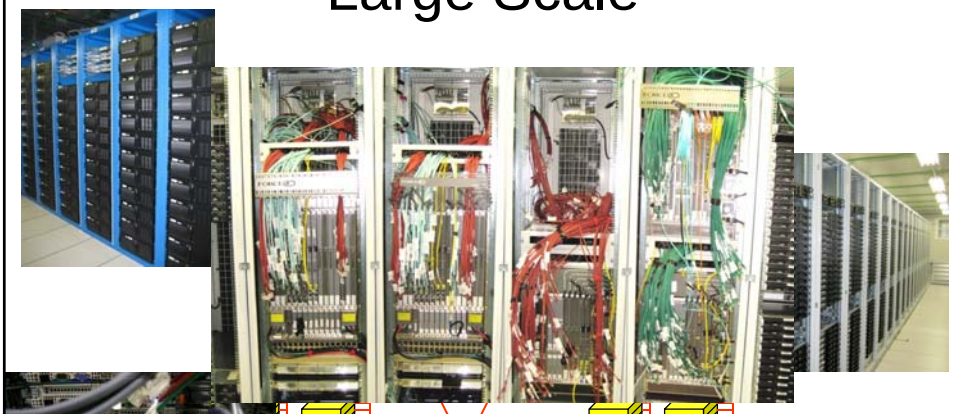
Monitoring the ATLAS TDAQ Network at CERN

Lucian LEAHU

Brasov, 15/01/2009

Large Scale



3000 nodes,
200 edge switches,
5 core routers
6000 ports

2

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

3

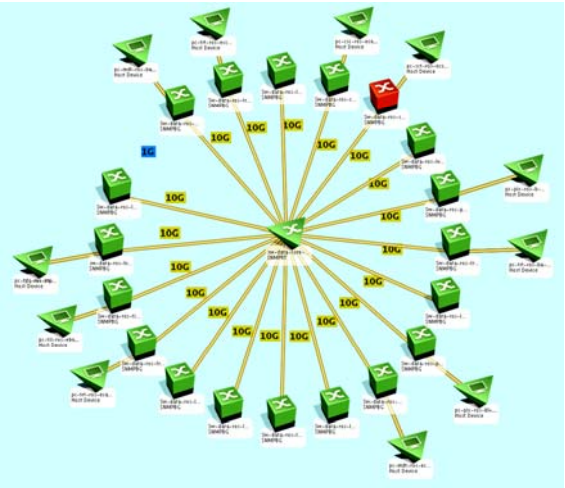
Commissioning

Basic question:
Does any of it work????
Monitor everything!

ICMP:
Internet Control Message Protocol
SNMP:
Simple Network Management Protocol

SPECTRUM

Tells you it’s alive
Tells you if it dies
Fetches status info..
- and hides it!



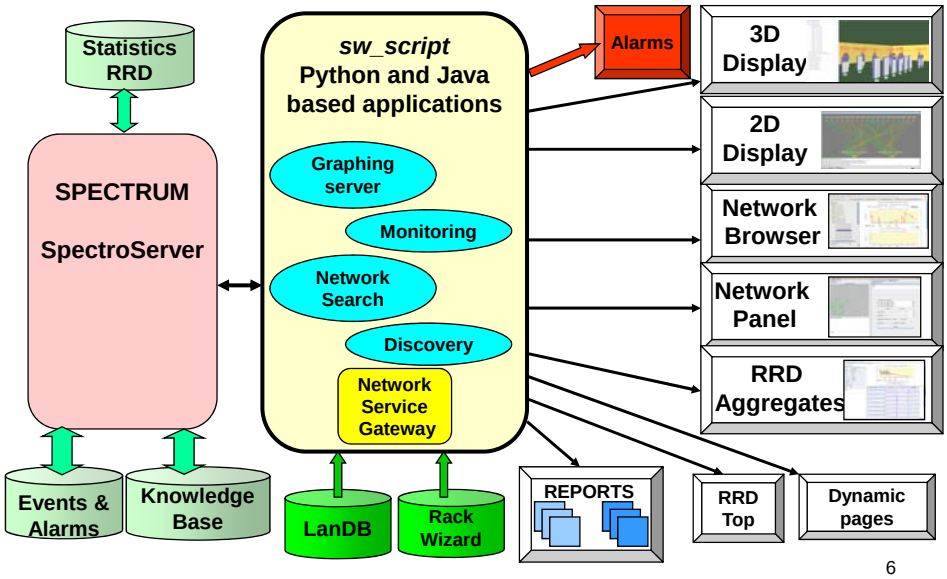
4

Commercial versus Proprietary

- Special needs:
  - Multiple networks per processor
  - Want to see whole picture for system analysis
  - Want to see all detail for component analysis
  - Want to see traffic volume visually
  - Want traffic/errors qualified
- Spectrum CORBA API clumsy
- Multiple requests hits the CPU hard

