

F. Chlebana
Dec 13 2006

Status of the CSL Upgrade

I. Bizjak, F.Chlebana, G. Guglielmo, (J.Lee), K.McFarland,
R.Sarkis, JJ.Schmidt, M.Shimajima,
R.Snider, G.Yu, D.Zhang, DataComm, Computing Division

Cast of Characters

Richard Sarkis: *CSL software development*

JJ: *Installation and planning*

Geum bong Yu: *Integration testing*

Ilija Bizak: *Monitoring software*

Tatsuya Masubuchi: *Hardware testing and configuration*

Makoto Shimojima, Kevin McFarland, Jedong Lee: *CSL experts*

Dehong Zhang, Gerald Guglielmo, Rick Snider: *Down stream logging*

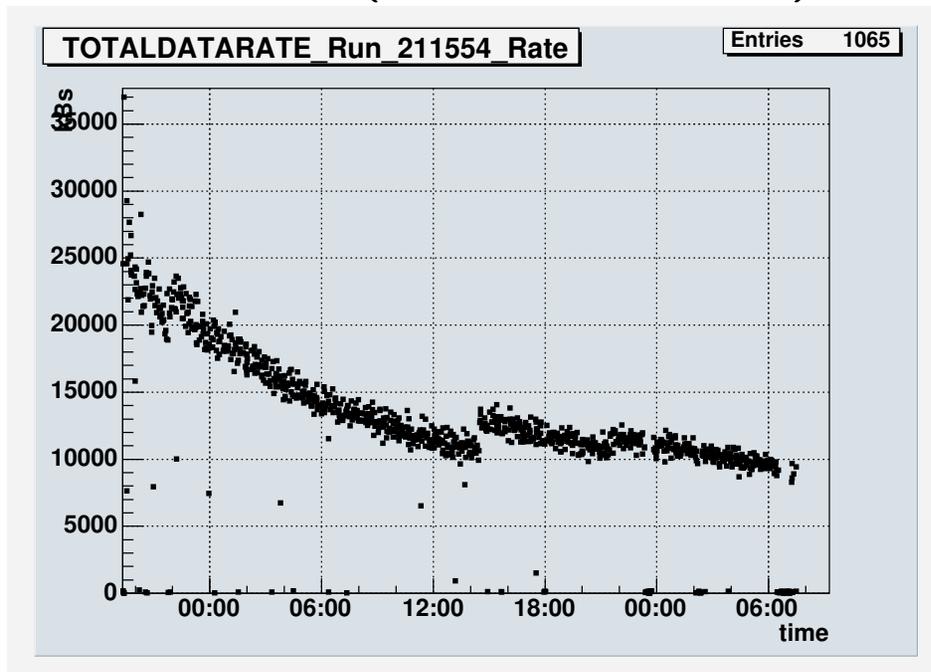
CD/Rex (Jason Hedden): *System configuration*

F. Chlebana: *Scallywag*

Why Upgrade the CSL

- Current CSL is limited to about 24 MB/s
→ Need to be able to handle higher logging rates

Run 211554 ($\mathcal{L} = 176 \times 10^{30}$)



What logging rate will we need?

250 KB/Event

1000 Hz

L3 rejection of 4-5 times

→ 50 - 63 MB/s

Currently can write to buffer disks at higher rates (45 MB/s), limited by tape writing speed...

CSL upgrade design target is 80 MB/s

EVB Rate	1000 Hz	1000 Hz	1200 Hz
Rejection at L3	4-5x -> 250-200Hz	4x -> 250 Hz	4x -> 300 Hz
Event size	200 KB	250 KB	250 KB
Required Bandwidth	40-50 MB/s	63 MB/s	75 MB/s
Datahandling limit	80 MB/s		

Provide enough logging capacity so that the CSL is not the bottleneck in the system

- Uses Silicon Graphics and IRIX
 - Becoming more difficult to support
 - Data handling (SAM) software not ported to IRIX
- Using aging hardware
 - “end of life” approaching
- Single Point of Failure
 - Dwindling SGI expertise

CSL Upgrades

Retain original software design and port onto new hardware

- Current software structure serves our purpose
- Limited resources for any major rewrite

