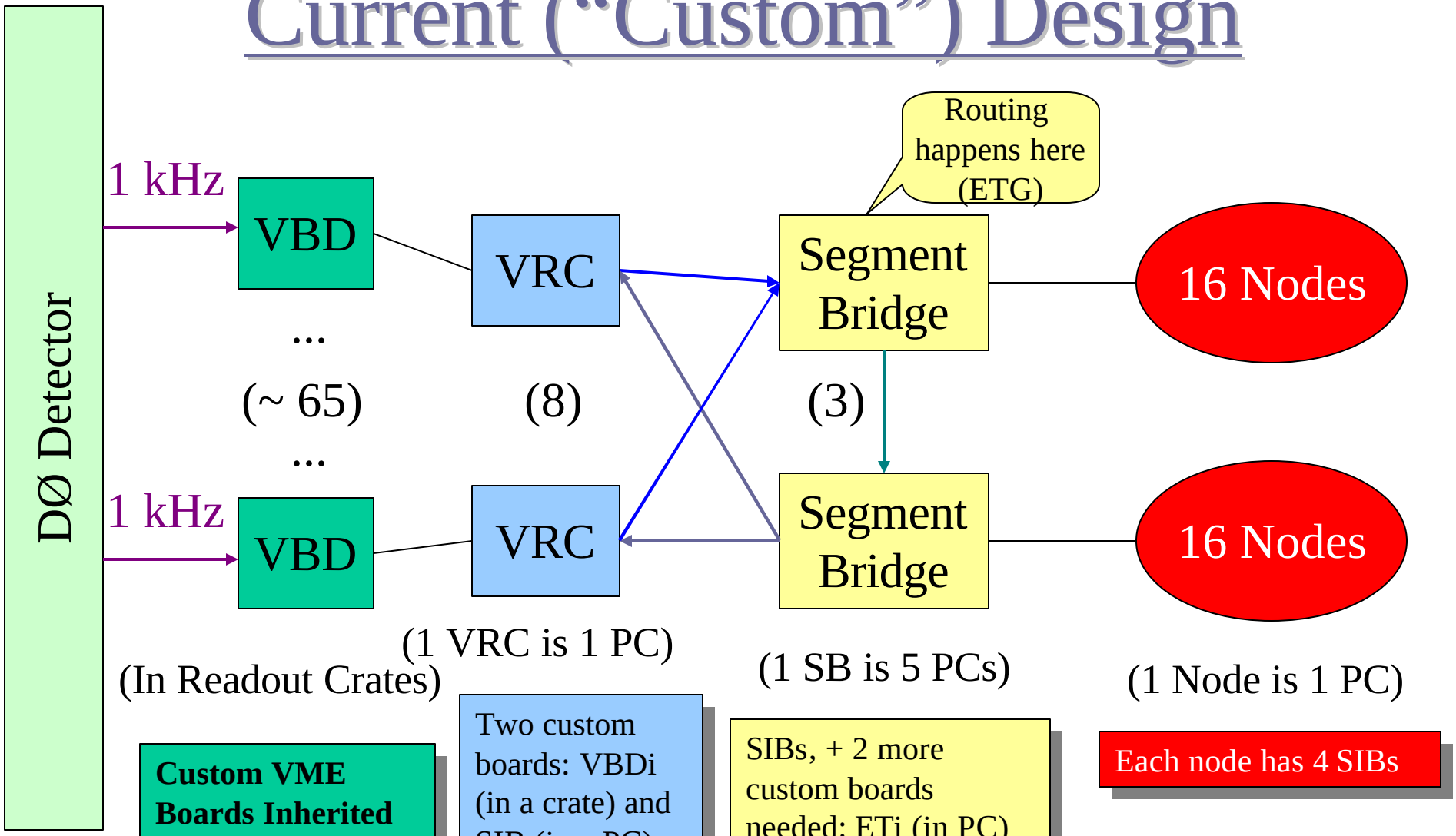
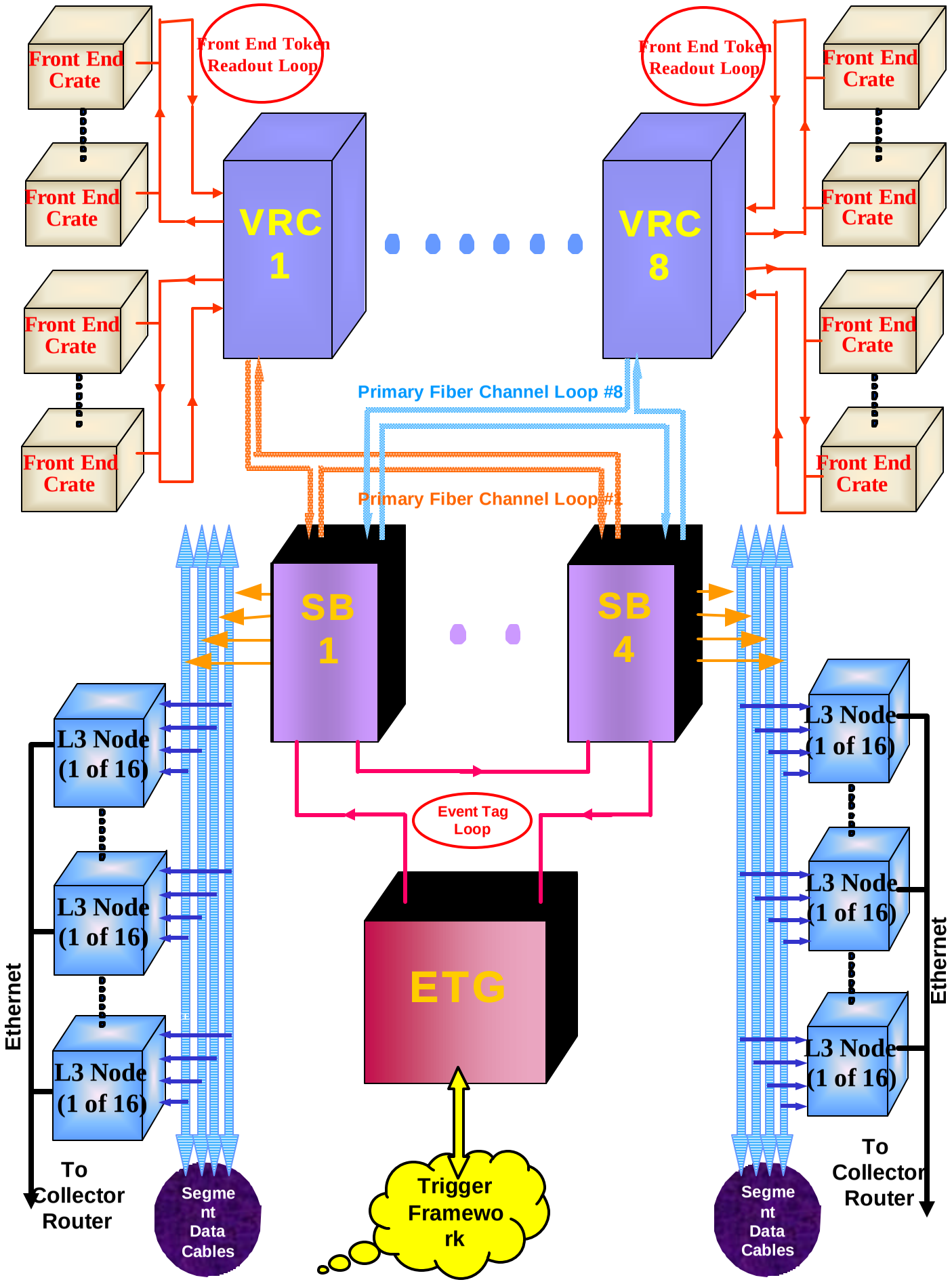
An Ethernet-Based Data Acquisition System For DØ

