DAQ and Online

WBS 1.3

S. Fuess, Fermilab

Introduction

- Outline
  - Organization
  - Structure of DAQ/Online system
    - Definition / scope
  - Issues & Motivations
  - Changes for Run IIb
  - Details of component upgrades
    - Requirements and plans
  - Cost, Schedule, Risk Analysis

Organization

D0 Run IIb Project
J. Kotcher, Project Manager
R. Partridge, Deputy; V. O'Dell, Associate; W. Freeman, Assistant
M. Johnson, Technical Coordinator
C. Yoshikawa, Budget Officer; T. Erickson, Administration

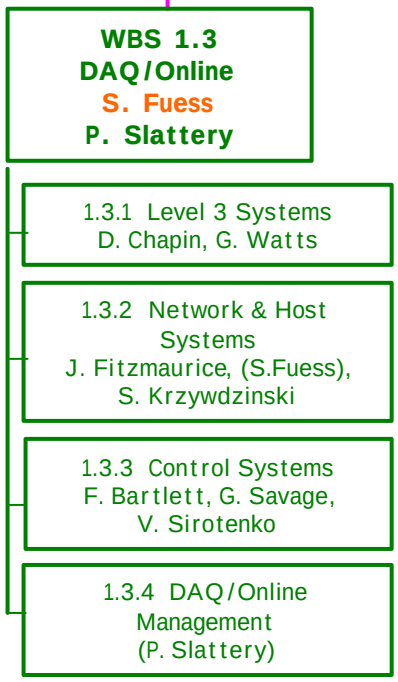
Online group has excellent record of providing working network, host, and control systems, fully operational since earliest days of RunIIa commissioning

- Much credit goes to Computing Division for their contributions

In 12/2001 group reorganized to include new DAQ (readout and event building) initiative

- Successfully designed and installed state-of-the-art commodity-based DAQ system now in active use
  - kHz capable system ready for RunIIb
- Special recognition due G. Brooijmans for leadership of project, Brown Univ., Univ. of Washington, and to Computing Division for critical support and active participation

