PNPI Status report
HEPIX 2010 spring

Alexander Lodkin
Anatoly Oreshkin
Tatiana Serebrova
Andrey Y Shevel - *reporter*

Quick overview

- PNPI structure
- HEPD CSD responsibility
    - Security, cluster status
- Consideration about small cluster future at PNPI HEPD
- Plan for small cluster upgrade at PNPI HEPD

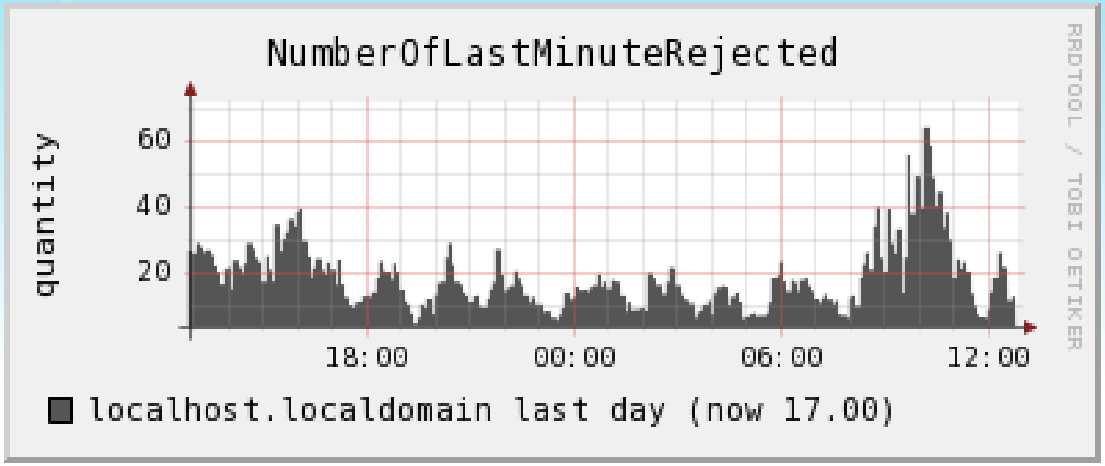
PNPI (www.pnpi.spb.ru) consists of ...

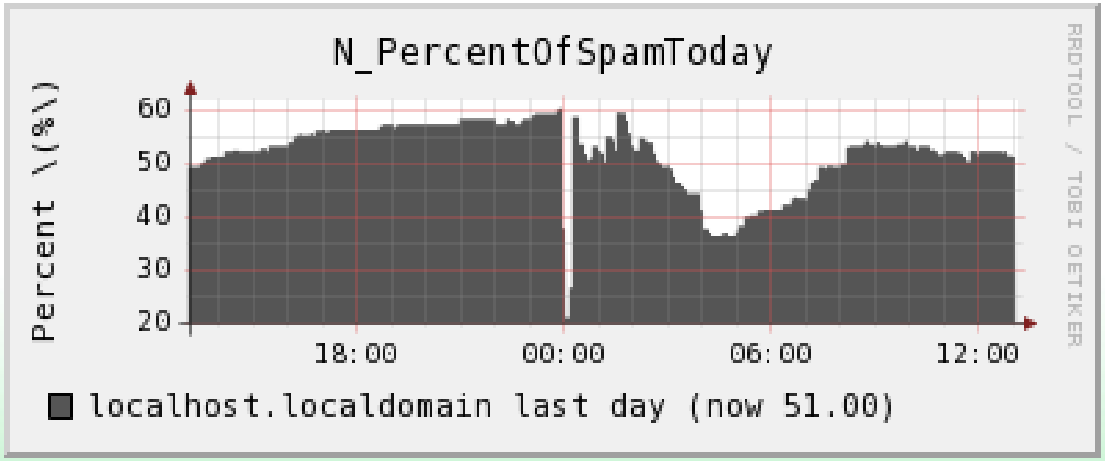
- High Energy Physics Division (**HEPD**)
  http://hepd.pnpi.spb.ru
  - Computing Systems Department (**CSD**)
    http://hepd.pnpi.spb.ru/CSD
- Neutron Physics Department
- Biology and Molecular Physics Division
- Theory Division

HEPD CSD responsibility...