Gustaaf Brooijmans - Fermilab

- Current “Custom” Design, Status
- Ethernet-Based Design:
  - Hardware
  - Software
  - Feasibility Test
- Conclusions

Current ("Custom") Design



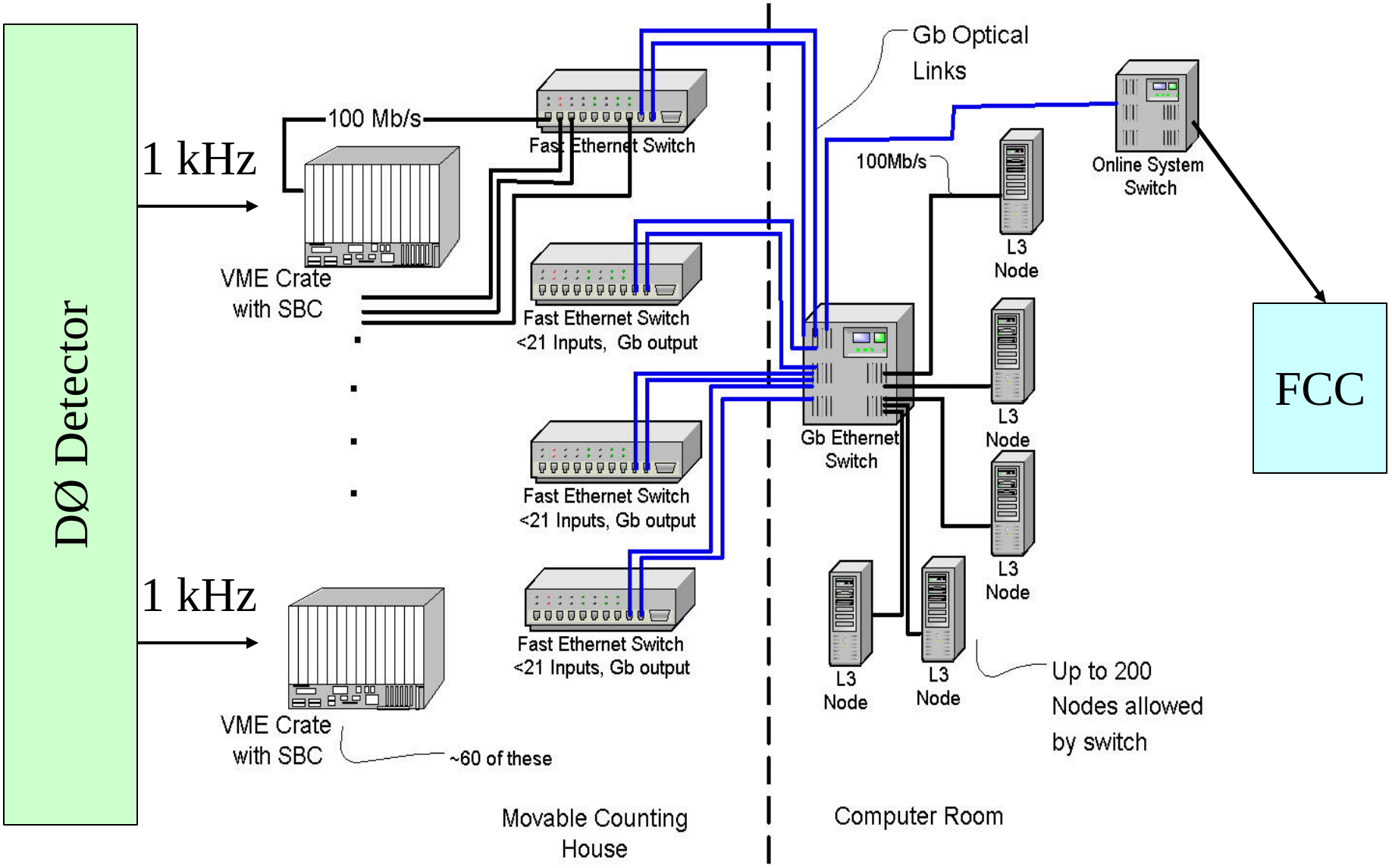

Front End Crate
Front End Token Readout Loop
VRC 1
VRC 8
Primary Fiber Channel Loop #8
Primary Fiber Channel Loop #1
SB 1
SB 4
L3 Node (1 of 16)
Event Tag Loop
ETG
Ethernet
To Collector Router
Segment Data Cables
Trigger Framework

Custom Design Status

- Four custom boards needed: VBDi, SIB, VBDi' and ETi
- Status:
  - Currently have prototype VRC's + software emulators, capable of reading out at 30 Hz - this supports detector commissioning
  - First production VBDi being tested at DØ, will get us to ~150 Hz
  - SIB going to the board manufacturer, available in January $\rightarrow$ 500 Hz
  - VBDi' and ETi design, layout remain to be done