Increase data throughput

- Use modular distributed logging architecture

Increased buffer capacity

- Buffer increased from 3.7 TB to 24 TB
- 80 MB/s → ~7 TB/day

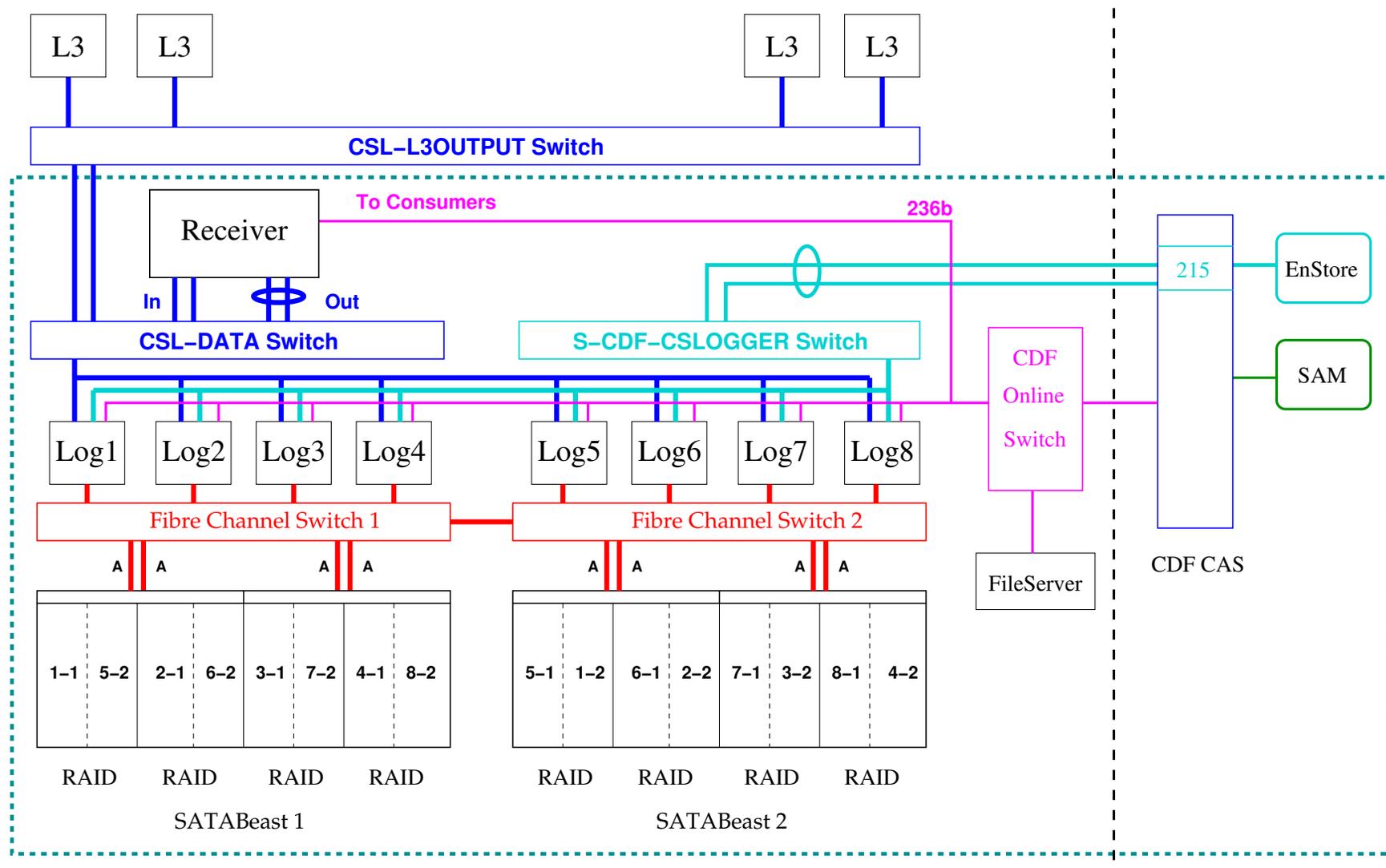
Improve maintainability and robustness

- Replace aging equipment with new servers/disk arrays
- Use Linux instead of IRIX
- Redundant hardware with hot spares

