

The Black Hole Comes To Brasov

Brian Martin
CERN

Childrens questions

- Why do I have to go to bed now
- Are we nearly there yet
- Where did I come from

- Why is the sky blue
- What am I made of

????????????????????????????????????

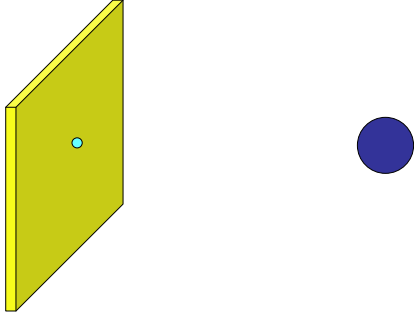
Brief History

- Can you keep cutting matter into smaller bits
- Greeks thought there was a limit
 - concept of the atom
- But the Greeks only thought about it
 - nothing new happened until people started to measure things
- Measurement and Classification
 - Compounds
 - Elements
 - Periodic table
 - Atomic weights
- **Yes but what are they made of, and why do they behave as they do**

Anything smaller than an atom?

- Experiments with electrical discharges in gases showed 'something' making light
 - That 'something' reacts to electro-magnetic fields
- And weighs much less than the lightest atom**
- Turned out to be an electron
 - You could easily steer and accelerate it.
 - Which means you can use them to smash into atoms and see if they will break

Fixed target collider



Twice the energy

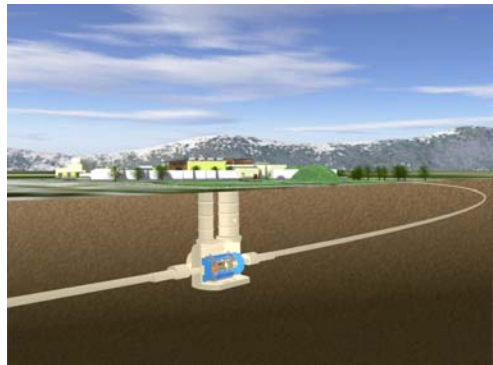
But now we have more childish questions like
How do you know where the bits went

and what they were
and what they do
and
and.....



CERN
site
and
the
LHC
ring

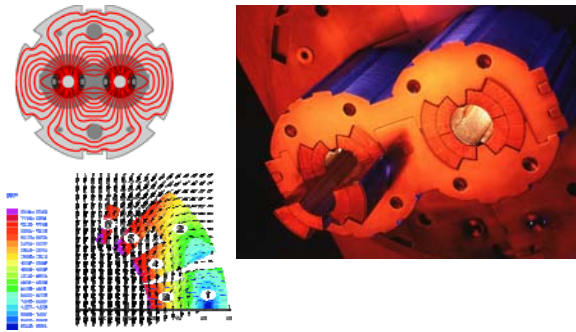
Cross-Section of LHC tunnel and Underground Experimental Area



