Managing the CERN LHC Tier0/Tier1 centre

Status and Plans

July 8th 2003

Tony.Cass@CERN.ch

The Problem

Summary of Computing Capacity Required for all LHC Experiments in 2007

source: CERN/LHCC/2001-004 - Report of the LHC Computing Review - 20 February 2001

(ATLAS with 270Hz trigger)

	CERN			Regional	Grand
	Tier 0	Tier 1	Total	Centres	Total
Processing (K SI95)	1,727	832	2,559	4,974	7,533
Disk (PB)	1.2	1.2	2.4	8.7	11.1
Magnetic tape (PB)	16.3	1.2	17.6	20.3	37.9

~6,000 PCs

Another ~1,000 boxes

.f. ~1,500 PCs and ~200 disk servers
t CERN today.

But! Affected by:

- Ramp up profile
- System lifetime
- I/O Performance
- …

Uncertainty factor: 2x

Issues

- **Hardware Management**
  - Where are my boxes? and what are they?
- **Hardware Failure**
  - #boxes × MTBF + Manual Intervention = Problem!
- **Software Consistency**
  - Operating system and managed components
  - Experiment software
- **State Management**
  - Evolve configuration with high level directives, not low level actions.
- **Maintain service despite failures**
  - or, at least, avoid dropping catastrophically below expected service level.

Hardware Management

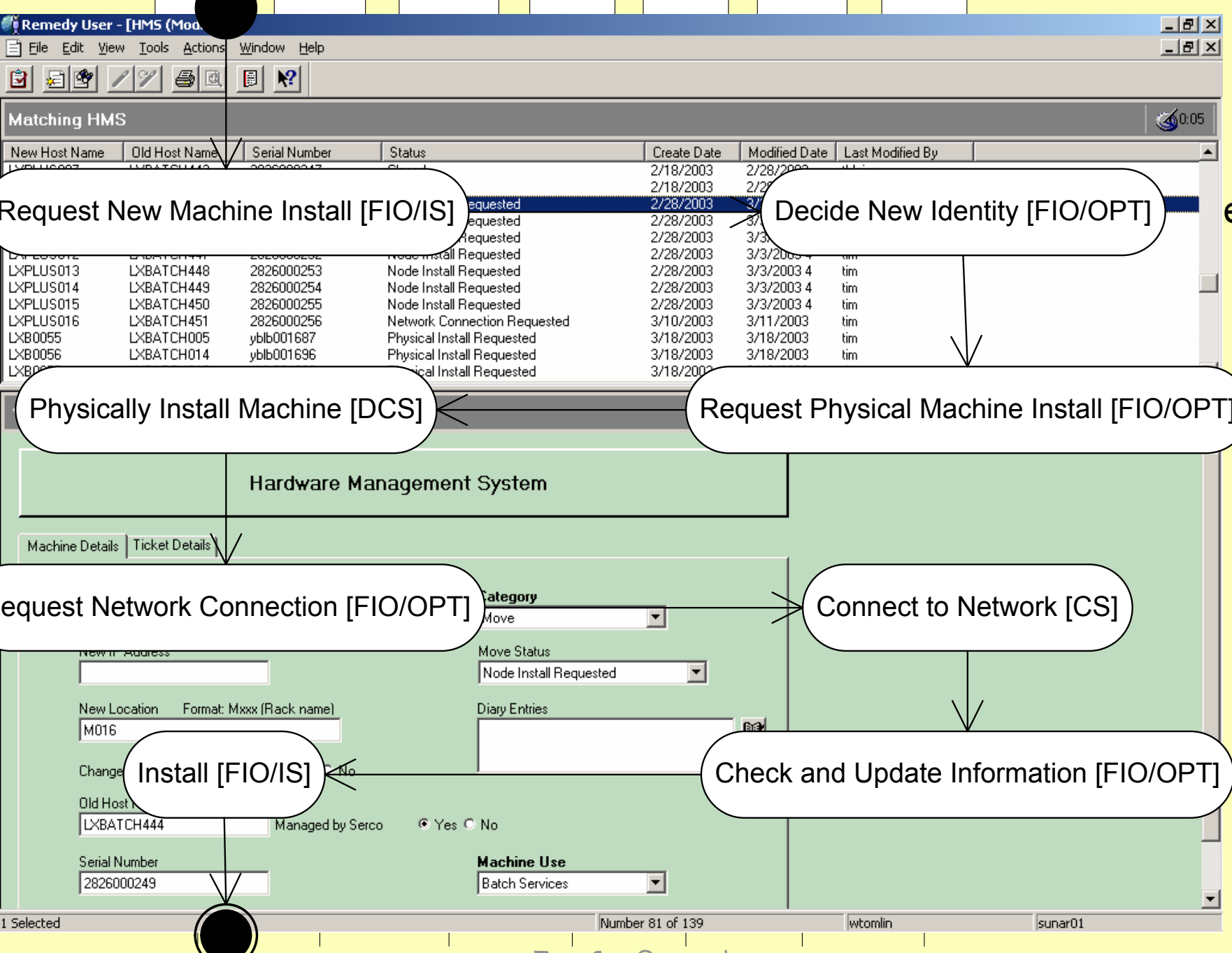
- We are **not used to handling boxes on this scale**.
  - Essential **databases** were designed in the '90s for **handling a few systems at a time**.
    - » 2FTE-weeks to enter 450 systems!
  - Chain of **people involved**
    - » prepare racks, prepare allocations, physical install, logical install
    - » and **people make mistakes**…

