EST MC Fellowship Midterm Report

ATLAS TDAQ Data-Flow

W.Vandelli
CERN – Physics Department/ATD

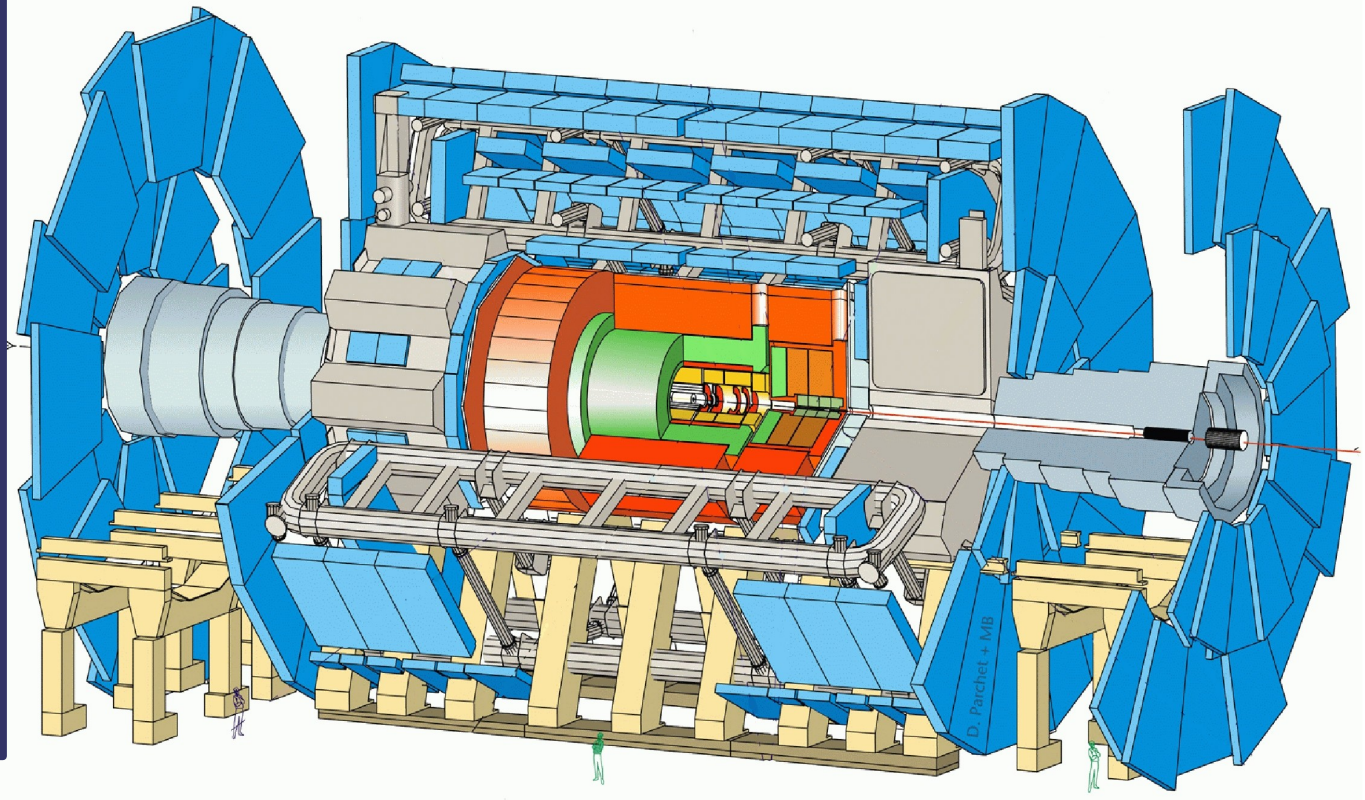
Introduction

- Started my appointment as EST fellow in January 2007
- Joined the ATLAS Trigger and Data Acquisition group at CERN