RF Accelerator Section



LHC: DIPOLE Construction

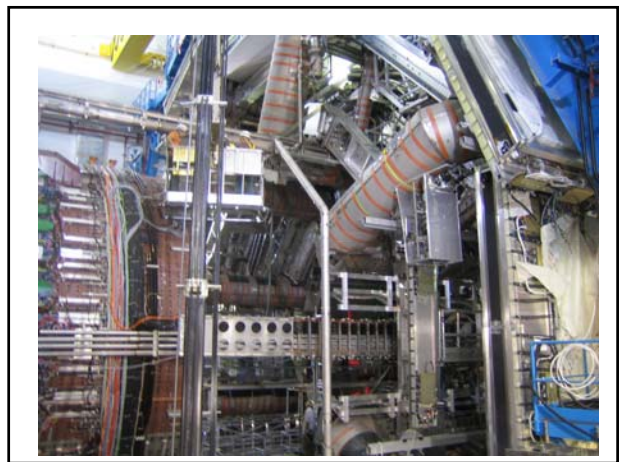
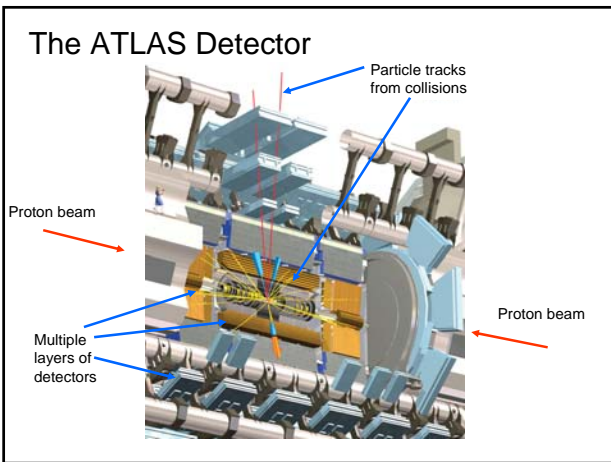
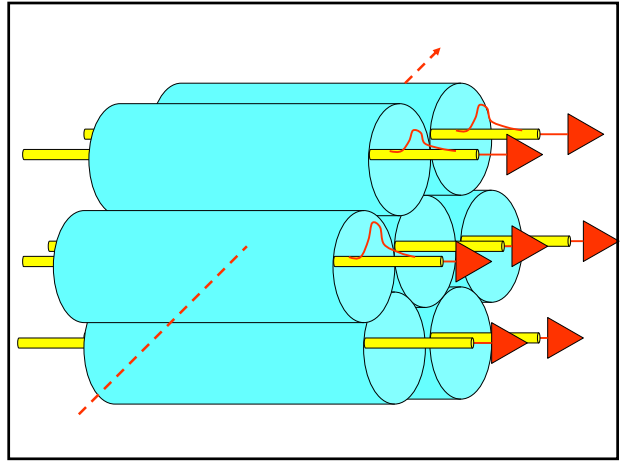
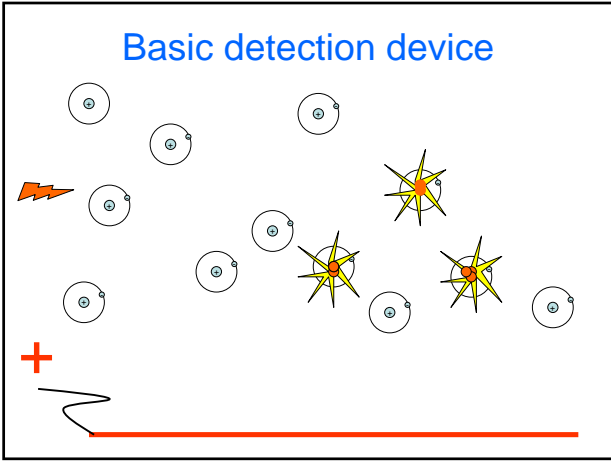


Lowering a Dipole into the tunnel



LHC Assembly in the tunnel

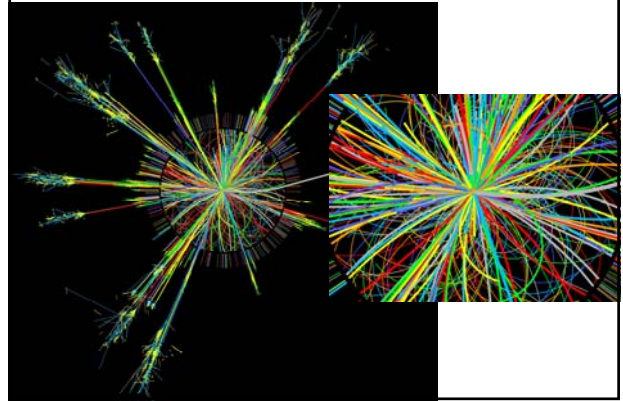




Reconstruction

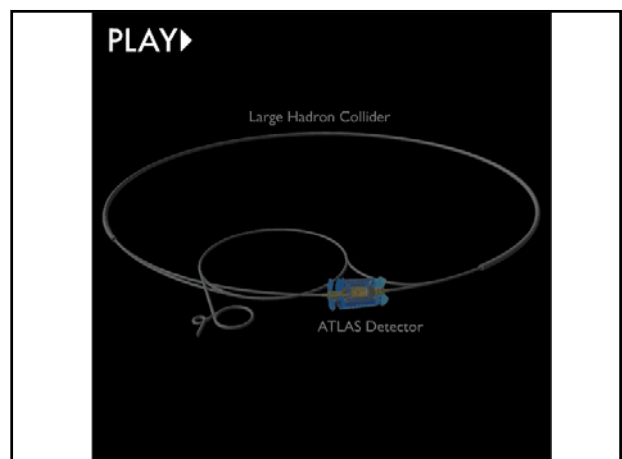
- By reading all the amplifiers
 - You obtain a set of 3D coordinates
- Need to associate them to reconstruct tracks
- Then see if all the tracks came from the same vertex
- Then see how bent the track was in the magnetic field
 - To calculate momentum
- Then decide if the energy measured is in the range of interest
- If so reconstruct the whole event and look for 'signature' characteristics