Expandable

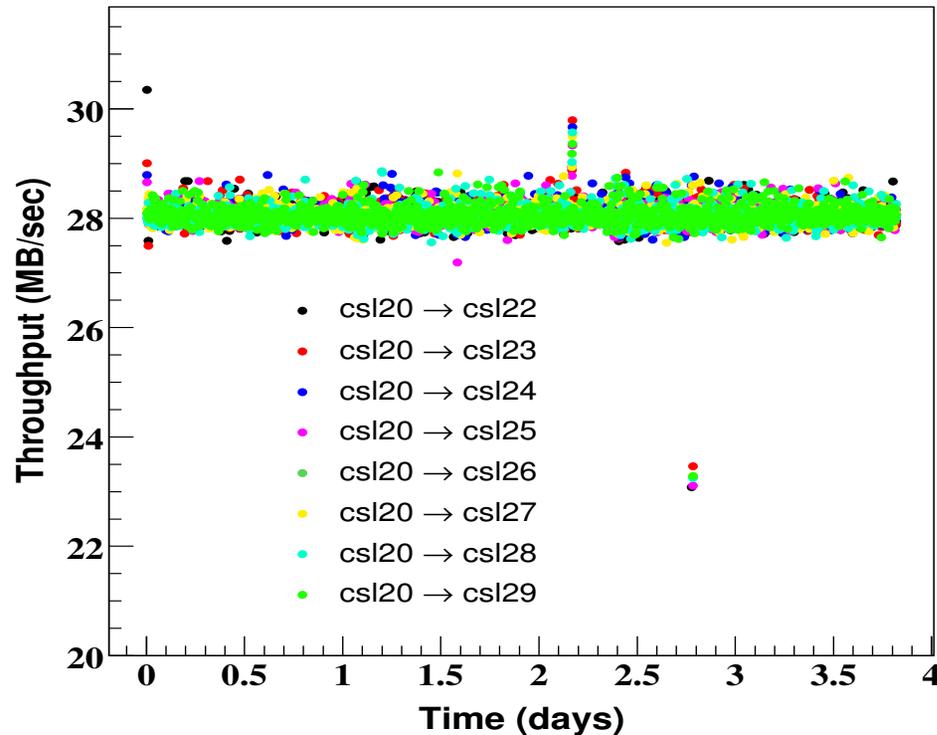
- Can increase number of logger nodes and disk buffer

Architecture Overview



The CSL Upgrade consists of everything in the dashed box...

Network Tests



Sending data from receiver node to *8 logger nodes*

Bandwidth is equally shared

Using two GigE links *bonded* together

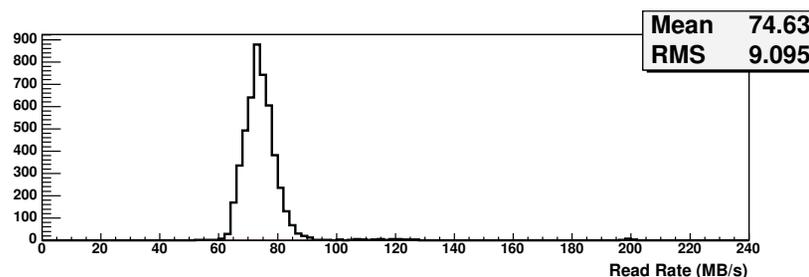
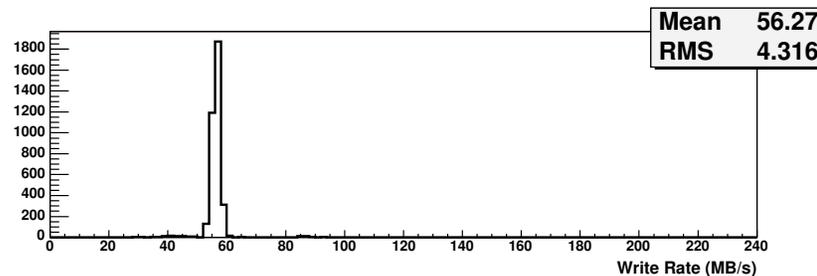
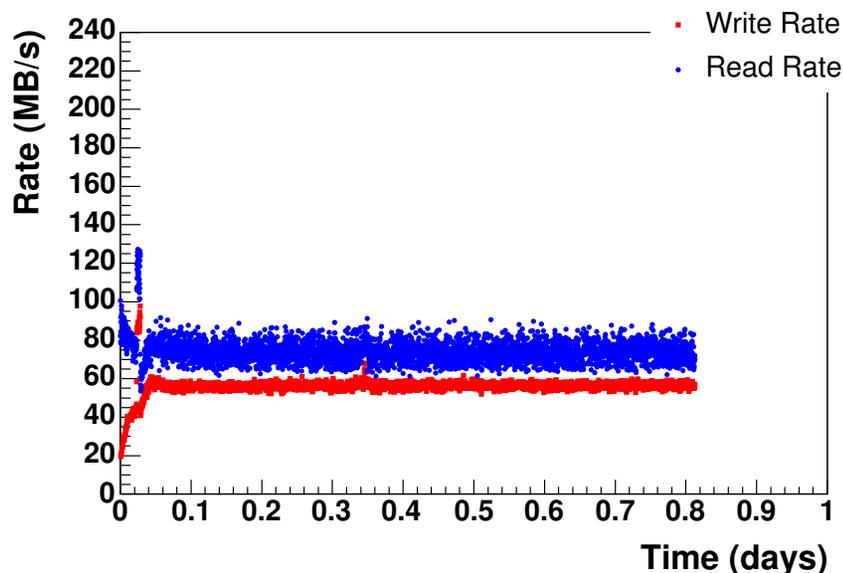
Total network rate: $8 \times 28 = 224$ MB/s

