NIBR : Where does it hurt?

Steve Litster

Manager of Advanced Computing Group

OGF22, Date Feb. 27th 2008

NOVARTIS

Agenda

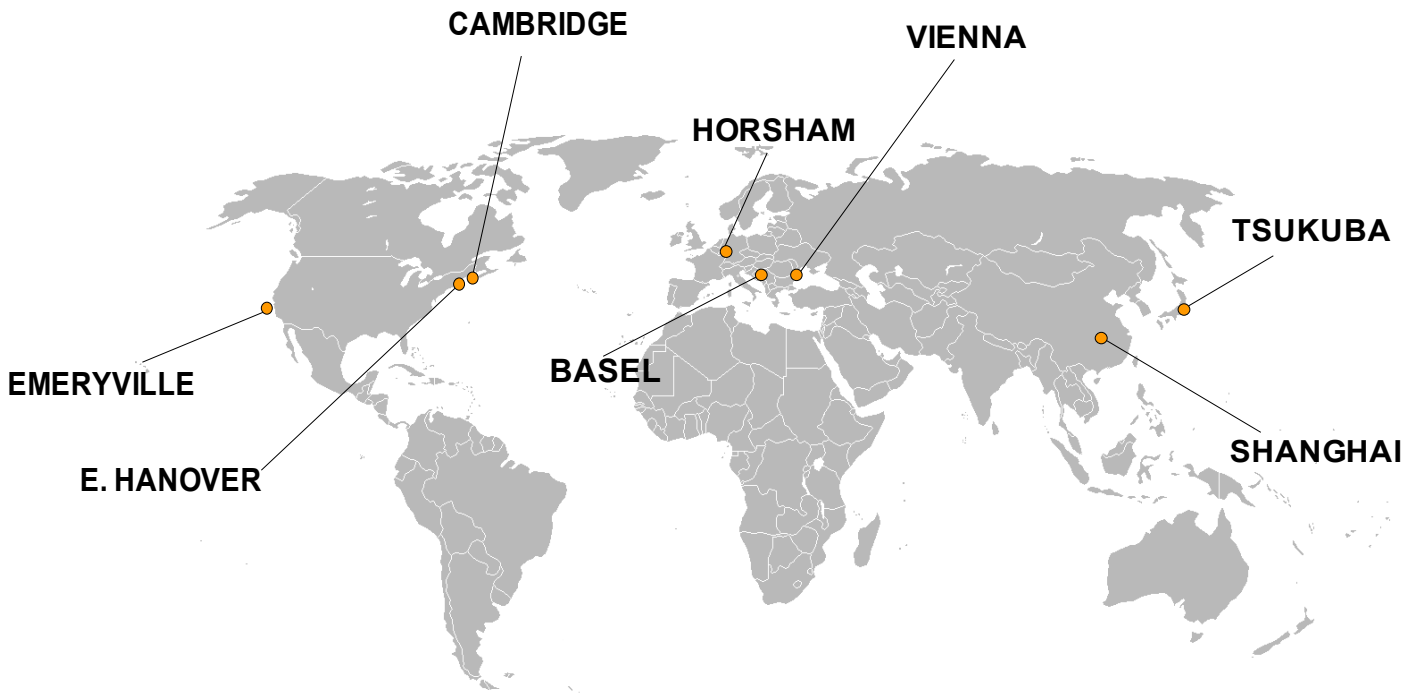
- Introduction to NIBR
- Build out of Cambridge Campus Grid
- Pain Points, Solutions and Lessons Learned

NIBR Organization

- Novartis Institutes for BioMedical Research (NIBR) is Novartis' global pharmaceutical research organization. Informed by clinical insights from our translational medicine team, we use modern science and technology to discover new medicines for patients worldwide.

