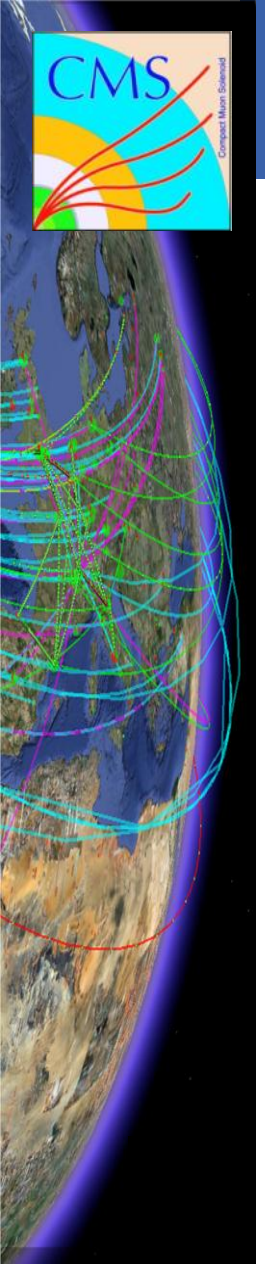
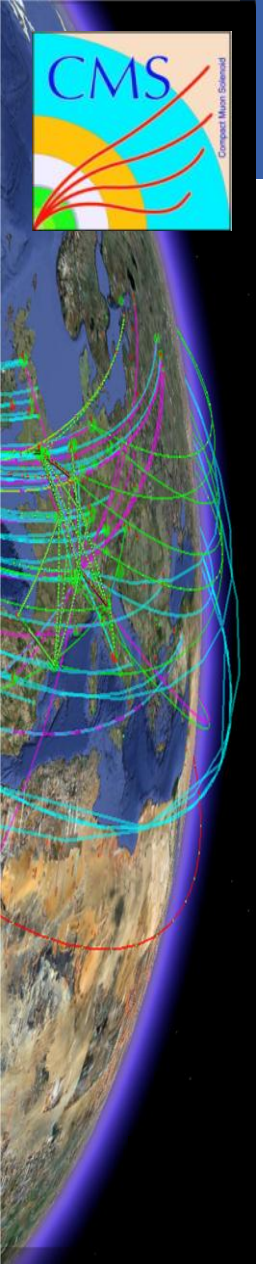


Information Technology in CERN

Vladimir Škarupelov

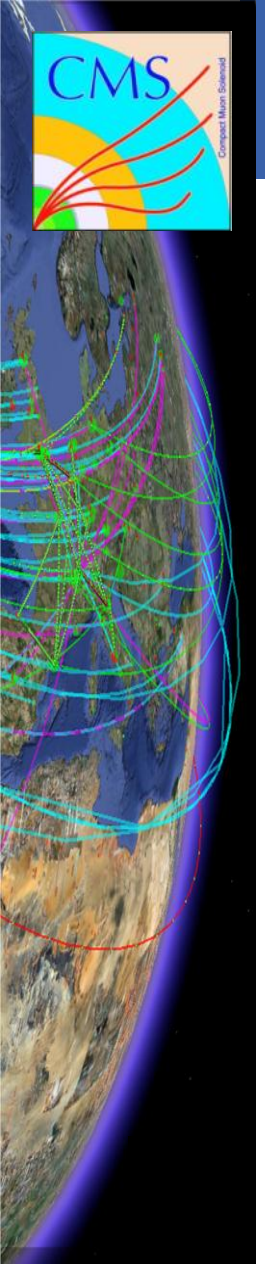
vladimir.skarupelov@cern.ch



Agenda

- Problems that IT solves in CERN
- Computing Model in CERN
- Operation systems in CERN
- Overview of my work
- Interesting facts
- Conclusion

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



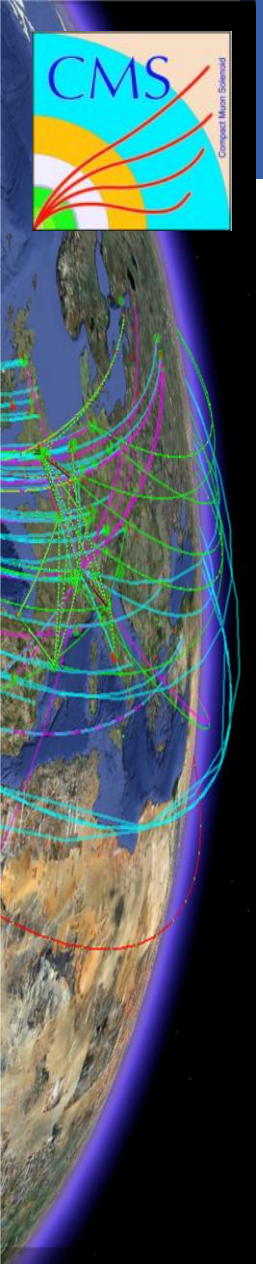
Problem

- The Large Hadron Collider (LHC) is producing roughly 15 petabytes (15 million gigabytes) of data annually – enough to fill more than 1.7 million dual-layer DVDs a year!
- Thousands of scientists around the world want to access and analyze LHC data
- Requirement of the great computing power