Verified that we get the expected full bandwidth when two nodes send to Receiver on separate input NICs while Receiver sends to two different Logger nodes on bonded output NICs

Disk IO Tests

Concurrent Read/Write tests on *all* 8 logger nodes

Bandwidth has been balanced now see the same IO performance across the 8 logger nodes.

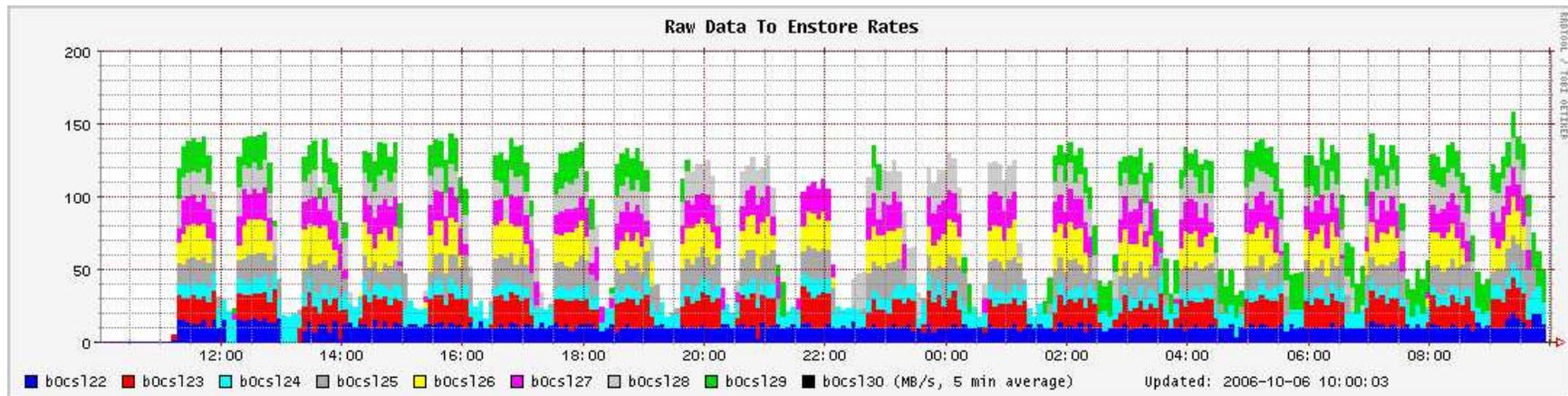


$$\text{Write} = 57 + 56 + 55 + 55 + 56 + 56 + 55 + 55 = 445 \text{ MB/s}$$

$$\text{Read} = 75 + 75 + 72 + 73 + 75 + 75 + 73 + 72 = 590 \text{ MB/s}$$

Dry Run Tests

Data was sent through the full CSL chain at about 100 MB/s continuously for more than 24 hours.



→ *We stressed the system well beyond our target of 80 MB/s and improved the performance and reliability.*

→ *During high rate testing revealed problems in the software that is already being used at D0.*

