype.cern.ch/dashboard/request.py/dailysummary



Still under development

- please report when you see something wrong
- or if you have suggestions

Group production monitoring

S. Jézéquel (LAPP)

Introduction

- New version of 'Historical view' monitoring for ATLAS Grid CPU activity

  http://dashb-atlas-job-prototype.cern.ch/dashboard/request.py/dailysummary
- Now usefull for group production monitoring
- Still under validation but many checks already done

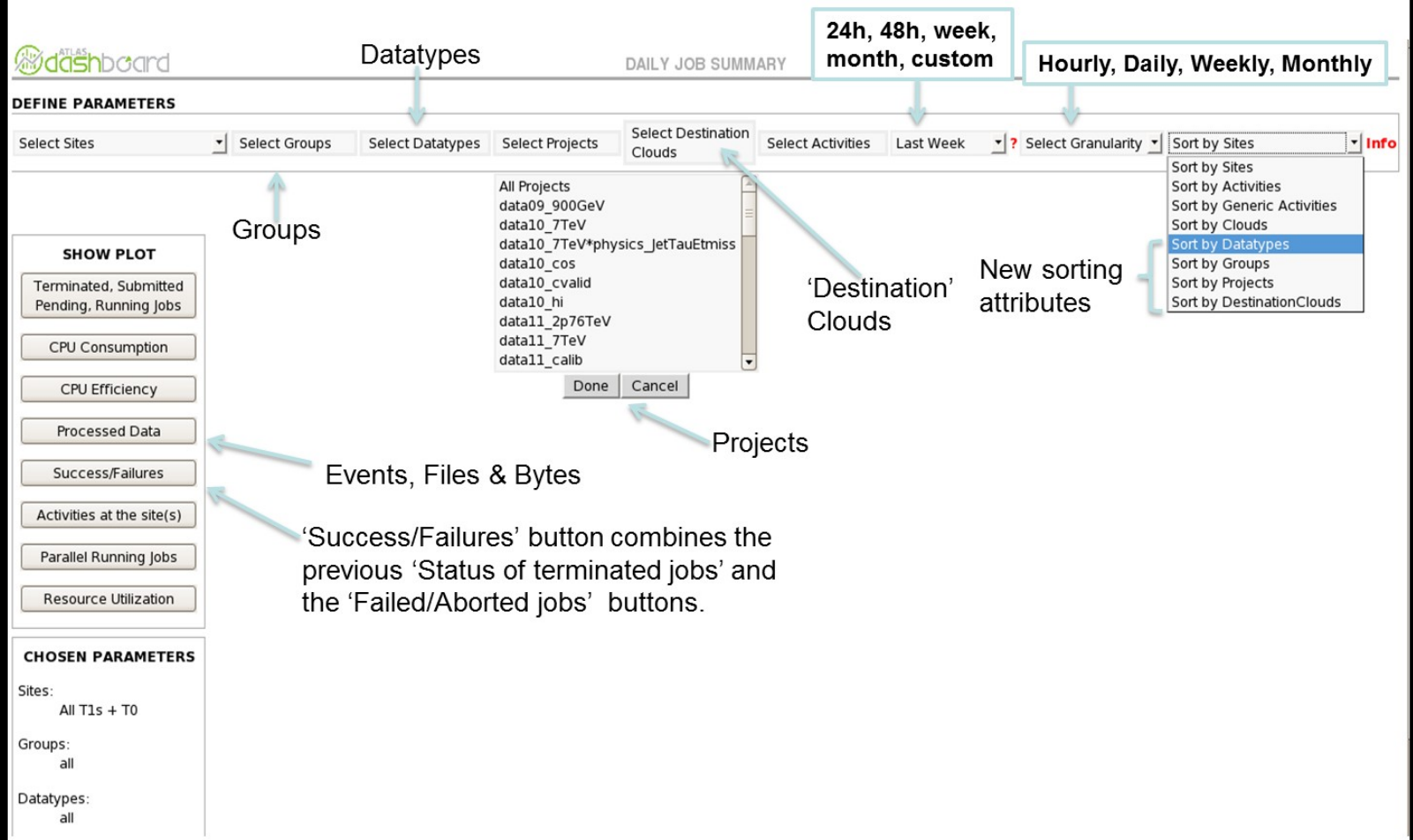
- Informations collected since 3 weeks

- Detailed presentation in 'Monitoring session' (in parallel to this session)