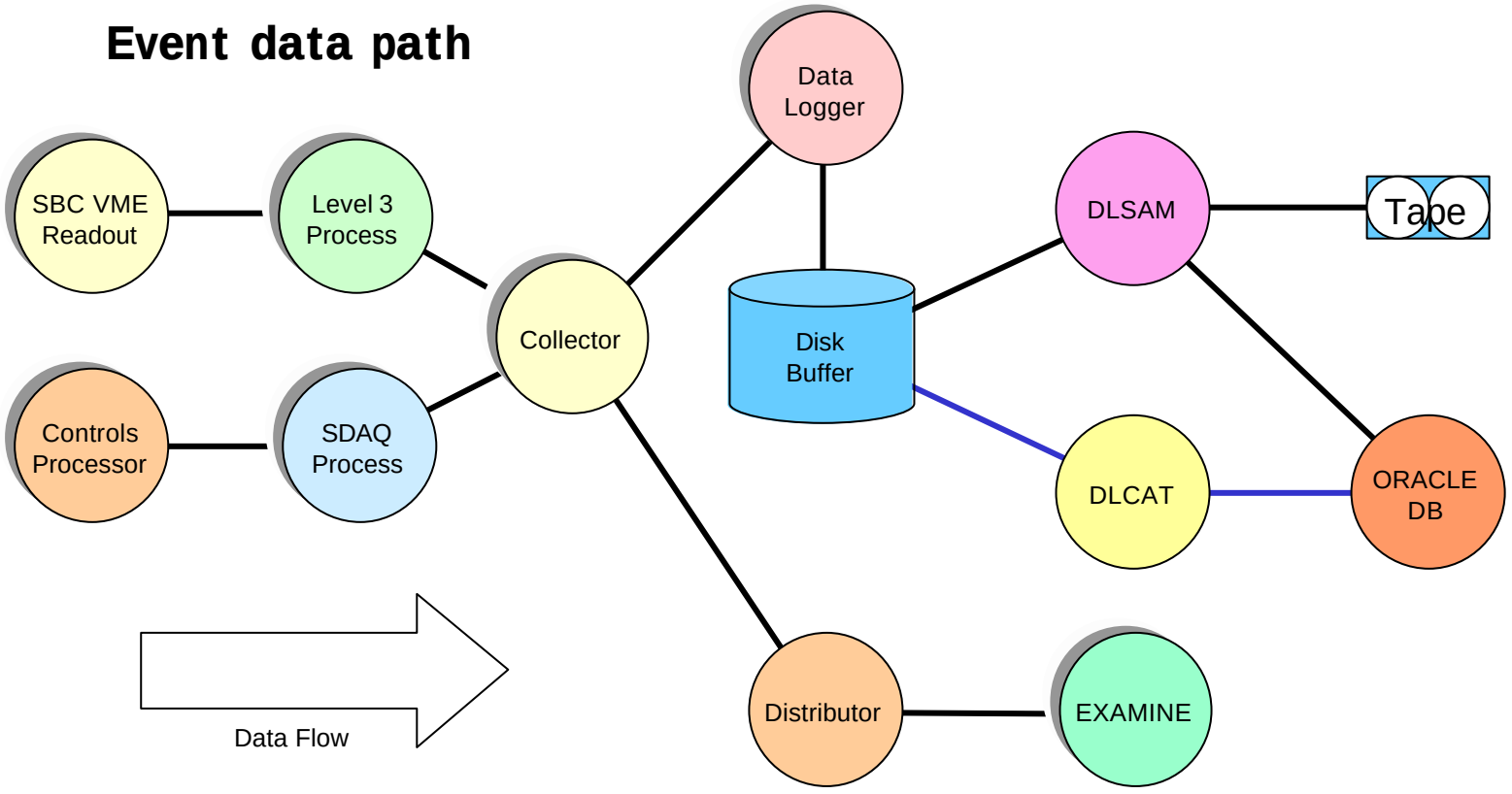
DAQ component of RunIIb efforts led and supported by this same group of experts