→ *Adapting solutions and unifying software*

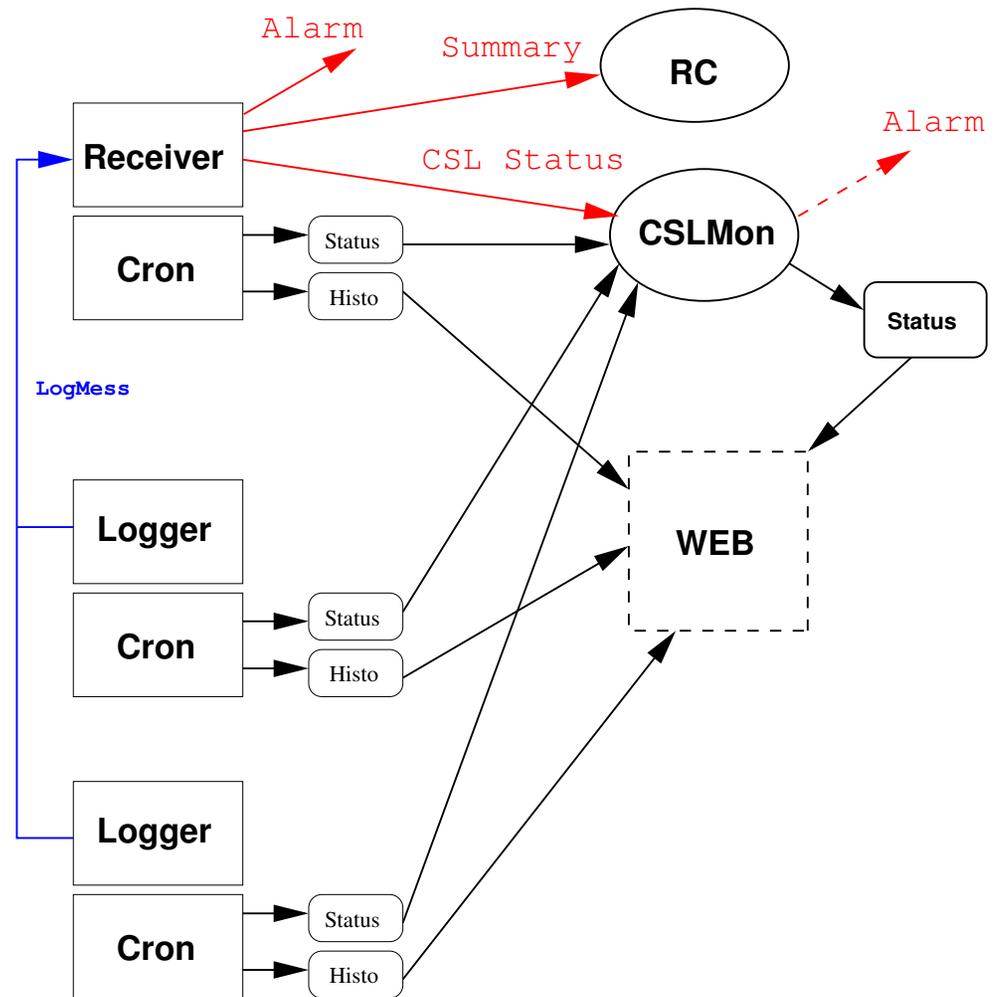
→ *Have not seen any hardware related issues since we finalized the configuration*

Monitoring

Same monitoring information that was available in the old system is available in the new CSLMon.

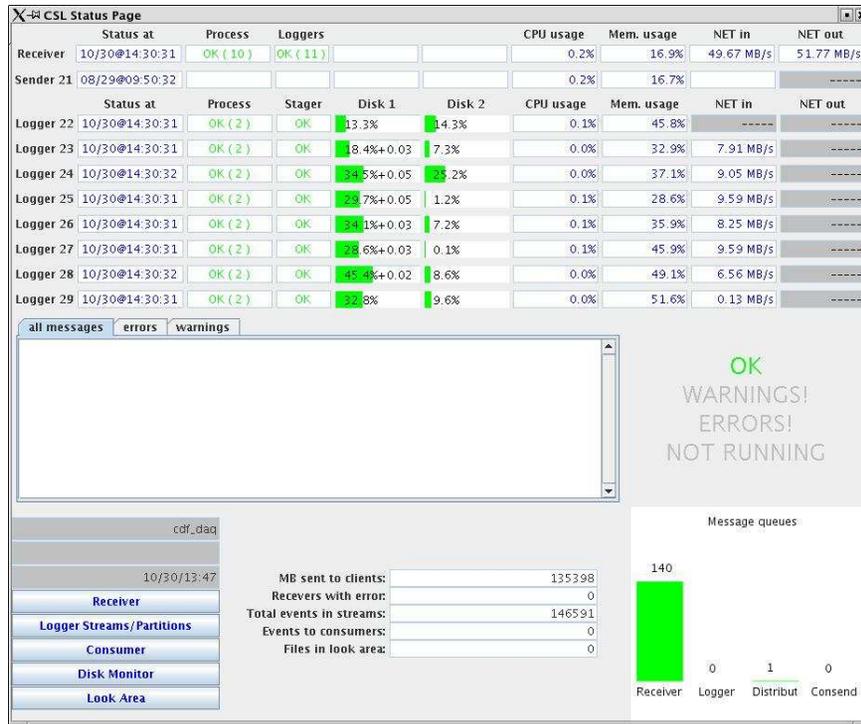
History monitoring plots available on the web

- Using rrdtools to generate plots
- Using ganglia for system monitoring



We have enough information to be able to monitor the performance of the CSL.