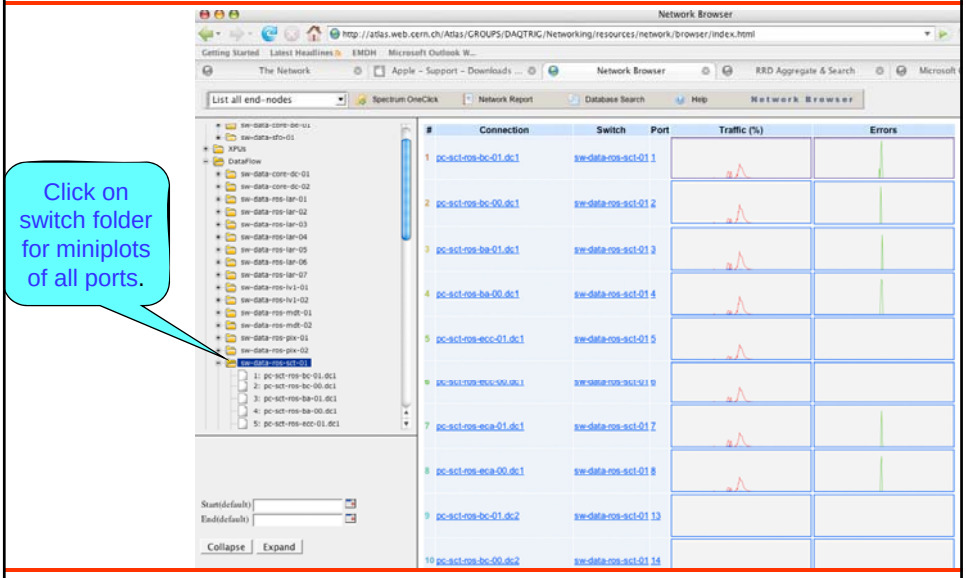
5

Current Infrastructure



6

Network Browser: Miniplots



7

Dynamic pages with real time traffic

Switch	Port Usage	Connection	Direction
swps-ctrl-core-01	0/21 162.5	UPLINK: swps-ctrl-l2-01 (0/24)	TX...
swps-ctrl-l2-01	0/24 159.7	UPLINK: swps-ctrl-core-01 (0/21)	RX
swps-ctrl-core-01	0/14 156.5	pc-preseries-l2sv-03.ctrl0	
sw-ctrl-core-01	0/0 52.8	UPLINK: D3178-2-RHPZL-1 (X 1)	
swps-ctrl-l2-01	0/13 39.6	pc-preseries-l2p-05.ctrl0	
swps-ctrl-l2-01	0/14 39.6	pc-preseries-l2p-06.ctrl0	

Real-time global top Most active connections in real time.

Class	Machines
DFM	pc-preseries-efp-17, pc-tdq-dfm-08, pc-tdq-xpu-0120
EF	pc-preseries-efp-17, pc-tdq-mon-13, pc-tdq-mo... pc-tdq-xpu-0067, pc-tdq-xpu-0068, pc-tdq-xpu-... pc-tdq-xpu-0074, pc-tdq-xpu-0075, pc-tdq-xpu-...
IS	
L2PU	pc-preseries-l2p-05, pc-preseries-l2p-06, pc-pre... pc-tdq-xpu-0004, pc-tdq-xpu-0005, pc-tdq-xpu-... pc-tdq-xpu-0011, pc-tdq-xpu-0012, pc-tdq-xpu-... pc-tdq-xpu-0018, pc-tdq-xpu-0019, pc-tdq-xpu-... pc-tdq-xpu-0025, pc-tdq-xpu-0026, pc-tdq-xpu-... pc-tdq-xpu-0033, pc-tdq-xpu-0034, pc-tdq-xpu-...

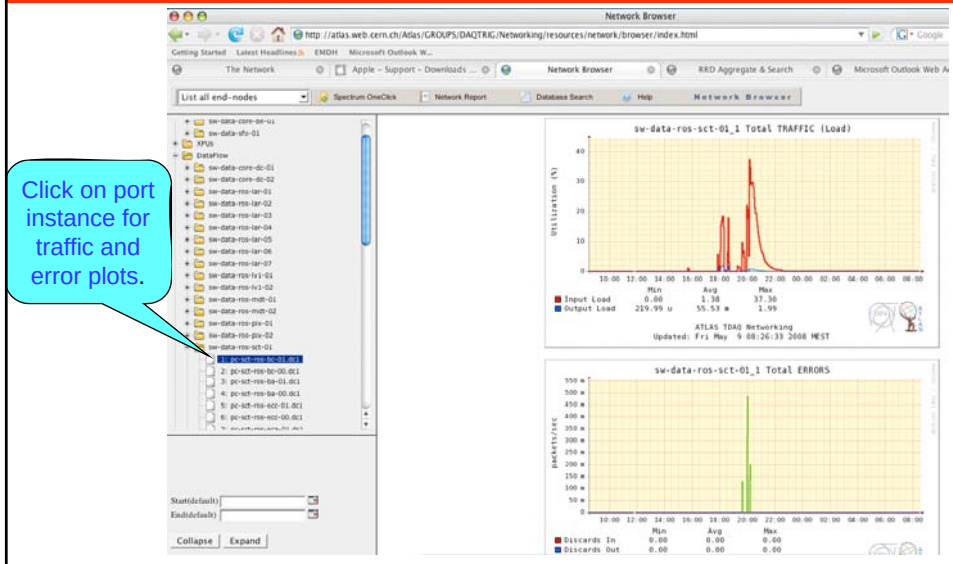
Current ATLAS applications running in the network.

port	connection	tx_mbit	rx_mbit	link_state	link_speed
0/0	UPLINK: D3178-2-RHPZL-1 (X 1)	12.8	55.3	up	1 G
0/1	pc-tdq-onl-01.ctrl0	0.0	0.0	up	1 G
0/2	pc-tdq-onl-02.ctrl0	0.3	0.2	up	1 G
0/3	pc-tdq-onl-03.ctrl0				
0/4	pc-tdq-onl-04.ctrl0				
0/5	pc-tdq-onl-05.ctrl0				
0/7	pc-tdq-onl-07.ctrl0	0.0	0.0	up	1 G
0/8	pc-tdq-onl-08.ctrl0	0.0	0.0	up	1 G
0/9	pc-tdq-onl-09.ctrl0	0.0	0.0	up	1 G

Connections and traffic per switch port.