- Approximately 5000 people worldwide

NIBR Locations

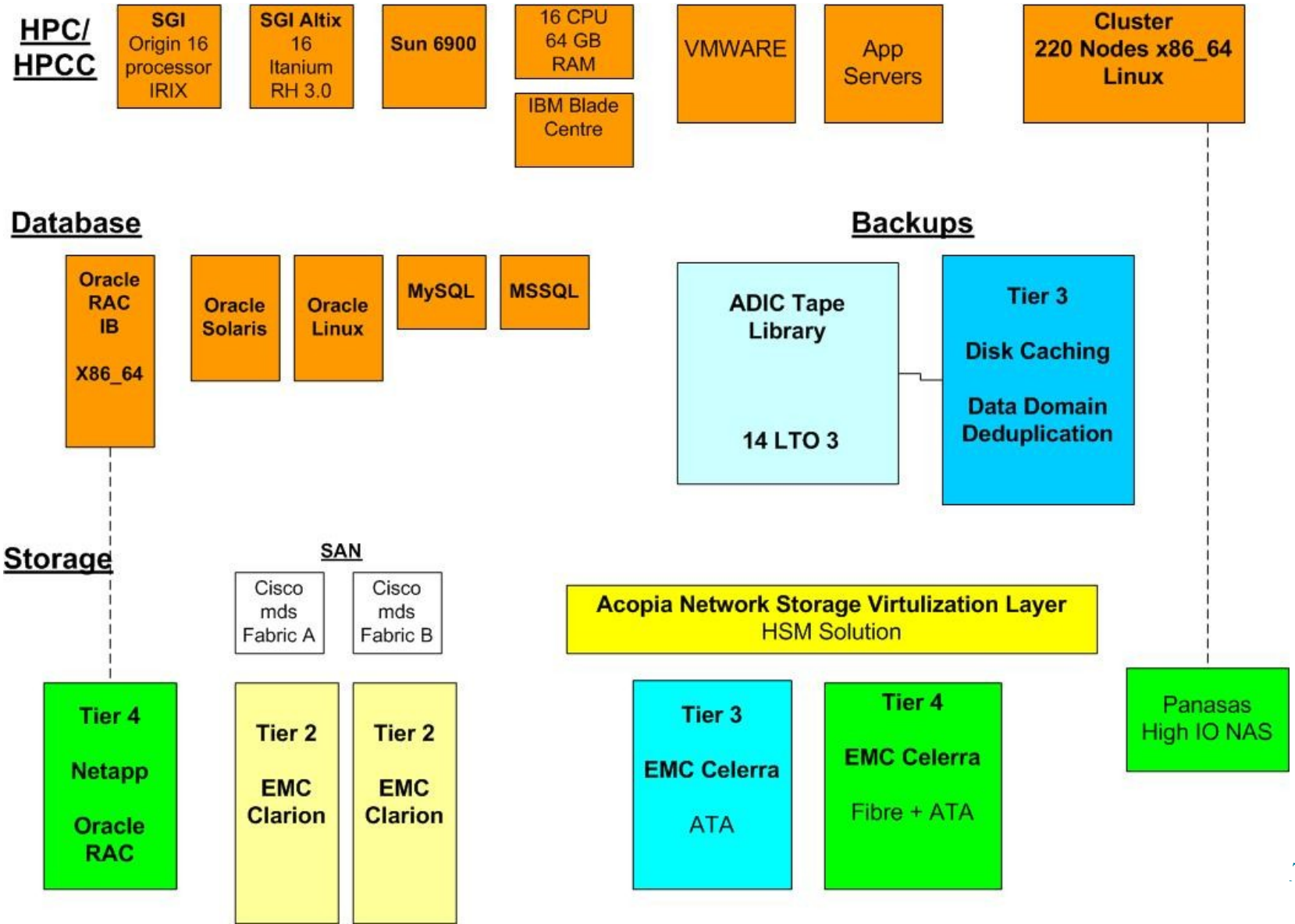


NOVARTIS

Cambridge Site “Campus Grid”

NOVARTIS

Cambridge Infrastructure



Cambridge Campus Grid: Highlights

- Compute Grid: 350 systems
  - 220 Node linux Cluster –housed in liquid cooled racks
  - SGE deployed on SMP, linux Cluster and Desktop systems
- Storage Grid: 200TB serving approx 800 Systems
  - Highly Scalable NAS and Virtualized storage environment.