The New CSLMon...



Important information is grouped on the main page.

Shift crew does not have to open additional windows (Avoids information overload...)

Same information is available as before.

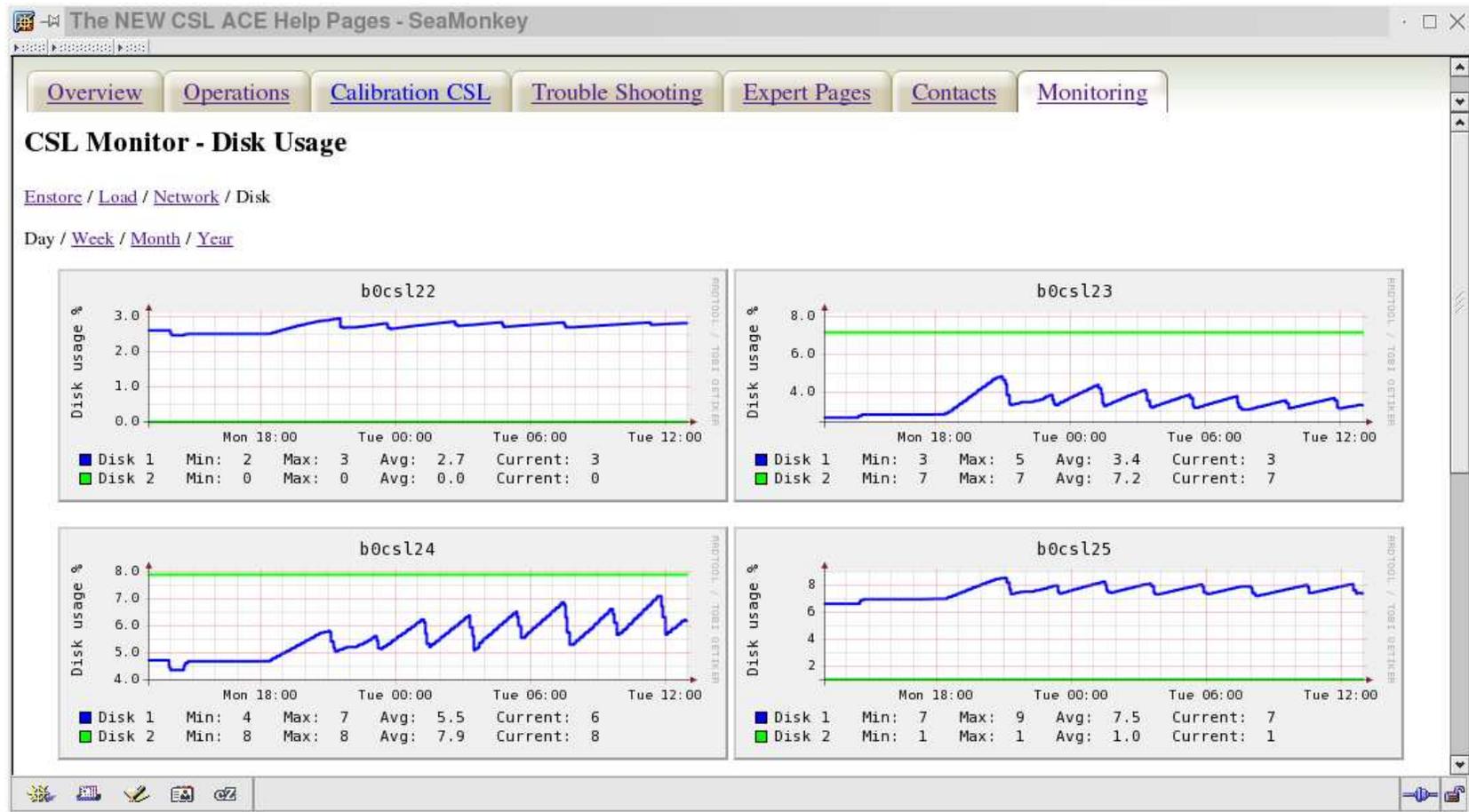
Other monitoring windows have the same look and feel as before.