8

Network browser: port stats



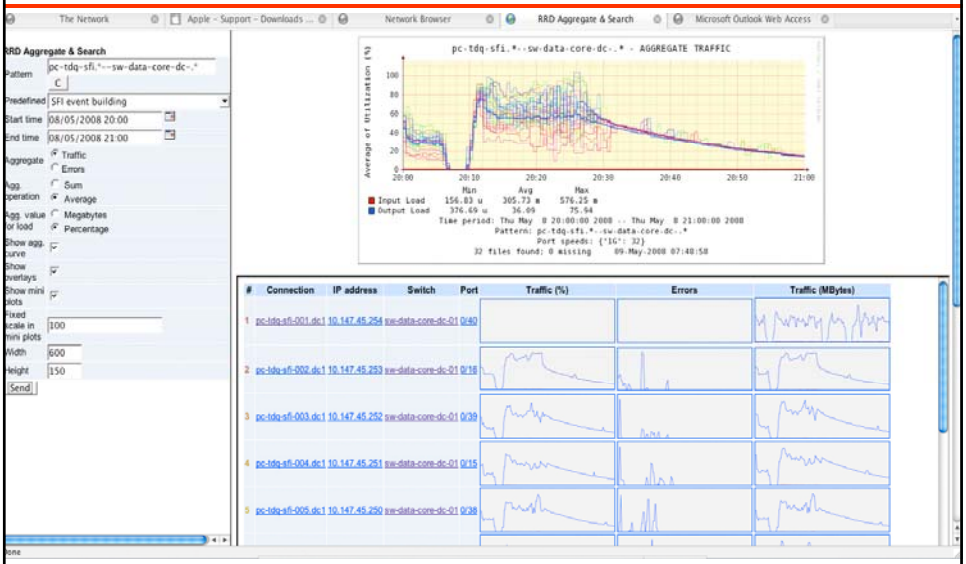
9

Affiliated mini-plots

#	Connection	IP address	Switch	Port	Traffic (%)	Errors
1	pc-tdq-sfo-01.dr0	10.152.32.253	sw-data-sfo-01	23		
2	pc-tdq-sfo-02.dr0	10.152.32.252	sw-data-sfo-01	22		
3	pc-tdq-sfo-03.dr0	10.152.32.251	sw-data-sfo-01	21		
4	pc-tdq-sfo-06.dr0	10.152.32.250	sw-data-sfo-01	12		
5	pc-tdq-sfo-07.dr0	10.152.32.249	sw-data-sfo-01	13		

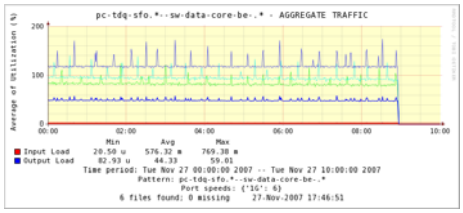
10

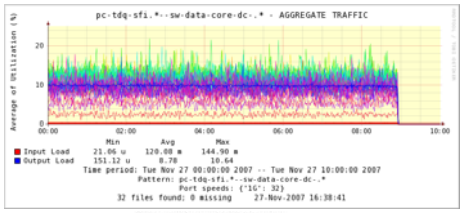
Searching and Aggregation

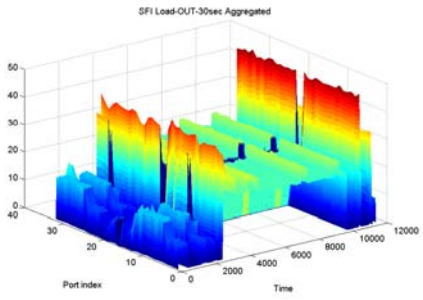


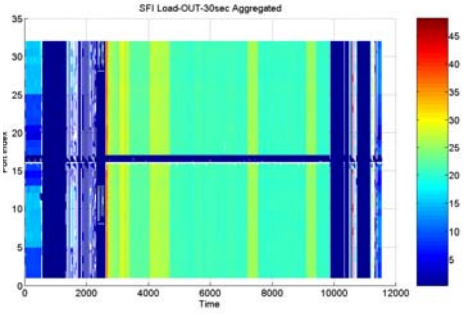
11

Scaling and aggregated plots



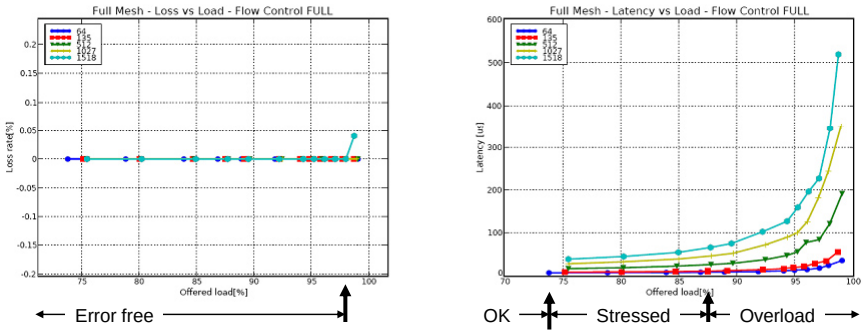






12

System Diagnostics



(a) Loss rate vs Load (b) Latency vs Load

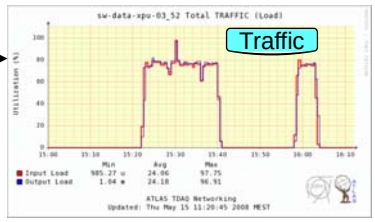
We have prior detailed measurements of losses and latencies against loads.
System dimensioned for a nominal operating maximum of 60% load on any port.
Naïve world view:
Flag Loads with high thresholds, above, say 85% throughput
Flag Errors with low thresholds, above, say, 0.05% or about 10pkts/sec at 1Gbps
Unfortunately it's not so simple

13

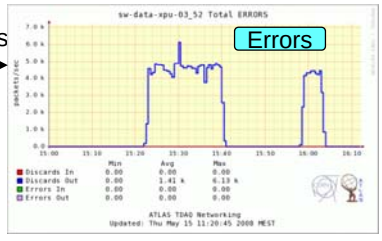
Where is the real error?