Unified home, data and application directories

In the beginning: NIBR

- **2001-2004**
  - **Computing Environment**
    - Built for Structural Biology, Computer Aided Drug Design and Bioinformatics
    - **United Devices PC Grid** : 2700 Desktop CPU's across NIBR
      - Primarily used for Molecular docking : GOLD and DOCK
    - **SMP Systems**:
      - Multiple. Departmentally segregated. OS: IRIX, Solaris, Linux
    - **Linux Clusters:**
      - (3 in US, 2 in EU)-departmentally segregated
    - **Desktops Systems**:
      - Standalone SGI and Linux

In the beginning: NIBR cont.

- **Storage Environment**
  - Multiple NAS systems and SMP systems serving NFS based file systems
  - TB's of local storage
  - Storage: growth-100GB/Month
  - Multiple home and data directories

- **Data Centres**
  - Approaching Capacity
  - Multiple Server rooms

- **Systems Management**
  - Majority owned and operated by scientific groups
  - 80% of time in reactive mode

Stabilizing the Environment

- **Project:**
  - Design and build a completely new infrastructure while maintaining existing services.
    - Consolidate the multiple compute environments
    - Design and implement a centralized storage environment
    - Centralize data stores and harmonize the directory structures
    - Power and cool it effectively
  - Time Frame : 5 months
  - Resources: 2.5 FTE's
  - There was no Data Centre!

Storage and Backup

- Issues:
  - No data or metrics around growth or utilization
  - Highly unstable storage infrastructure
  - Backups were unreliable
  - We had no idea where the data or information was!
  - 40% time spent dealing with storage issues
  - Inconsistent data formats
  - Meta-data not being utilized
  - **Massive amounts of unstructured data**

Storage and Backup cont.

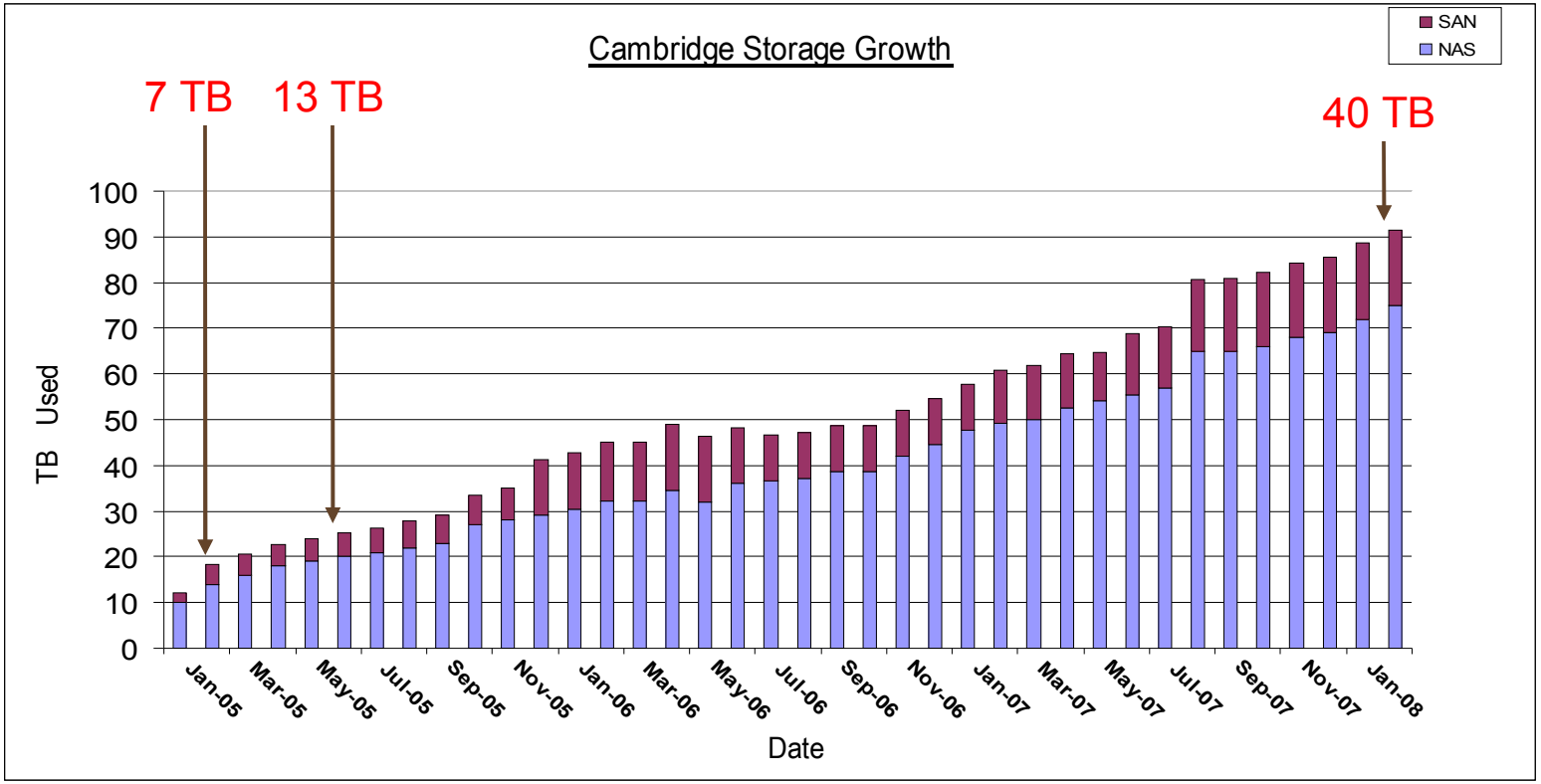
- Centralizing the Storage:
  - Implemented fully redundant SAN Environment
    - Removed all local storage arrays that were constantly failing
    - Provided flexible, highly available, high performance disk system
  - Implemented Highly Available, High Performance NAS Solution
    - Centralized Lab-Data (INBOX), Home, Data and Application directories
    - Consolidated data from numerous lab and local storage arrays
    - Implemented a consistent naming scheme e.g.. /usca/home
    - Monitor, predict and backup data