- So, staged “road” to 1 kHz:
  - Now: 30 Hz
  - With VBDi: 150 Hz
  - With SIB (January): 500 Hz
  - Full system (Segment Bridges, ETG - July): 1+ kHz
- But, technical problems and poor schedule performance have prompted exploration of alternative solution - rest of this talk

A Commodity Readout System



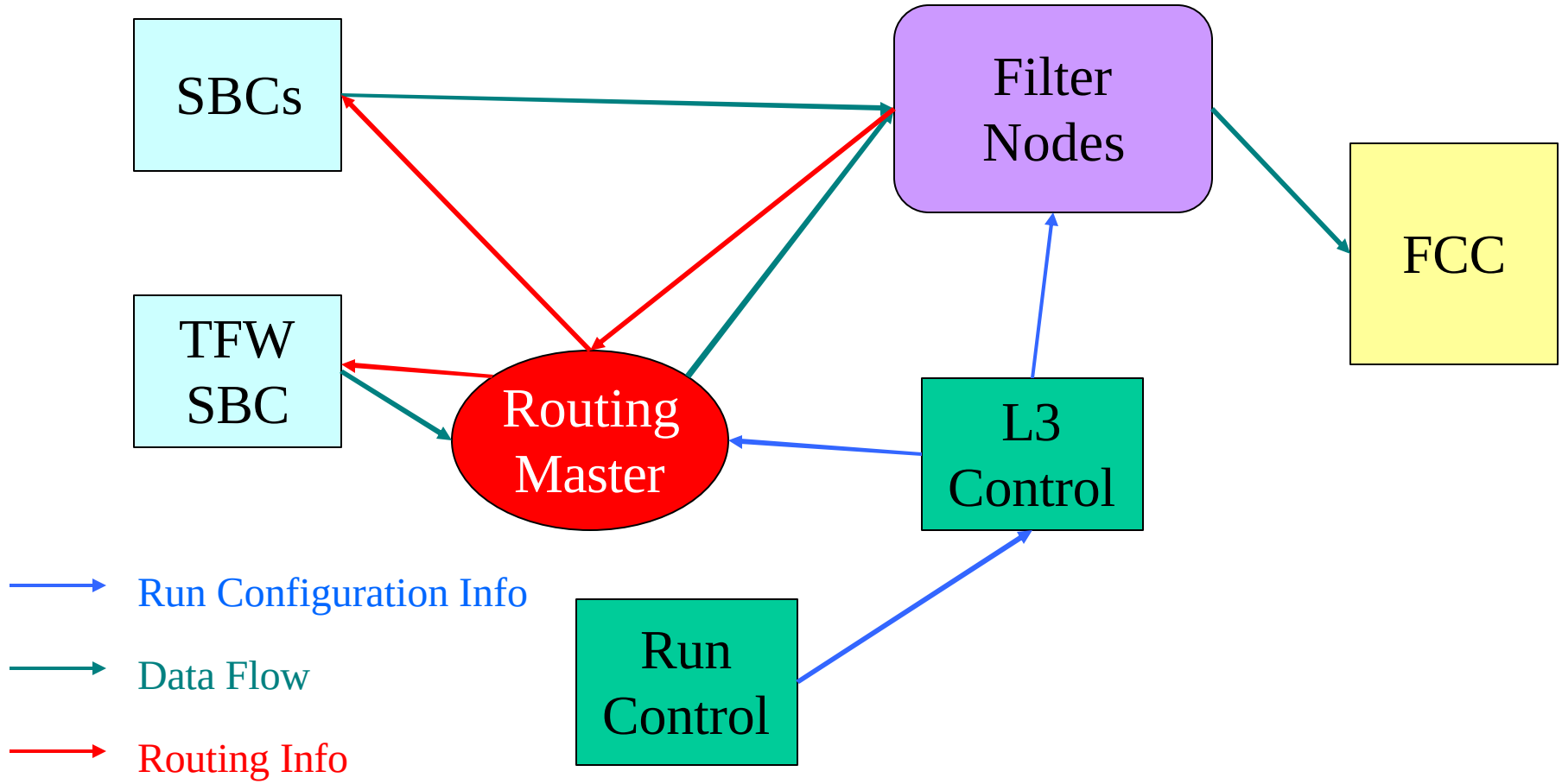
- Single Board Computers (SBC) read the event data over the VME bus, and send it to a Level 3 node according to routing instructions received from the *Routing Master*
- The Routing Master program runs on the SBC reading out the Trigger Framework (TFW) crate (so it has access to the event number and the triggers each event satisfied)
- Event building is done in the Level 3 nodes.

