CONVERGING LAB AND CLASSROOM

MULTIPURPOSE WORKSTATION CLUSTER: GOALS AND DESIGN

2024, skoltech.ru

This presentation available at: http://ftp.sleepgate.ru/sci/

GOALS

R&D activities
1. To provide interactive SMP workstations for proprietary scientific and engineering software.
2. To provide interactive SMP workstations for open source scientific and engineering software.
3. To provide interactive workstations for scientific software development.
4. To provide interactive workstations for scientific visualization, pre- and post-processing.
5. To be user-friendly for operators.

Educational activities
1. To provide interactive platform for supervised learning.
2. To provide means for interactive scientific and engineering computation, visualization, pre- and post-processing.
3. To be user-friendly both for supervisor and students.

Infrastructure objectives
1. Data safety.
2. Easy of administration (to be user-friendly for sysadmins).

WHAT IS “TO BE USER-FRIENDLY” MEAN?

For operators, learning supervisors and students

Familiar desktop experience.

I. Windows explorer:
   1. Pre-Windows 7 (should be considered obsolete).
   2. Windows 7 (still used in production).
   3. Windows 10 and onwards.

II. X window managers and desktop environments:
   1. Gnome 3 and onwards, KDE 4 and onwards.
   2. Lightweight: MATE (Gnome 2), Enlightenment (pre-E17, E17 and onwards), lxde, lxqt, xfce, etc.
   3. Dozens of other window managers (twm, FVWM, mwm, WindowMaker, etc.)

Familiar software, either proprietary or open source.

For system administrators

Easy of control and routine task automation.

I. Consistent init system across all infrastructure (to systemd or not to systemd?)
II. Conventional infrastructure services (ISC DHCP, ISC BIND, Samba CIFS, Samba DCs, PXE+TFTP, FTP, NFS, etc.)
III. Centralized control.

THE ONLY WAY: OU-LEVEL CLOUDS

As for now, there are many way to utilize the advantages of virtualization at the level of OU.

All workstations are serve as virtualization hosts and connected to the backbone network.

I. Pros:
   1. Reduced administration efforts at OU level.
   2. No need for extra hardware.

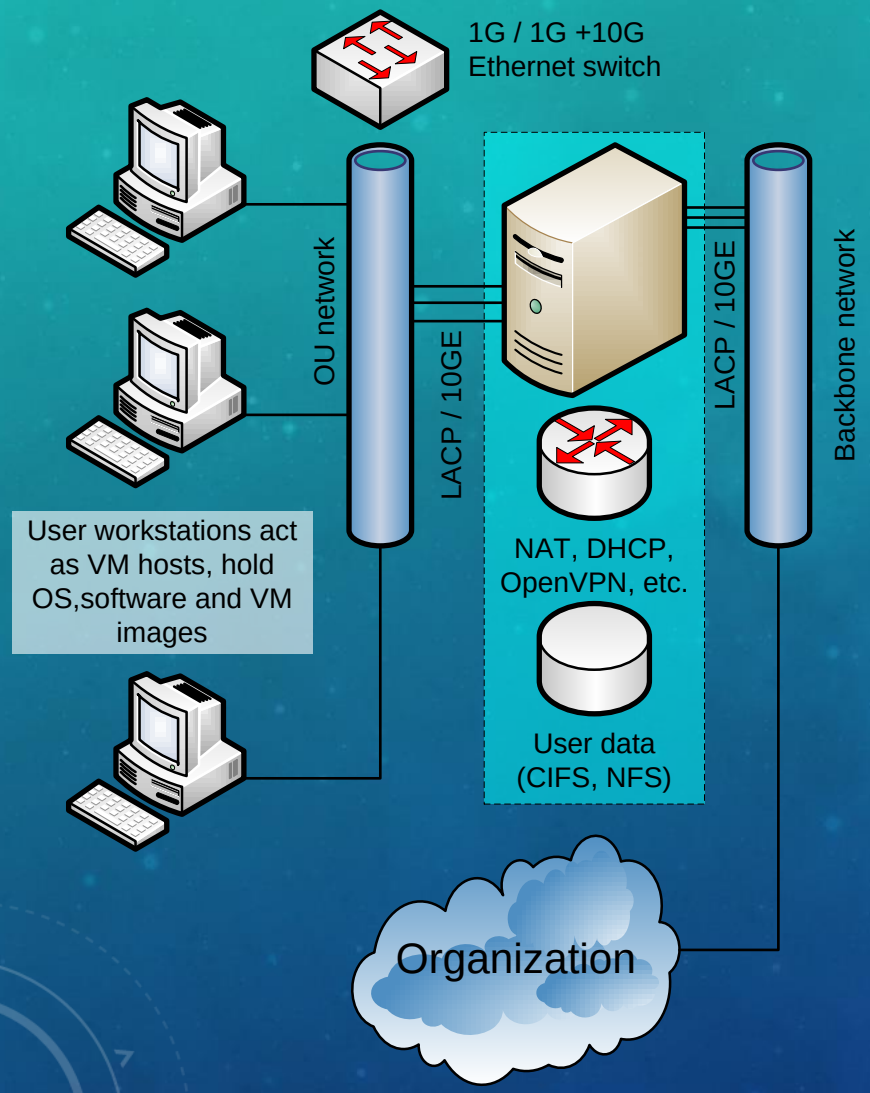
II. Cons (mostly due to lack of flexibility):
   1. Limited administration abilities at OU-level (no OU-level centralized storage, backups and updates).
   2. Limited network flexibility at OU-level.