- Receiver
- Logger Streams/Partitions
- Consumer

We would like to make this available on the web... but requires additional development

History monitoring plots available on the web

- Using rrdtools to generate plots
- Using ganglia for system monitoring

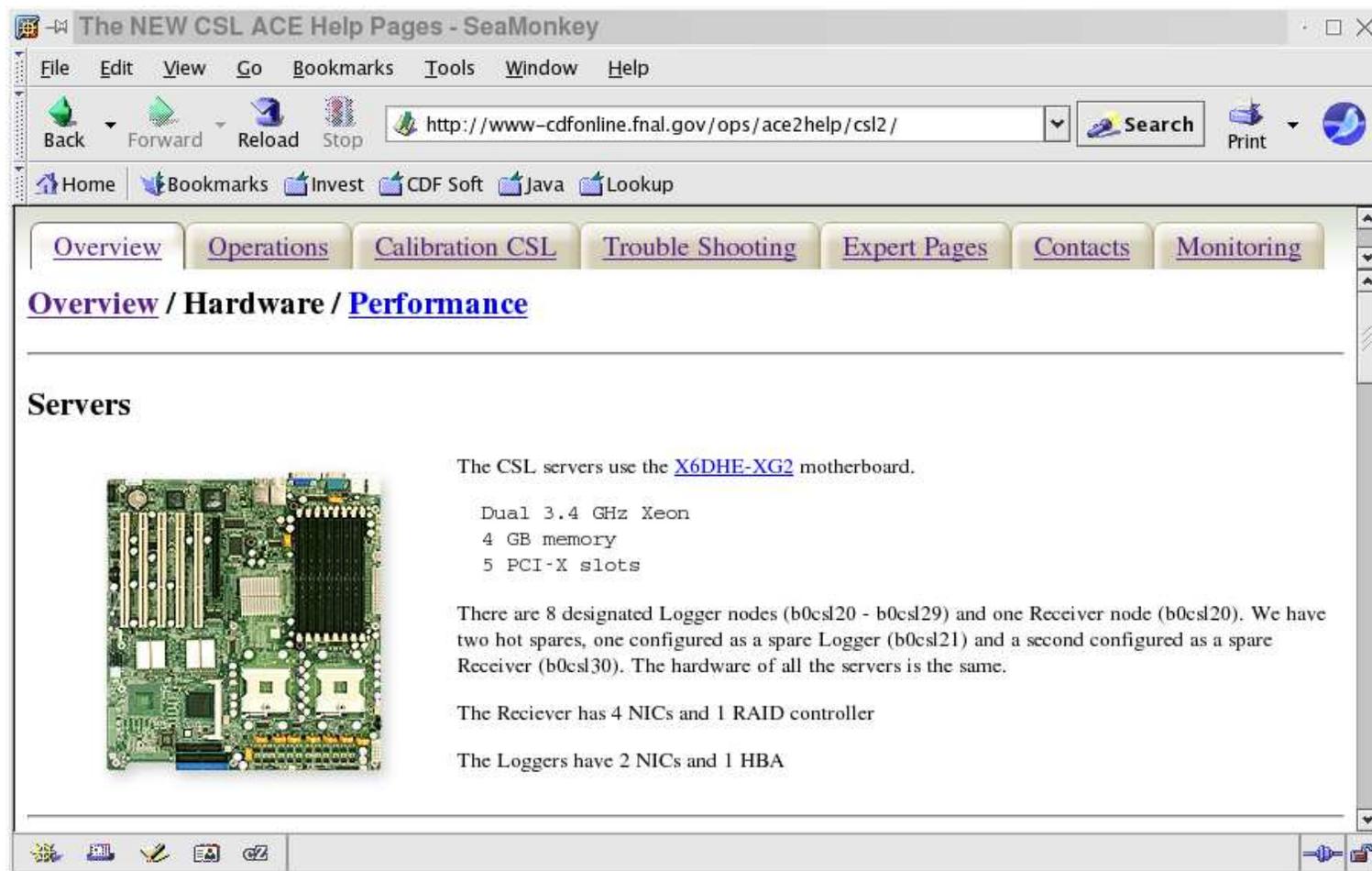


<http://www-cdfonline.fnal.gov/ops/ace2help/csl2/>

→ *Monitoring tab*

Documentation on the “Interwebs”

We updated operating instructions and “Expert Pages” with shift instructions as well as notes for ourselves.



The NEW CSL ACE Help Pages - SeaMonkey

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://www-cdfonline.fnal.gov/ops/ace2help/csl2/> Search Print

Home Bookmarks Invest CDF Soft Java Lookup

[Overview](#) [Operations](#) [Calibration CSL](#) [Trouble Shooting](#) [Expert Pages](#) [Contacts](#) [Monitoring](#)

[Overview](#) / [Hardware](#) / [Performance](#)

Servers



The CSL servers use the [X6DHE-XG2](#) motherboard.