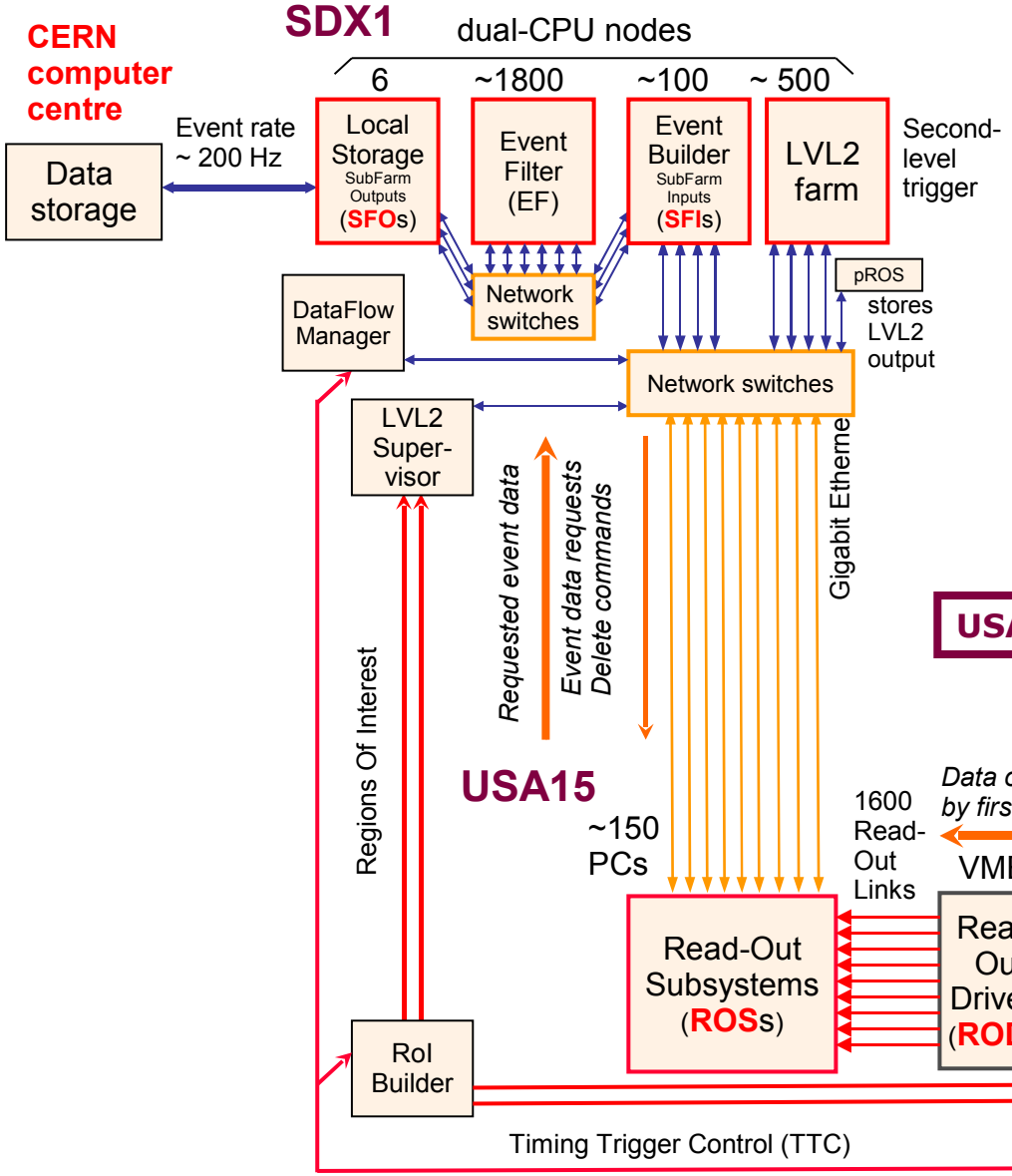
ATLAS is one the LHC experiments

During the last two years, a lot of efforts were dedicated to the completion and commissioning of the detector

