K
I
P
T

Grid Middleware Configuration at the KIPT CMS Linux Cluster

S. Zub, L. Levchuk, P. Sorokin, D. Soroka

Kharkov Institute of Physics & Technology, 61108 Kharkov, Ukraine

http://www.kipt.kharkov.ua/~cms
stah@kipt.kharkov.ua

Why do the LHC experiments (such as CMS) need the Grid?

- LHC nominal luminosity of $10^{34}cm^{-2}s^{-1}$ corresponds to 10^{9} proton-proton collisions per second
- In CMS detector about 10^{-7} of the total event flow will be selected by a multi-level trigger for the off-line event processing and analysis
- Data should be archived in a high performance storage system with the rate of 100 Hz
- Size of one CMS event is 1 Mbyte $\Rightarrow$ more than 1 Pbyte annually
- Typically, ~10^{9} CMS events (or 10^{15} bytes of information) have to be (remotely) processed and analyzed in order to find manifestations of “new physics”
- Grid technology is the best solution for this task!

Main distinguishing features of Grid technology

- New level of encapsulation for data, analysis programs and computing resources
- Uniform interface between client and network distributed computation environment
- Coordinated and dynamic resource sharing
- New opportunities for scientific cooperation via multi-institutional virtual organizations
- Middleware for easy and secure communication

What is KIPT CMS Linux Cluster (KCC)?

Performances:

11 nodes (22 CPU's, 38 Gflops), 2.5 TB HDD, 9 GB RAM

System Software and Middleware:

SLC, NFS, NIS, PBS etc., Firewall and LCG-2_4_0

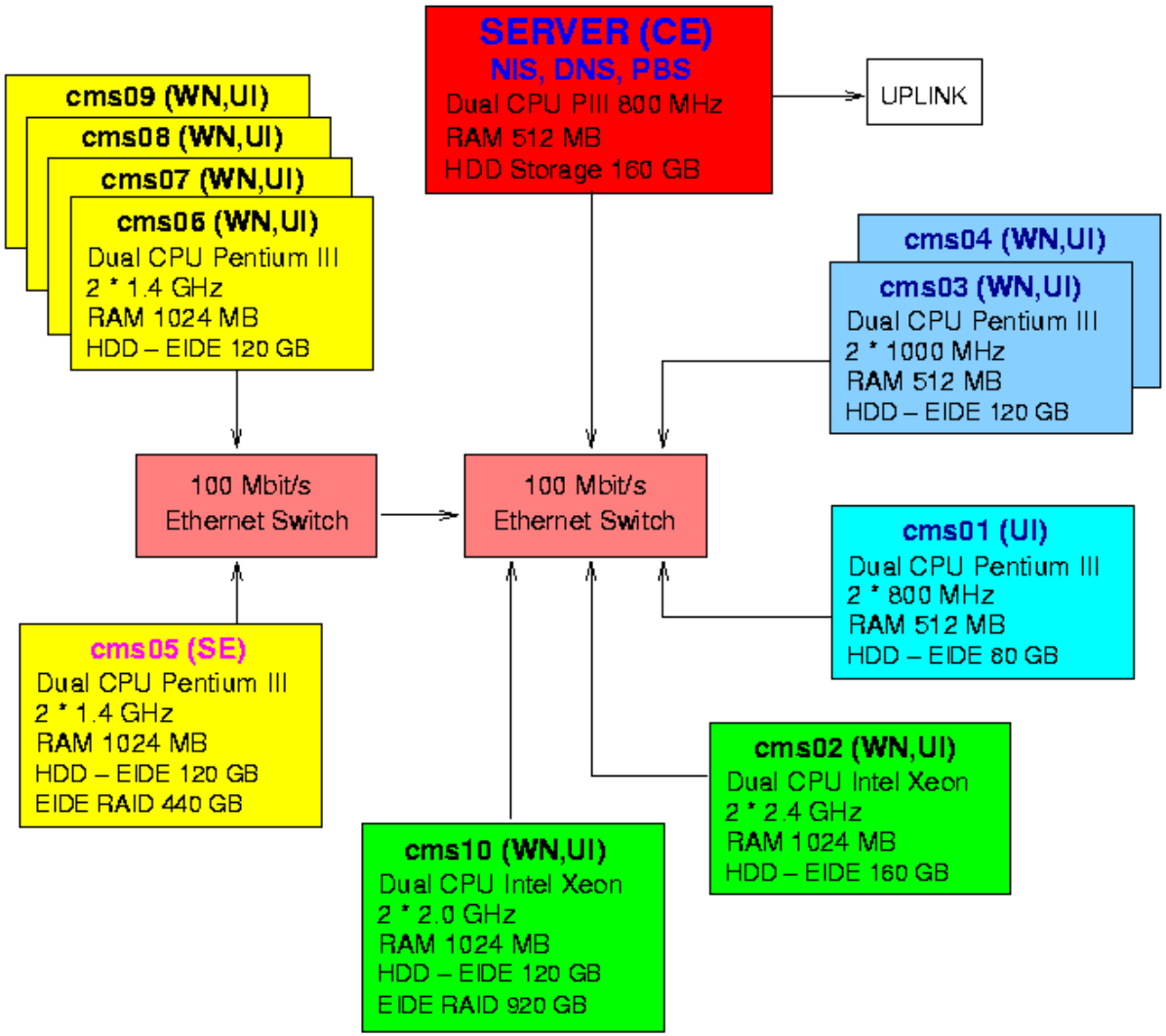
CERN and CMS Software:

CERNLIB (including PYTHIA and GEANT) , ROOT, GEANT 4, CMKIN, CMSIM and OSCAR, ORCA etc.