Only OU-level server for iSCSI/CIFS/NFS (user data, VM) and infrastructure services is connected to the backbone.

I. Pros (mostly due to great flexibility):
   1. OU-level server may act as NAT and isolate OU network from the backbone (we also reduce a burden for organization-level network administrators).
   2. Centralizing control and backup.
   3. Many infrastructure services can operate on OU server, which opens the possibility for implementation of different virtualization schemas for user workstations; if powerful enough, OU server can act as VM host.

II. Cons:
   1. Careful planning must precede implementation (here we are!).
   2. Extra hardware and administration needed (see above)

SCHEMA 1: OU SERVER FOR NAT AND USER DATA



OU server isolates OU from backbone and provides centralized storage for user data

Pros:

I. NAT:
   1. Isolates OU from up-level network, eliminates possible IP conflicts with backbone network.
   2. Makes it possible to run OU-level AD-like authentication services (preferably Samba).

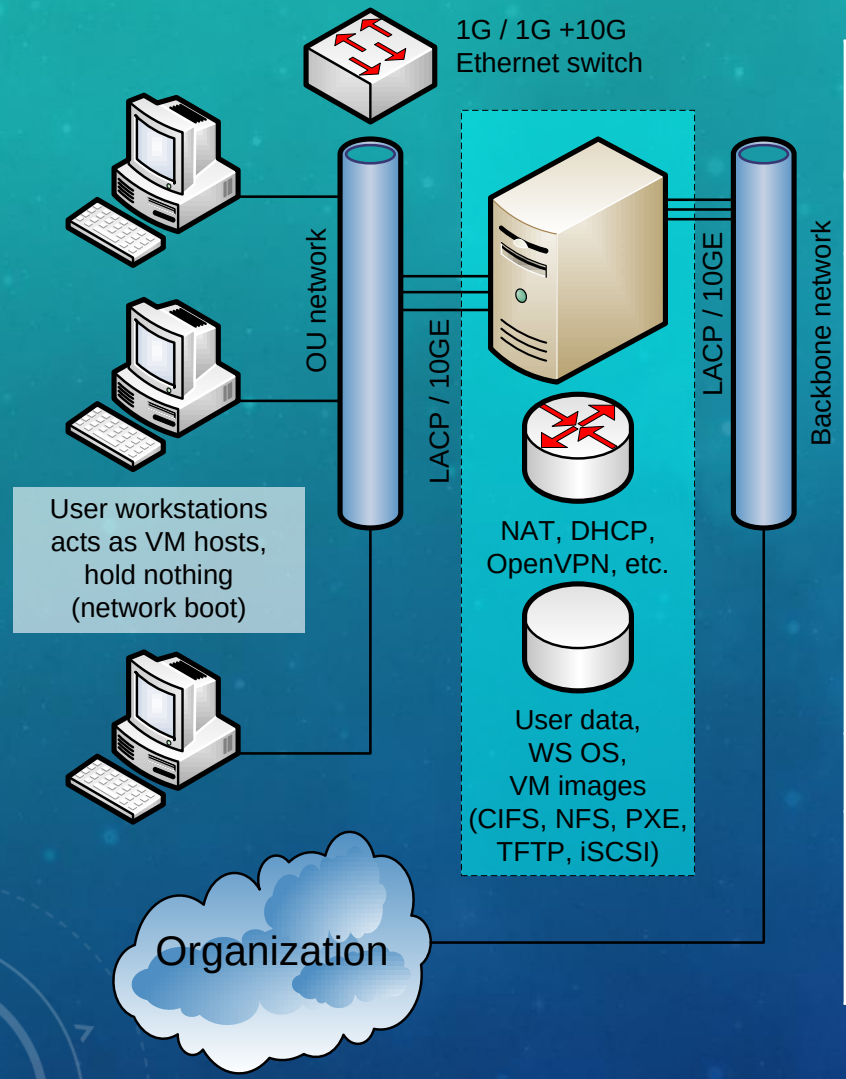
II. OU-centralized storage:
   1. Allows to implement complex software and hardware data storage technologies (ZFS, hardware RAID, hardware SAN) and data backups (including incremental ones).