TCP

80% →

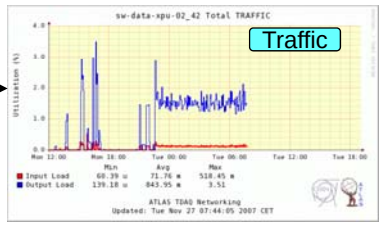


5K pkts/s →

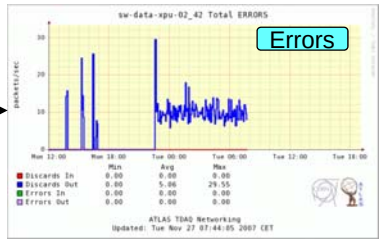


UDP

<2% →

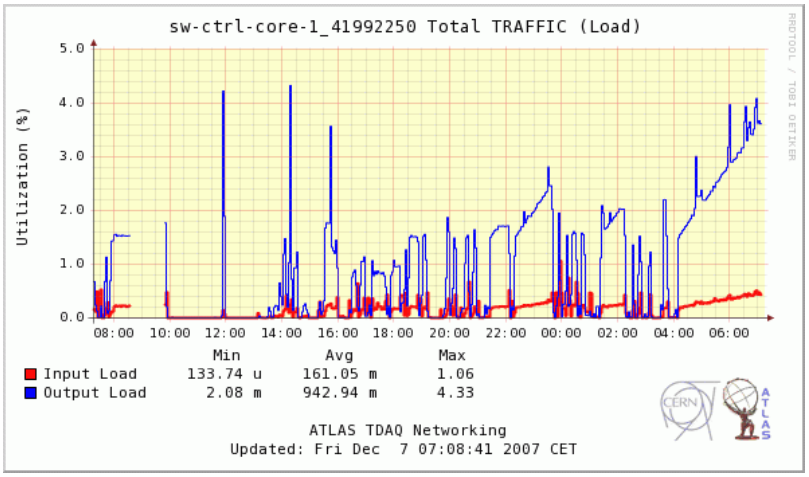


10 pkts/s →



14

Anomalous traffic



…and from here to rules based expert systems

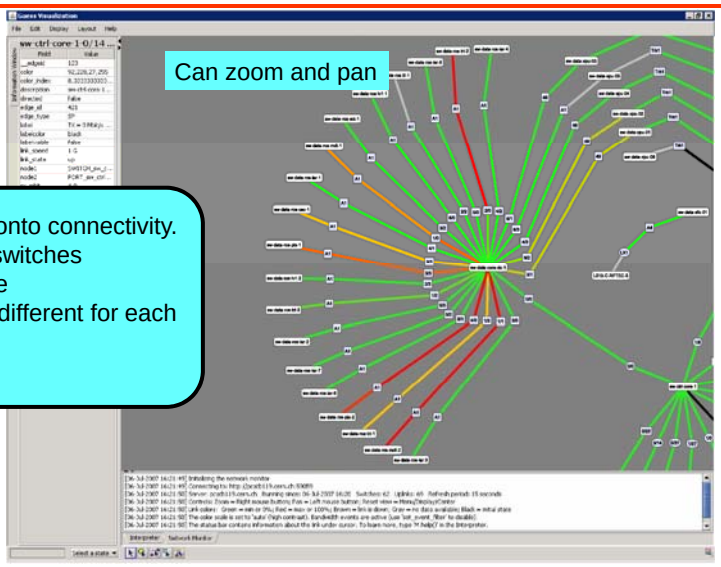
15

Transition to displays

- So far have concentrated on low level port and switch monitoring and diagnostics.
- Identified scaling issues
- Want to have a display with a system view
- Want to retain architectural model
- Want to exploit real time stats

16

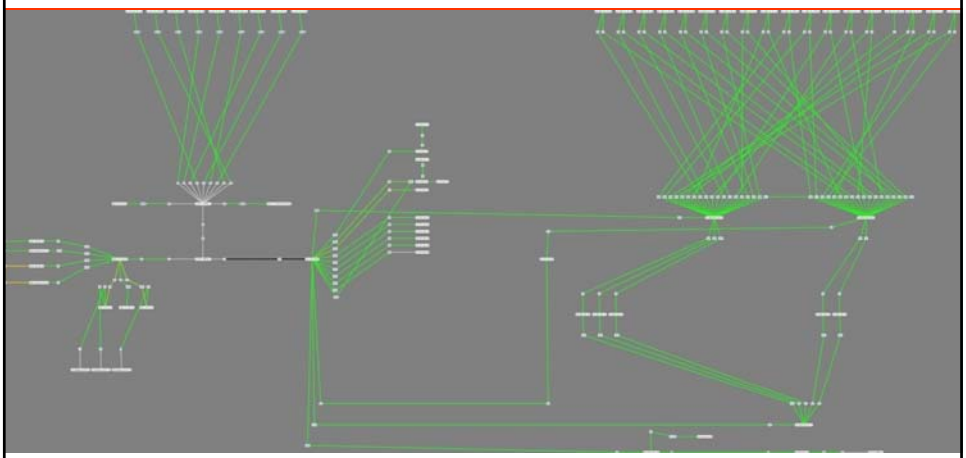
GUESS 2D Visualisation



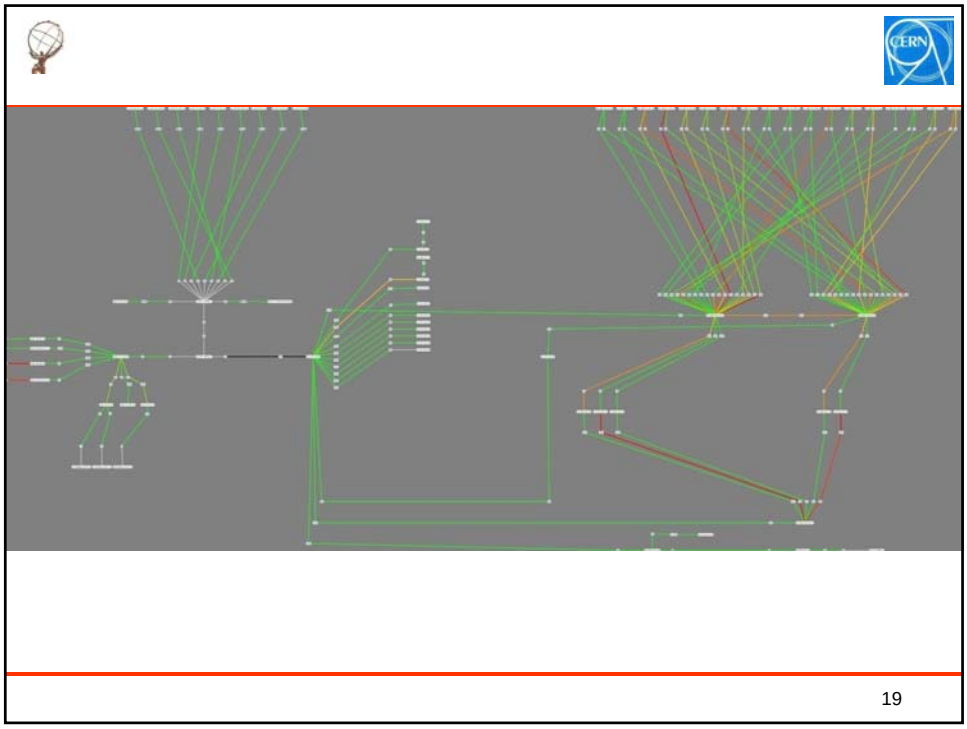
- Traffic overlaid onto connectivity.
- No hosts, only switches
- Only 1 data core
- Autoplacement different for each discovery