Solution

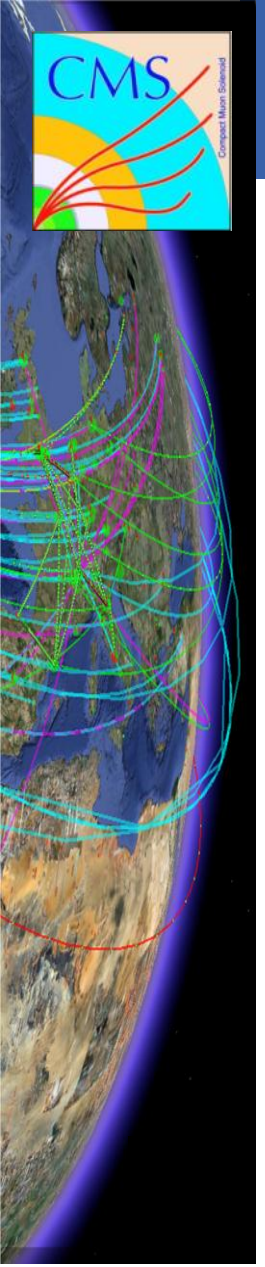
Solution

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



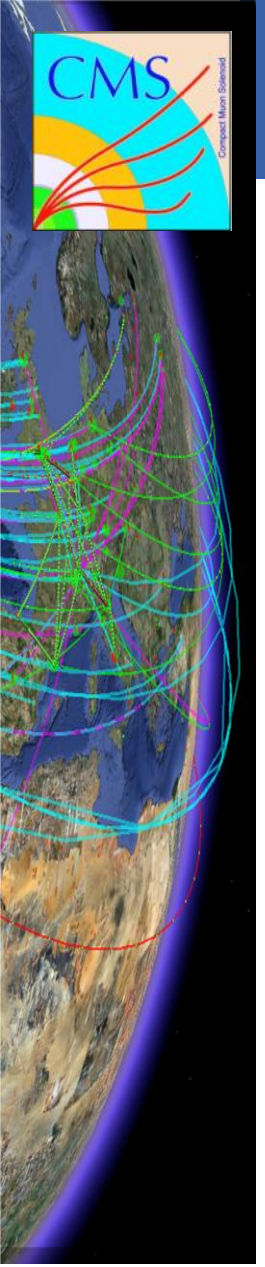
Solution

- Worldwide LHC Computing Grid
- Large data volumes are spread among different storage disks and even computing elements.
- Strong backup facility
- Good Data Access
- Scalable computing power



LHC Grid Tiers

- Tier-0 (T0) @CERN
  - Storage of RAW data
  - Initial data reconstruction
  - Copy to T1
- Tier-1 (T1) @large CMS collaborating centers (i.e. in FNAL and RAL)
  - RAW data backup
  - Re-reconstruction
  - Analysis Object Data (AOD) extraction
  - Copy to T2



LHC Grid Tiers

- Tier-2 (T2) @ "small" centres at universities:
  - services for local communities
  - grid-based analysis for the whole experiment (Tier-2 resources available to whole experiment through the grid)
  - Monte Carlo simulation for the whole experiment