- Institute (PNPI) mail server
  - SL4x, ~1K accounts, anti{**virus**,**spam**}, several scrips to reduce **spam** volume;
- HEPD http server and http/socks proxy
- HEPD Small computing cluster
- Computer network (~400 hosts located/distributed over 6 buildings)
- HEPD WiFi network
- HEPD Video room

Typical spam rejection view





Examples of rejection substrings (~100):

adsl
cable
kabel
pool
broadband
…

- rejected ~23K-25K per day
- ~75% of rejected relay hosts have no DNS records

SPAM distribution on top domains in March 2010

16.4%	ru
15.7%	com
15.4%	net
7.6%	pl
3.7%	cz
3.2%	fr
2.7%	de
2.2%	it
2.2%	gr
1.8%	ro
1.7%	lt
1.4%	ua
1.4%	nl
1.3%	pt

In total ~150 domains

Centralized servers features

- Centralized (mail, http, proxy) **servers features**:
  - Each local disk on all centralized servers is implemented as program RAID-1;
  - Each centralized server (mail, http, proxy) is connected to separate UPS;
  - The small computing cluster is equipped by two UPSes;
- Last year acomplished upgrades
  - video-conference room was equipped with **Tandberg EDGE 95 MXP** video/network hardware + other appliances.
  - WiFi network has been deployed for **HEPD**.

HEPD small computing cluster

- **Small computing cluster** (6 physical nodes:Xeon i5520/2.4/1333/EM64/Quard and older, 18 virtual nodes, host OS SL5.3/**Xen**)
  - batch system = **SGE**;
  - home made set of scripts to keep cluster up;
  - **NFS**, **NIS**, autofs, **AFS**.
  - **SL4x** and **SL3x** as guest OSes
  - 4 TB disk space, home made backup system
  - **around 140 registered users**, 5-15 every day
  - Quite old batch cluster (started in 1998 with **CODINE**, passed through many upgrades).
- **Do we need for such small cluster in the future? Let us take a look around.**

Another example of small cluster …

- In 2000 (!) I and Jerom Lauret have deployed the small cluster (30+machines, tape library, 6 TB disk space) for **Nuclear Chemistry Group** at **SUNY** (Stony Brook). Cluster runs until now(!). Physicists in the group are involved into **PHENIX** (presumably now into **ATLAS** as well). The cluster is working for ~20 users without regular maintenace (rare minor correction over Internet). Web site is in **cloud** http://sites.google.com/site/ramdata2009.
- Dark sides of small cluster without regular maintenance are obvious.
- Well, **what is the future of small clusters in small physics groups, laboratories** ?
    - Roles/Aims.
    - Size.

Cluster's roles/aims in small physics group/laboratory

- Main aim is to use for
  - Development of new algorithms/programs;
  - Analysis of small portion of the data (~ 200 TB) not only for LHC;
  - Also for small laboratory the cluster might be served as pool of spare machines in case of emegency.

Which is good cluster size for small laboratory?

- About dozen+ physicists who involved into real data analysis (runs jobs, got new analysis results)
- it has to be taken into account contemporary tendencies:
  - cloud computing technology (it leads to understanding that cheapest computing is possible on huge computing installations like **google**, **azure**, **amazon**, may be **CERN**, **Tier 1s**, etc);
  - grouth of computing power per unit (server);
  - understanding that with growth of a number of servers in cluster we got less computing power per watt;
- All above reasons helped us to recognize that **small cluster** (~12-24 nodes) is best solution (*it is not expensive, easy to reconfigure to fit the concrete task needs, easy to maintain, easy to use as gateway to large or huge computing facility*)