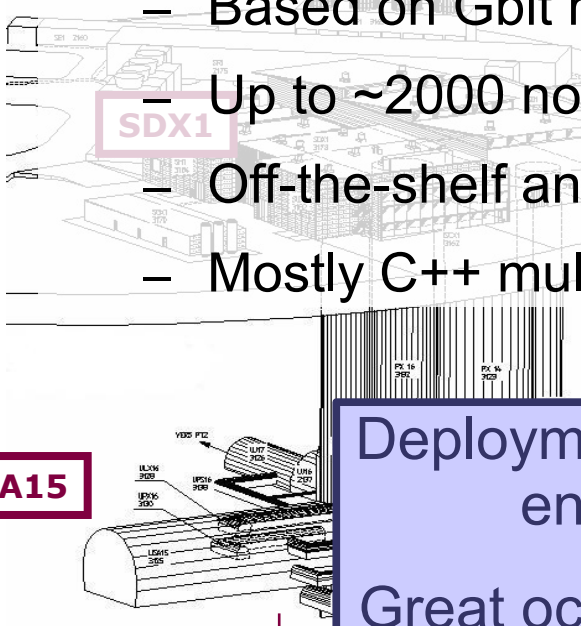
The TDAQ played a fundamental role in this process



ATLAS TDAQ



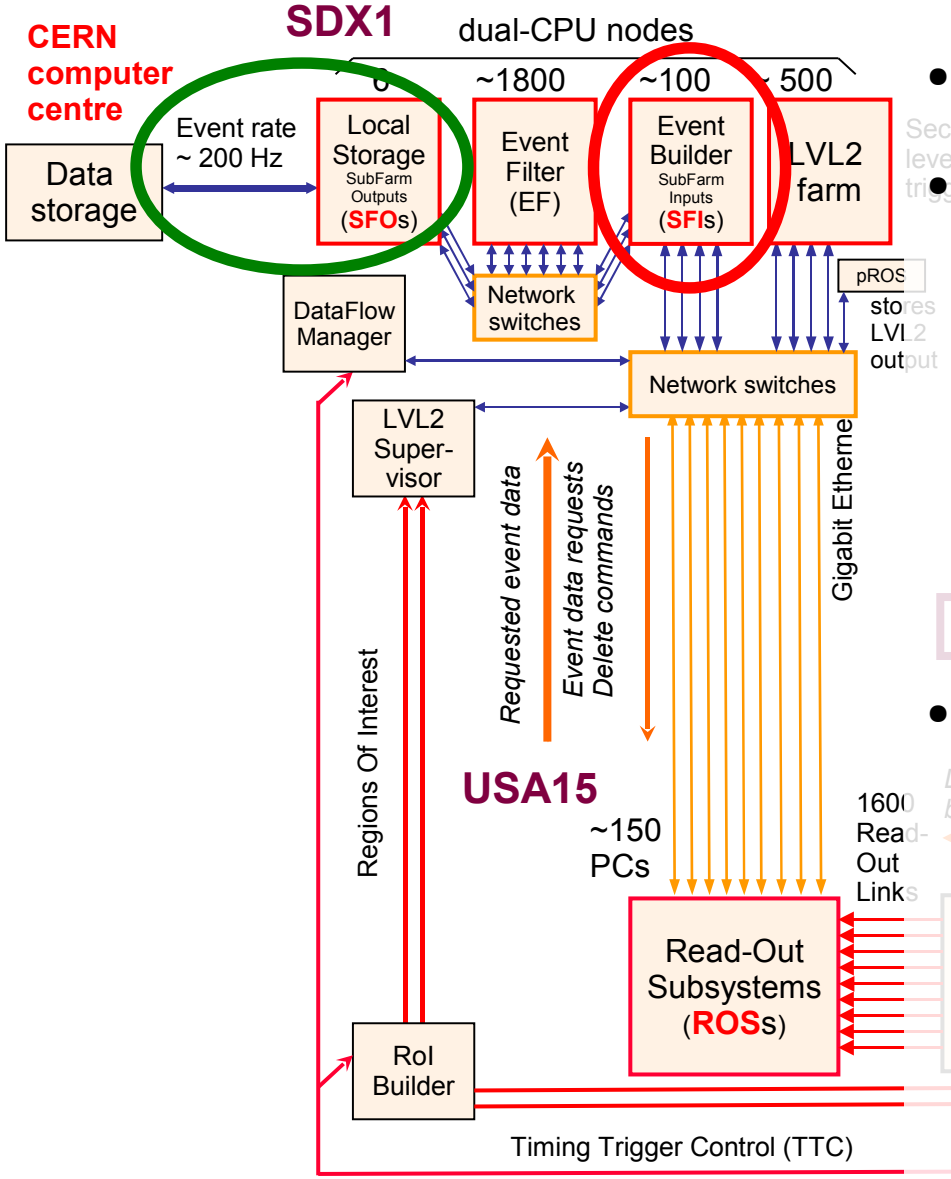
- Highly distributed system
  - Based on Gbit networks
  - Up to ~2000 nodes
  - Off-the-shelf and custom hw
  - Mostly C++ multi-threaded sw