Storage and Backup cont.

- Original Design: A single NAS File System
  - **2004** 3TB (no problems, life was good)
  - **2005** 7TB (problems begin)
    - Backing up file systems >4TB problematic
    - Restoring data from 4TB file system-even more of a problem
      - Must have a “like device”, 2TB restore takes 17hrs
    - Time to think about NAS Virtualization.
  - **2006** 13 TB (major problems begin)
    - Technically past the limit supported by storage vendor
      - Journalled file systems do need fsck'ing sometimes

Storage Infrastructure cont.

- Storage (Top Priority)
  - Virtualization
  - Growth Rate 2TB/Month (expecting 4-5TB/Month)

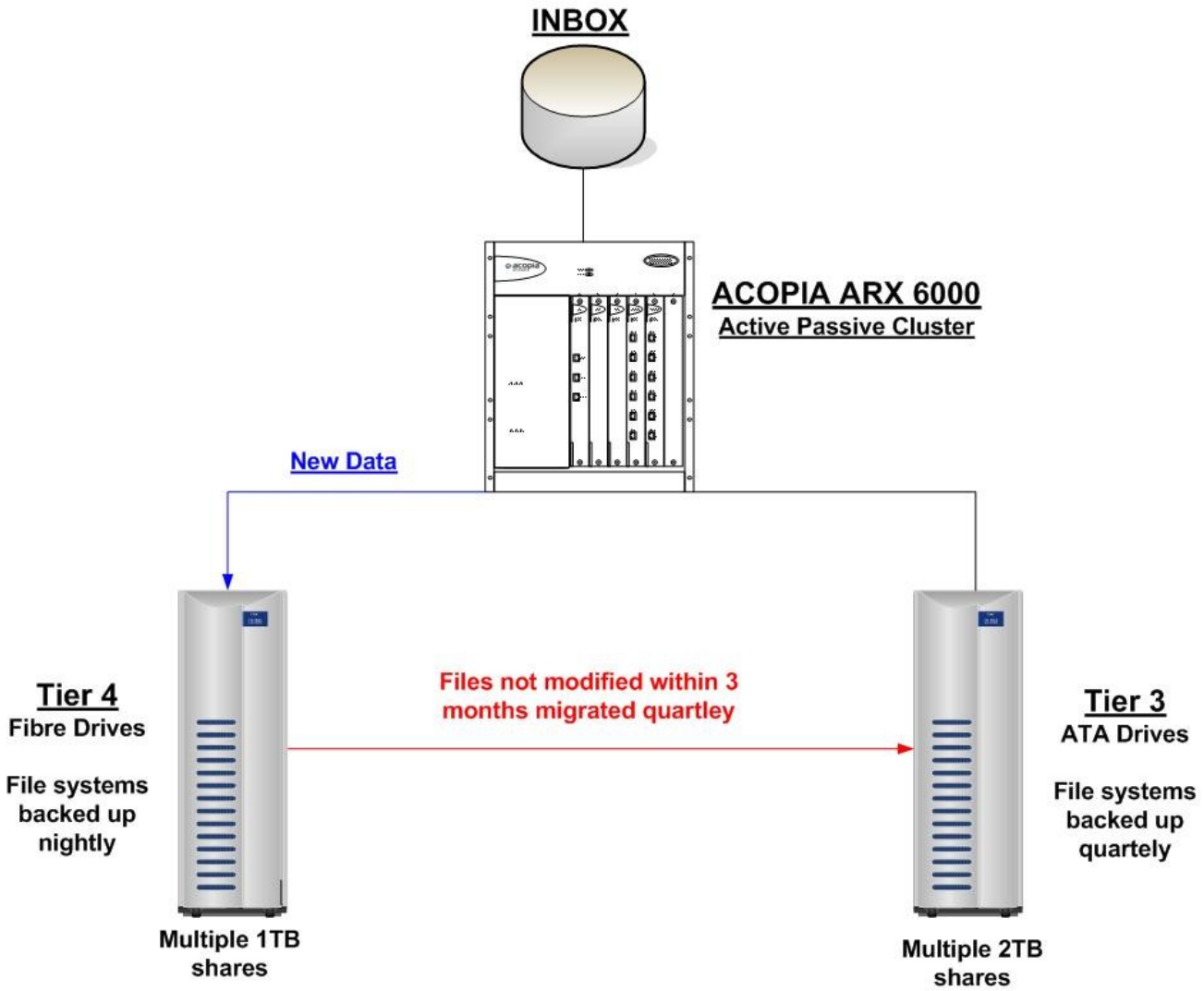


NOVARTIS

NAS Virtualization

- Requirements:
  - Multi-protocol file systems (NFS and CIFS)
  - No “stubbing” of files
  - No downtime to users due to storage expansion/migration
  - Throughput must be as good as existing solution
  - Flexible data management policies

NAS Virtualization Solution



NAS Virtualization Solution cont.

- Pro's
  - Huge reduction in backup resources (90-30 LTO3 tapes/week)
  - Less wear and tear on backup infrastructure (and operator)
  - Cost savings : Tier 4 = 3x, Tier 3 =x
  - Storage Lifecycle –NSX example
  - Much less downtime (99.96% uptime)

- Cons
  - Can be more difficult to restore data
  - Single point of failure (Acopia meta data)
  - Not built for high throughput (TBD)

Storage and Backup cont.

- Future requirements:
  - Storage
    - Global namespace (extend what to the other sites?)
    - Peta-byte scale storage solution
    - Must be capable of high I/O and bulk storage ?
  - Data Handling
    - Implement ILM
    - Drive information from the data.

Storage and Backup cont.

- Avoid having to create a technical solution?
  - Discuss data management policy before lab instruments are deployed
  - Start taking advantage of meta data from day 1.
- In the meantime:
  - Monitor and manage storage growth and backups closely
  - Scale backup infrastructure with storage
  - Scale Power and Cooling –new storage systems require >7KW/Rack.

