

## INFCARE8 PROJECT - BACK-OFFICE SYSTEM

The application of the robust back-office solution is regarded to be essential in case of integrated systems. However, we include it in a separate development section in this project proposal because a qualitative and quantitative technological development can be experienced and can be expended in near future in case of medical information science systems. For example the spread of and the output data quantity of imaging systems means not only a quantitative problem but also requires a different qualitative management since the simplest archiving, retrieval and search functions have changed both regarding their type and performance demand. The application of GRID technology based systems is expected to appear in near future both regarding calculation and storage. On the other hand, the logical and legal management of medical data is under change by the creation of network-like organizational systems (for instance the principle of „store it where it is originally created” can no longer be kept), which has not become evident in Hungary during the development of IKIR system that ends now. Thus the establishment and modelling of such back-office system concept is essential, which shall be flexible even above the today stipulated and in near future expected demands, and this gives the real challenge.

### PLANNING AND OPTIMIZING HPC CONFIGURATION BASED UPON PROBLEM BASED THINKING

Medical attendance is more and more shifting towards the transparent, accountable, technology-based approach. An increasing proportion of the chronic ill children's parents require special services that adapt to their requirements, lifestyle and life quality. In our INFCARE8 R&D project we elaborate and realize the system of the integrated mechatronic and IT environment at prototype level.

This system consists of the following subsystems:

- 24-hour video supervision system, in a portable design, with simple local installation and an automatic data storing and alarm system (including taking of a full-value recording of the covered patient),
- AAL technology-based intelligent bed modules for three functions (controlled medicine storage and medication; patient-nurse-physician communication; supporting of learning-culture-entertainment),
- Development of mobile diagnostic head unit ( the wireless EEG "helmet"),
- Integrated middleware framework to realise the critical messaging,
- HPC based robust and reliable back-office architecture

System requirements based on problem based thinking:

- Possibility of converge different information source like CT, MRI and EEG
  - Large memory can help
- High speed data coming (40-50 MB/m) for a long time (around the clock)
  - Large common high speed storage
  - High availability (no single point of failure)

HPC architecture based on problem based thinking:

- Large fat node cluster
- Common large cluster storage based on Infiniband
- High availability (no single point of failure) means everything should be duplicated
- Széchenyi University can't set-up a super computer alone
  - Cooperation with Hungarian Academy and other Universities
- We can setup three centers one mini in Győr (cc. 1TFLOP) a small at Budapest (cc. 5 TFLOP) and a larger in Szeged (cc. 14 TFLOP)
- All three configuration has different technology
  - Mini – C3000 blade Infiniband DDR, Intel 2 socket
  - Small – SL165 tray Infiniband QDR, AMD 2 socket
  - Large – C7000 blade Infiniband QDR, AMD 4 socket

### Optimization: Problem & Solution

**Problem:** Cooling in small configuration SL165 servers seems to be "tired"

**Solution:**  
HP Watercool rack optimization

12:54:04 START SLEEP (300)					
START: 12:59:40:00	WC01L2C2	103679	144	15	48 227.40
START: 13:03:37:00	WC01L2C2	103679	144	15	48 261.68
START: 13:08:08:00	WC01L2C2	103679	144	15	48 304.75
13:13:02 START SLEEP (360)					
START: 13:19:22:00	WC01L2C2	103679	144	15	48 201.19
START: 13:22:53:00	WC01L2C2	103679	144	15	48 254.79
START: 13:27:17:00	WC01L2C2	103679	144	15	48 297.07
13:32:02 START SLEEP (420)					
START: 13:39:24:00	WC01L2C2	103679	144	15	48 192.43
START: 13:42:46:00	WC01L2C2	103679	144	15	48 245.48
START: 13:47:01:00	WC01L2C2	103679	144	15	48 293.42
13:52:00 START SLEEP (480)					
START: 14:00:06:00	WC01L2C2	103679	144	15	48 190.08
START: 14:03:26:00	WC01L2C2	103679	144	15	48 238.08
START: 14:07:33:00	WC01L2C2	103679	144	15	48 284.87

**Problem:** High available boot in diskless environment

In diskless environment we use NFS as a central OS image store with CMU. The NFS server can be put in high available configuration. The compute nodes can be hang if NFS failover occur.