Experiment Support

UI of the Prototype

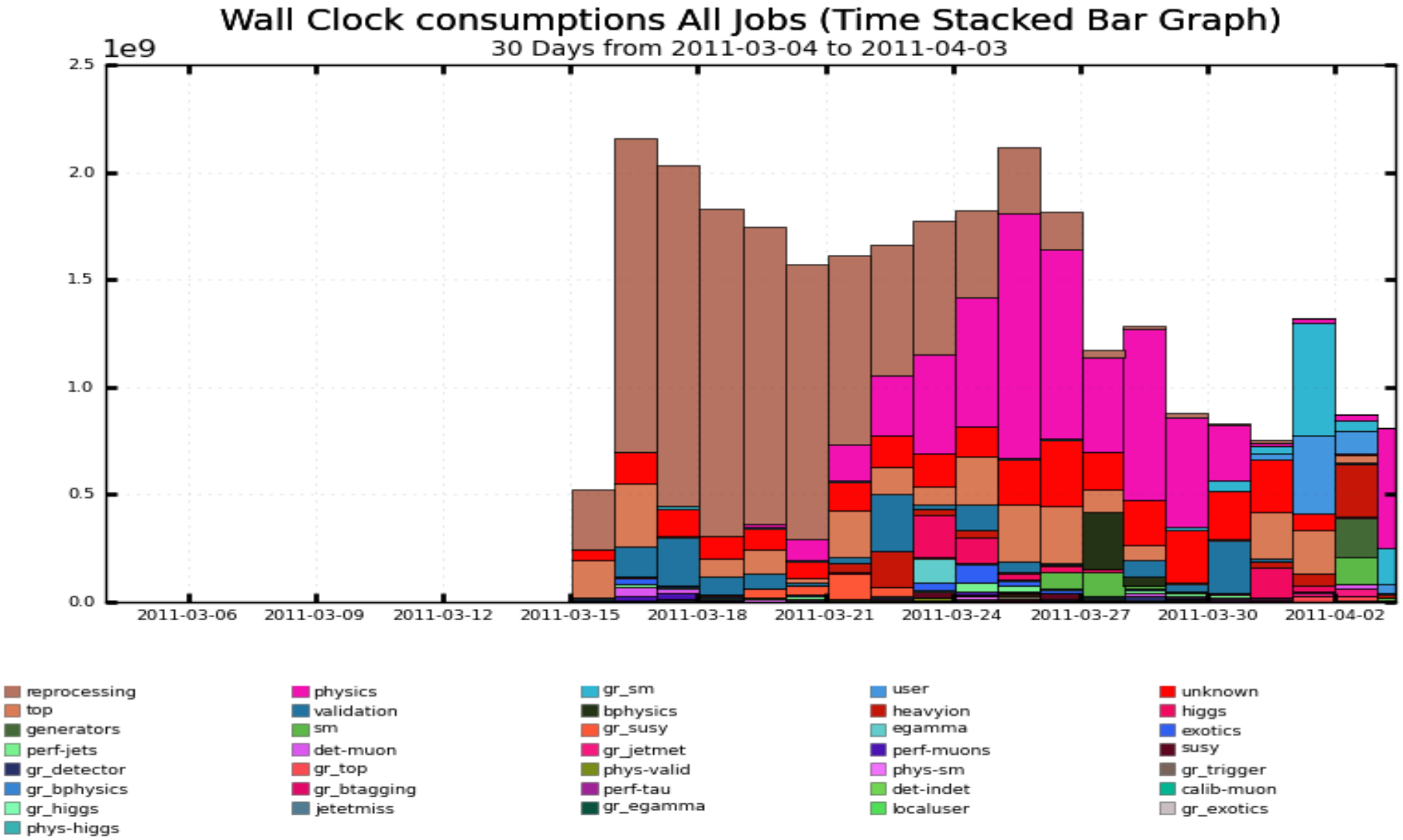
CERN IT Department



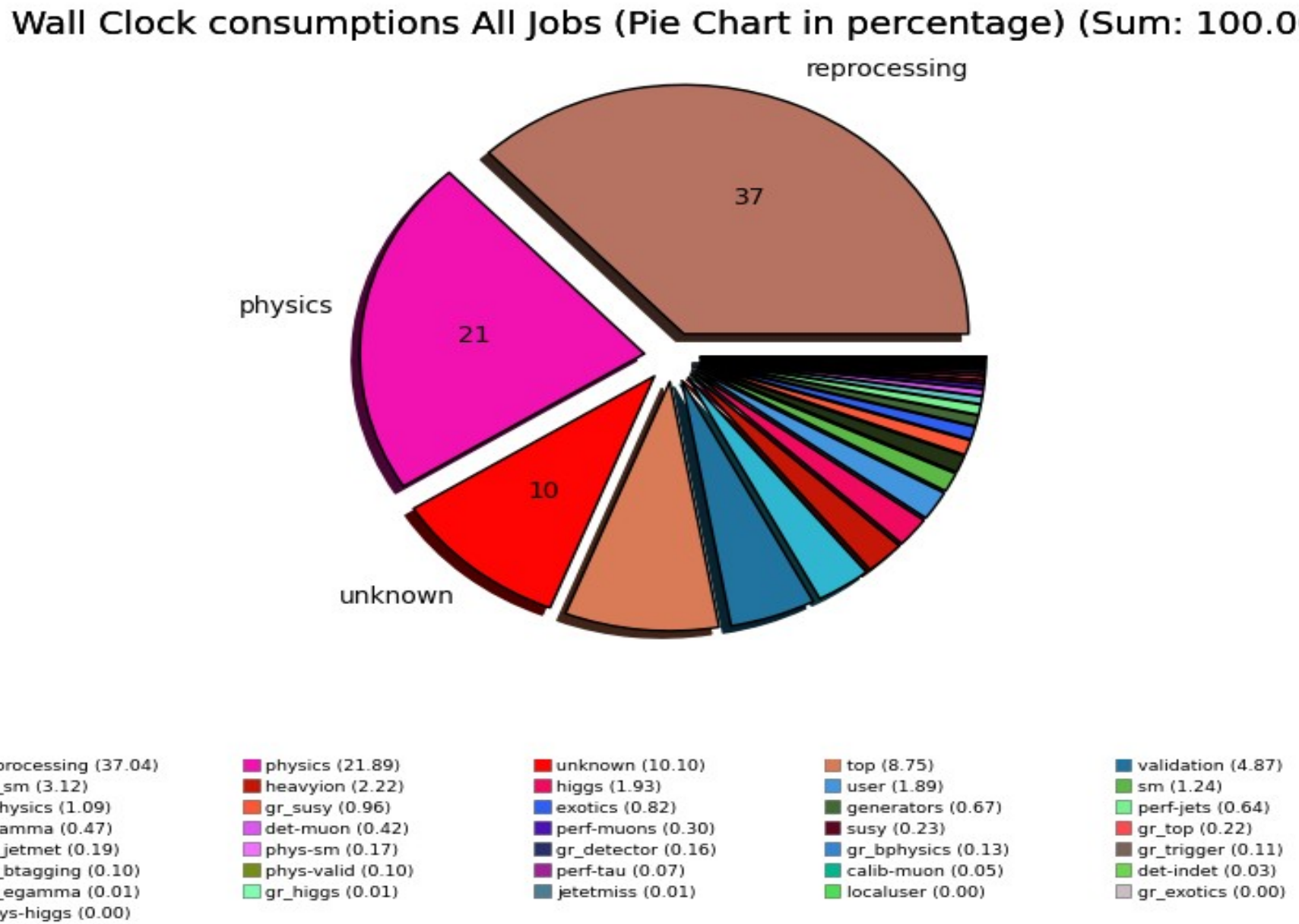
Nomenclature for group monitoring

- Naming convention for group activity
  - Nomenclature defined in Panda
- Related to voms group:phys-*,perf-*,...
  - Pathena jobs with group priority
- Starts with gr_
  - Central production for groups :
- Localuser/user : pathena/ganga jobs without specific priority
- Remaining part :
  - Validation, reprocessing, Physics, … : Central production