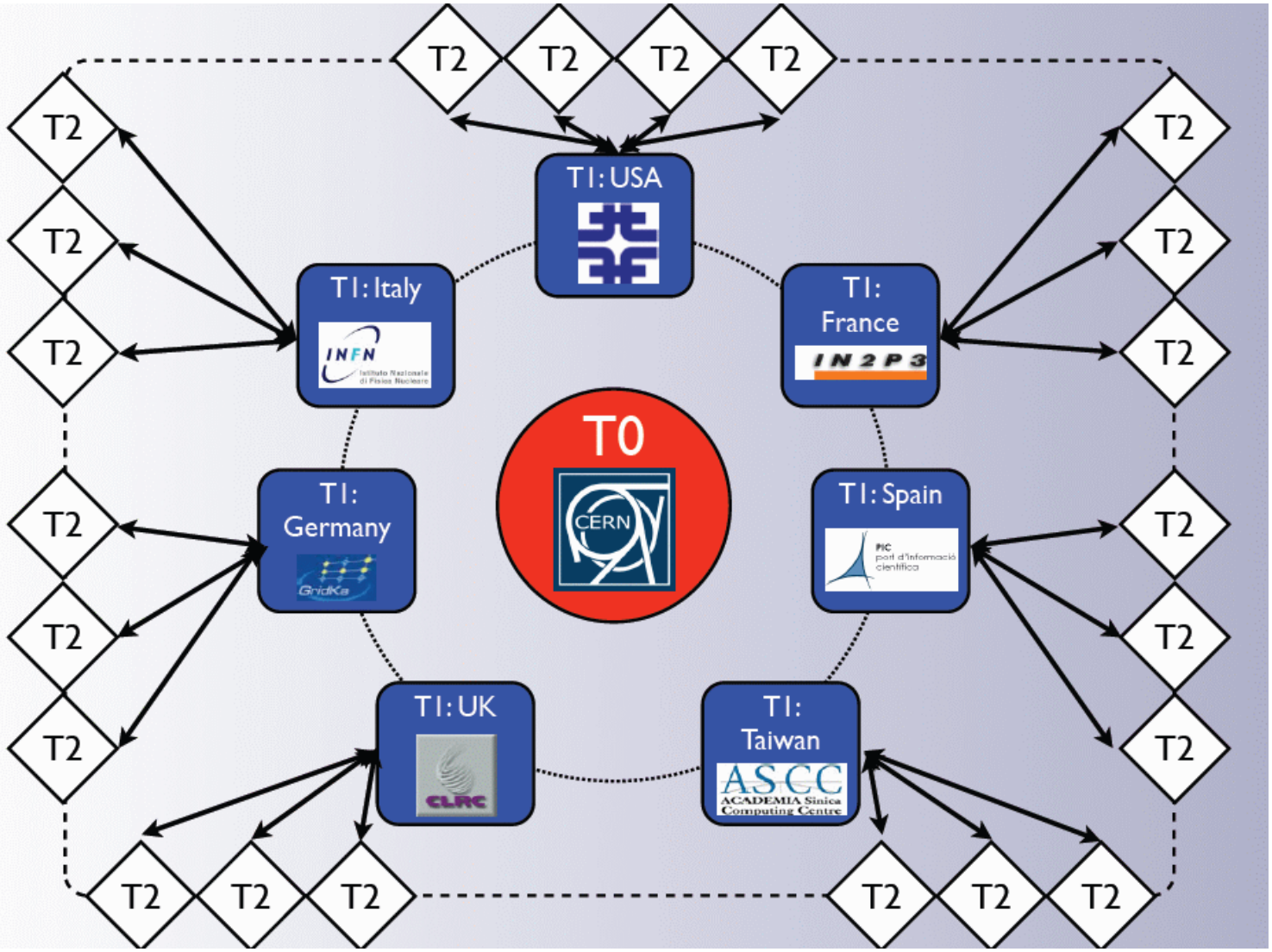
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

T0 Computing Center @CERN

CERN IT Department

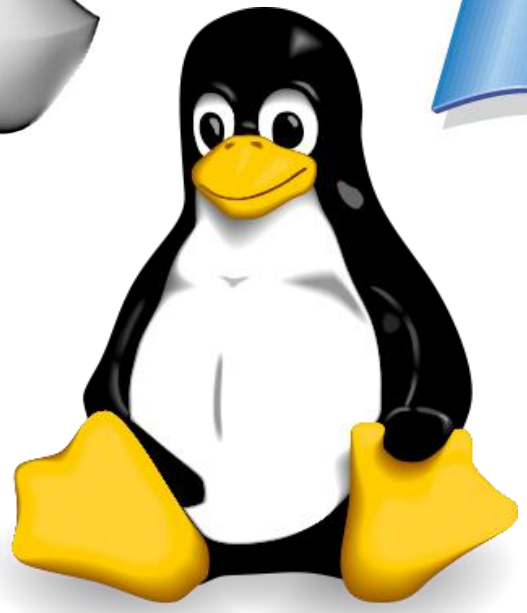
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Tier Structure



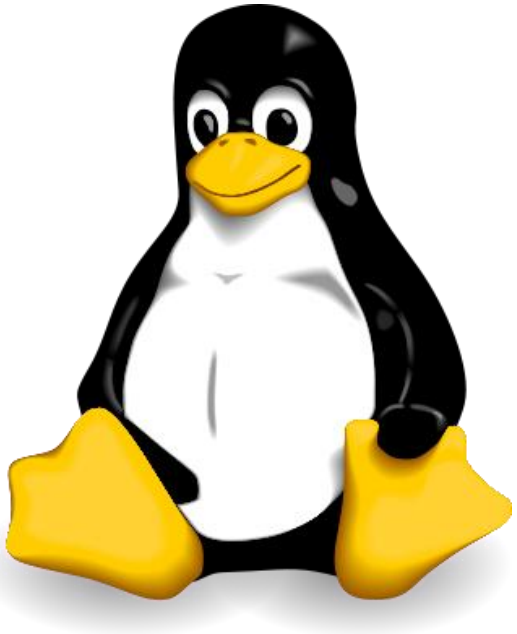
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