17

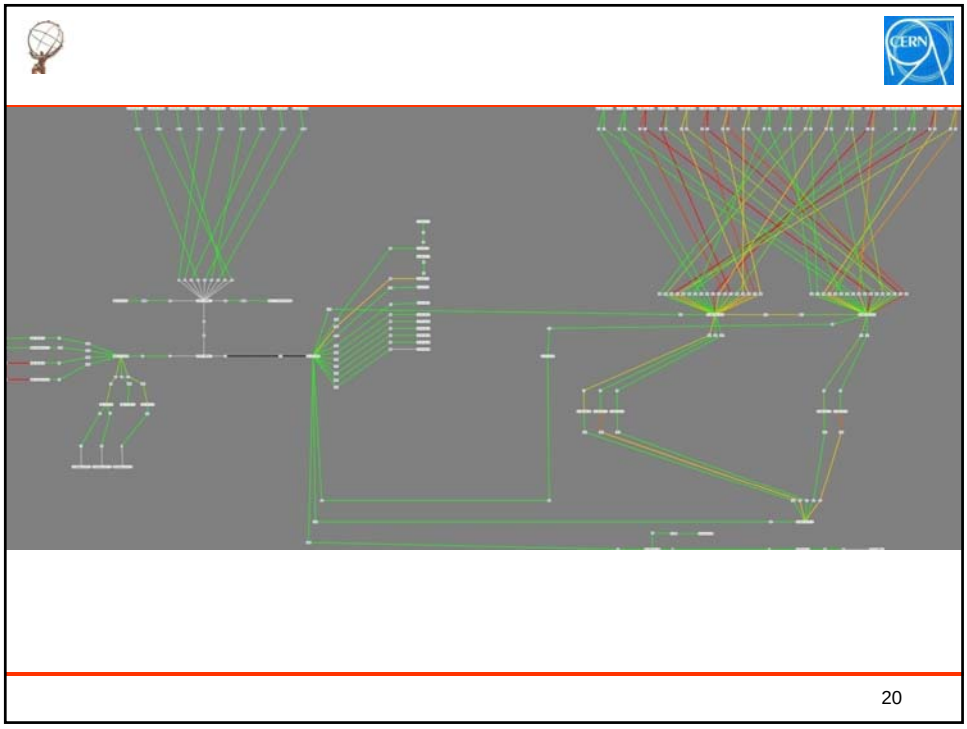
GUESS: Guided Placement



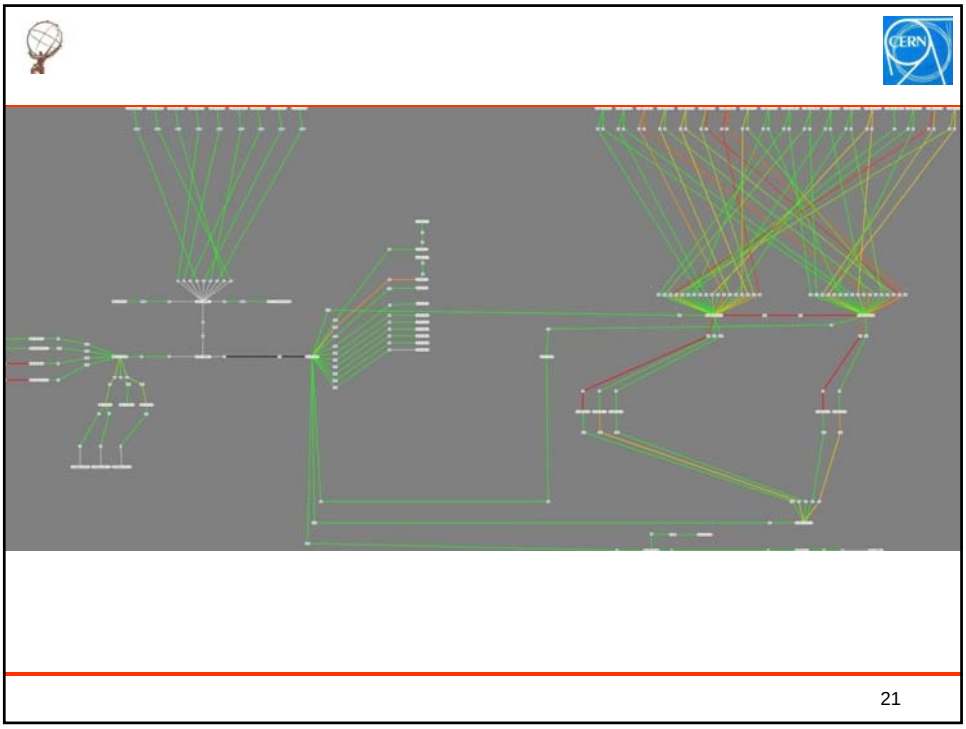
18



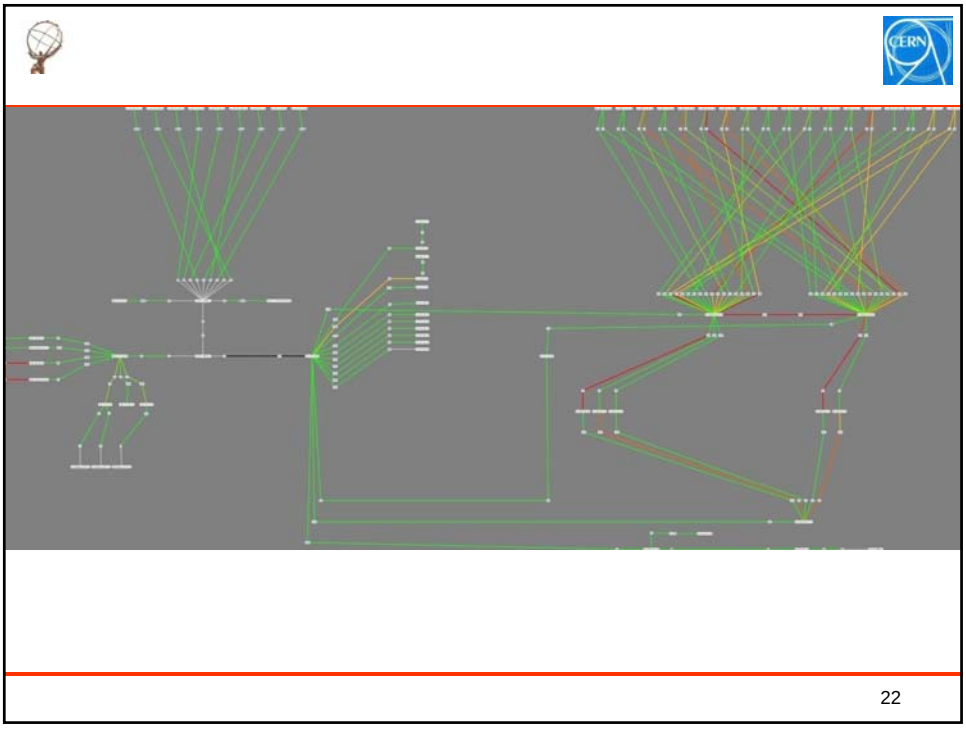
19



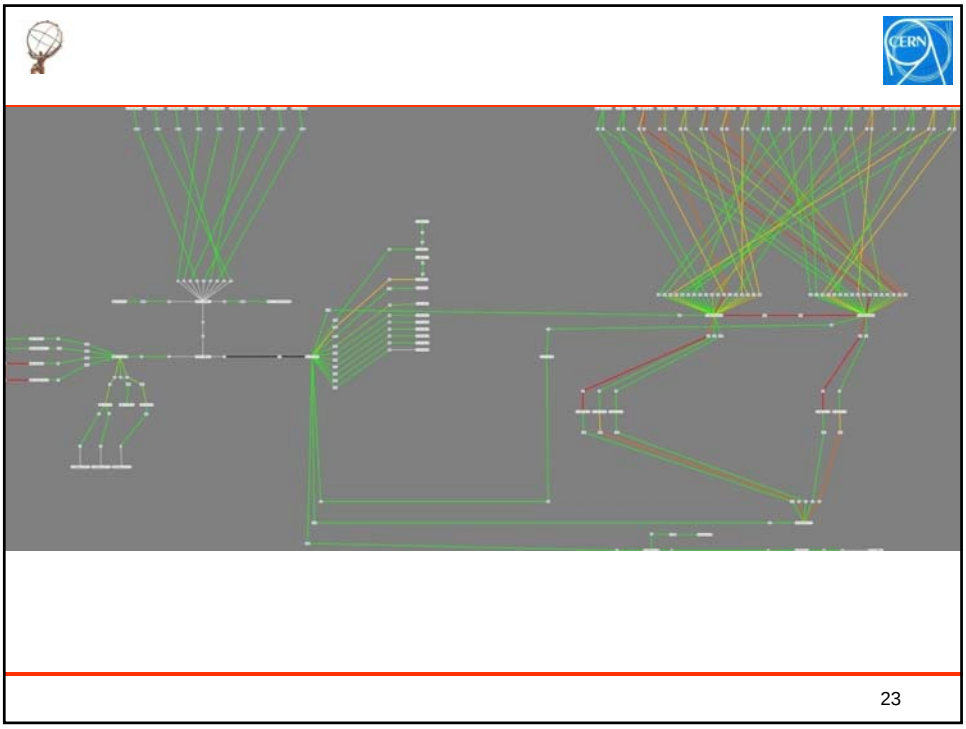
20



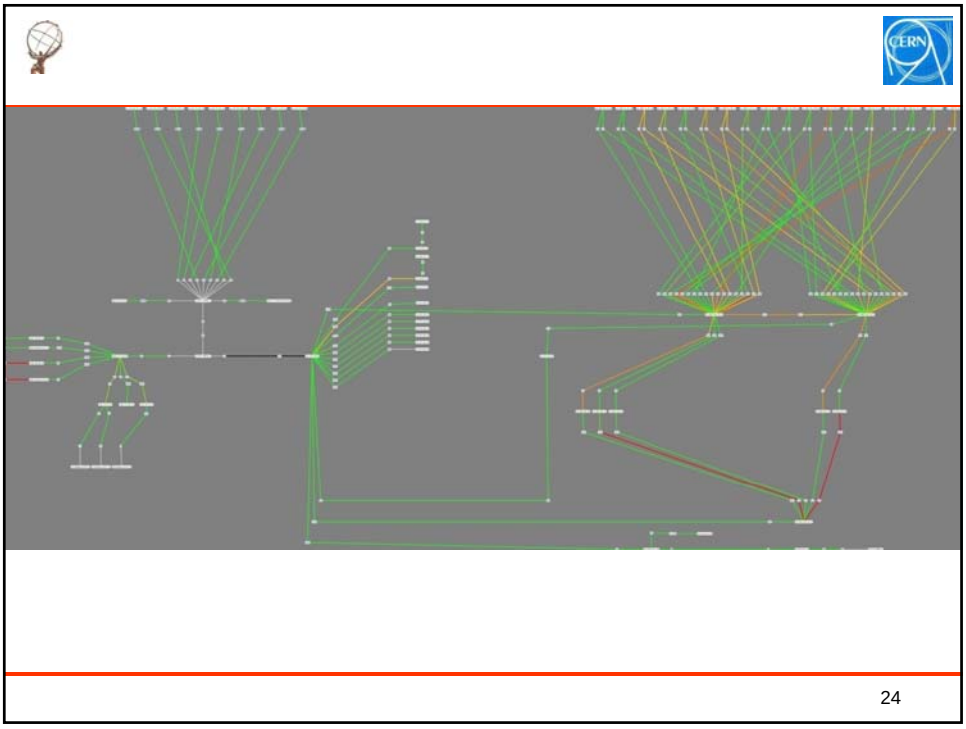
21



22



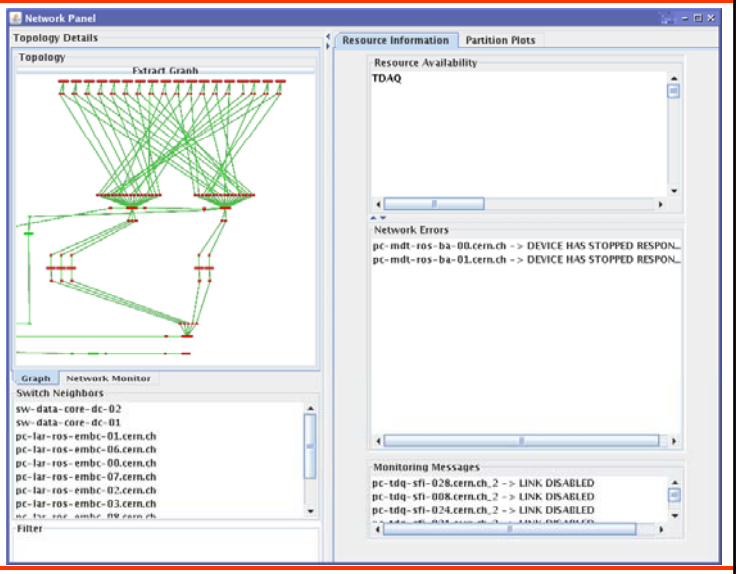
23



24

Atlas Users Network Panel

For operators we provide a summary of status per application set.



25