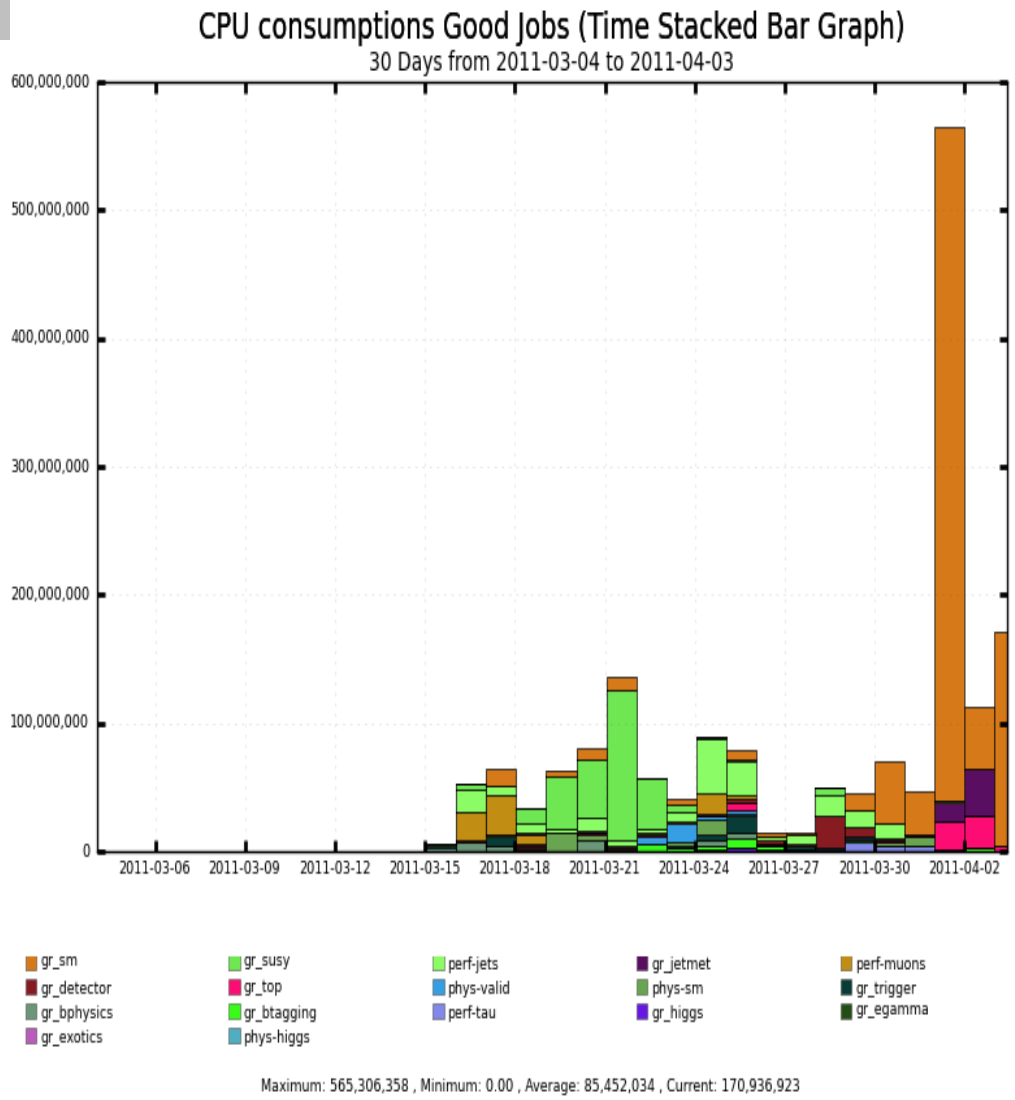
Global results(1)