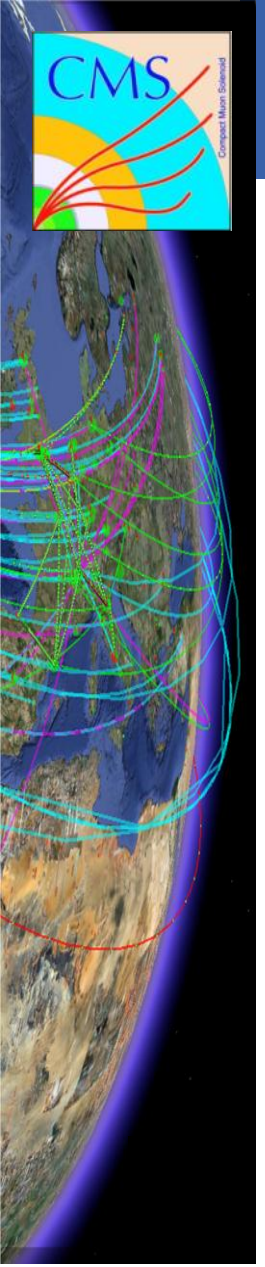
Operation System



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

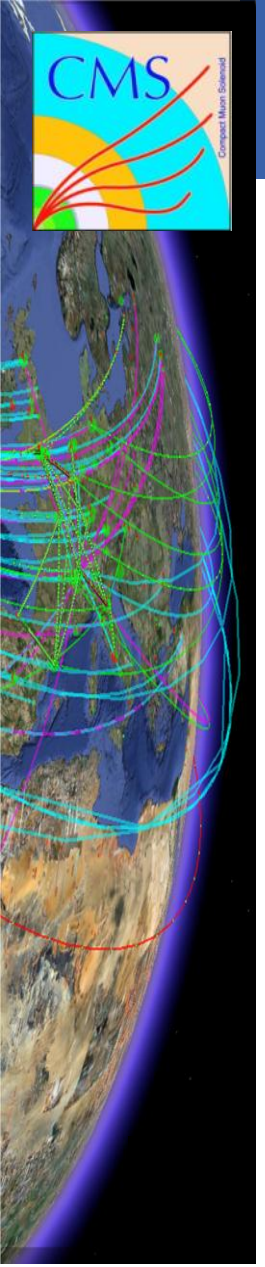
Scientific Linux





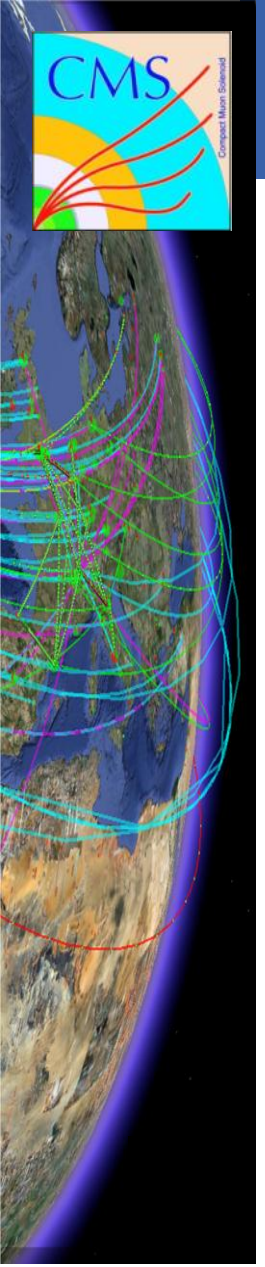
Operation System

- https://www.scientificlinux.org/
- SL is a Linux release put together by CERN, Fermilab and various other labs and universities around the world.
- reduces duplicated efforts of the labs
- has a common install base for the various experimenters



Scientific linux

- Distributions are created from RedHat (very popular linux for enterprices)
- Added specific software like openAfs, Alpine, lpadmin-cern, etc...
- Most populat versions in use:
  - Scientific Linux 3 (quite Old)
  - Scientific Linux 4
  - Scientific Linux 5
  - Scientific Linux 6



Other Operation Systems

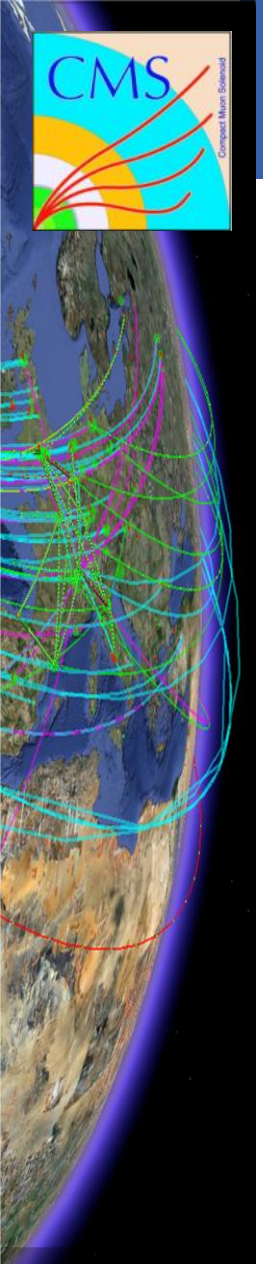
- Microsoft Windows XP,7
- Mac OS
- Android
- Nokia Symbian
- ...

