Unified Software Framework for Upgraded Belle DAQ System

R. Itoh, T. Higuchi, M. Nakao, and S. Y. Suzuki
KEK

presented by Ryosuke Itoh

Computing in High Energy and Nuclear Physics 2006
Online Computing Session
Feb. 13, 2006
Mumbai, India
Outline

1. Introduction

2. Belle DAQ upgrade

3. Software Design

4. Implementation
   - Frontend Readout Module
   - Readout PC
   - Event Builder
   - Reconstruction Farm

5. Summary
1. Introduction

**KEKB Accelerator**

- **KEKB B–Factory**: 8.0 GeV e–, 3.5 GeV e+
- **Belle Detector**
  - Electromagnetic Calorimeter: CsI(Tl) 16$X_0$
  - TOF counter
  - Si Vertex Detector: 4 layer DSSD
  - Central Drift Chamber: Tracking + dE/dx, 50-layers + He/C$_2$H$_4$
  - Aerogel Cherenkov Counter: $n = 1.015$~$1.030$

**Belle Experiment**: B-factory experiment at KEK in Japan to study CP violation in B meson decays.
- Precise determination of unitary triangle
- Discoveries of new particles: X(3872), Y(3940)
- New physics search in radiative decays: $B^0 \rightarrow K^*l^+l^-$, $b \rightarrow d\gamma$
Recorded luminosity $\sim 550$ fb$^{-1}$
(as of Feb. 07, 2006)

Peak Luminosity $> 1.6 \times 10^{34}$ cm$^{-2}$ sec$^{-1}$
Daily luminosity $> 1$ fb$^{-1}$
2. Belle DAQ Upgrade

Belle DAQ system until 2005 summer

* Q-to-T conversion + FASTBUS TDC, Switchless Event Builder
* Ave. Trigger Rate ~ 500Hz, Event Size ~ 40KB
  Intrinsic dead time ~ 4% ( + 3.5% due to injection VETO for continuous injection)
* RFARM for real-time event reconstruction
DAQ intrinsic deadtime (injection VETO(3.5%@10Hz) is NOT included)

Crab cavity installation is scheduled in Mar-Apr 2006.

L will reach \( > 2\sim3 \times 10^{34} \text{ cm}^{-2}\text{sec}^{-1} \) in a few years

\[ \text{Trigger rate} > 1\text{KHz} \]

\textbf{DAQ deadtime} > 20% which is unacceptable.

+ SuperKEKB upgrade => up to 30KHz!

1. Upgrade of readout system to \textit{pipeline-based system}
   * Unified pipeline readout module has been developed.

2. Upgrade of Event Builder/RFARM also.
   * To remove 650Hz barrier
Global Design of DAQ upgrade (toward SuperKEKB)

3-stage event building

Pipeline Readout Modules

Stage-1  Stage-2  Stage-3

Sub detectors

Readout modules  > 1000
Readout PC  > 50
Event Building units  > 10
HLT farms

Transfer Network Matrix

100base-TX

PC

mass storage
Unified Pipeline Readout Module (COPPER)

Digitizer cards (implemented as daughter cards)

Form factor = VME 9U

CPU card (operated by Linux)

RadiSys EPC-6315
- Intel PentiumIII 800 MHz w/ 512 MB memory.
- Network booted
- RedHat Linux 9

two100base-TX ports
(for control and data flow)
Transfer Network Matrix

Readout PCs (~50)  Transfer Network Matrix

Event builders (~10)
Partially upgraded Belle DAQ (2005 fall - )

COPPER readout for EFC
(for the system test in beam)
EFC COPPER crate
COPPER TDC:
made compatible with LeCroy 1877S using FINESSE cards with AMT chips.

The COPPERs are already on operation from exp#45.

Before Q/T cabling
Transfer Network Matrix + Event Builder 1 and 2
2\textsuperscript{nd} unit of RFARM

Each server houses dual Xeon in 1/2U
\rightarrow 4 Xeon CPUs in 1 U

40 x Dual Xeon servers (3.4GHz)
3. Software Design

[DAQ system from software viewpoint]

* point-to-point TCP/IP based data transfer everywhere.
* Linux operated PC is used in all DAQ components (inc. readout modules).
* 3-stage event building ... repeated use of similar processing

"Building block" construction of DAQ software

readout crate  readout PC  event building farm  recon. farm

readout program  evt. build. program  transmitter  receiver
Considerations

Variety of complicated processing is performed on each node.