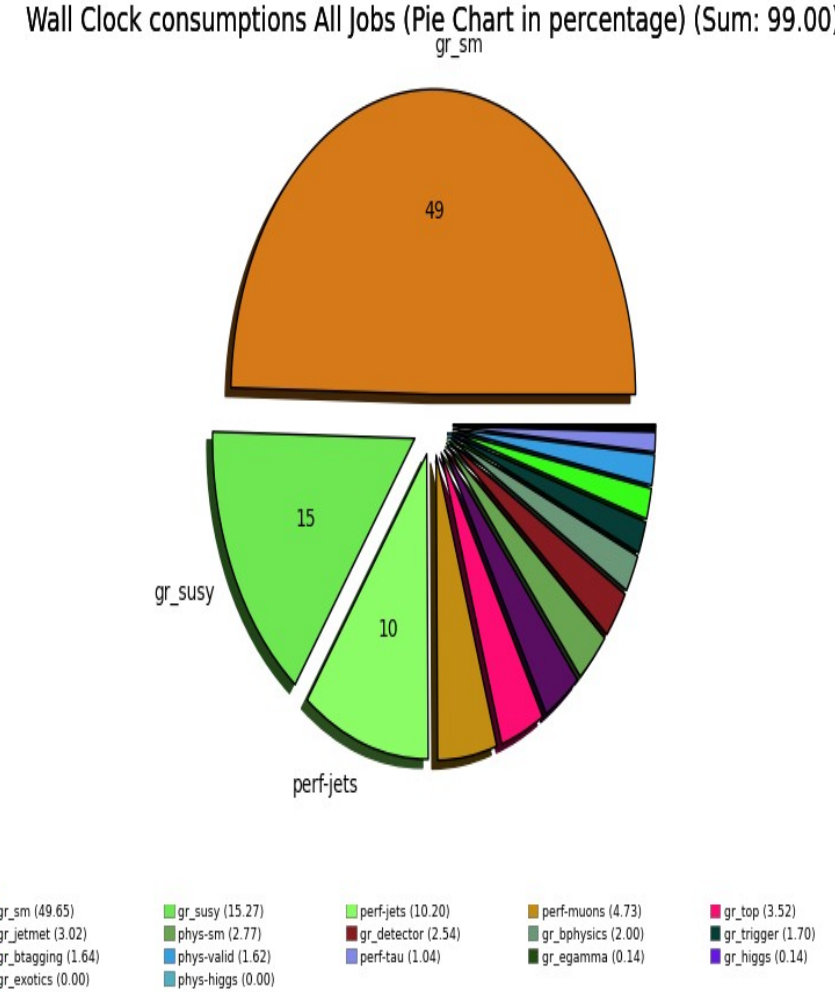


Global results