Simulated Event



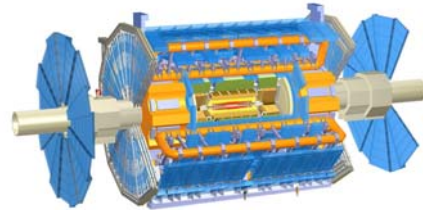
And the Black Holes??

- Accelerating particles to near the speed of light gives them mass.
- Colliding them creates , for a very short time a high energy density
- want to see what new particles could emerge from this density
- could create small volume of density sufficient to trap photons
- But it would be unstable
- And evaporate into a puff of Hawking radiation
- If not would sink to center of earth
 - But even there it doesn't have critical mass, needs many more
 - Tens of kilos
- Cosmic rays do this all the time
- If what they produced was stable would have several tons by now
- Black hole in Wall Street





Building A Really Big Digital Filter!

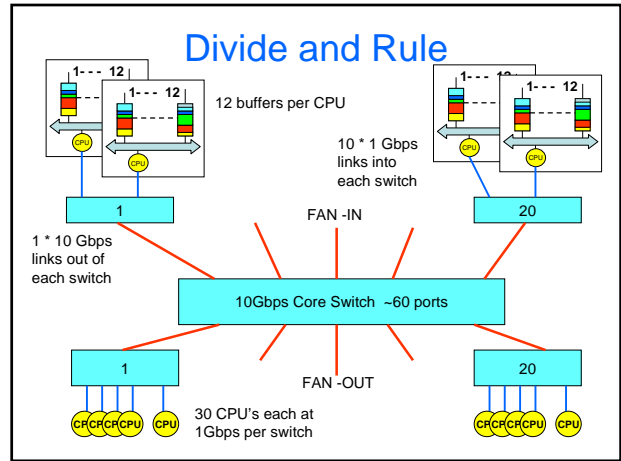
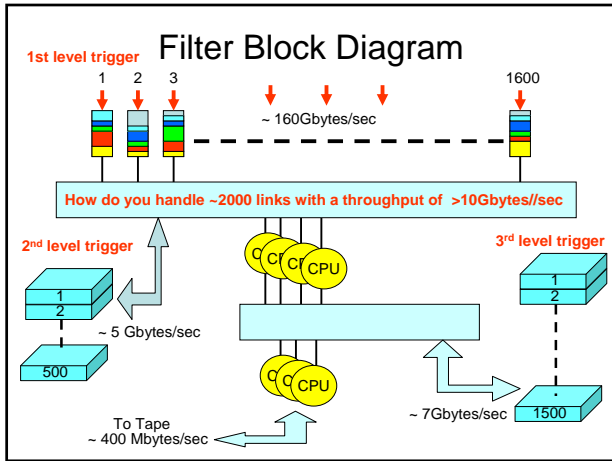


The signal source

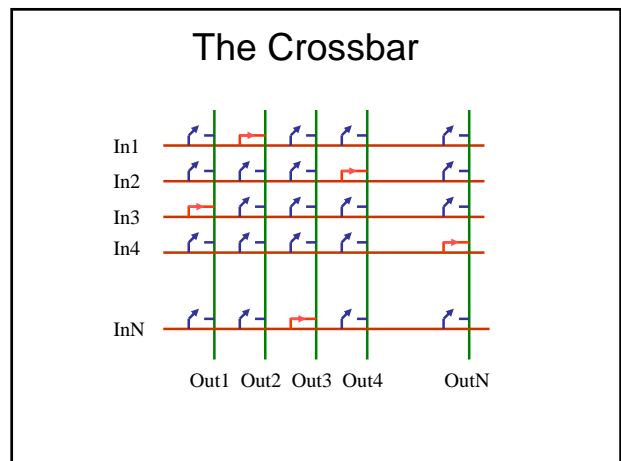
- Bunches cross at 40Mhz or 25ns
- 1st level trigger cuts it to 75 KHz (100KHz possible)
- 1600 channels
- 160Gbyte/sec (40DVD's per second)
- Very little of it is 'interesting'