Computing and Applications

- Issues
  - Multiple clusters –different OS's, job schedulers, hardware
  - SMP systems-75% CPU utilization serving NFS
  - Different OS's, and multiple application stacks
    - Modeling : linux and IRIX
    - Bioinformatics group: linux and Solaris 8
  - PC Grid-Utilization falling due to complexity of application on boarding
  - SMP systems operating as Prod, Dev. and Test on the same system
  - Multiple applications written in every language

Computing and Applications cont.

- Standardize Operating Systems and Queuing Systems
  - RH Linux and Sun Grid Engine
- Consolidate Clusters
  - Introduced 210 Node Linux Cluster
  - Departmentally segregated through SGE queues
- Consolidate SMP systems
  - Introduced scalable SMP systems e.g.. Sun 6900
- Introduction of VMWARE ESX clusters
  - Physical servers hosting 130 application servers
  - ESX 3:VI, HA/DRS

Computing and Applications cont.

- Consolidate Application space
  - Introduction of NAS based application repositories
    - Available across all systems (even worldwide e.g.. /usr/prog)
- Create the “Campus Grid”
  - Linux Desktops, cluster, SMP systems: Utilizing SGE
  - All windows and *NIX systems accessing common data stores

Computing and Applications cont.

- Future (highly dependent on open standards)
  - Tighter integration of the GRIDs (storage and compute) with:
    - workflow tools –Pipeline Pilot, Knime
    - home grown and ISV applications
    - windows based systems and applications
  - Standardize on one or two programming languages
  - Application Virtualization (windows environment “on boarding”)
  - Portals/SOA –SGE Integration
  - Global-Grid Computing? (currently no requirements)

Environmental Constraints

- Power, Cooling and Floor Space
- Issues: We didn't have enough of any of them
  - Costly renovations to server rooms
    - UPS Upgrades: 65kV to 180kV system
    - CRAC Unit Upgrade: 2x 20 ton units
    - Installed and upgraded Redundant Power circuits.
    - 5 server rooms no room for expansion.
    - Cluster rooms >110F in less than 15 minutes
- Time to plan a new data center

Original Server Room

Environmental Constraints

- Data Centre Project
  - Typically 2 years (financial approval through build-out)
  - NIBR DC: 8 months (renovation of existing space)
    - Lights-out facility (remote power, KVM, console access)
    - 15 minutes from campus
    - Space for 60 racks (currently occupying 20)
    - No raised floor
    - Liquid cooled racks for High Density equipment
- Future: Disaster recovery

New Data Centre

NOVARTIS

Acknowledgements

- Bob Cohen, OGF
- Novartis ACS Team:
  - Jason Calvert, Bob Coates, Mike Derby, James Dubreuil, Chris Harwell, Delvin Leacock, Bob Lopes, Mike Steeves and Mike Gannon.
- Novartis Management:
  - Gerold Furler, Ted Wilson and Remy Evard (CIO)

Additional Slides :NIBR Network

Note: E. Hanover is the North American hub, Basel is the European hub

All Data Center sites on the Global Novartis Internal Network have network connectivity capabilities to all other Data Center sites dependent on locally configured network routing devices and the presence of firewalls in the path.

Basel; Redundant OC3 links to BT OneNet cloud

Cambridge; Redundant dual OC 12 circuits to E. Hanover. Cambridge does not have a OneNet connection at this time. NITAS Cambridge also operates a network DMZ with 100Mbps/1Gbps internet circuit.

E. Hanover;

Redundant OC12s to USCA

OC3 to Basel

T3 into BT OneNet MPLS cloud with Virtual Circuits defined to Tsukuba and Vienna,

Emeryville; 20 Mbps OneNet MPLS connection

Horsham; Redundant E3 circuits into BT OneNet cloud with Virtual Circuits to E. Hanover & Basel