2D Display Limitations

2D Display is very good for switch to switch traffic visualisation
No 'level of detail' feature as you zoom in or out
Can't incorporate host details - - -would overwhelm the plot

BUT..

I want all the detail when something goes wrong
I want neighbouring views when examining individual behaviour
I want the system wide view for visual correlation
I want different levels of detail depending on my viewpoint
I want fly-through, and navigation and easily visible errors
I want … I want …

I want GOOGLE Earth for my network.

26

3D Display

Google earth as inspiration
Variable detail as a function of viewing distance

Variable viewing angles
Intuitive navigation

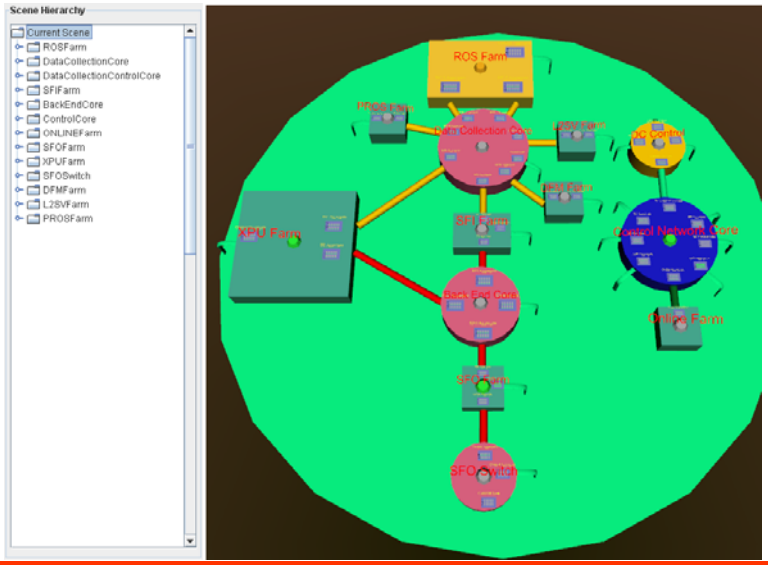
Unfortunately Google doesn't cope with our dynamic update requirements

So we went looking for display software that does.