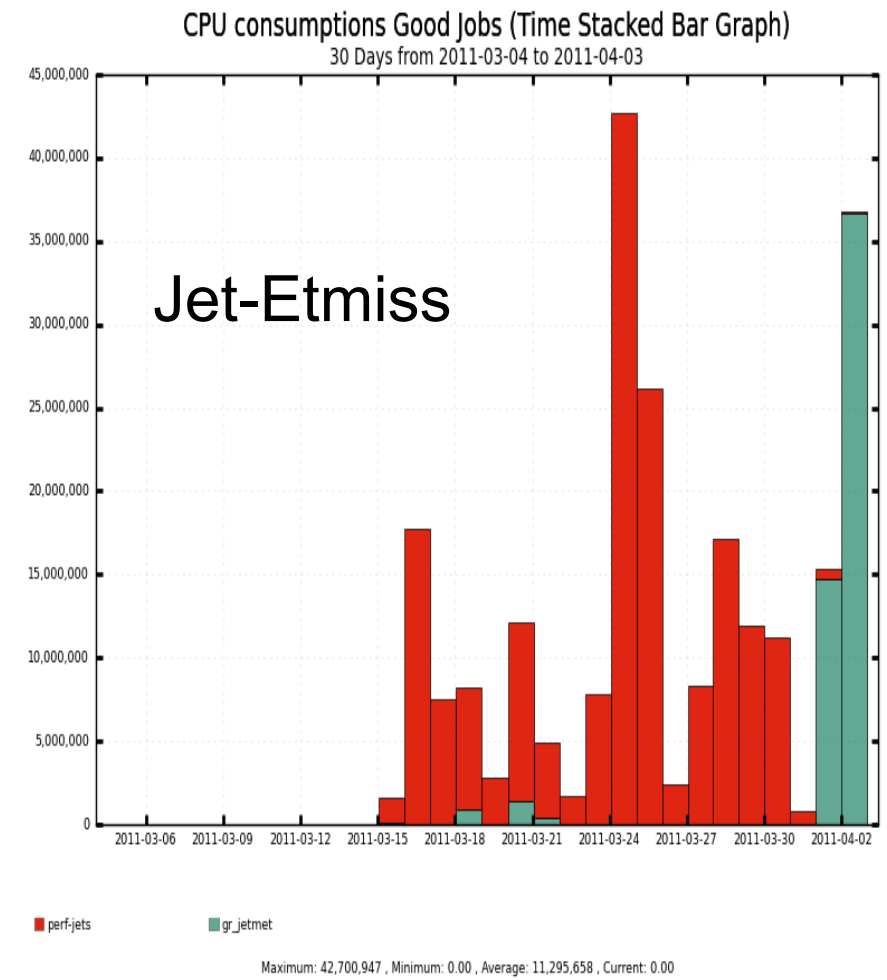


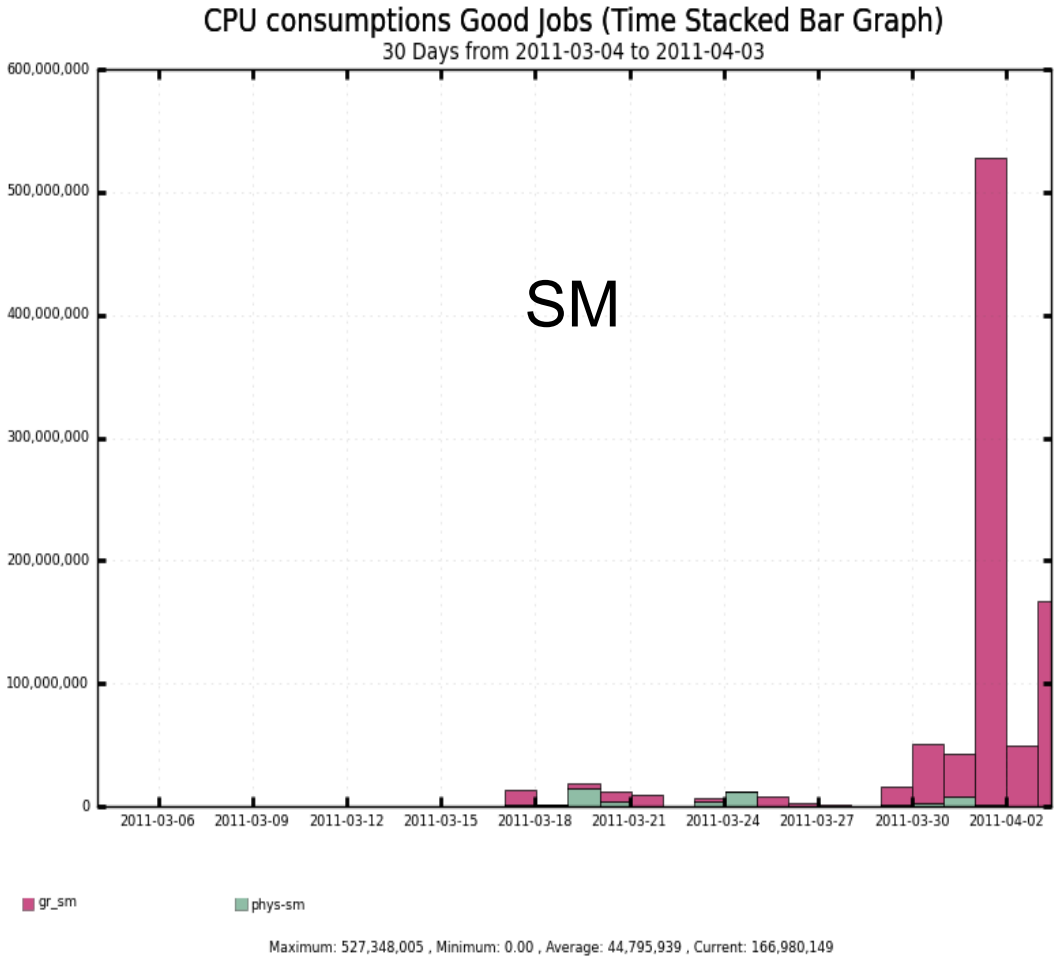
Group activity