Plans for nearest cluster future …

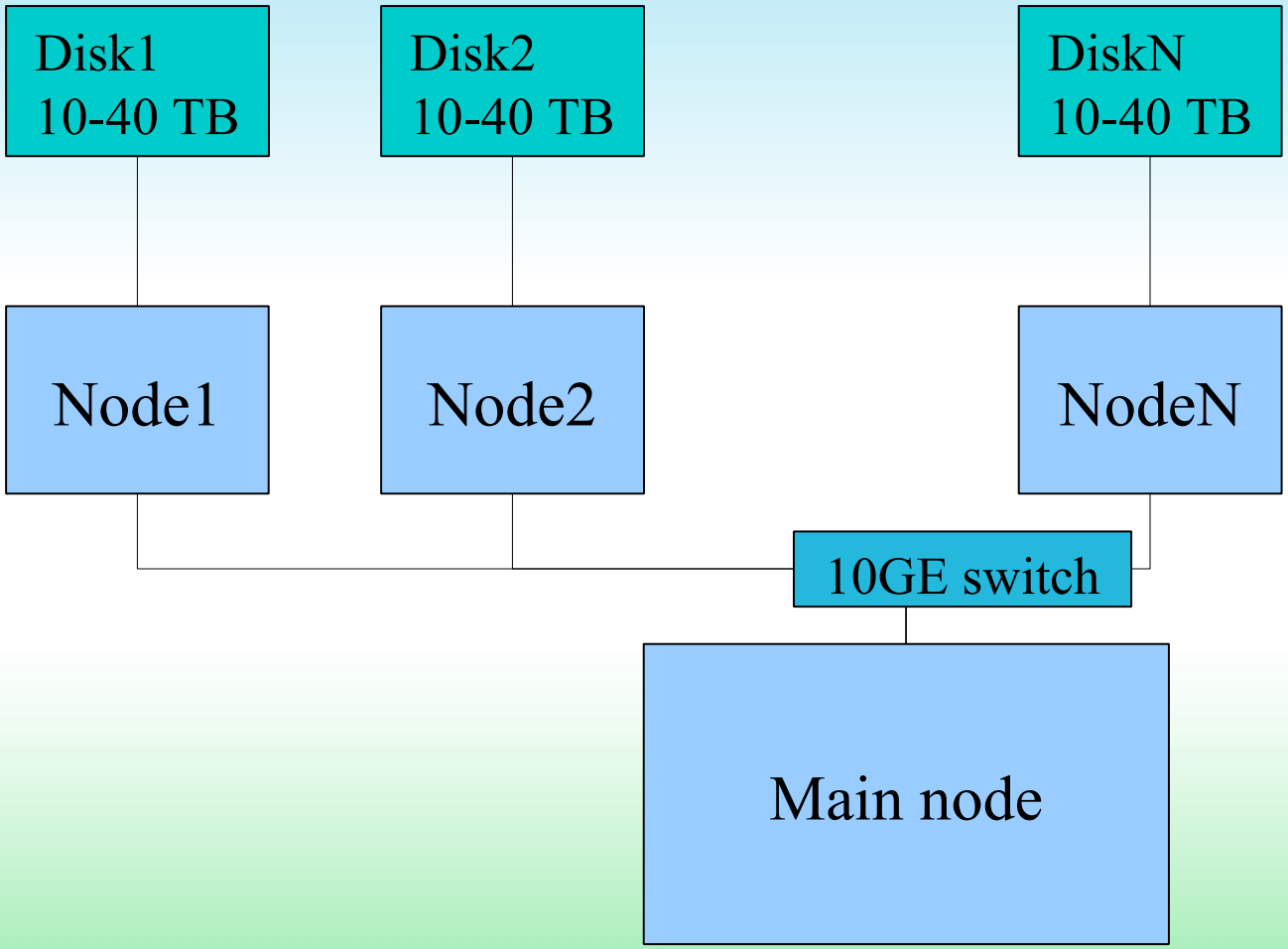
- **Convert the cluster into agile private cloud** with taking into account:
  - minimize a manual interventions to the cluster during operation;
  - make possible to implement **cloud** cluster **Tier3;**

– hardware plan
  - ~200 TB of disk space in total
  - total number of physical nodes ~20 units

- **Main aim** is to use for:
  - develop of algorithms for physics analysis on data from **LHC** and related needs;
  - debug new analysis programs;
  - analyse of local data;
  - use as gateway to large clusters (outside **HEPD**).

Planned cluster scheme

In total ~200 TB of disk space



JBOD

e.g. SAS 9200-16e HBA

Program RAID6

Quick tech overview of the cluster requirements …

- each cluster node has to have locally connected significant disk storage (10-40 TB); update the drives every five years (if we need for ~200 TB, it means ~40 TB has to be bought each year; with $300/TB we have to spend ~$12K per year;
- **KVM** over **IP** (hardware);
- update each node every three years; if planned ~12-24 nodes means we need to buy 4-8 nodes each year; with $3K/node we have $12K-$24K per year
- In total for hardware we have to spend ~**$24K-$36K** each year.
- virtualization; **SGE** as batch system in **cloud cluster**; simple cluster system (**cfengine**?) for host cluster; **CERNVM** on **cloud cluster** (also see «Contextualization in Practice:The Clemson Experience» - http://indicoprev.cern.ch/contributionDisplay.py?contribId=100&sessionId=10&confId=59397);

Thank you ! Questions?