X3D (enhanced VRML)
Octaga Player

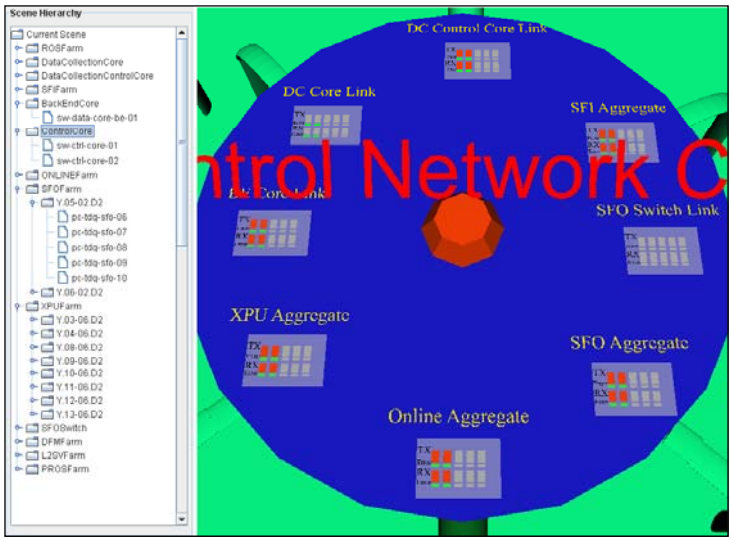
27

3D Top Level View



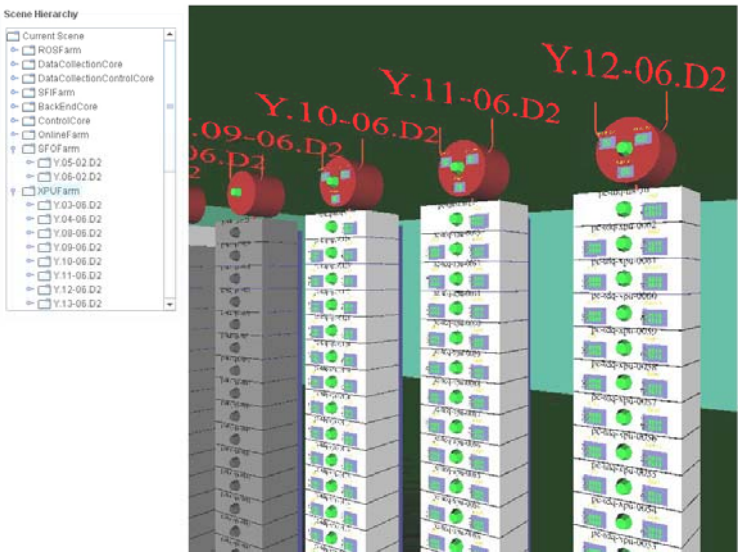
28

3D Top Level View



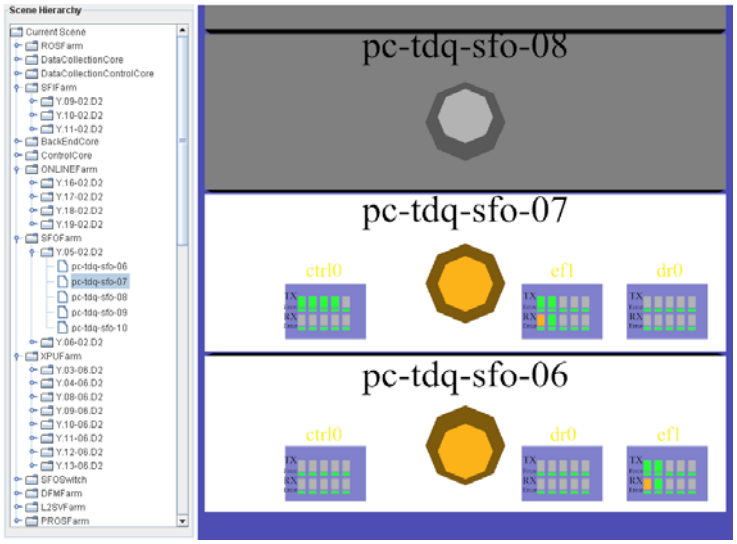
29

Detail of processor farm



30

Detail of a Processor



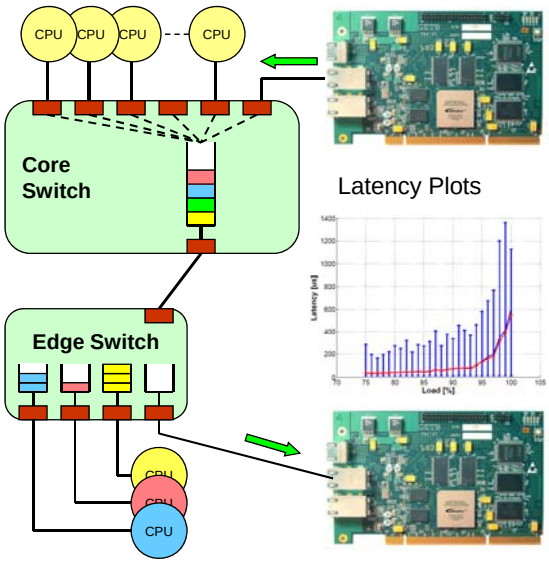
31

Diagnostic Tools: 1

- YATG (Yet Another Traffic Grapher
  - High speed SNMP-based traffic monitoring (the switch is the limiting factor)
  - Fine time granularity statistics for selected device interfaces
- ATLAS-like traffic bandwidth measurements
  - Distributed applications replicate the transactional request-response transfer protocol
  - Demonstrate maximum achievable bandwidth

32

Diagnostic tools :2



Latency Plots

Dynamic queue growth measurement

33

Diagnostic tools :3 SFLOW (1)

- sFlow is the standard for statistical packet sampling
  - Each network port → sampling system
  - All packet samples → central location (software)
  - Analysis → information about the *content* of the traffic

- By collecting packet samples, the packets can be classified into flows
  - A flow ~ network conversation between two applications
  - The bandwidth occupancy for each flow can be estimated

- We developed an sFlow analysis application in order to study the technology

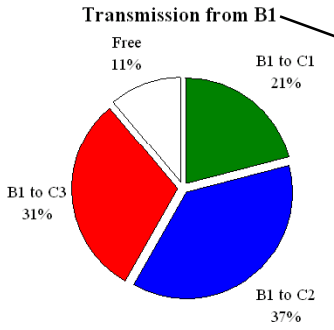
34

Diagnostic tools :3 SFLOW (2)

For one switch port

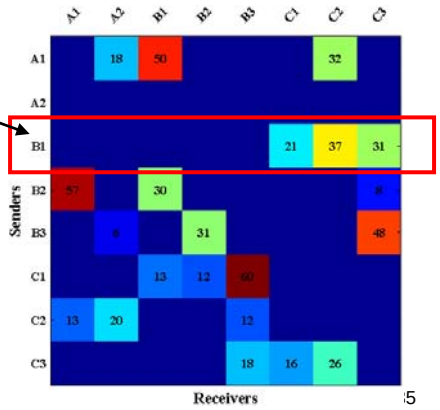
Using SNMP → "bandwidth usage on port B1 is 89%"

Using sFlow → Pie chart with traffic distribution



For the entire network

Traffic matrix with all senders and receivers



35

Summary

- large scale network
- monitor, diagnose
  - commercial tools
  - "in-house" built software
  - plots for 6000 ports
  - different levels of visual abstraction
- 24/7 operation
- errors not welcome

36