Deployment started at the end of 2006.

Great occasion to join the TDAQ commissioning and contribute to the detector integration

Event Builder and Data Logging



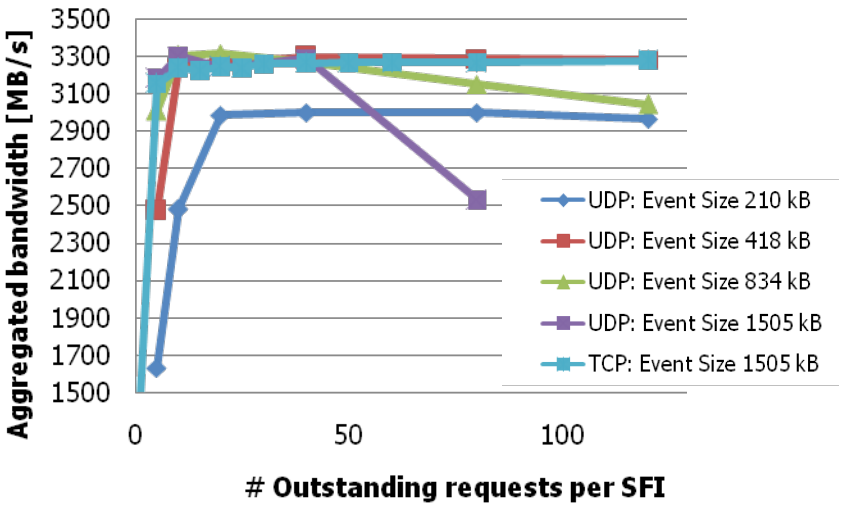
- Involved in the DataFlow community
- Event Builder (SubFarm-Input)
  - Responsible to assemble full events upon acceptance of the second-level trigger
  - Design predicted a 100 node farm
  - Target 1.5kHz@1.5MB = O(5GB/s)
- Data-Logging (SubFarm-Output)
  - Data are temporary stored on local disks and transferred to the permanent storage
  - 5 node farm
  - Target 200Hz@1.5MB = 300MB/s

TDAQ Commissioning

- TDAQ technical runs
  - Dedicated testing period on the final system
- ATLAS Milestone weeks
  - Detector integrated data-taking
- Both provided major feedback for TDAQ, in terms of functionalities and performance

- Active participation in commissioning activities
  - Proposing performance measurements
  - Following the development based on the commissioning feedback
  - Learning how to best utilize and tune the system