Structure of DAQ/Online

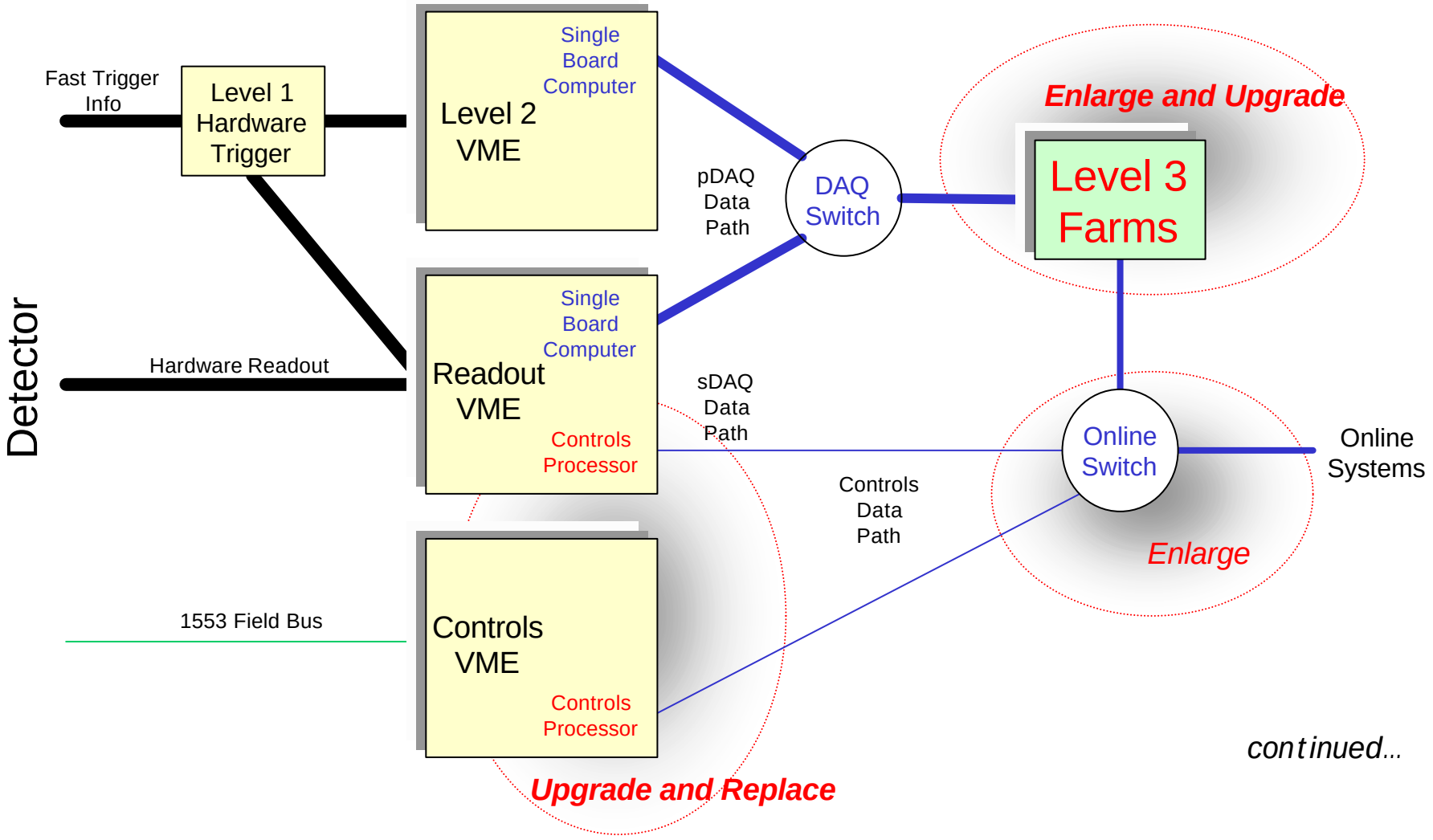
- **Scope**
  - **The DAQ/Online system consists of:**
    - **Event path:**
      - DAQ and Online network
      - Single Board Computers (SBCs) in VME readout crates
      - Level 3 software filter farm
      - Host systems for data transport and logging
    - **Control and Monitoring:**
      - Control room systems
      - Event data monitoring systems
      - Slow control system, including VME processors
    - **Infrastructure:**
      - Database servers
      - File servers
    - **Software associated with the above**
  - **See supporting documents (TDR, BOE, Schedule) also**

