
# Notes on power, space and networking interfaces

Interfaces between SURF conventional facilities and DUNE DAQ

Draft version

compiled by Giles Barr

21/9/2015

# Overview

1. Pictures of caverns etc. – from collaboration meeting.
2. Architecture variations
3. Location variations
4. Input from ‘Reference design estimate’ DOCDB 208.
5. Input from WA105/Dual-phase
6. Input from MicroBooNE – Georgia
7. Input from CERN Felix tests
8. CPU time tests

9. Summary of where we are; what happens next.

# Architecture variations

- Have tried seven different architectures:

1. RCE system – ZS and buffer in FPGAs, separate L1 trigger, full data for big triggers
2. Same as RCE system, but move buffering into computer over Ethernet/PCIe (will require changed architecture on COB)
3. Pipe all data to computer, use hardware acceleration way of doing ZS in computer
4. Pipe all data to computer, use software for ZS in computer (This is the Felix way)
5. Pipe all data to computer and across to trigger nodes, do all processing in trigger node (Like NOvA/SK)
6. Step back from mode 5, do lossy filtering near front end, then send all filtred data to trigger, like MINOS, NOvA, SK. (Lose ability to have big events non-zero suppressed
7. Like 3, but hardware gices ZS hints to computer to prevent it needing to read all the data.

Three of these are the best to study as representative: 1,(2,3,7), 4 [Ignore 5 and 6 as worse physics, no change in power consumption).

# Location Variations

# DOCDB 208

LATEST INFO fron DAQ - Information must be worked into plan!

This sheet from Giles

| Preliminary List of Under Ground Control Room Equipment-Per 10kt Detector (reference design) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| # of racks | Location | Rack Name | Rack Components | Space (U) | Quantity | Power (VA) | Power Tot (kVA) | Notes |
| 75 | FD | ChimneyDAQ[01:75) | SLAC RCE System | 5 | 1 | 160 | 12 | 2018 version of RCE (1COB/APA = 10 links/RCE) |
| 75 | FD | ChimneyDAQ[01:75) | Boardreader computer | 2 | 1 | 250 | 18.75 | |
| 75 | FD | ChimneyDAQ[01:75) | SSPs | 1 | 20 | 40 | 60 | Complete guess these SSP numbers |
| 75 | FD | ChimneyDAQ[01:75) | Slave Timing Unit | 1 | 1 | 40 | 3 | Assume rebuild |
| 75 | FD | ChimneyDAQ[01:75) | 24 Port 1G Switch | 1 | 1 | 100 | 7.5 | Spec sheet says 58W max for 24-port, BTU more like 100W |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| 1 | | Centre-DAQ | Trigger Farm Nodes | 2 | 20 | 300 | 6 | Assume 10 cores each |
| 1 | | | Disk nodes | 3 | 2 | 300 | 0.6 | Read Out Supervisor |
| 1 | | | 48 port 1G Switch | 2 | 4 | 100 | 0.4 | |
| 1 | | | 24 port 10G switch | 2 | 2 | 1000 | 2 | Guess |
| 1 | | | RC/SC/server nodes | 2 | 10 | 300 | 3 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | **Total (kVA)** | **21** | | | **113.25** | |

# DOCDB 208

| Preliminary List of Under Ground Control Room Equipment-Per 10kt Detector (PCIe/FELIX rough estimate) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **# of racks** | **Location** | **Rack Name** | **Rack Components** | **Space (U)** | **Quantity** | **Power (VA)** | **Power Tot (kVA)** | **Notes** |
| 75 | FD | ChimneyDAQ[01:75) | OpenCL/PCIe cards | 0 | 6 | 200 | 90 | Each rack is 80 links/APA. 32 channels/link. Each link 1.2GBit/ |
| 75 | FD | ChimneyDAQ[01:75) | Boardreader computer | 5 | 2 | 250 | 37.5 | 3 cards/computer |
| 75 | FD | ChimneyDAQ[01:75) | SSPs | 1 | 20 | 40 | 60 | Complete guess these SSP numbers |
| 75 | FD | ChimneyDAQ[01:75) | Slave Timing Unit | 1 | 1 | 40 | 3 | Assume rebuild |
| 75 | FD | ChimneyDAQ[01:75) | 24 Port 1G Switch | 1 | 2 | 100 | 15 | Spec sheet says 58W max for 24-port, BTU more like 100W |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| 1 | | Centre-DAQ | Trigger Farm Nodes | 2 | 20 | 300 | 6 | Assume 10 cores each |
| 1 | | | Disk nodes | 3 | 2 | 300 | 0.6 | Read Out Supervisor |
| 1 | | | 48 port 1G Switch | 2 | 4 | 100 | 0.4 | |
| 1 | | | 24 port 10G switch | 2 | 2 | 1000 | 2 | Guess |
| 1 | | | RC/SC/server nodes | 2 | 10 | 300 | 3 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | | | | | 0 | |
| | | | **Total (kVA)** | **19** | | | **217.50** | |

# DOCDB 208

**electronics power consumption computed for a single 10 kton double-phase module.**
**from email - Dario Autiero (June 25, 2015)**

-MicroTCA crates with digitization boards (10 AMC boards/crate, 64 ch/board, 640 channels/crate): 240 units x 0.6 kW/unit = 144 kW
-Front-end: 20W x 240 = 4.8 kW
-Back-end OPENCL cards+PCs: 20 units x 1 kW/unit = 20 kW

Grand total: 144 kW (DAQ) + 4.8 kW (FE) + 20 kW (BE) = **168.8 kW**

# CPU testing

- There are two critical parts for the CPUs
  1. In some schemes, the entire data is received in the CPU cores and the zero suppression is done there. The CPU time is likely dominated by the initial read and decoding of the data. We are timing this:
     a. Justo Martin-Albo (New postdoc at Oxford) Measured 1GB/s reading speed on laptop – will try the timing program on DAQ computers in parallel today
     b. Tom Junk 2GB/s reading speed on desktop
  2. Trigger farm: More tricky, probably mostly hit counting in a sliding window from a list of hit count data, so read-in operation will still be important, but more computation needed as well.

So we should be capable of firming up the numbers on the number of CPU nodes needed soon

# Summary

- Still uncertainty on the number of CPUs, especially in the designs that use CPUs heavily.
- It is already clear that the designs with FPGAs use less power by a big factor than designs that rely heavily on CPUs
- The surface is an easier environment than the CUA and the CUA is easier than on top of the cryostat.
  - Of the seven architecture options we have, only three are dramatically different from the power point of view
  - There are nine combinations of locations, grade them by ease of providing the power/cooling
- Make table of 'architecture' vs 'location': List the physics capabilities that are lost, and cost difference for each one.
  - It is fairly clear-cut for the networking that provided we don't send all the non-ZS data to the surface, the 96-fiber cable is sufficient. But we should do a properly documented version of this argument.
  - Power and central facility space is much more problematic to decide. Need guidance from LBNF of cost differences.