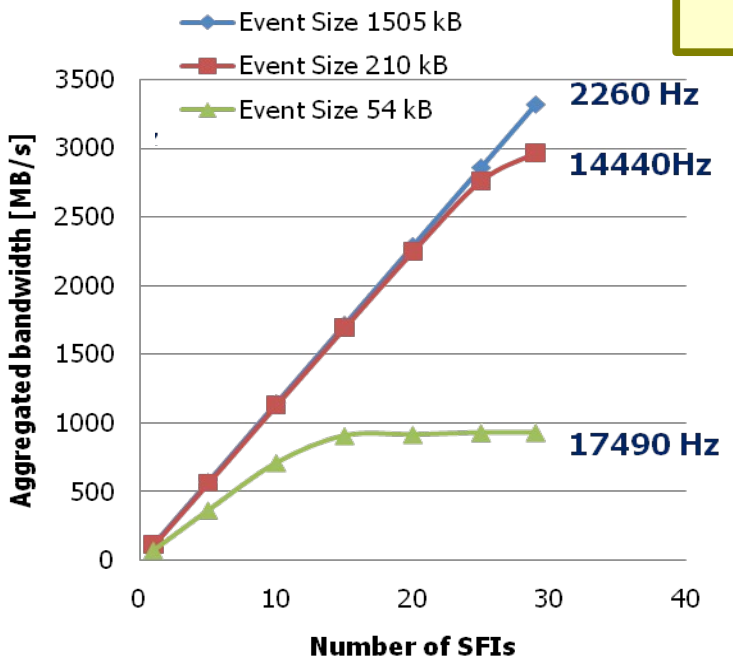
Event Builder

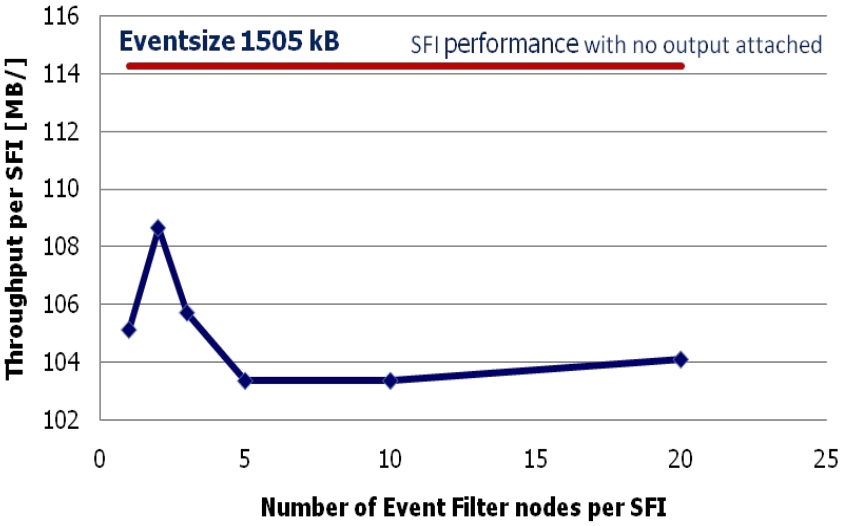


Many measurements performed on the first batch of nodes

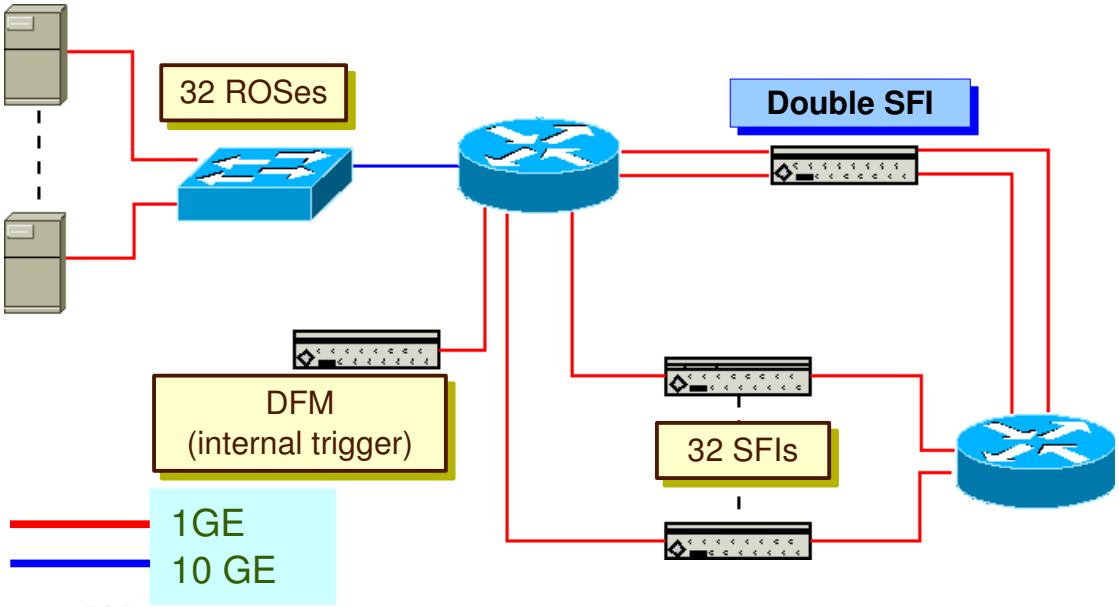
Validation of the design and performance assessment