**Solution:** Boot sequence of RHEL should be change (initrd)

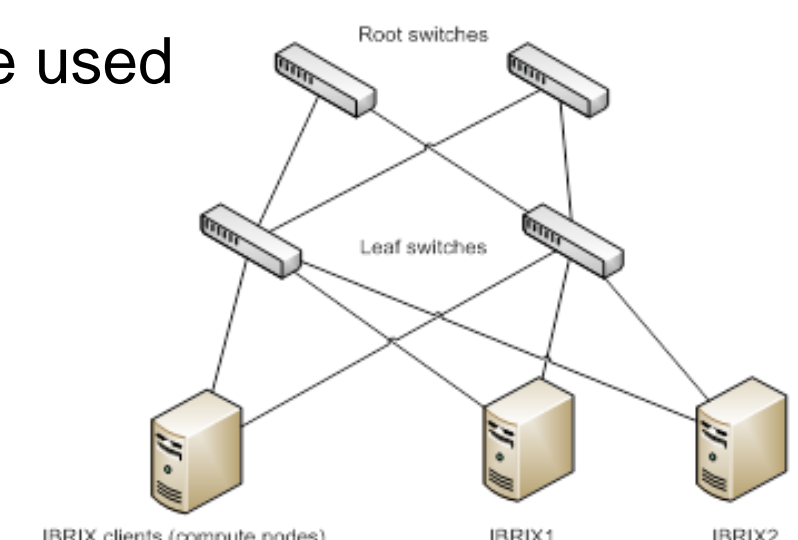
- First we have to put the Ethernet port under a bond interface
- We have to mount the NFS share over UDP instead of TCP (which is the RHEL5 default)
- Big lessons learned Redhat NEVER mention NFS over TCP is NOT state less ergo not fail safe!

**Problem:** High available cluster filesystem

We use HP X9320 storage solution (IBRIX) on top of Infiniband. The high availability on Infiniband has limitations, what we didn't know.

**Solution:** Just only one fat-tree topology can be used

- IBRIX using Linux bonding for HA
- IB bond driver support just mode 1
- Two separate IB network run well with MPI but not with IPoIB bond



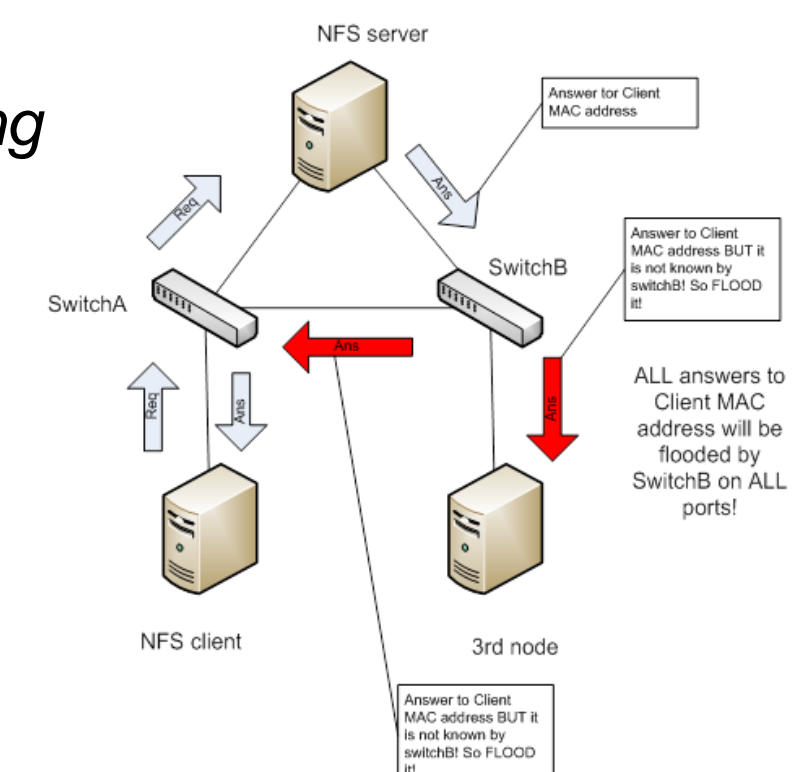
**Problem:** Unsolicited unknown unicast flood

HP cluster platform using Procurve switches, IBRIX solution using bond mode 6 on Ethernet. This situation can cause large amount of unknown unicast flood!

### Problem understanding

**Solution:**

Use 3com switches with IRF instead of Procurve or send a broadcast on every compute node in every 300 second



**Problem:** High failing rate of disk drives in HP X9320

After 6 month of operation we can imagine a very high failure rate of disk drives in cluster file system (cc. 60 drives left the RAID set from 192 in 6 month!) HP stated out 2TB midline SAS drives should not have more than 40% I/O load, but who switch of a 240 TB file system every day?

**Solution:** There is a problem fixing project now to change all MDL SAS drives to enterprise category one.