Triggering

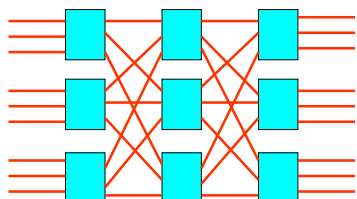
- 1st level trigger gives a pointer to a region of interest
- 2nd Level: only need to read the tracks in that region to decide if it is worth reading the whole event
 - Only a subset of the data : 100KHz at 20Kb
- Reject about 97% of events, 3.5KHz accepted
 - Best guess need for up to 500 CPU's using 8Ghz technology
- 3rd level; read the full data for the accepts
 - 3.5 KHz with 2Mb events
 - 200Hz accept rate
 - Best guess need for up to 1500 CPU's using 8Ghz technology



- ### Bus versus links and switches
- Bus technology favoured for previous experiments
 - 'private' domain, well controlled
 - Bounded latency
 - BUT
 - Shared resource
 - Bandwidth limitations
 - Link and switch scales, and is cheaper
- So : how does it work???



A multistage network



More than one path from any input to an output
A center stage must connect to all other chips.

N times the Bandwidth?

First the good news: Yes you do get N times the bandwidth
Then the bad news: Only if you have communicating pairs

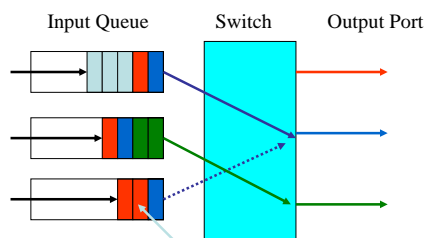


Max throughput

Major problem

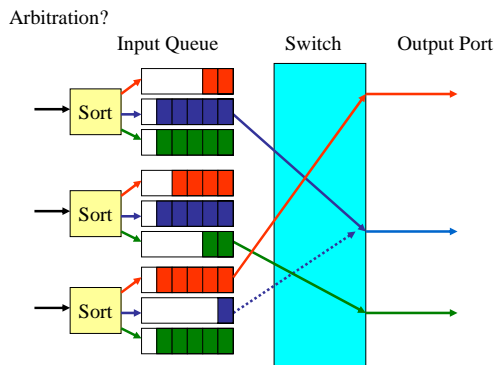
If more than one input wants to talk to the same output
Conflict is what you get.
Either you throw away conflicting packets (bad for public relations)
Or you can queue packets at the input

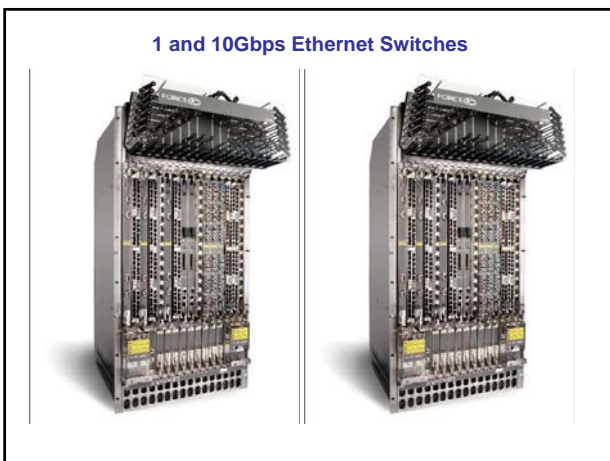
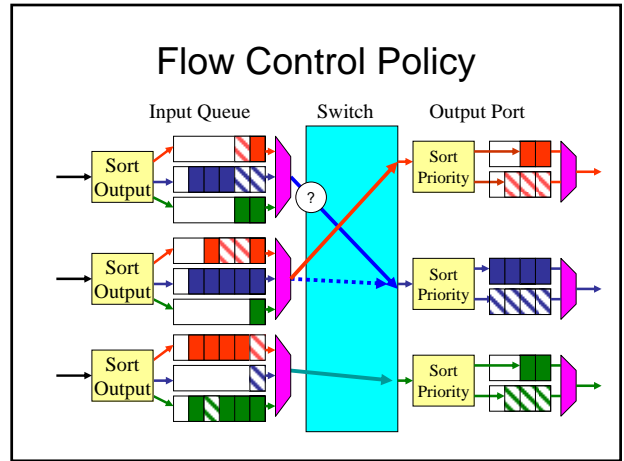
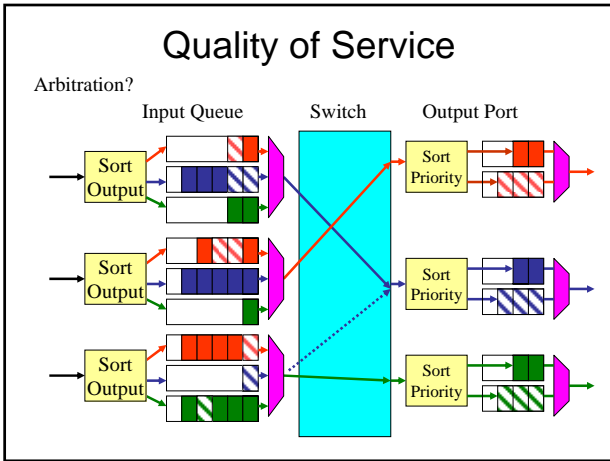
Input Queueing: Not a salvation