III. Every workstation holds all except long-term user data.
   1. Fast boot and software load.
   2. Reduced traffic between workstations and OU server.

Cons:

I. Hard to implement centralized software backup and update.
II. Workstations have to be powerful enough.

SCHEMA 2: SCHEMA 1, WS OS AND VM IMAGES



OU server not only isolates OU from backbone, but holds all software and data

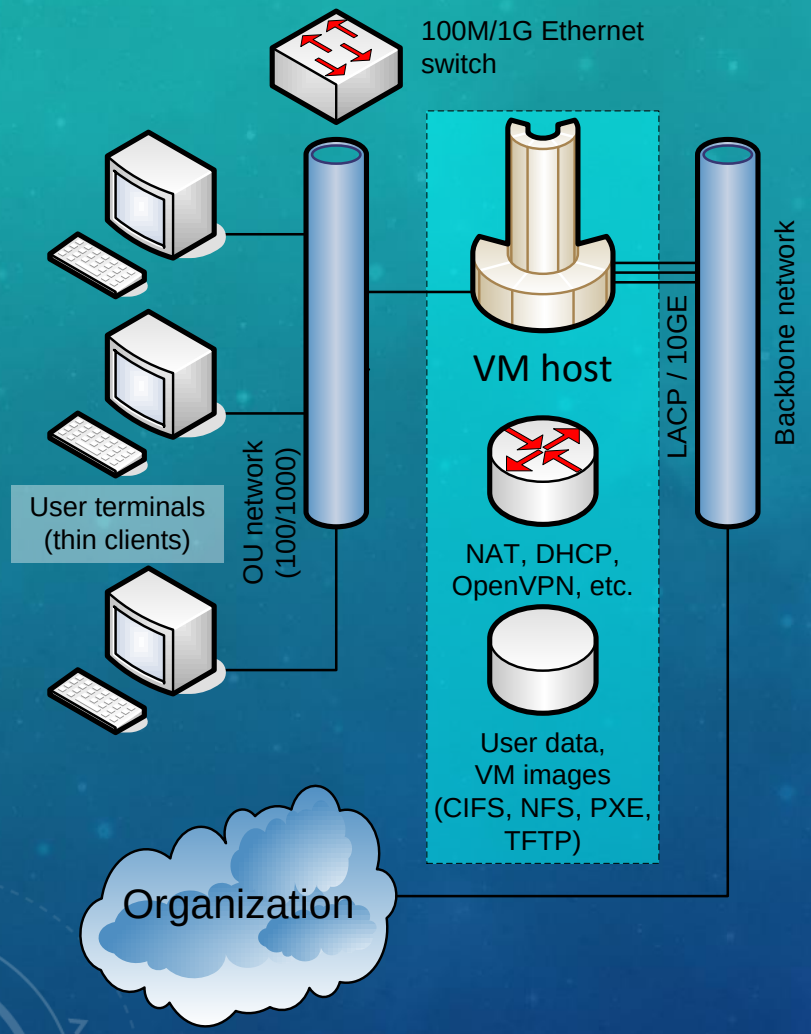
Pros:

I. NAT: just the same features as for Schema 1.
II. OU-centralized storage:
    1. Also allows to implement complex data storage technologies.
    2. With carefully planned maintenance schedule this schema allows centralized backups not only user data, but also user operating systems and software (including incremental backups).
III. Workstations can be diskless.

Cons:

I. Heavy traffic between workstations and OU server.
II. OS boot and software load may be quite slow.
III. Workstations have to be powerful enough.