Hardware Components

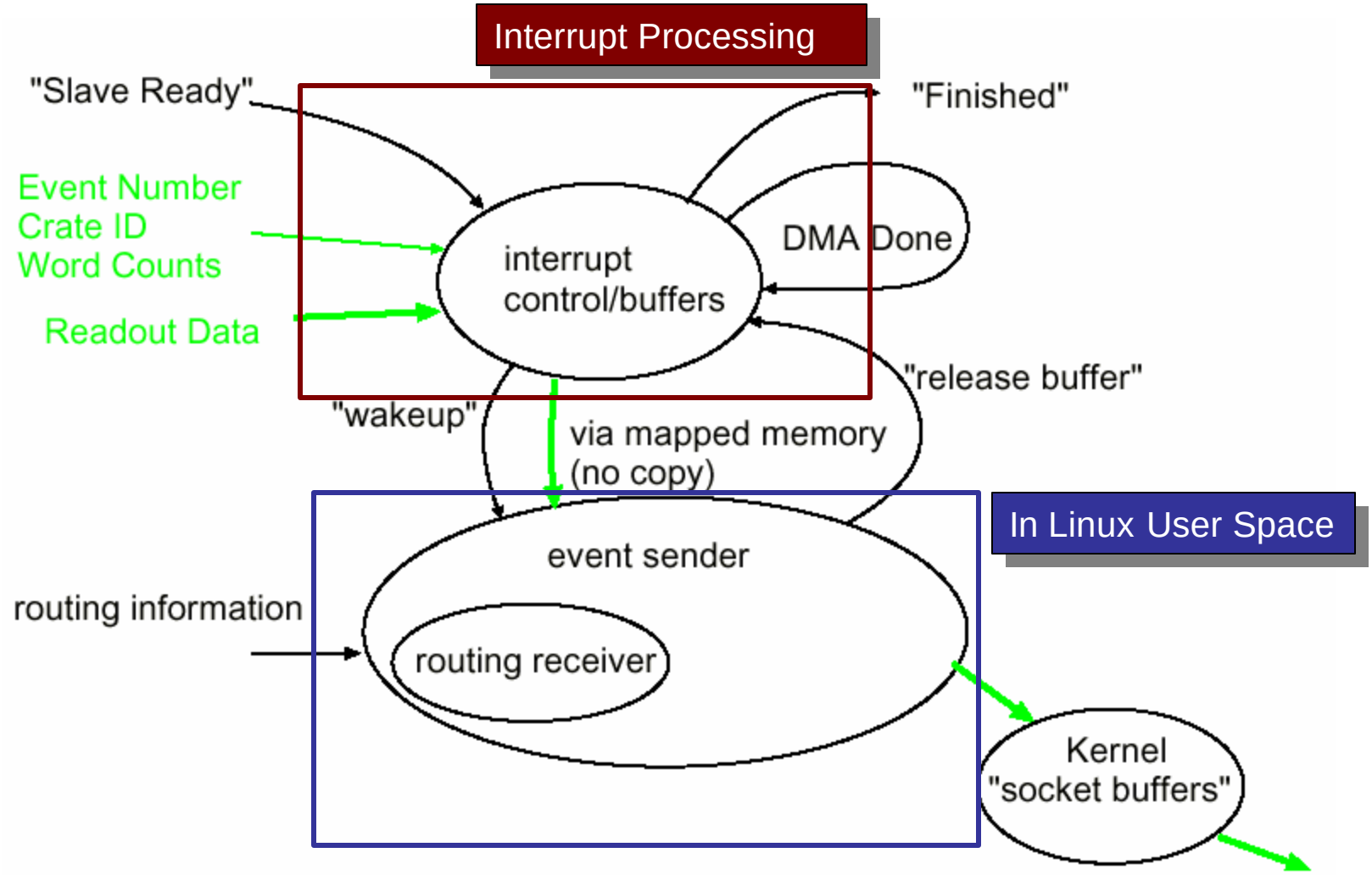
- Single Board Computers:
  - require good processor speed, enough memory, Ethernet, fast VME access, two channels of digital I/O (readout handshakes).
  - using VMIC 7750 for tests: 933 MHz PIII, 128 MB RAM, 2 Ethernet ports, Tundra Universe II for VME-PCI interface. Added Acromag PMC470 digital I/O card.
  - One custom component: passive extender board to fit 6U SBCs in 9U crates.