Any other operation system can be used for connecting to cern front-end (lxplus.cern.ch) using ssh channel

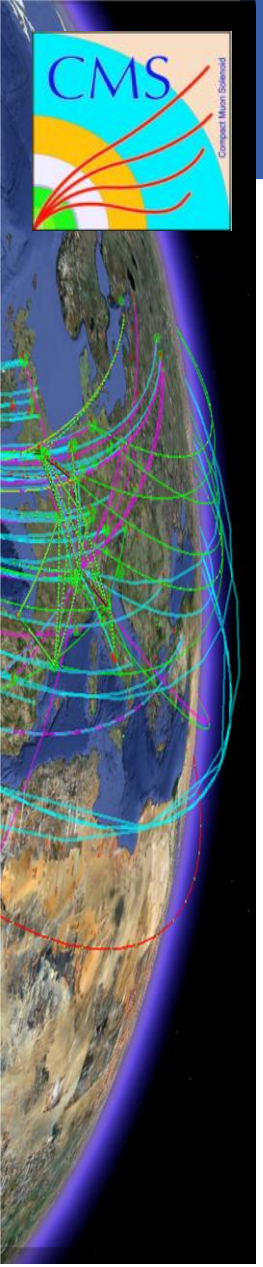
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Overview of my work

CERN IT Department

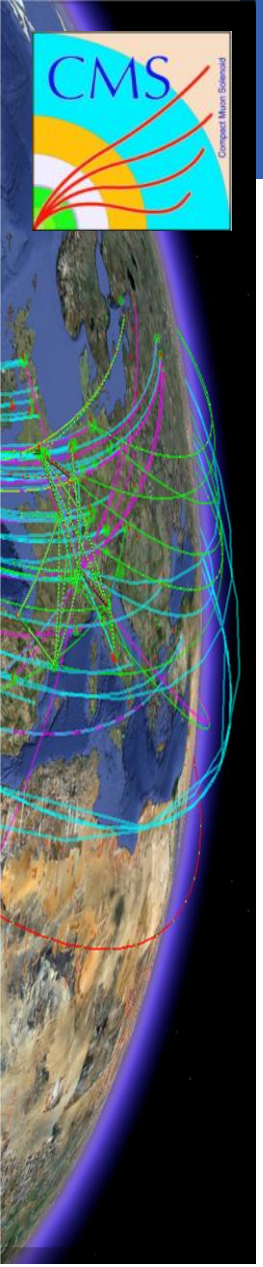


CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



Overview of my work

- The Core Software group is a project in the Offline area of CMS.
- This group, between the different responsibilities, is responsible for the integration and testing of the CMS Software.
- CMS runs twice daily "integration builds" (IB's) in which the latest version of our software is built
- Number of software QA and basic runtime crash and performance tests are done to look for problems



CMSSW IB Page

CERN IT Department

CMSSW integration builds

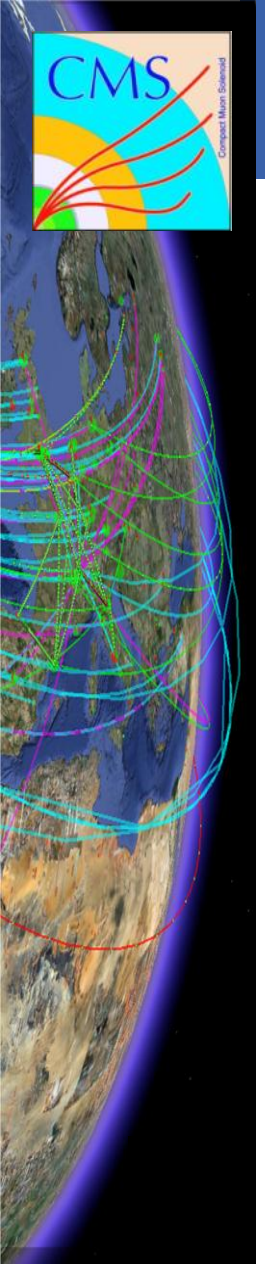
Click on the "summary" links to get some summary information for the build status and/or relval status.

Click on the alert icon ( ) to see information provided by the release manager for that specific IB.

Release cycle 5.0 -- back to top of page

day	IB	platforms	builds	RelVals	OtherTests	Q/A page
sun	CMSSW_5_0_X_2011-10-23-1400	slc5_amd64_gcc434	summary details	summary details	summary details	Q/A info Strip Charts
		slc5_amd64_gcc451	summary details	summary details	summary details	Q/A info Strip Charts
		slc5_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts
		osx106_amd64_gcc421	summary details	summary details	summary details	Q/A info Strip Charts
		osx106_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts
sun	CMSSW_5_0_X_2011-10-23-0200	slc5_amd64_gcc434	summary details	summary details	summary details	Q/A info Strip Charts 15 valgrind errors
		slc5_amd64_gcc451	summary details	summary details	summary details	Q/A info Strip Charts
		slc5_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts
		osx106_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts
sat	CMSSW_5_0_X_2011-10-22-1400	slc5_amd64_gcc434	summary details	summary details	summary details	Q/A info Strip Charts 31 valgrind errors
		slc5_amd64_gcc451	summary details	summary details	summary details	Q/A info Strip Charts
		slc5_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts
		osx106_amd64_gcc421	summary details	summary details	summary details	Q/A info Strip Charts
		osx106_amd64_gcc461	summary details	summary details	summary details	Q/A info Strip Charts