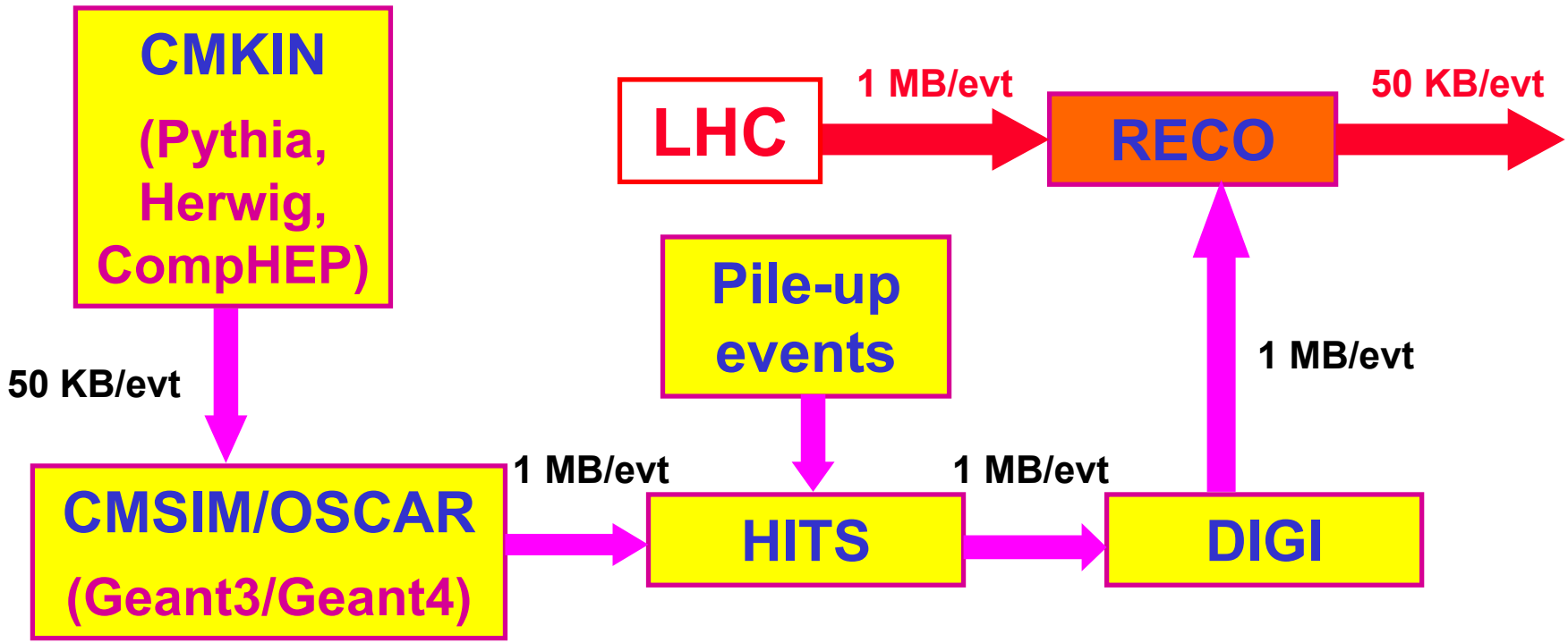
KCC is a part of the MOSCOW Distributed RC

KCC resources are allocated to CMS jobs exclusively

NSC KIPT CMS Linux Cluster
(current status)



Complete cycle of the CMS event generation/reconstruction



McRunjob – complex set of scripts on *Python, C, Perl*

Moreover: *BOSS, MySQL, …*

Results of KCC participation in CMS Monte-Carlo event production over 2002–2005

K I P T

Data Set	Assignment ID	Number of events generated	Data type
Hi02_photon100_110	1829	1000	CMSIM
Hi02_qcd80_90	1644	5000	CMKIN, CMSIM
Hi02_qcd100_110	1645	1000	CMKIN, CMSIM
eg02_BigJets	1958, 1959, 1960, 1961	85250	CMKIN+CMSIM, HITS, DIGI, RECO
Bt03_qcd120_170tth	2245	80000	CMKIN
Mu03_tt2mu_Mll1000	2549	500	CMKIN
eg03_Wenu_calibration	2840	99500	CMSIM
Jm03_Wjets_20_50	3024	100000	CMKIN
eg03_Zee	3047	100000	CMSIM
Tt_ch_140_tb20	5344, 5345, 5481	150000	CMKIN, CMSIM, HITS

Participation in production with LHC Grid ⇒

Providing communication channel with CERN > 10 Mbit/s (by 2007 ~ 100 Mbit/s)

CMS cluster of NSC KIPT by 2007 (plan): 100 Gflops CPU, 30 TB HDD

What is our specificity?