- Switches:
  - Cisco 2948G (“concentrators”): 100 Mb/s (copper) input from SBCs, Gb (fiber) output. Limit to 10 SBCs per fiber → no congestion possible (100% contingency allows doubling these if needed)
  - Cisco 6509: capable of handling 16 GB/s (average expected rate is 250 MB/s). Gb fiber in (from 2948G’s), 100 Mb/s (copper) out to Level 3 Nodes, with 112 MB of buffering in the switch per 48 nodes.

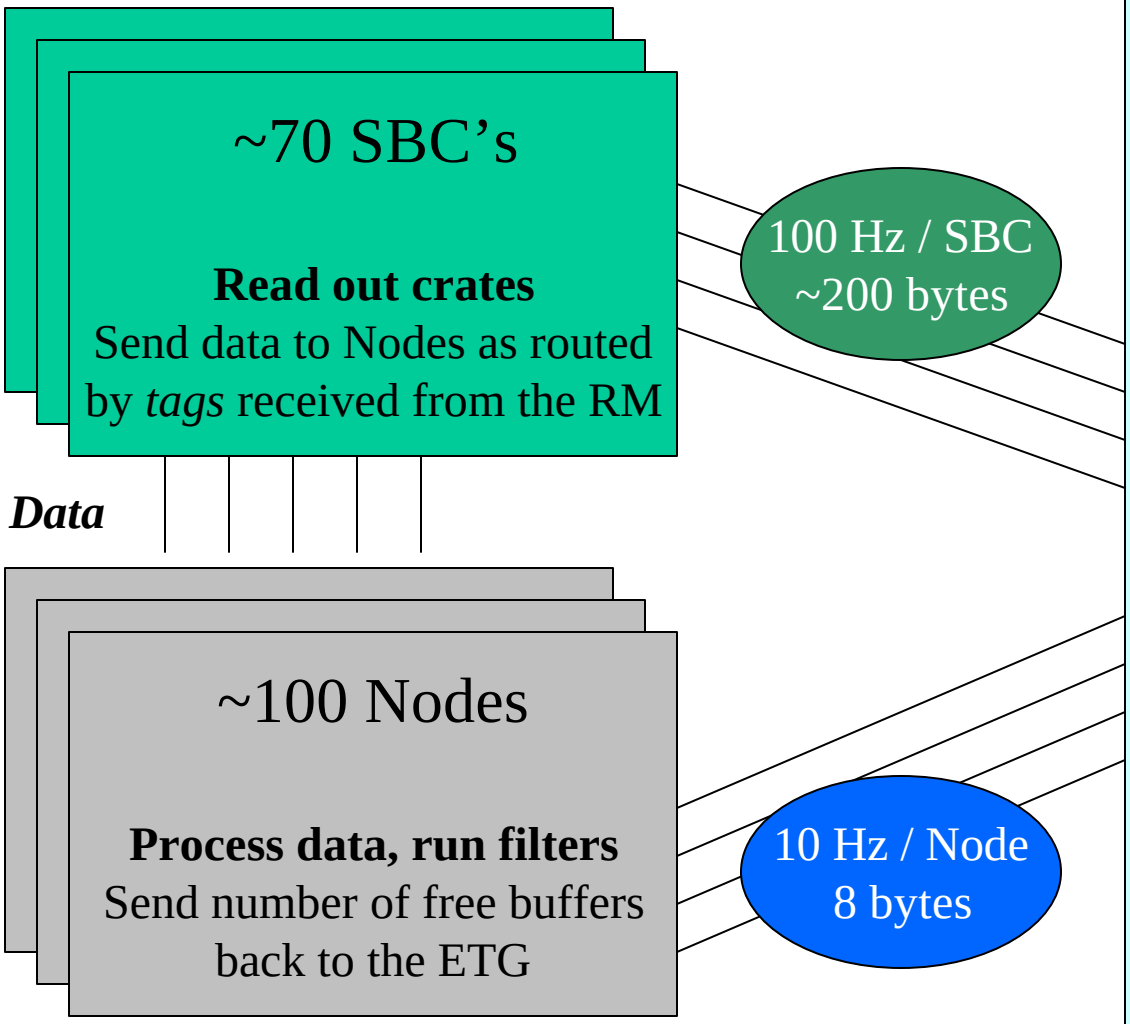
Software Components



Single Board Computer Software



Routing Master



Routing Master (RM)