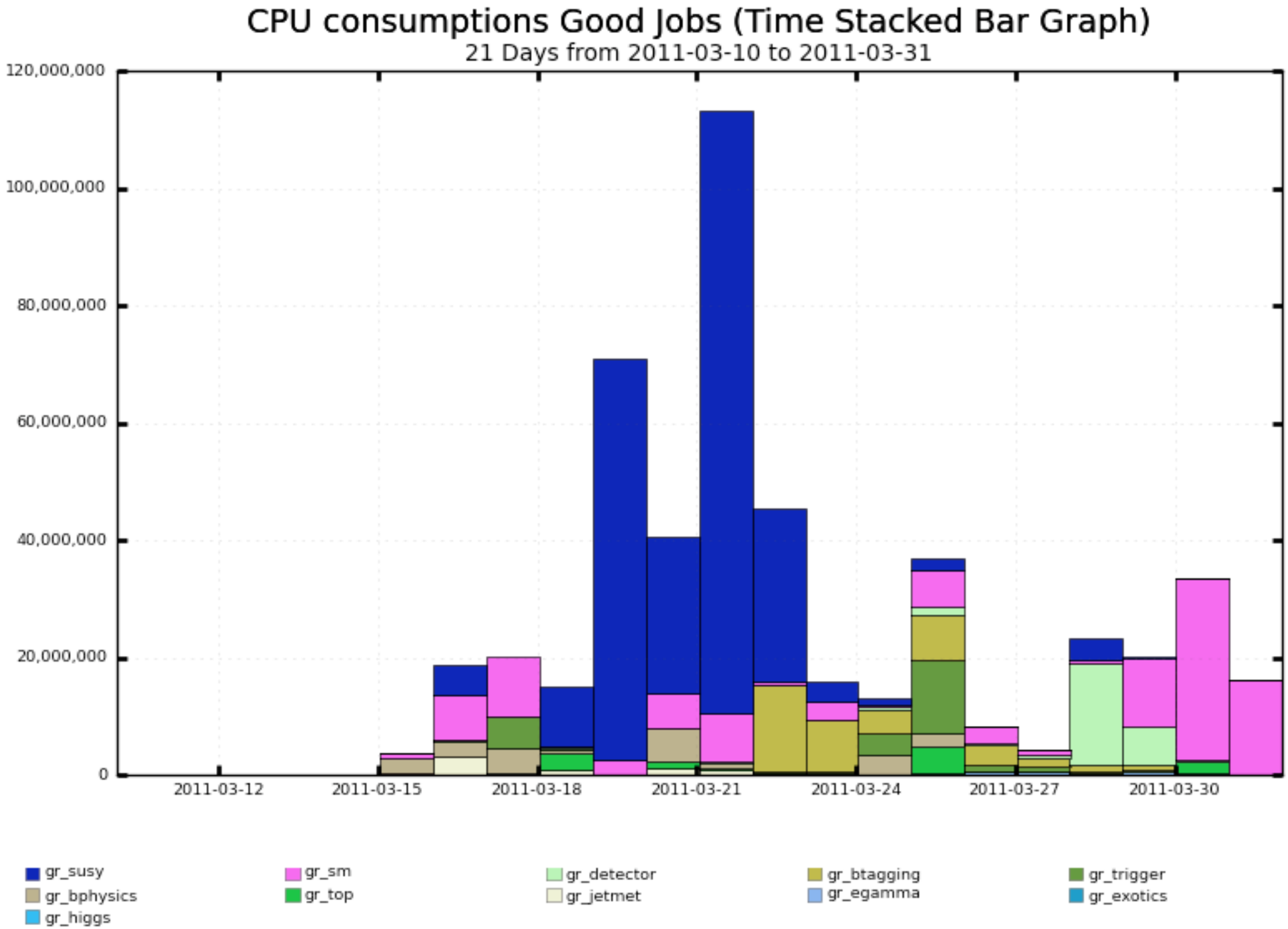
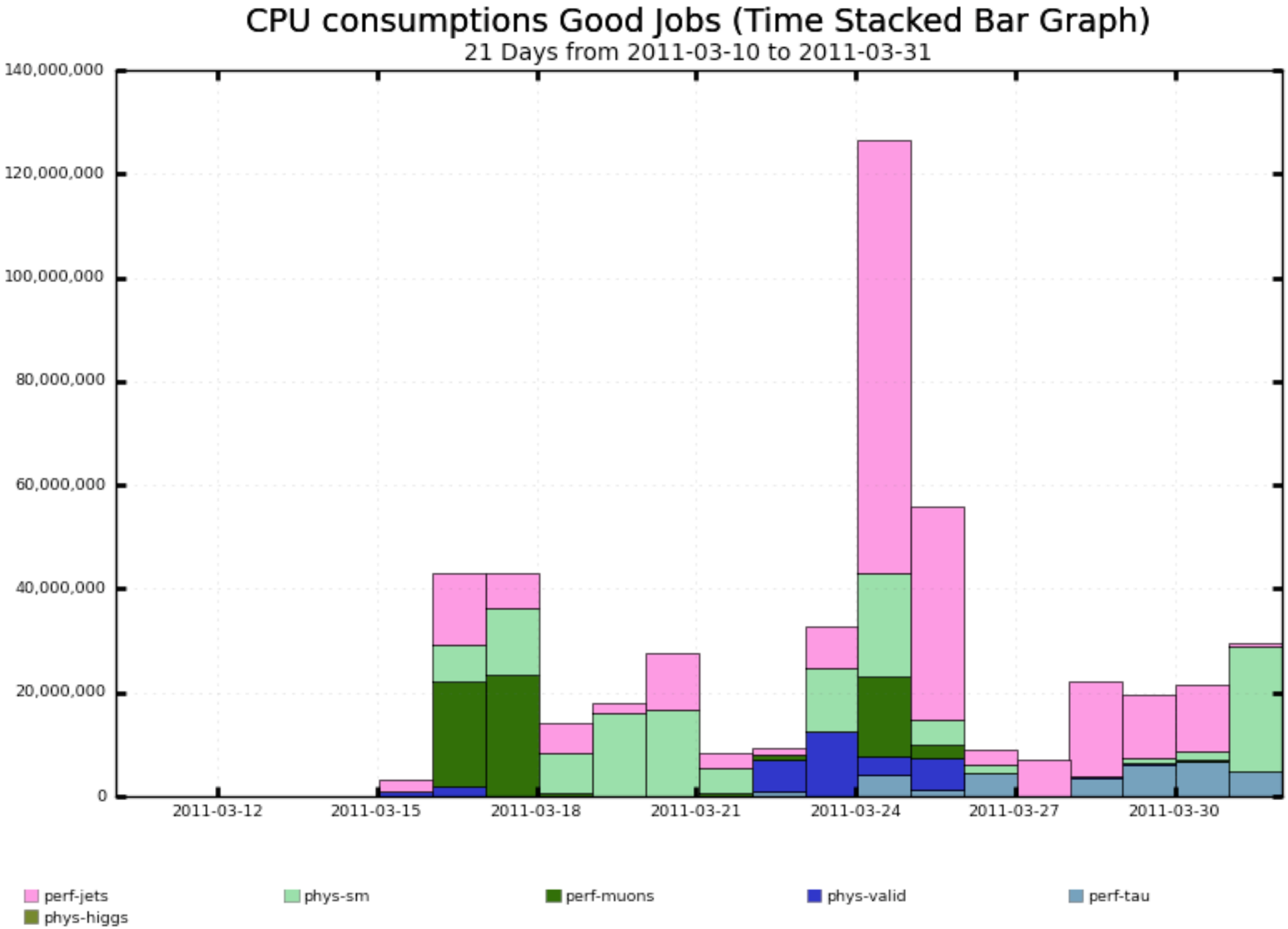


Specific groups





‘Group Production’ activities



Pathena group jobs