- Small PC-farm (KCC)
- Small scientific group of 4 physicists, combining their work with system administration
- CMS tasks orientation
- No commercial software installed
- Self-security providing
- Narrow bandwidth communication channel
- Limited traffic

K
I
P
T

Installation and Configuration of Grid Middleware

- *Ssh_up.sh*
- *Pbs_up.sh*
- *Grid_up_*.sh*
- *Hosts_up.sh*
- *Firewall_up.sh*

⇨

Welcome
to download and tuning!

http://www.kipt.kharkov.ua/~cms

Our requirements!

- Minimal acceptable but functional set of Grid Middleware
- Compatibility of Grid Middleware elements like a CE&WN and WN&UI
- Independence of Grid Middleware installation of the quality of the communication channel
- Balancing of traffic used by the network services to provide installation and functionality

First step (Installation)

- Download and install 2 RPMs on all nodes:
  - j2sdk-1_4_2_08-linux-i586.rpm
  - lcg-yaim-2.4.0-3.noarch.rpm
- Download all needed RPMs and create the hierarchy of directories like at CERN web – server. Create and configure our own web – server which provides catalog structure similar to that of CERN server.
- Correct the update lists and switch off auto-update.
- The yaim packages selected for the nodes installation:
  - *CE:* lcg-CE-torque (CE_torque)
  - *SE:* lcg-SECLASSIC (classic_SE)
  - *WN:* lcg-WN-torque (WN_torque)
  - *UI:* lcg-UI (UI)
- CE installed on the gateway machine; firewall provides masquerading for the other nodes of KCC. So, we only need 2 real IP addresses for CE and SE to provide Grid Cluster. Moreover, CE is located in the DMZ of KIPT firewall. All this provide the acceptable level of security
- Running script that uses yaim functions to realize the plan of installation

Second step (Configuration)

- Home directories of the pool-account users are provided by sharing of home directories from CE through NFS and NIS
- Torque needs configuration (next slide)
- Work of the batch system supposes using of scp- commands. So, the passwordless authentication is needed. We use hostbased authentication to provide the execution of the batch job on KCC (over next slide) .

Third step (Optimization)

- Make a symbolic link of the /etc/grid-security/certificates in the NFS shared directory (on CE). Then make a symbolic link of this directory to the /etc/grid-security/certificates on the all other nodes. Even for our small PC-farm this saves ~1GB monthly.
- Keep the working ntpd (time server) only on CE. Other nodes periodically ask them time using ntpdate (in cron). Its provides a good stratum (~2-3) on all nodes.
- Look at other hints on our web page.

Torque configuration

- Use script which configures Torque automatically (to be run on all cluster nodes)
- Make sure that hostbased ssh access is configured properly (next slide)
- Using of transit query (feed) makes batch system more stable and automatically solves problems with CPUT consuming jobs.
  - ➔create queue feed
  - ➔set queue feed queue_type = Route
  - ➔set queue feed route_destinations = medium
  - ➔set queue feed route_destinations += veryshort
  - ➔…
- Setting the additional properties to the nodes (file nodes) on Torque server provides further flexibility of the system
  - ➔cms01 np=2 cluster prod kiptcms cms01
  - ➔…

OpenSSH configuration

- Use script which automatically configures sshd (sshd_config and ssh_config) for hostbased authentication (start it as a root on all nodes)
  - sshd_config
    - → …
    - → HostKey /etc/ssh/ssh_host_key
    - → HostDSAKey /etc/ssh/ssh_host_dsa_key
    - → HostbasedAuthentication yes
    - → …
  - ssh_config
    - → …
    - → HostbasedAuthentication yes
    - → EnableSSHKeysign yes
    - → PreferredAuthentications hostbased,publickey,password
    - → …
- Script collects the open keys from nodes included to the file shosts.equiv (and hosts.equiv) and forms proper files (ssh_known_hosts and ssh_known_hosts2). Then restarts service.

Summary

- An enormous data flow expected in the LHC experiments forces the HEP community to resort to the Grid technology
- The KCC is a specialized PC farm constructed at the NSC KIPT for computer simulations within the CMS physics program and preparation to the CMS data analysis
- Further development of the KCC is planned with considerable increase of its capacities and deeper integration into the LHC Grid (LCG) structures
- Configuration of the LCG middleware can be troublesome (especially at small farms with poor internet connection), since this software is neither universal nor “complete”, and one has to resort to special tips
- Scripts are developed that facilitate the installation procedure at a small PC farm with a narrow internet bandwidth