The D0 Level 3
Data Acquisition System

Doug Chapin
Brown University

Outline
• Overview
• Monitoring

For the D0 L3DAQ Group
Brown University
Fermilab
University of Washington
D0 Data Acquisition System

Detector Component Data

Analog L1 buffer

Digital L2 buffer

Level 1 Trigger

Level 2 Trigger

Custom hardware

Offline Storage

Col./Router Datalogger

Level 3 Trigger

Level 3 DAQ

250kB / event

7.6MHz

5kHz

1kHz

50Hz

Commodity processor farm

Commodity hardware

Accept

Accept
**L3DAQ Requirements**

**Total of 63 VME Readout Crates**
- 1-10 vme modules / crate
- 1-20kB total / crate / event

**L3 Trigger Processor Farm**
(currently 82 nodes)

1KHz  250kB/event
Commodity-Based System

CISCO 6509 Ethernet Switch

All ethernet TCP/IP

Supervisor CPU

COOR

Farm CPU

Farm CPU

Collector/Router

Currently 82 farm nodes

SBC = single board computer

ROC

SBC

ROC

SBC

ROC

SBC

ROC

SBC

ethernet concentrator

ethernet concentrator

ethernet concentrator

Gb fiber

100Mb

Global Trigger Framework (TFW)
Communication Flow

Features
- Apply backpressure through TFW interface
- Multiple simultaneous runs (D0 requirement)
- Farmnode degeneracy

Software
- Linux
- TCP/IP implemented via ACE communication and utility library
- Open and multi-platform

Doug Chapin (Brown University)  DAQ Jamboree - 4 Aug 2003
Network Switch and Farmnodes

Network
- CISCO 6509
  - Really fast backplane
  - 1MB buffer per output port
- Concentrators
  - 100Mb -> Gb fiber
  - 100Mb/s ethernet in SBCs

Farmnodes
- Dual Processor
  - AMD 2000 and PIII 1GHz flavors
- Dual ethernet (100MB/s)
  - Connections to L3DAQ switch and Online switch
VME Readout Crate

Typically one to ten data cards per crate (1 to 20kB total)

Data Card

Data Card

VME block transfer(s)

SBC-initiated

SBC

ctrl reg programming (VME)

CPU Card

Controller Card

L2 Accept

J3 Slave Ready

Done

Data card, CPU, and Controller cards specific to crate type
**Single Board Computer**

- VMIC-7750
- PIII 933MHz
- 128MB RAM
- Dual 100Mb/s ethernet (~24MB/s)
- Custom 9U extender
- 128MB Flash Disk
- Tundra Universell PCI<>VME
- PMC Digital IO Card

**Custom Driver**
- IO card
- J3 handshake
- VME transfers
- Event fragment buffering (12MB)
Supervisor Operation

Primary function
L3DAQ interface to COOR

Supervisor is only active during:
STOP/START
PAUSE/RESUME
DOWNLOAD

• Passes L3 trigger config to Filtershell/Scriptrunner
• Determines which nodes are assigned to a run
• Passes L1/L2 trigger-crate-farmnode association to RM
• No communication with SBC or EVB
• Automatically reconfigures RM/nodes if nodes fail/recover
SBC Operation

- Match event numbers in head Fragment and head Route Tag
- Send to appropriate nodes if match
- Discard fragment (or Route Tag) if mismatch
  - Automatic re-sync

Event fragments from VME → Event Fragment Queue (50) → SBC Process → Route Queue (100) → Route Tags From RM

Fragment queue and VME operation handled by driver module

Route Tag contains:
- event number
- list of dest. nodes
EVB Operation

- Fragments combined, keyed on event number
- Expected crate list received from RM
- Event is incomplete after 1 sec timeout
  - Associated crates and nodes appear **RED** in fuMon/uMon
- Free buffer count sent to RM when necessary
  - 20 total buffers, but only max of 3 advertised to RM
Routing Master

Operation
• Dedicated SBC
• Routing decision
  • Based on run configuration and L2 trigger decision
• Sent to SBCs and target nodes
• Apply backpressure to TFW when too few farmnode buffers

Trigger Framework Interface
• VME cards
• Hardware FIFO
  • Event number
  • L2 accept trigger bit mask
  • L1-disable bit registers

Routing Info to SBCs: event# and destination farmnode(s)