Structure of DAQ/Online



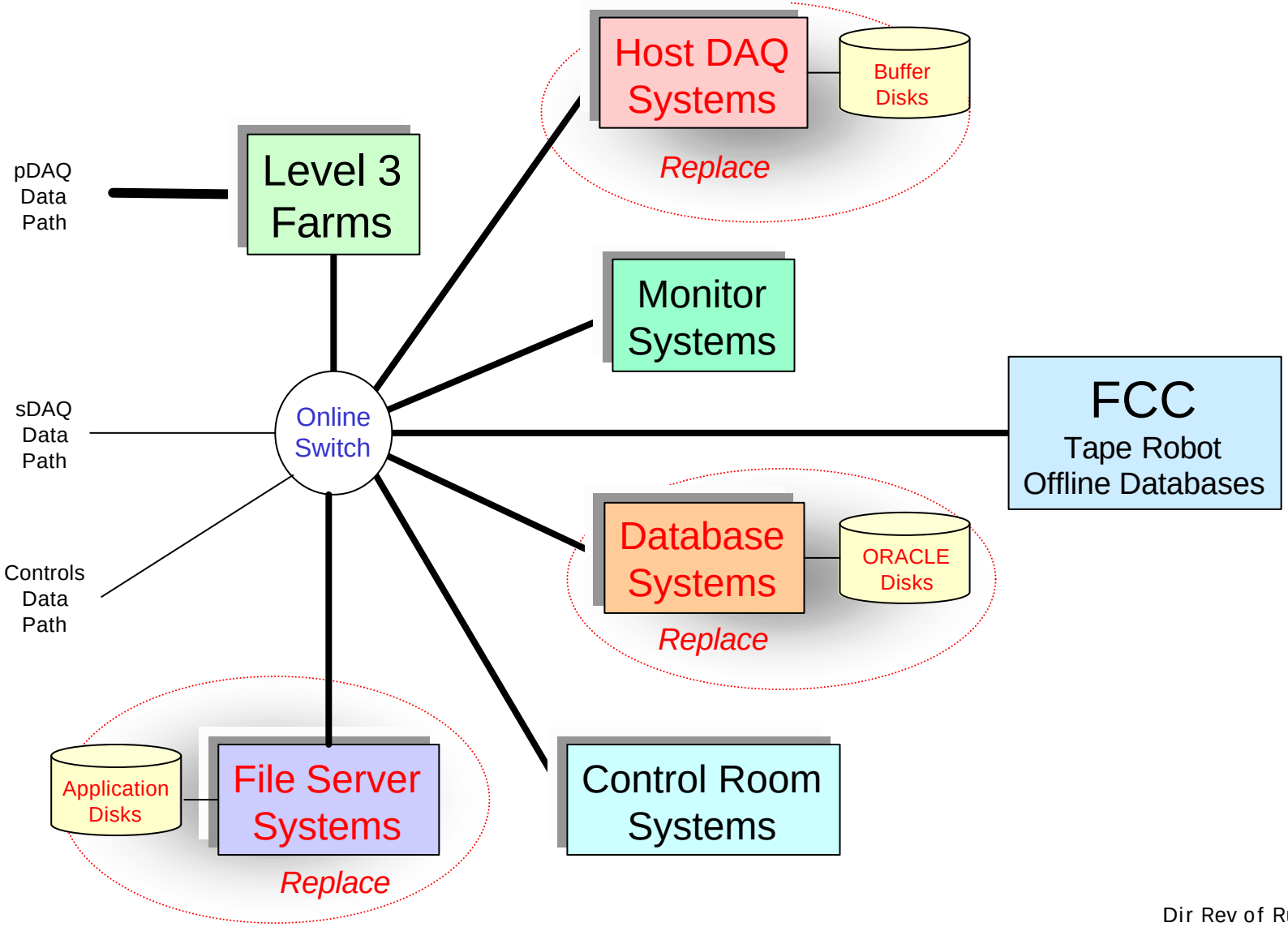
SOFTWARE architecture is unchanged for Run IIb

Structure of DAQ/Online



HARDWARE <u>architecture</u> is unchanged for Run IIb
But many <u>components</u> change

Structure of DAQ/Online



Issues & Motivations

- **Meet RunIIb performance requirements**
  - More difficult trigger environment will require increased Level 3 processing (more later…)
  - Greater capacity to manage Offline data may invite higher peak recorded event rate (up to 100Hz instantaneous)
- **Hardware maintenance**
  - Many Run IIa systems will be reaching end of life expectancy
    - Host / Database / File Server systems purchased between 1997 and 2000
  - Maintenance more costly than replacement
  - Replacement also addresses capability needs

Issues & Motivations

- **Software support**
  - Operating system support for Run IIa Host systems a potential problem
    - Currently use ~~DEC Compaq~~ HP Tru64 cluster
  - ORACLE support by Computing Division consolidating
  - Great care already taken in Run IIa software development to be OS independent
    - Linux is the strategic platform

Issues & Motivations

- **Reliability!**
  - **Target is > 99% system availability**
    - **Shown to be possible in Run I**
    - **Compare system costs to cost of lost beam**