* Wave form sampling / feature extraction on readout modules.
* “Level 2.5” trigger on first nodes of event building farms.

Data reduction with node-by-node optimized algorithm on all nodes.

Readout program / Event building program:

→ Should be capable of accepting variety of data processing codes.
  - Such codes are written by many people some of those are not familiar with DAQ software.

Needs an offline-like programming environment for DAQ which does not require DAQ-specific knowledge.
Idea: DAQ software based on offline analysis framework

Offline analysis framework: B.A.S.F. (Belle AnalysiS Framework)

* Well established in Belle and everyone is familiar with it.
* User code can be linked with the framework as modules by a dynamic link.
* Users don't have to take care of data I/O.
  ← data handling through data management package

Readout program / Event Building program:
* B.A.S.F. with DAQ modifications.

Data flow between nodes:
* Socket I/O implemented in B.A.S.F.
Green boxes are linked using dynamic link as well as "modules".
* Data handling is done through "Panther" package.
B.A.S.F. with DAQ mods.

Socket I/O between nodes: based on unified “d2packet” protocol
Slow control

NSM (Network Shared Memory) is used.

* Capable of
  - shared memory handling over network
  - message passing between nodes
    ↦ asynchronous handling by hooked-up action functions
* DAQ control is done through message passing from one MASTER node to many client nodes.
* Support for hierarchical network structure through functional master.
Unified DAQ Software Framework

Input flow:
- d2packet
- event building package
- shared mem
- output server w/ tx's

Output flow:
- receivers
- shared mem
- spawn/terminate

NSM Interface

Control Net
4. Implementations

a) Readout Module (COPPER)

- Special version of B.A.S.F. is used, "cbasf" : light-weight B.A.S.F.
  * parallel processing capability is removed.

- FINESSE(digitizer) cards are accessed through a device driver with standard system calls (open(), close(), read(), write(), and ioctl()).

- Data from digitizers are formatted into Panther at the very beginning.

```
cbasf

**control net**

**module chain**

```

- Initialization of special devices are handled by modules.
b) Readout PC

Dual Xeon server (RH9 Linux) → SMP with 4 CPUs

COPPER crate (9U VME)

- Event fragments from COPPERs are received by a set of rx processes and placed in a buffer on a shared memory according to event number.

- Buffers with all event fragments filled are sent to event process by the “event server” by passing pointers.

- “Output server” sends processed outputs to event builders in turn through sockets.
c) Event Builder

- Repeated use of software for Readout PC: **Unified framework**

Dual Xeon server (RH9 linux)

Software on single node

Event Builder Unit

Level 2.5 software trigger is implemented as a B.A.S.F. module
d) Reconstruction Farm

Nodes are dual Athron or Xeon servers operated by Linux (RH7.3/9 and WBL3)

Event Builder

sock2rb:
receive an event from socket and place it on ringbuf

rb2sock:
pick an event from ringbuf and send it to socket

RingBuffer:
Ring buffer on shared memory

rb2file:
save data in a file

input distributor

processing nodes

output collector

up to 85 nodes

Each node: controlled by NSM in a closed net thru. functional master on a control node.
5. Summary

* Upgrade of Belle DAQ system is on going to keep up with a luminosity increase in coming years.

* A unified DAQ software framework is developed based on Belle's offline analysis framework combined with a socket I/O and event building package.

* The framework is used even on every front-end readout module.

  provides easy and consistent environment for the development of DAQ software

* The partially-upgraded Belle DAQ system is now being operated with the unified software framework and is working stably.
Backup Slides
EFC COPPER Integration

1 COPPER for CDC
6 COPPERs for EFC

KEK-VME 9U crate

Readout PC
Trigger Timing Distribution

Trigger distribution timing from TTRX to FINESSE is tunable a.l.a HPGG.

Now, 3\(\mu\)s additional latency.