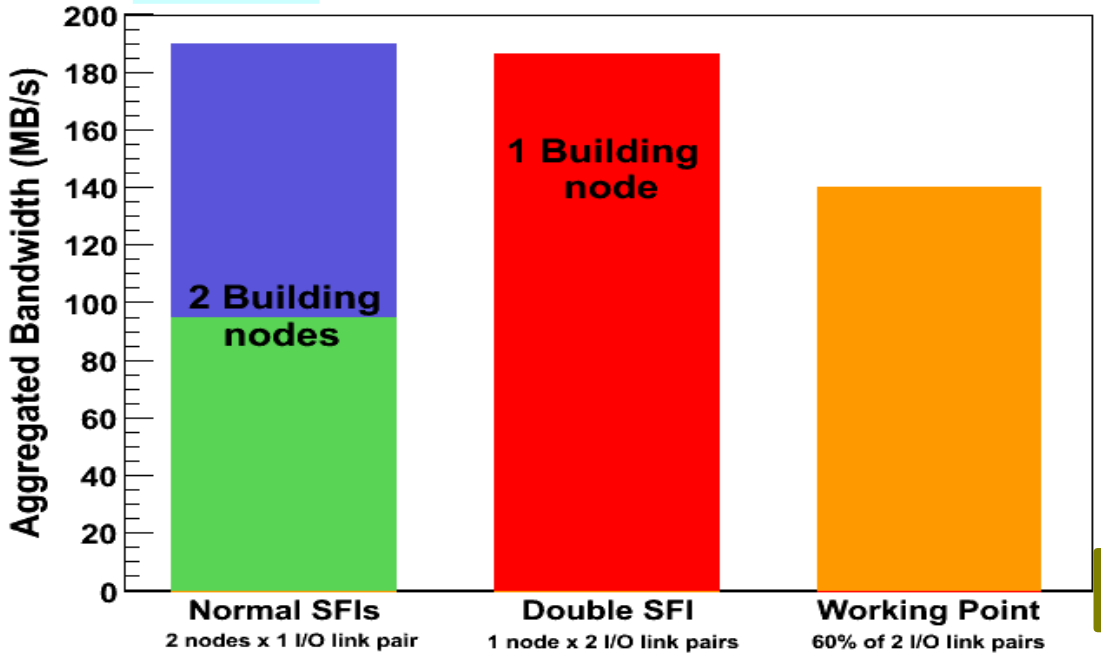
Presented at IEEE Real Time Conference 2007 and published