- Dual 3.4 GHz Xeon
- 4 GB memory
- 5 PCI-X slots

There are 8 designated Logger nodes (b0csl20 - b0csl29) and one Receiver node (b0csl20). We have two hot spares, one configured as a spare Logger (b0csl21) and a second configured as a spare Receiver (b0csl30). The hardware of all the servers is the same.

The Receiver has 4 NICs and 1 RAID controller

The Loggers have 2 NICs and 1 HBA

Recovery from Hardware Failure

Very redundant and self repairing

→ In most cases recovery simply requires restarting the run

Deep buffers (24 TB → 3 days at 80 MB/s) *Plenty of time to react to down stream problems*

Disk Array (failed disk, controller, chassis)

RAID automatically rebuilds

Path to secondary LUN would still be available

Logger *Switch to hot spare (b0csl21)*

Receiver *Switch to hot spare (b0csl30)*

SANBox *Can withstand single box failure*

GigE Switch *Move to hot spare*

Tested switching to secondary LUN when primary LUN is full

Started using the new CSL on Oct 31 2006 (Store 5052)

Since then we lost 0.9% (1 pb) of data due to CSL related issues

- 438/1040 nb CSL testing while code developer is at the lab
→ *Requires starting stopping runs*
- 189/1040 nb Logging directory removed
→ *Reverted back to SGI CSL during diagnosis*
- 100/1040 nb COT data corruption lead L3 filters crashing and CSL stopped processing data
→ *Restart procedure failed because monitoring was holding shared memory segments*
- 93/1040 nb CSL logger message queue was full and stopped we handling data
→ *Restart procedure took a long time*

Working Through a Few Operations Issues

- Have a couple of problems requiring restarting the run

Bad Events (fixed?)

MAX Client Connection (related to “Bad Events” issue?)

→ *Developed a robust and fast recovery script*

- Orphaned Files, not closed and left in the .open area

→ *Mostly fixed but still an issue when the CSL is restarted...*

- We sometimes get truncated “db files”.

→ *Developed script to reconstruct the db file from the data. Cron job fixes the files and moves the data files to an output area so the stager can pick it up and copy to Feynman.*

- Min/Max event per Run Section Table not filled

→ *Need to update script and back fill table entries. In progress...*

Additional Development

There are a couple of **non-critical** outstanding issues:

- Want to get “ether channeling” to work between L3 Output and CSL Data switch, *boosts bandwidth from 120 to 240 MB/s.*
- Eventually write data to the new tape robot *In 2007...*

Long Term Support Issues

Rochester

- 1 Onsite Graduate Student
- 1 Remote Computer Professional (part time)

Tsukuba

- 1 Onsite Graduate Student

University College London

- 1 Remote PostDoc (part time)

Computing Division

System hardware support CD/Rex

(System Admin that did most of the configuration is now gone)

Stager/Logger

Data handling group

Not sufficient to provide ongoing support

SPL (*Willis is the acting SPL*)

Coordinate activities and address operational issues

CSL pager carriers

Should have at least 3

CSL Software support

No resident expert

Monitoring software support

No resident expert

→ I expect that this will be the place we will continued development as we gain additional operational experience

Stager/Logger pager coverage

Will need point of contact

System support

CD/Rex - *I think this should be ok...*

Summary

The new CSL is being used in production

→ Review committee has signed off and we have declared the system operational

Hardware Performing Well and is Robust

→ Can operate well above the target of 80 MB/s

→ Redundant system, quick recovery from hardware failures

Monitoring framework in place and provide enough information to understand performance issues

<http://www-cdfonline.fnal.gov/ops/ace2help/csl2/>

→ *Monitoring tab*

This is an area that would benefit from ongoing improvements

Documentation and Operating Procedures are in place

A Few Remaining Operational Issue are Being Addressed

- *Orphaned file checking (Implemented)*
- *Handle bad events (Fixed?)*
- *Truncated db file (Repair procedure)*
- *Fill missing database information*

Support Issues

- *No resident software experts*
- *Need to develop resident expertise for the monitoring software*
- *Need at least 3 people for pager support*

Remaining Tasks

- *Couple of non-critical tasks that need to be finished up.*

Propose to set Jan 8 as the day we disconnect the old SGI boxes from the internet

→ *Wait one week and then decommission the old CSL equipment*

Can now think of taking advantage of the increased bandwidth

What is the current Level 2 Accept limit?

→ *Is it 900 Hz? 1000 Hz? 1200Hz?*

Several downstream improvements have been made (VRB read-out, EVB...) *We are still limiting the trigger....*