- **Simplicity**
  - **Systems need to be maintained for Run IIb era**

Changes for RunIIb

- Systems to be discussed
  - 1.3.1 Level 3 filter farm processors
  - 1.3.2 Host systems
    - 1.3.2.4 Storage systems
    - 1.3.2.6 DAQ hosts
    - 1.3.2.7 ORACLE database systems
    - 1.3.2.8 File servers
  - 1.3.3 Control systems
- TDR and Schedule also track continuing RunII operations
  - Useful for resource (effort) allocation
  - Otherwise not mentioned in this talk
- Note:
  - Any costs shown are unburdened current FY $

Component Details

- **Level 3**
  - **Processing time requirements:**
    - **At $2x10^{31}$ $cm^{-2}s^{-1}$ current filter code, including global tracking, requires ~300 msec/event on 1 GHz PIII**
      - **Maps to 250 msec/event for single interaction, equivalent to 7.5 SpecInt95 secs/event**
    - **Extrapolation depends upon complexity of event, a function of luminosity and bunch crossing time**
      - **For L3 filter code, processing most dependent upon occupancy. Assume linear rise in requirements with number of interactions.**
      - **For 1 primary + 5 background interactions per crossing and 1 KHz L3 input rate, predicted CPU requirement is 45K SpecInt95**
    - **Precise requirements difficult to predict**

Component Details

- **Level 3 (cont'd)**
  - **Acquisition plan:**

	Run IIa			Run IIb	
Year	FY01	FY02	FY03	FY04	FY05
Added nodes	48	32	32	32	64
Node type	Dual 1 GHz PIII	Dual 1.67 GHz AMD	2003 model 1.6x previous	2004 model 1.6x previous	2005 model 1.6x previous
Clock factor	1.00	1.67	2.67	4.28	6.84
Equivalent Dual 1 GHz PIII	48	53	85	136	437
Retired nodes				48	32
Node type				Dual 1 GHz PIII	Dual 1.67 GHz AMD
Clock factor				1.00	1.67
Equivalent Dual 1 GHz PIII				48	53
Net number of nodes/ports	48	80	112	96	128
Net equivalent Dual 1 GHz PIII	48	101	186	274	658

    - **658 Dual 1 GHz PIII → 39K SpecInt95**
  - **96 nodes @ ~$2.2K = $210K**
    - **Staged in FY04 (32), FY05 (64)**

Component Details

- **Host systems and storage**
  - **Architecture**
    - **Fibre channel SAN with RAID arrays for Database and File Servers, JBOD for data buffer disks**
      - **1.3 TB RAID**
      - **4 TB JBOD (48 hours of buffer)**
      - **FC fabric or AL interconnect**
    - **Redundant (primary/secondary) small servers (2-4 CPU) for each of:**
      - **Data logging system**
      - **ORACLE database system**
      - **File server system**

Component Details

- **Host systems and storage (cont'd)**
  - **Acquisition plan:**
    - **Storage**
      - **SAN infrastructure** **$20K**
        JBOD buffer disk and chassis **$40K**
        RAID disk arrays and controllers **$43K**
      - **Early R&D (FC on Linux) effort, then purchase FY04**
    - **Servers**
      - **Secondary servers & infrastructure (and Primary File Server)** **$20K each**
        Primary servers & infrastructure **$32K each**
      - **Secondary server purchase FY04**
        Primary server purchase FY05
        - **As late as possible to get best performance**

Component Details

- Control systems
  - Processor upgrades
    - Run IIa systems include ~100 embedded VME processors for downloading, control, monitoring, and diagnostic readout of the detector
      - Motorola 68K and PowerPC family
    - Concerned with longevity and maintenance / replacement costs
    - Strategy is to replace ~50% of processors
      - Target those with largest processing loads
      - Remainder become spares
      - 53 nodes @ $2.8K each = $148K

Component Details

- Control systems
  - Acquisition plan:

Detector subsystem	# of Processors	Processor types	Replacement plan
Control and Monitoring	12	(6) 16MB PowerPC (5) 64MB PowerPC (1) 128MB PowerPC	Replace, use old processors for HV or spares
High Voltage	29	(29) 16MB PowerPC	Retain, with new and spare needs met from other replacements
Muon	~40	(40) 4MB 68K	Replace (17) in readout crates, retain as spares
Tracker readout	24	(12) 64MB PowerPC (12) 128MB PowerPC	Replace, use old processors for HV or spares
Test stands	~7	mixed low end	Use available

Component Details

- Test stand support
  - 1% and 10% test stands at SiDet are “mini” DAQ/Online systems
    - Require SBC (readout) and Host/Control/Monitor Linux node (to be furnished)
    - Require control system processor (to be supplied from operational spares)
  - DAQ/Online group will supply experts to assist in installing systems and supporting operations
    - Requires fraction of experts’ time during installation and operation
    - Loaded into schedule

M&S Cost Summary

Equipment	M&S Cost	Contingency %	Contingency Cost	Total Cost
Level 3 filter nodes	**210**			**357**
Nodes, racks, power supplies, cables	210	70	147	357
Fibre Channel SAN	**60**			**90**
Hubs, cables	20	50	10	30
Fibre Channel JBOD buffer disk	40	50	20	60
RAID storage array	**43**			**65**
FC RAID controllers	10	50	5	15
Disk crates, rack	15	50	8	23
Hot-swappable disks	18	50	9	27
DAQ HOST system	**52**			**78**
Primary high-end server	32	50	16	48
Secondary mid-range server	20	50	10	30
Database System	**52**			**78**
Primary high-end server	32	50	16	48
Secondary mid-range server	20	50	10	30
File Server system	**40**			**60**
Primary mid-range server	20	50	10	30
Secondary mid-range server	20	50	10	30
Control System	**148**			**185**
Processor upgrades	148	25	37	185
Total	**605**		**308**	**913**

Contingency guidance from Project Management Plan

Schedule

- **Purchases**
  - **Level 3**
    - Staged purchases in FY04 and FY05
    - Follow CompDiv farm node purchasing process
  - **Host systems and storage**
    - Typically a development/secondary system in FY04 followed by a production/primary system in FY05
    - “From the catalog” component purchases
  - **Controls processors**
    - Development systems in FY04, production systems in FY05

Schedule

- **Effort**
  - **Level 3**
    - Monitor and follow annual CompDiv farm node evaluation, purchasing, and commissioning process
  - **Host systems and storage**
    - Continuing system and market evaluation effort
    - Commissioning and integration effort following purchases
    - Followed by longer term operations responsibilities
  - **Controls processors**
    - Follow EPICS community progress on adapting new processors
    - Evaluation and application testing within D0 environment
    - Deployment of new nodes following purchase

Schedule

- **Effort (cont'd)**
  - **Loaded into schedule**
  - **Summary:**
    - **Mostly "fractions of" personnel who are otherwise dedicated to Run II operations**
      - **Largest components:**
        - **Strategic planning (PHYSF)**
        - **System installation and management (COMPF)**
        - **ORACLE database migration (COMPF)**
        - **Control system upgrades (PHYSF,COMPF)**
    - **Net technical labor cost ~$150K**

Risk Analysis

- **Highlighted elements**
  - **Cost**
    - Required number of Level 3 filter nodes a function of the interaction environment, detector performance, and software performance. 70% contingency to address concern.
    - Details of future VME processors unknown. Address risk with moderate 25% contingency for unit costs plus possible scope reduction.
  - **Schedule**
    - No significant risks
  - **Scope**
    - No significant risks
  - **Technical**
    - No significant risks

DAQ / Online Summary

- Building upon fundamentally solid hardware and software system architectures
- System constructed of commodity components
- Project addresses requirements for RunIIb capabilities and system longevity