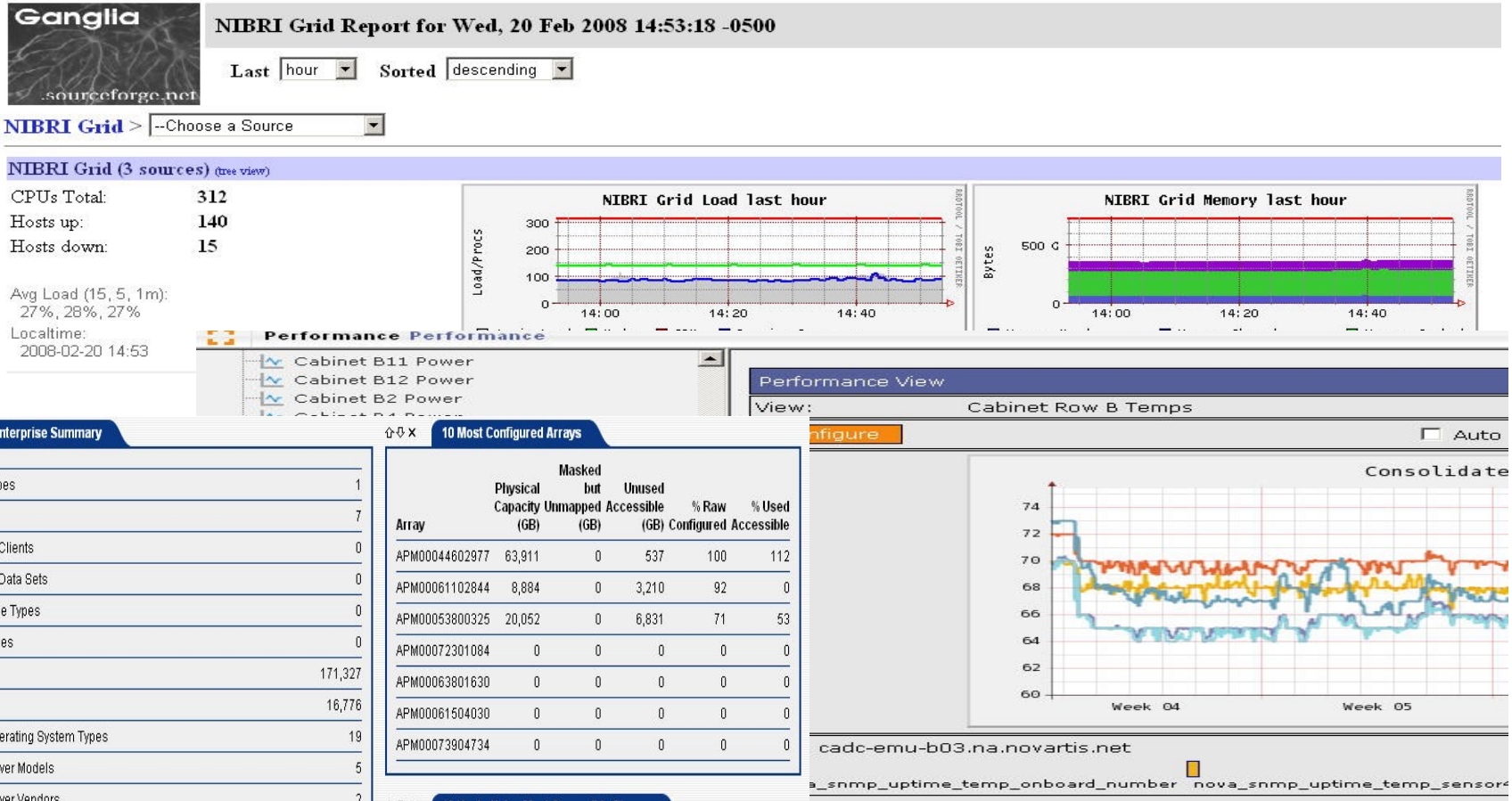
Shanghai; 10Mbps BT OneNet cloud connection, backup WAN to Pharma Beijing

Tsukuba; Via Tokyo, redundant T3 links into BT OneNet cloud with Virtual Circuits defined to E. Hanover & Vienna

Vienna; Redundant dual E3 circuits into the BT OneNet MPLS cloud with Virtual Circuits to E. Hanover, Tsukuba & Horsham

Cambridge Infrastructure-Monitoring and reporting

- Systems Monitoring and Alerting



NOVARTIS