Double Builder Node





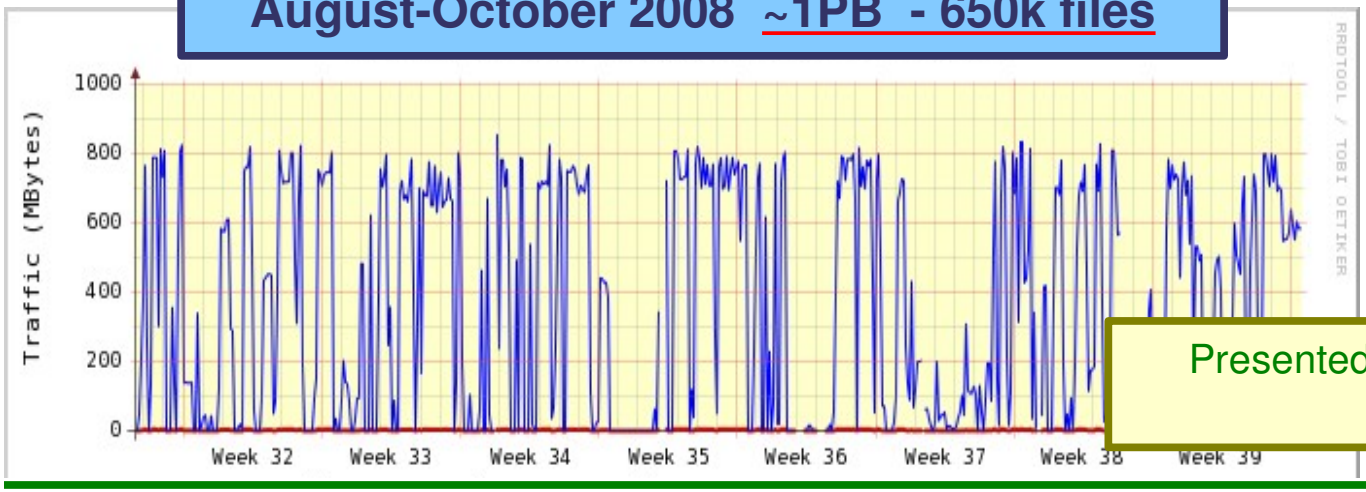
- Exploit modern multi-core processors, running more than one application per node
  - Extended network capabilities
- Required the study of dedicated operating system and network configurations
- Currently deployed in the ATLAS TDAQ system
  - 63 nodes running 94 applications

Presented at IEEE NSS MIC 2007 and published

Data Logging: squeeze out the hardware

- Farm installed at the beginning of 2007
- Several improvements allow to efficiently use the hardware
  - Filesystem and operating system tuning
  - Communication thread pool to handle up to 2000 clients
- 550MB/s sustained aggregated I/O rate on top of 50TB local storage

August-October 2008 ~1PB - 650k files

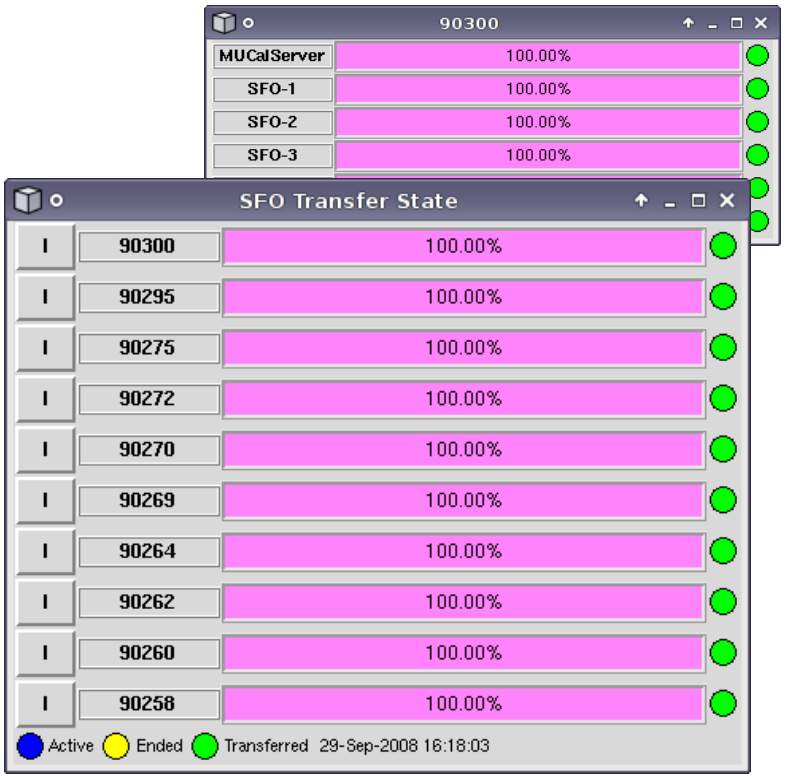


The only TDAQ sub-system regularly operated at (or even beyond) the design

Presented at IEEE NSS MIC 2007, 2008 and published

Online-to-Offline Transfer

- Did not have bookkeeping and communication tools between the online and offline worlds
  - SFO-Tier0 handshake database