SCHEMA 3: ALL ARE CENTRALIZED



OU server is responsible for everything

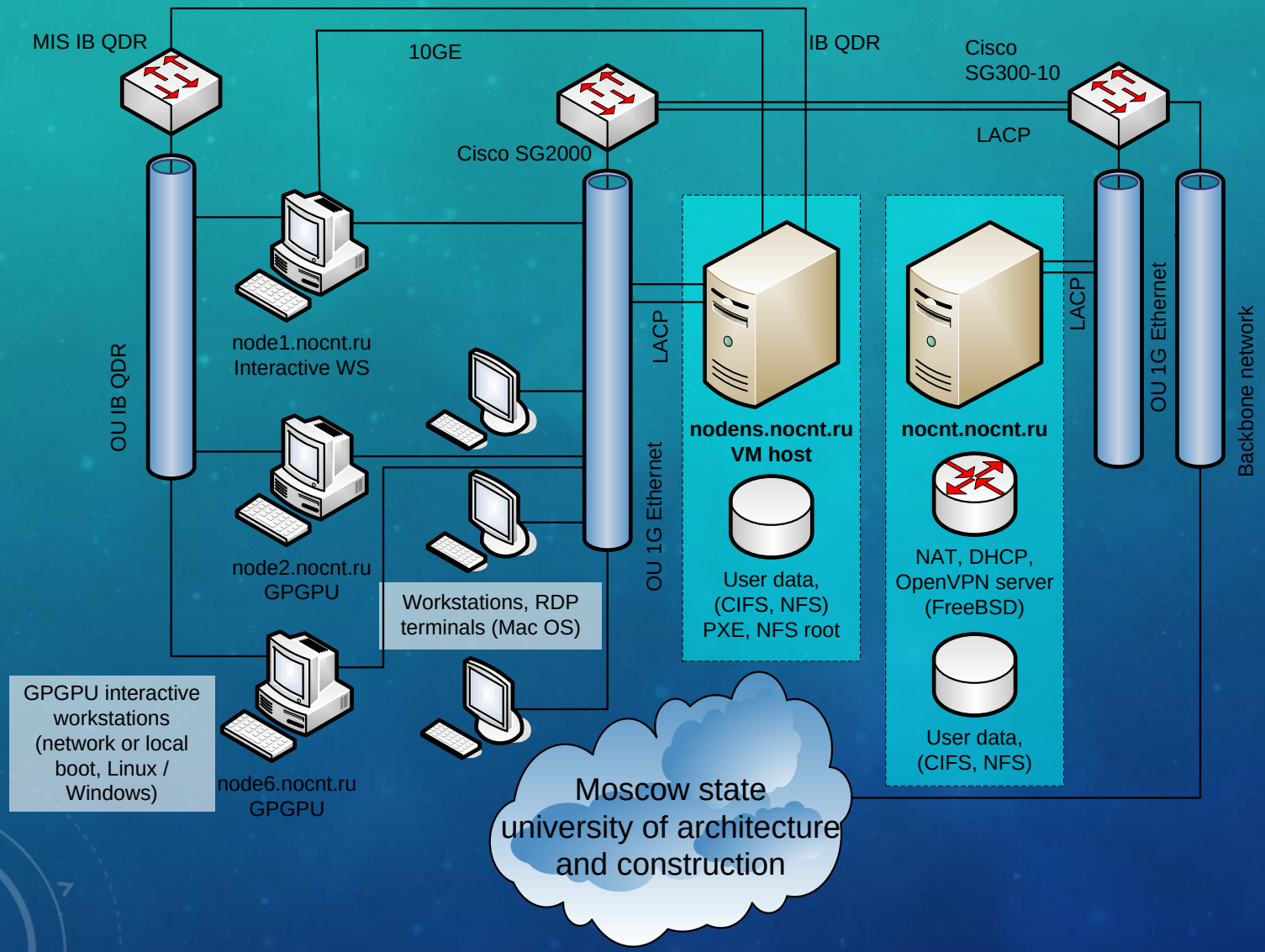
Pros:

I. Can act as NAT, as in Schemas 1 and 2.
II. Fully centralized.
III. Workstations can be diskless and lightweight (e.g., ARM-based).
IV. Moderate network traffic in OU network.
V. Zero time for OS boot.
VI. Fast software boot and data access (all data exchange takes place inside central hub).

Cons:

I. Central hub must be very powerful:
   1. Should consider using cluster with 100 or more cores.
   2. Heavy memory requirements.
   3. Should consider using internal fast cluster interconnect with low latency, e.g. InfiniBand.