Connection Management

oxes and cables must all be in the correct place or our network
anagement system complains about the MAC/IP address association. One
r two errors not unlikely if 400 systems are installed. Correct? Or
orrect database?
Or buy pre-racked systems with single 10Gb/s uplink. But CERN doesn't
ave the money for these at present…)

Hardware Management

- We are not used to handling boxes on this scale.
  - Essential databases were designed in the '90s for handling a few systems at a time.
    - » 2FTE-weeks to enter 450 systems!
  - Chain of people involved
    - » prepare racks, prepare allocations, physical install, logical install
    - » and people make mistakes…
- Developing a **Hardware Management System** to track systems
  - 1st benefit has been to **understand what we do**!


Request New Machine Install [FIO/IS]
Decide New Identity [FIO/OPT]
Physically Install Machine [DCS]
Request Physical Machine Install [FIO/OPT]
equest Network Connection [FIO/OPT]
Connect to Network [CS]
Install [FIO/IS]
Check and Update Information [FIO/OPT]
Remedy User - [HMS (Mod
File Edit View Tools Actions Window Help
Matching HMS
0:05
New Host Name
Old Host Name
Serial Number
Status
Create Date
Modified Date
Last Modified By
LXPLUS013 LXBATCH448 2826000253 Node Install Requested 2/28/2003 3/3/2003 4 tim
LXPLUS014 LXBATCH449 2826000254 Node Install Requested 2/28/2003 3/3/2003 4 tim
LXPLUS015 LXBATCH450 2826000255 Node Install Requested 2/28/2003 3/3/2003 4 tim
LXPLUS016 LXBATCH451 2826000256 Network Connection Requested 3/10/2003 3/11/2003 tim
LXB0055 LXBATCH005 yblb001687 Physical Install Requested 3/18/2003 3/18/2003 tim
LXB0056 LXBATCH014 yblb001696 Physical Install Requested 3/18/2003 3/18/2003 tim
Hardware Management System
Machine Details
Ticket Details
Move
Move Status
Node Install Requested
New Location
Format: Mxxx (Rack name)
M016
Diary Entries
LXBATCH444
Managed by Serco
Yes
No
Serial Number
2826000249
Machine Use
Batch Services
1 Selected
Number 81 of 139
wtomlin
sunar01

Hardware Management

- We are not used to handling boxes on this scale.
  - Essential databases were designed in the '90s for handling a few systems at a time.
    - » 2FTE-weeks to enter 450 systems!
  - Chain of people involved
    - » prepare racks, prepare allocations, physical install, logical install
    - » and people make mistakes…
- Developing **Hardware Management System** to track systems
  - 1st benefit has been to **understand what we do**!
  - **Being used to track systems** as we migrate to our new machine room.
  - Would now like **SOAP interfaces to all databases**.