- Reads TFW crate
- Gets trigger bits fired
- Decides which crates need to be read out
- Tells SBC's which node(s) to send data to
- Sends TFW block to correct node(s)
- Enables/disables triggers
- Heart of monitoring / run configuration / error reporting

Feasibility Test

- Decision to proceed with test made at the end of August
- Ultimate goal: demonstrate feasibility of an Ethernet-based DAQ for DØ
- Also valuable source of information to estimate
  - implementation schedule
  - installation approach and associated disruption to data taking

- Risks to be investigated:
  - VME integration problems with some of the readout modules
  - potential performance limitations by some network components
  - generation of coherent noise in the calorimeter ADCs
  - insufficient understanding of the software needs
  - bugs in existing and new software (as much as possible)
  - unanticipated issues (as far as possible)

Test Description

- Aim is to get a “slice” of the system working at the end of the October-November shutdown :
  - 10 SBCs
  - Network equipment: 1 intermediate switch (Cisco 2948 G) + high performance ports for main switch (Cisco 6509)
- Software development started in September/October (depending on component)

Test Results

- Stable readout of calorimeter, muon and tracking test crates at 500+ Hz for days (rate limited by current system - have to read out TFW crate)
- Stable readout of the (real) Trigger Framework crate over the last 3+ weeks
- Stable, coordinated readout of the TFW crate and a calorimeter crate, with routing master, and receiving of event fragments in a linux node
- Event builder crashes when including more crates, understood, fixing it

Preliminary Test Conclusions

- While tests are ongoing, early conclusions are that:
  - DØ readout crates of many types can be read out using SBCs:
    - Level 1/tracking, muon and calorimeter ok
    - Working on Level 2 now
  - The amount of software needed to coordinate the readout (connections, routing) is manageable
- These two statements address the two key issues that are *specific to DØ* when it comes down to demonstrating that an Ethernet-based DAQ is *feasible* for this experiment

Cost Estimate Summary

WBS	1.1			
WBS	ITEM		CONTINGENCY	
1.1	Commodity DAQ	M&S		TOTAL
		TOTAL	%	Cost
1.1.1	*Switches*	95,188	49	141,866
1.1.1.1	Cisco 6509	75,600	36	102,690
1.1.1.2	Cisco 2948G	19,588	100	39,176
1.1.2	*SBC's*	278,140	30	361,582
1.1.3	*Level 3 Nodes*	35,000	100	70,000
1.1.4	*Cables, patch panels, etc.*	35,000	50	52,500
1.1.5	*Network diagnostic equipment*	20,000	30	26,000
1.1.6	*Teststand*	19,000	46	27,650
1.1	**Total**	**482,328**	**41**	**679,598**

482,328 679,598

Schedule

- Three driving components:
  - Hardware procurement
  - Software development
  - System Integration

  Proceed in parallel
- Hardware schedule based on feasibility test + lab experience
- Software and integration estimates based on fine segmentation of tasks, and experience with the Level 3 linux farm (software part of the schedule was made in mid-October, we are ahead)

- Note that we have a (partially tested) plan to gradually convert DØ → minimal disruption
- Main schedule features:
  - Decision to implement on January 2, 2002
  - Software integration complete on March 19, 2002
  - SBCs and switches available April 16, 2002 (start gradual system conversion)
  - Full system operational July 18, 2002

Conclusions

- Custom solution is progressing, but concerns remain
- Investigating a commodity solution:
  - Design complete (TDR)
  - Results from feasibility tests very good
  - A few things left to test
  - Working on installation plan that will be minimally disruptive to DØ (no significant downtime required)
  - Feasibility tests confirm schedule