http://cmssdt.cern.ch/SDT/html/showIB.html

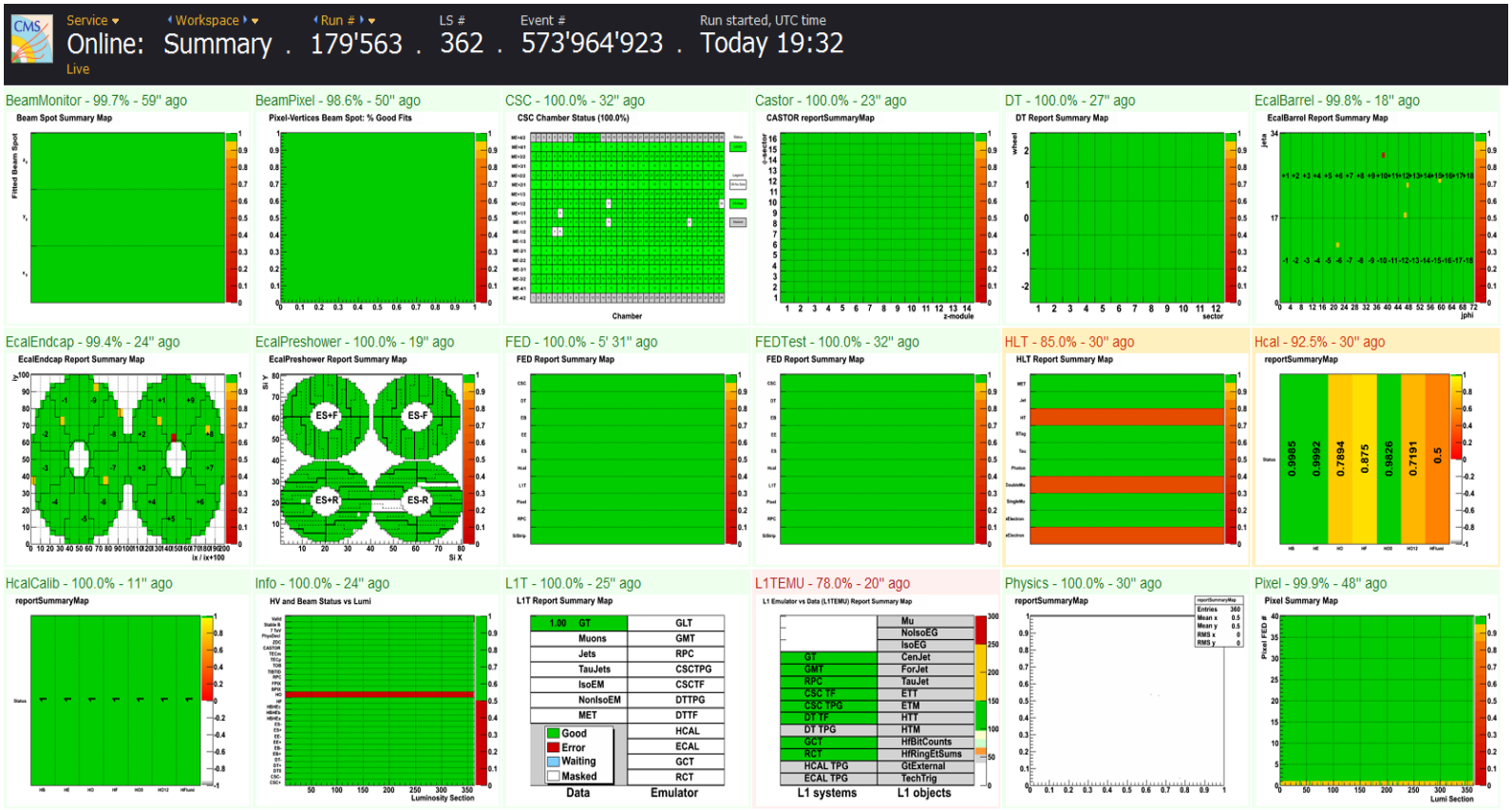


Overview of my work

- At the moment these IB's to first order test only technical "software" aspects, and do not verify even at a basic level that the data output by the applications is sensible in terms of physics quality.
- CMS Data Quality Monitoring (DQM) offline group have also invested a significant amount of effort in developing an infrastructure for DQM for the large scale monitoring of physics quantities from the output of our applications