SFOTZ Interface

Table: Run LumiBlock File

Output: SFOID RUN# LB# StreamType StreamName Event # Size (GB)

Selections: > Run # >= Out of P1

> LB # >=

C 1/11/2008 >= Date >= 1/8/2008 C

QUIT Execute Clear

SFOID	RUN #	Total Streams	Opened Streams	Closed Streams	Transferred Streams
SFO-1	93024	5	0	0	5
SFO-2	93024	5	0	0	5
SFO-1	92842	4	0	0	4
SFO-1	92825	3	0	0	3
SFO-1	92821	3	0	0	3
SFO-1	92812	1	0	0	1
SFO-1	92743	3	0	0	3
SFO-1	92724	2	0	0	2
SFO-1	92721	1	0	0	1
SFO-1	92718	1	0	0	1

- A set of Oracle tables, filled online, keeps track of
  - Run and Lumiblock statuses
  - File statuses, from creation to deletion
  - File properties (e.g. checksum, #events, ..)
- Drives the offline processing and provides bookkeeping and monitoring functionalities
  - Essential when handling O(10k) files per day

On the battlefront

- Acquired a large experience in bringing and keeping in operation the system as a whole
  - TDAQ representative in data processing and management large scale tests
  - Responsible for several TDAQ testing periods
  - TDAQ support for several detector weeks
  - Run Control shifts
  - On-call expert
- Might sounds like a lot of service work

I think is also a training experience

Organize, Coordinate, Decide, Negotiate, Report

Conclusions

- Contributed in the assessment and optimization of ATLAS TDAQ data-flow performances
- A lot of developments, in different areas
  - Network communications
  - Database exploitation
  - Python scripting
- Supporting the system operations involved the interaction with many ATLAS members and with different scientific backgrounds
- Results have been presented (either by me or colleagues) in workshops and conferences

Conferences and Workshops

- EPS HEP 2007, Manchester, England
  - "Cherenkov Light Contribution in Lead Tungstate Crystals"
- ATLAS TDAQ Workshop 2007, CERN
  - "DataFlow – Recent Results and Outstanding Issues"
- IEEE NSS 2007, Honolulu, USA
  - "The ATLAS Event Builder"
  - "The Data-Logging System of the Trigger and Data Acquisition for the ATLAS Experiment at CERN" - Poster
- IPRD 2008, Siena, Italy
  - "Readiness of the ATLAS Trigger and Data Acquisition system for the first LHC beams"

Internal Documents and Presentations

- "The Message Format of the ATLAS TDAQ DataCollection", v2.4 & v2.5

- Many (many) presentations and reports in several ATLAS forums
  - DAQ/HLT Commissioning meeting
  - DAQ/HLT Coordination meeting
  - TDAQ DataFlow meeting
  - Monitoring Working Group meeting
  - Run Coordination meeting

Publications

- H.P. Beck at al, “Performance of the Final Event Builder for the ATLAS Experiment”, IEEE Trans. Nucl. Sci., 55-1 (2008), p.176
- W. Vandelli et al, “Cherenkov Light Contribution in Lead Tungstate Crystals”, 2008 J. Phys.: Conf. Ser. 110, 092034
- A. Battaglia at al, “The Data-Logging System of the Trigger and Data Acquisition for the ATLAS Experiment at CERN”, IEEE Trans. Nucl. Sci., in publication
- W. Vandelli et al, “The ATLAS Event Builder”, IEEE Trans. Nucl. Sci., in publication

Training

- 2nd CERN-Fermilab HCP Summer School, CERN, Jun 2007

- “C++ - Hands-on Introduction”, 24h course, Apr 2008
- “Python - Hands-on Introduction”, 24h course, Oct 2008
- “Python: Advanced Hands-on”, 32h course, Dec 2008

- “French language”, 2x60h courses, Oct-Dec 2007, Jan-Mar 2008