Head Of Line blocking: these packets can't get to idle output port
Efficiency drops. Random traffic is 59% of maximum.

Virtual Output Queuing





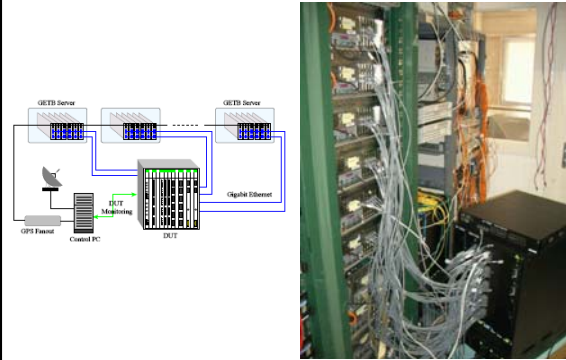
- ### Which one to buy!
- IS it blocking or non-blocking
 - How well does the arbitration work
 - Is it fair
 - Is performance affected by packet size
 - What is the latency
 - How well does flow control work
 - What monitoring facilities are there
 - TEST TEST and TEST Again

GETB

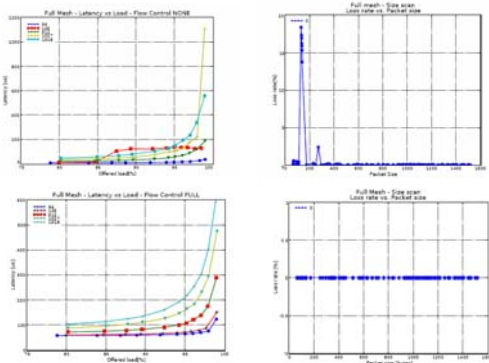


Programmable traffic flow generator

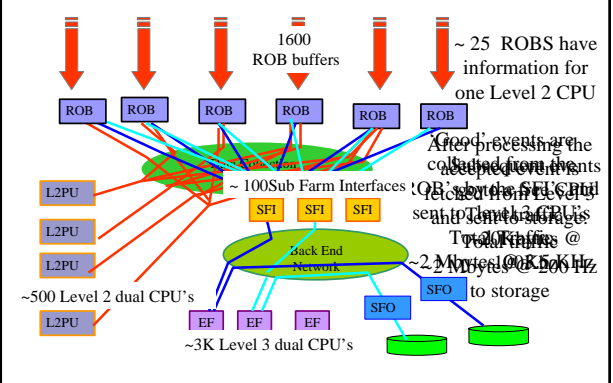
Switch Test Bench



Performance Analysis



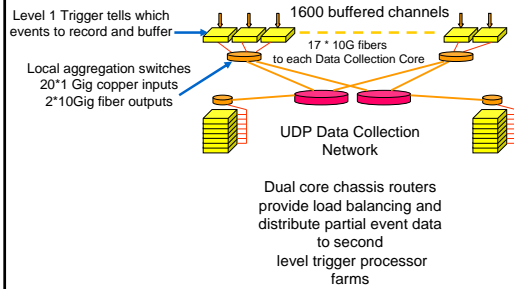
Trigger And Event Data Flow



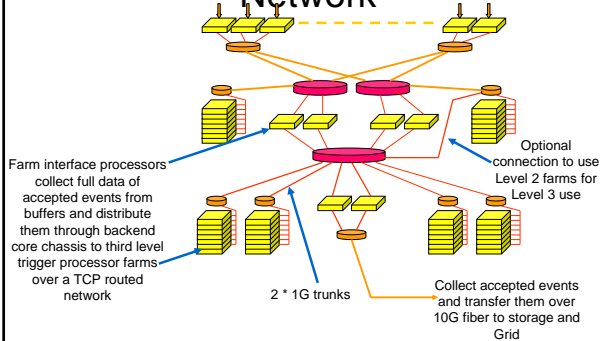
Implementation

- **Defined requirements**
- **'Paper' Models to 'estimate' traffic**
- **Evaluated existing technologies**
- **Predicted a cost model**
 - Telco crash
 - ATM vs Ethernet
 - 10G availability
 - looking ahead 3 years??
- **Now to the design**

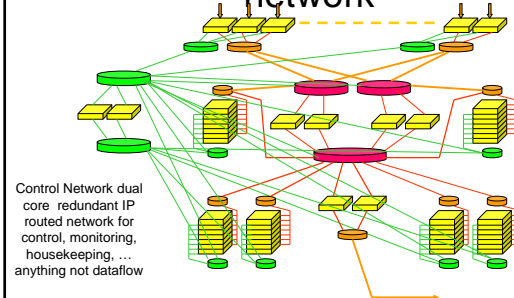
UDP Data Collection



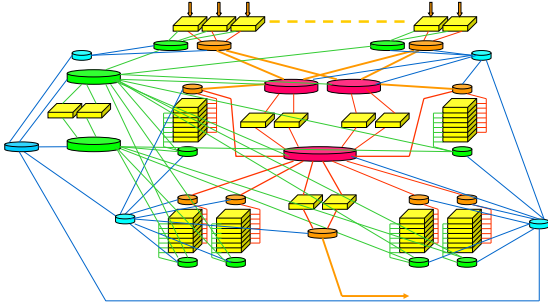
Plus IP routed Back End Network



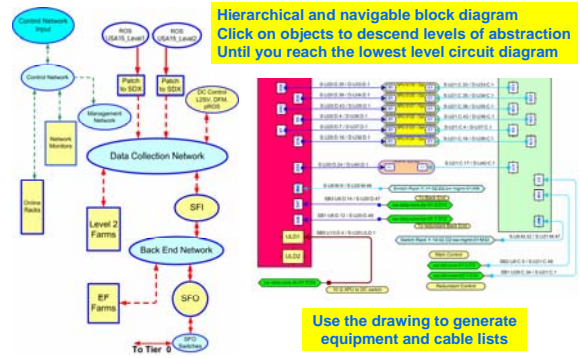
Plus dual redundant control network



Plus traffic monitoring