DQM GUI Page

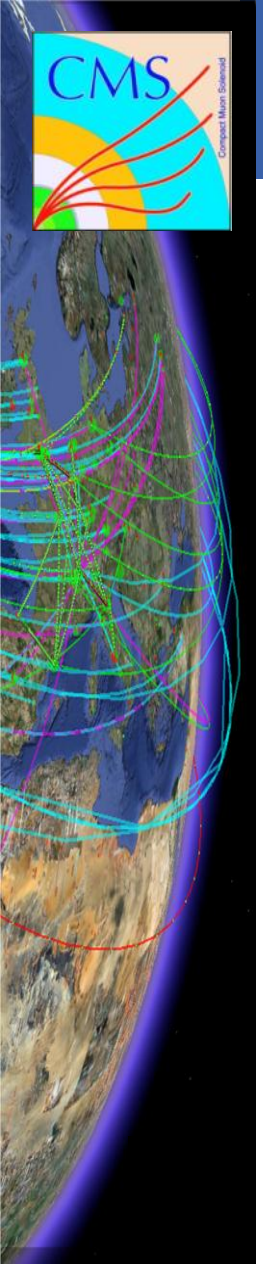
CERN IT Department



https://cmsweb.cern.ch/dqm/online/

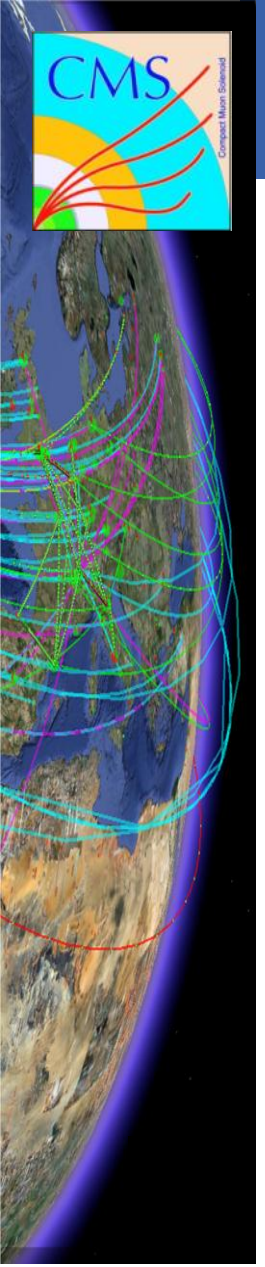
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Vladimir Škarupelov, IT in CERN



My role

- add automated monitoring of physics data quality to CMSSW integration builds
- reuse large scale DQM infrastructure
- work on the increased automation of the validation activities done for the large scale production and testing
- release management of CMSSW
- administration of Estonian T-2 GRID



Automation results

CERN IT Department

CMSSW Integration Build test log viewer

Logs available in /afs/cern.ch/cms/sw/ReleaseCandidates/slc5_amd64_gcc434/www/sun/5.0-sun-02/CMSSW_5_0_X_2011-10-23-0200/pyRelValM

Integration Build pyRelValMatrixLogs -- Back to IB portal

Step	total	passed	failed	timeout	notRun
step1	137	137	0	0	N/A
step2	120	120	0	0	32
step3	72	72	0	N/A	32
step4	13	13	0	N/A	32
Comparison	2	1	1	0	0

Summary log (runall.log)

Failed tests : (0)
back to top

RelVal tests: (137 tests)
back to top

#	workflow	step1	step2	step3	step4	Comparison
1	1.0 ProdMinBias	log cmd	log cmd	log cmd		
2	2.0 ProdTTbar	log cmd	log cmd	log cmd		
3	3.0 ProdQCD_Pt_3000_3500	log cmd	log cmd	log cmd		
4	4.17 RunMinBias2011A	log cmd	log cmd	log cmd	log cmd	report
5	4.22 RunCosmics2011A	log cmd	log cmd	log cmd	log cmd	
6	4.23 ValSkim2011A	log cmd	log cmd			
7	4.24 WMuSkim2011A	log cmd	log cmd			
8	4.25 WElSkim2011A	log cmd	log cmd			
9	4.26 ZMuSkim2011A	log cmd	log cmd			

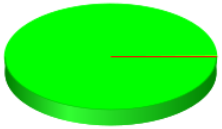
Automation results

Summary

60967 COMPARISONS:

- SUCCESS: 99.8% (60862)
- NULL: 0.2% (99)
- FAIL: 0.0% (6)

To the DQM GUI...



Sample:

GLOBAL

Run1 and Run2:

165121 - 165121

Releases:

- CMSSW_5_0_X_2011-10-23-0200
- CMSSW_5_0_X_2011-09-29-0200

Statistical Test (Pvalue threshold):

- CHI2 (1E-05)

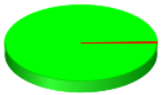
Sub-Directories

Top...

SiStrip

835 COMPARISONS:

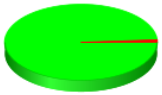
- SUCCESS: 99.5% (831)
- FAIL: 0.5% (4)



EcalPreshower

97 COMPARISONS:

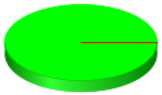
- SUCCESS: 99.0% (96)
- FAIL: 1.0% (1)



JetMET

2318 COMPARISONS:

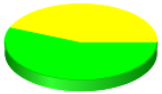
- SUCCESS: 100.0% (2317)
- FAIL: 0.0% (1)