EXAMPLE: NOCNT.RU OU NETWORK



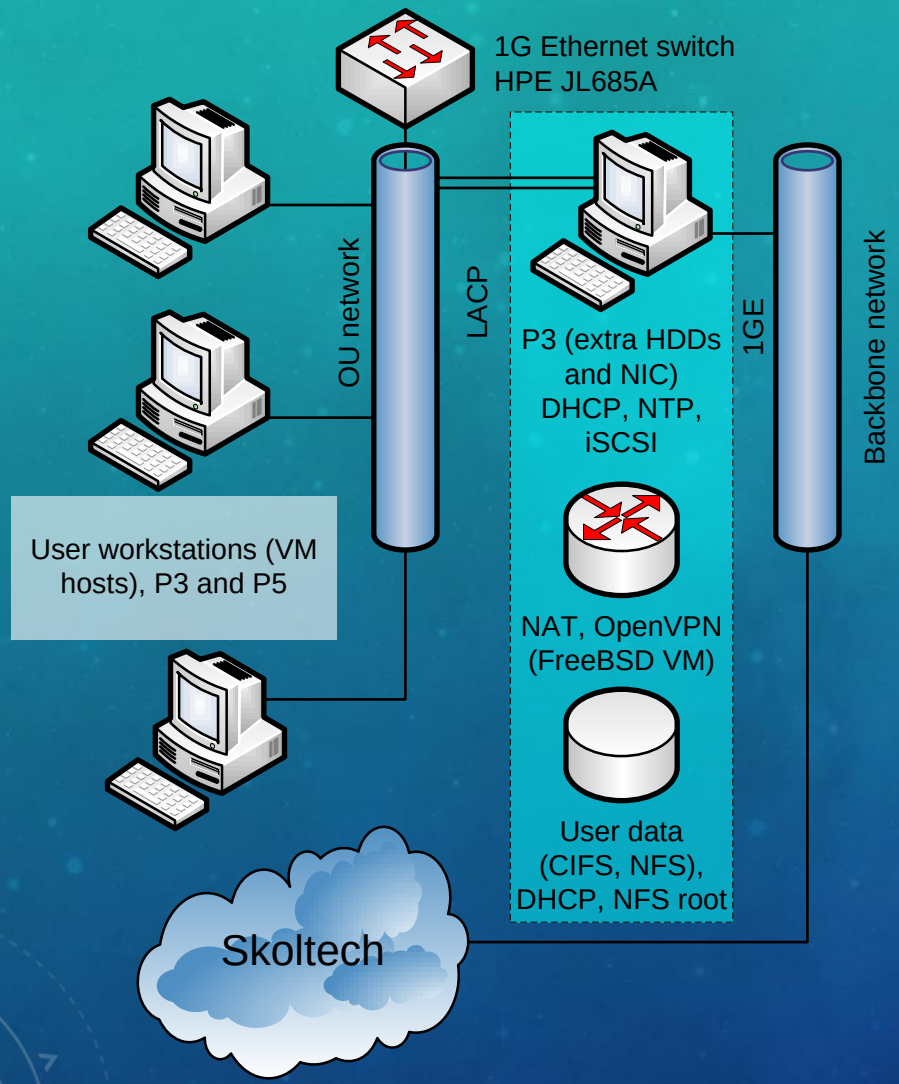
Hybrid 1G and 10G Ethernet, QDR IB, WS and compute nodes

Flexible:

I. Can be used as terminal station for MGSU primary cluster.
II. Can be used as 6-node HPC cluster with GPU capable of FP64.
III. Can be used during classes.
IV. First node can effectively be used for interactive work.
V. Hybrid boot: either from OU network (NFS root for all nodes is on nodens.nocnt.ru) or from local HDDs on nodes.

Operates since 2013!

PROPOSITION: BASED ON PURCHASED HARDWARE



Selected schema: schema 1 with optional network boot

One P3 can be used as NAT and non-volatile storage server.

I. This P3 should operate 24/7.
II. The host OS preferably be RHEL-like, e.g. AlmaLinux.
III. ipfw-based NAT should be started in FreeBSD VM.
IV. Extra hardware should be bought:
    1. 2-port network interface card, preferably Intel chip (e.g. Intel PRO 1000 PT), about 20 th. Ru.
    2. Extra HDDs: (4+3)x6TB, preferably HGST or WD, about 140 th. Ru.
    3. Second-level gigabit smart switch, e.g. HPE JL685A, about 80 th. Ru.

OS for Windows VM.

This must be further discussed.