Visio Circuit Drawing



Geographical Location

Property	Value
Rack	T-03-04-03
Detector	DAG
Usage	EF rack
Responsible	DAG
Comment	Validated
Last change	3/15/2007 8:28:52
Rack Height (ft)	47
Weight (kg)	0
Total weight (kg)	487.58
Power absorbed by T.	0
Power supported (kw)	14.87
Power dissipation (kw)	3.72
Power dissipation to	3.6
Power connection type	Regular power
Status	
Electrical Reception	
Slots on Front	47
Slots on Back	47
Slots on Back door	47

The data handling area is full of racks.

Each rack carries processors and their local switches.

The rack is defined in X and Y for any given rooms floor plan.

The rack contents are defined in horizontal rows

Equipment Data Base

Top Assembly Folder: Properties

Top Assembly Identifier: 200QUC13000001
Other Identifier: 10106
Description: Trigger & DAQ - Computers

Property Values	Internal Value	Unit
PC Model	Surfex V302	
PC Serial Number	0052010042	
Budget Code		
PC Warranty Period	872637-6	
Manufacturer Serial Number	00.05 - 00.08	
CPU 1 Serial Number		
Memory 1 Serial Number		
Memory 2 Serial Number		
MAC Address 1		
MAC Address 2	FC1-0	
Bus Type 1		
Bus Type 2	none	
No. of supported add ins	2	
Number of USB ports	1	
Number of Serial Ports	1	
Number of Parallel Ports	1	
Function		