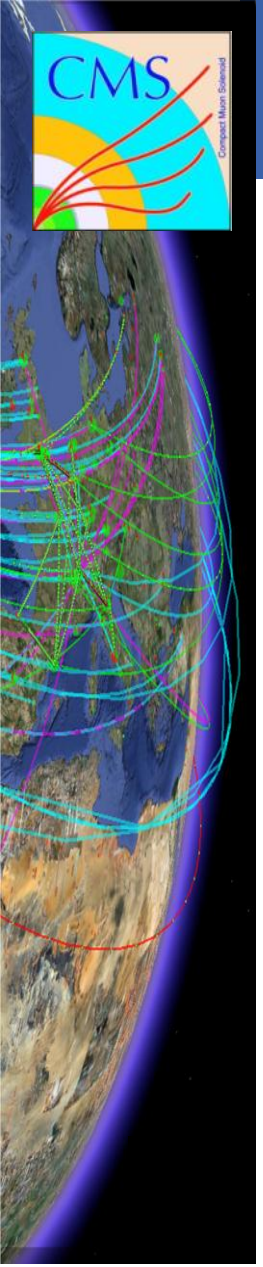


Tracking

110 COMPARISONS:

- SUCCESS: 56.4% (62)
- NULL: 43.6% (48)

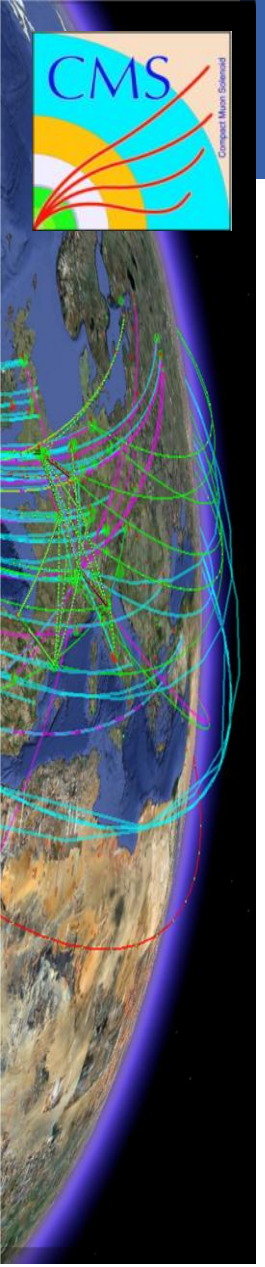




Estonian T2 Cluster

- Located in Tallinn (hardware change is done on site)
- Using planetar system items names in domain names, i.e:
  - mars.hep.kbfi.ee
  - moon.hep.kbfi.ee
  - sun.hep.kbfi.ee
- Can be administrated remotely (even switched off/on)
- Installation and configuration of new virtual machines and software

This is intersting, isn't it ?



What is it?

CERN IT Department

World Wide Web

The WorldWideWeb (W3) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary of the project, Mailing lists , Policy , November's W3 news , Frequently Asked Questions .

What's out there?
Pointers to the world's online information, subjects , W3 servers, etc.

Help
on the browser you are using

Software Products
A list of W3 project components and their current state. (e.g. Line Mode ,X11 Viola , NeXTStep , Servers , Tools , Mail robot , Library)

Technical
Details of protocols, formats, program internals etc

Bibliography
Paper documentation on W3 and references.

People
A list of some people involved in the project.

History
A summary of the history of the project.

How can I help ?
If you would like to support the web..

Getting code
Getting the code by anonymous FTP , etc.

Original first web page!

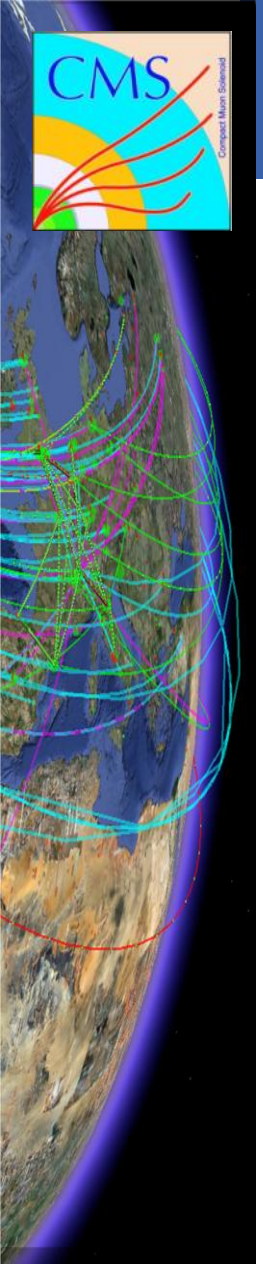
- http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html
- The World Wide Web began as a CERN project called ENQUIRE, initiated by Tim Berners-Lee

Interesting facts

- This NeXT Computer used by Sir Tim Berners-Lee at CERN became the first Web server.

- This Cisco Systems router at CERN was probably one of the first IP routers deployed in Europe.

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



Conclusion

- CERN is using GRID to store and operate data ✓
- The tier structure of the GRID ✓
- The Scientific Linux is used as main operation system in CERN ✓
- My role in CERN ✓
- Intersting discoveries at CERN ✓

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Questions?

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it