Hardware Failure

- **MTBF is high**, but **so is the box count**.
  - 2400 disks @ **CERN today: 3.5×10^6 disk-hours/week**
    - » **1 disk failure per week**
- Worse, these **problems need human intervention**.
- Another role for the **Hardware Management System**
  - **Manage list of systems needing local intervention**.
    - » Expect this to be prime shift activity only; maintain list overnigh and present for action in the morning.
  - **Track systems scheduled for vendor repair**
    - » Ensure vendors meet contractual obligations for intervention
    - » Feedback subsequent system changes (e.g. new disk, new MAC address) into configuration databases.

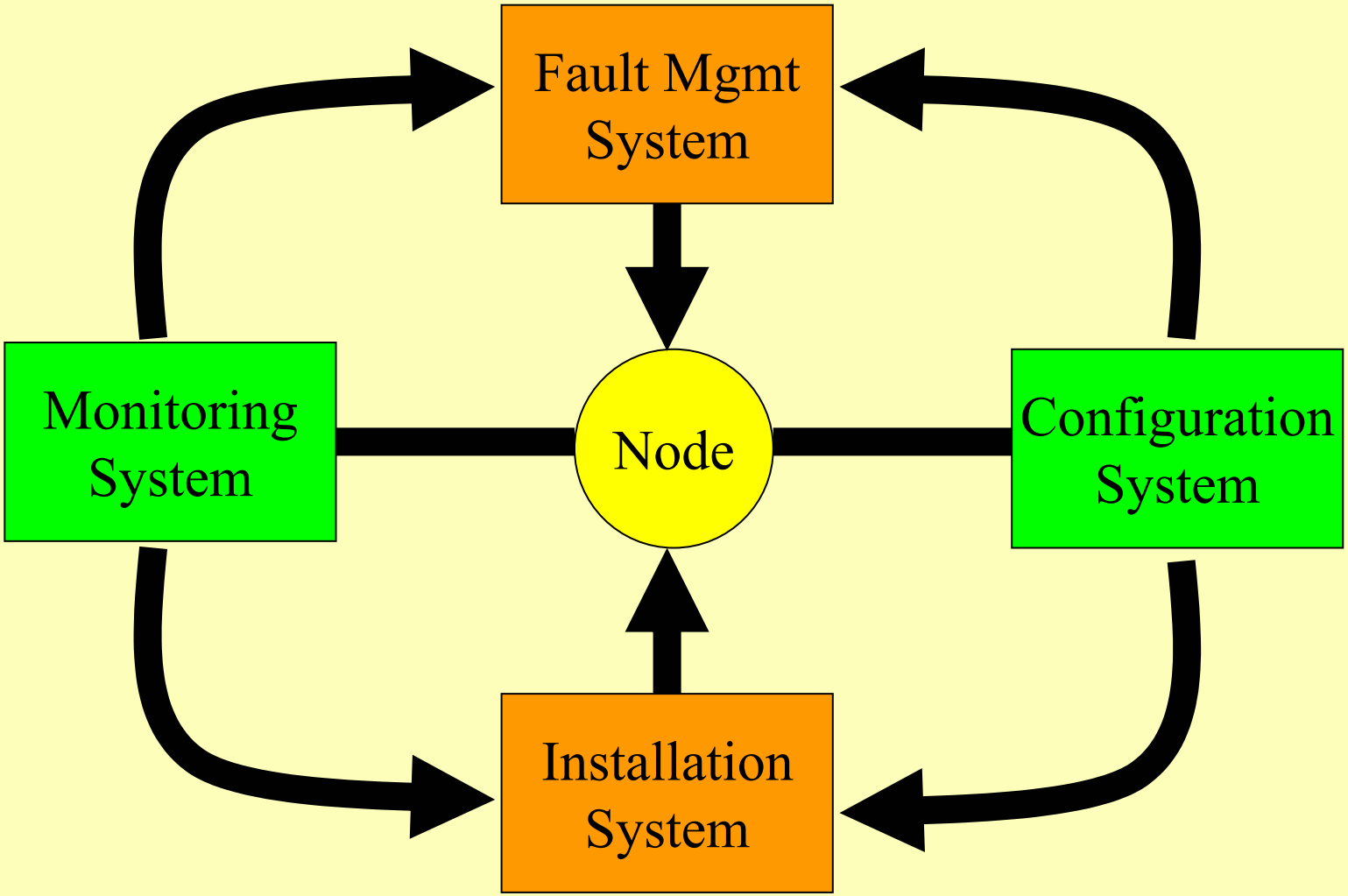
Software Consistency - Installation

- **System Installation requires knowledge of hardware configuration** and use
  - ☹ There will be **many different system configurations**
    - » different functions (CPU vs disk servers) and acquisition cycles
    - » Hardware drift over time (40 different cpu/memory/disk combinations today)
  - ☺ Fortunately, there are **major groupings**
    - » Batch of 350 systems bought last year; 450 more this year
    - » 800 production batch systems should have identical software
- Use a **Configuration Management Tool** that **allows definition of high level groupings**
  - EDG/WP4 CDB & SPM tools are being deployed now
    - » but much work still required to integrate all config information and software packages.

Software Consistency - Updates

- Large scale software updates pose **problems**
  - **Deployment Rapidity**
  - **Ensuring consistency** across all targets
- Deployment Rapidity
  - **EDG/WP4 SPM** tool enables **predeployment of software packages to local cache** with delayed activation.
    - » reduces peak bandwidth demand—good for networks and central software repository infrastructure.