Every piece of equipment is registered in an inventory database

Cable Installation Data Base

CABLE: 2263600

Identifier	2263600
Status	-
Detector	DAQ
Sub-system	DAQ/HLT-I
Function	sw-data-core-dc-013 to Patch Panel 5
Type	Multimode optical fibers
Standard Field	D-Q/Y 14-02 D2/OM3/Y 20-06 D2
User Field	SB3 U13 D.2 / D U29 D 21-22
Diameter	18
Comment	

Enter ITEM_ID, LABEL or SERIAL NUMBER

Show properties

Every cable is registered with its:

- Number
- Type
- Source location rack and row
- End location rack and row
- Function

LAN Database

Device Information

- Device Name: D1125 2V IP25 SHIPM 0716 [Last Operation]
- Location: 3125 2-0000 (Zone: RACK Y 07-16 A2)
- Manufacturer: HP
- Model/Type: PROCURVE 3400CL24 J4805A
- Generic Type: SWITCH
- Description:
- Tag: IT/
- Serial Number: S0631SFQDP
- Operating System: NONE -Version: UNKNOWN
- CERN Inventory number:
- Network Interface Card(s): — PRIVATE INFO — [Why?]
- Responsible for the device: NETWORK INFRASTRUCTURE IT INFRASTRUCTURE NETWORK@CERN.CH
- Main User of the device: SYSADMINS ATR PH ATR SYSADMINS@CERN.CH
- HCP Response: This system CAN obtain an IP address automatic
- Last changed: 17-01-2007 (16:06)

Interface(s) Information

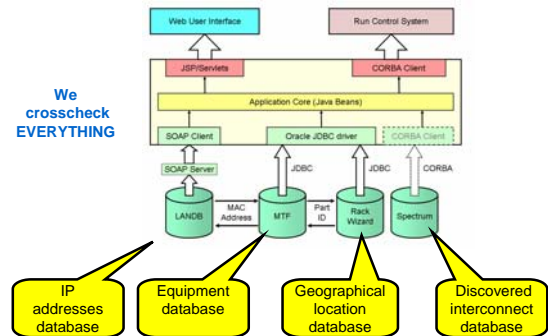
Every device in the network must be registered:

- Security
- IP address management
- Error diagnostics

Database Consistency

- Multiple databases
- Data entered by different people
 - At different times, or not at all
- Things get moved and not recorded
- Only LANDB 'must' be up to date (no IP)
- Inventory management problem
- How do we keep track of it all?

Better definitions for better monitoring



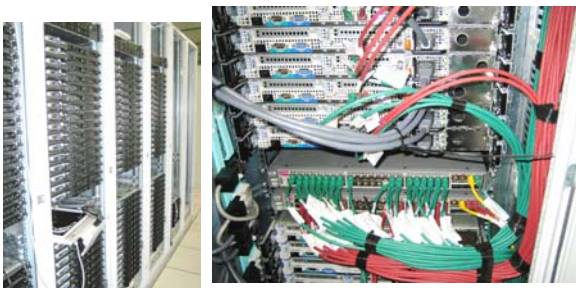
Installing cables



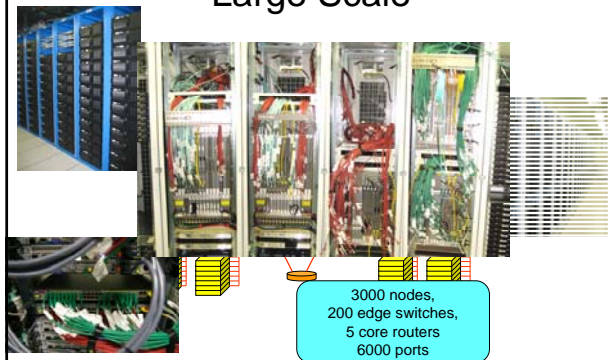
Installing Core Switches



Installing Edge Switches



Large Scale



Plus physicists!

- Network dimensioned to meet 'requirements'
- Maximum average link occupancy <60%
- **Should** mean peace of mind for Network Support
- **Actually** seen as a challenge by physicists
 - 40% for free! Turn up the wick until something breaks!
- Continuous running out of spec!
- Must distinguish between 'real' and 'self inflicted' problems

MONITOR EVERYTHING