Expected crate list to farmnodes

Free buffer tokens from farmnodes

Routing Master

FIFO

TFW Interface

Event number/Trigger bit fired mask

Individual trigger bit disable
Routing Master Disable Logic

Global Disable
- Disable all 128 L1 triggers
- Even if triggers are in SDAQ run!
- Occurs when RM gets programmed by Supervisor
- START/STOP/PAUSE/RESUME transitions
- Usually short (few seconds)

Routing Group Disables
- Disable all L1 triggers associated with a run
- When <16 total free buffers available on farmnodes in that run
- Re-enable when >24 buffers available
- Since farmnodes only advertise max 3 free buffers
- **Your trigger list should assign at least 10 nodes to your run**
  - Exception is CAL calibration run
  - Otherwise disables affect all other PDAQ runs
Event Buffering

Routing Master
- Buffer 10 event tags (routing info) before sending to each SBC
  - Minimize ethernet overhead
  - Without buffering: 63 crates * 1kHz = 60,000 packets/sec

SBC
- Buffers up to 50 event fragments (before routing)
- 10 for RM event tag buffer + 40 for TFW FIFO depth
- Large (1MB) TCP send buffer

Farmnode Event Builder
- 20 buffers (event processing)
- Only send 3 max free buffer tokens to RM
  - avoid overflowing SBC send buffer (max frag size is 256kB)

6509 Switch
- 1MB output buffer per port
- Farmnode TCP receive window set to 10kB
  - 10kB * 80 connections = 800kB (less than buffer size)
- Avoid dropped packets
Monitoring

Dedicated monitor server
• Pulls info from clients (SBC,EVB,RM) and caches it

Main shifter displays
• **uMon**: SBC and RM info
• **fuMon**: EVB and some Filtershell info
• Common color scheme
  • **Red**: incomplete events
  • **Yellow**: no monitoring information
    • crashed process?
    • Network problem?
• **White**: OK

Read the help webpages for **uMon** and **fuMon**

uMon even has a help button
<table>
<thead>
<tr>
<th>Node 16 Details</th>
<th>CPU: 35%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run 176930</td>
<td></td>
</tr>
<tr>
<td>Im: 3.3Hz</td>
<td></td>
</tr>
<tr>
<td>Out: 0.3Hz</td>
<td></td>
</tr>
<tr>
<td>&lt;188k&gt;</td>
<td></td>
</tr>
<tr>
<td>IP: 0</td>
<td>AD: 3</td>
</tr>
<tr>
<td>COM: 0</td>
<td>FB: 20</td>
</tr>
<tr>
<td>Fail: PS[0]</td>
<td>PS[1]</td>
</tr>
<tr>
<td>Total: 44313</td>
<td>54331</td>
</tr>
<tr>
<td>Failed: 44313</td>
<td>54331</td>
</tr>
<tr>
<td>Passed: 44313</td>
<td>54331</td>
</tr>
<tr>
<td>Sent: 44313</td>
<td>54331</td>
</tr>
<tr>
<td>SendErr: 0</td>
<td>0</td>
</tr>
<tr>
<td>StopReq: 8</td>
<td>8</td>
</tr>
<tr>
<td>Restarts: 0</td>
<td>0</td>
</tr>
<tr>
<td>ExitS: 0</td>
<td>0</td>
</tr>
<tr>
<td>State: WAIT</td>
<td>WAIT</td>
</tr>
</tbody>
</table>

Out/Pass (Reject) 100Hz
36Hz/10% (9.6)

CPU Utilization

L3DAQ Farm Über Monitor
fuMon Details

In-Process buffers

Complete buffers

Advertised buffers

WAIT

FILTER

FLATTEN

SEND
Please read the fuMon and uMon documentation on the L3DAQ webpage
L3DAQ Webpage

http://d0ol/www/groups/l3daq/

- What-to-do-When (WTDW)
  - Common problems and solutions
- Logfiles for SBCs
- uMon and fuMon documentation (read these!!!)
- SBC Manager Page
  - Reset SBCs at the click of a button
Common Scripts

l3xdaq_reset
  • Restart SBC processes (and RM)

l3xreset / l3xstop
  • Restart / stop EVB and Filtershell processes
  • l3xstop useful for removing flaky nodes

More at the WTDW page
Summary

D0 L3DAQ is best operating DAQ in HEP?

Please

• Visit the webpage
• Read uMon and fuMon documentation
  • Phone Doug or Mike with questions
• Read WTDW before encountering common problems
• Email general suggestions/complaints to d0daq@fnal.gov