- Consistency
  - Similar to installation, accurate configuration information needed.
  - But, also **need tight feedback** between **configuration database**, **monitoring tools** and **software installation system**.

Keeping nodes in order



State Management

- **Clusters are not static**
  - OS upgrades
  - reconfiguration for special assignments
    - » c.f. Higgs analysis for LEP
  - load dependent reconfiguration
    - » but best handled by common configuration!
- Today:
  - **Human identification** of nodes to be moved, **manual tracking** of nodes through required steps.
- Tomorrow:
  - **Give me 200, any 200. Make them like this. By then.**
  - A **State Management System**.
    - » Development starting now.
    - » Again, needs **tight coupling to monitoring & configuration systems**.

Grace under Pressure

- The pressure:
  - **Hardware failures**
  - **Software failure**
    - » 1 mirror failure per day
    - » 1% of CPU server nodes fail per day
  - **Infrastructure failure**
    - » e.g. AFS servers
- We need a **Fault Tolerance System**
  - Repair simple local failures
    - » and tell the monitoring system…
  - Recognise failures with wide impact and take action
    - » e.g. temporarily suspend job submission
  - **Complete system** would be **highly complex**, but we are **starting to address simple cases**.

Conclusions

- The **scale** of the Tier0/Tier1 centre **amplifies simple problems**.
  - Physical and logical installation
  - Maintaining operations
  - System interdependencies.
- Some **basic tools are now being deployed**, e.g.
  - A **Hardware Management System**
  - EDG/WP4 developed **configuration and installation tools**
- **Much work still to do**, though, especially for
  - a **State Management System**, and
  - a **Fault Tolerance System**.