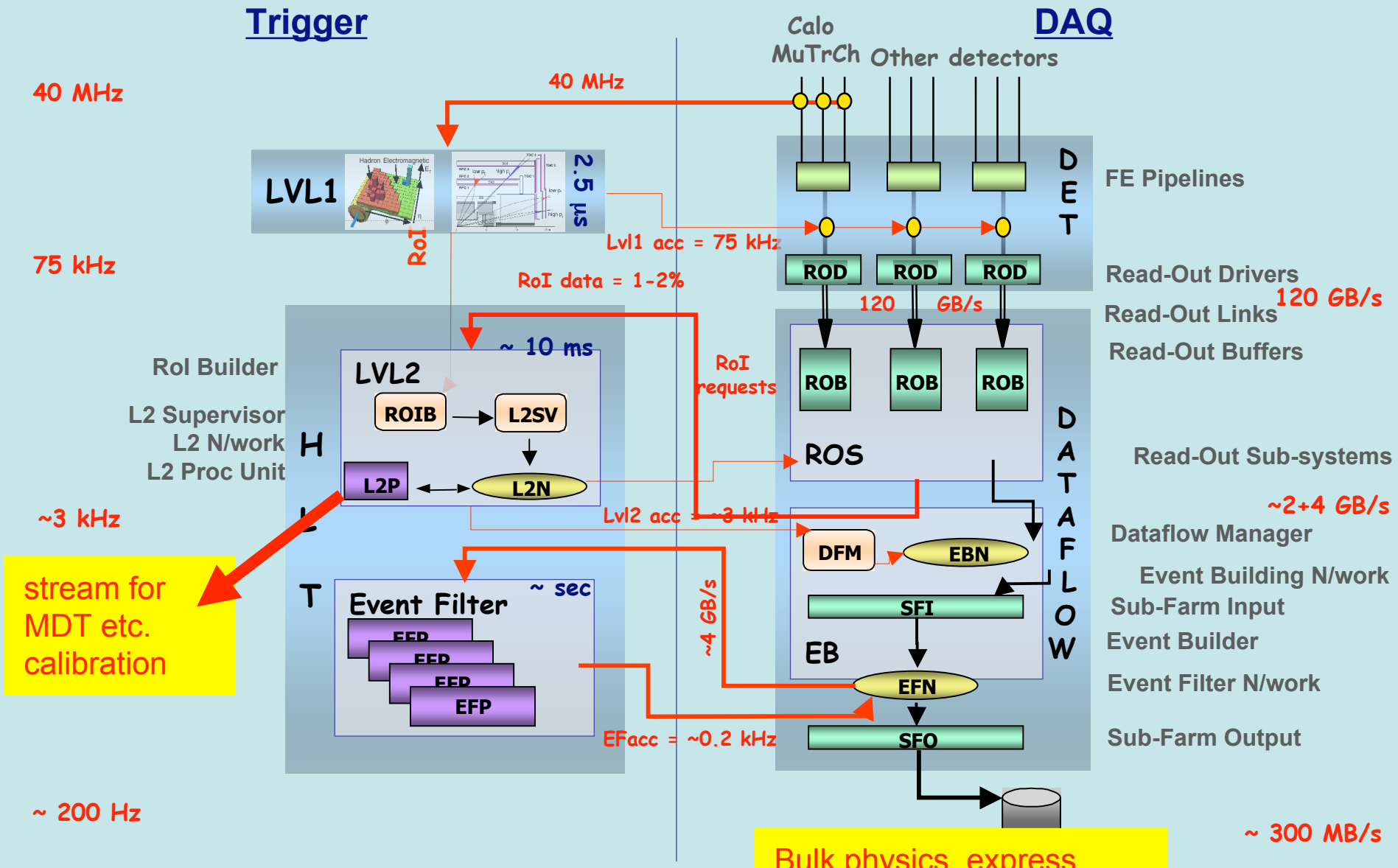
Update on Data Streaming *

- Charge to the Study Group
- Baseline and Context
- Data flow from HLT to Tier0
- Studies with physics streams
- Number and size of files to Tier0
- Calibration streams and express stream
- Overlapping vs. exclusive streams
- Status

* Jean-Francois Arguin, Sebastien Binet, Szymon Gadomski, Marjorie Shapiro, Ian Hinchliffe, Jack Cranshaw, David Malon, Richard Hawkings, Tom LeCompte, Roger Jones, Gordon Watts, Zhongliang Ren, Sander Klous, RD Schaffer, Wouter Verkerke, Ivo van Vulpen, Fabiola Gianotti, Dario Barberis, Nick Ellis, David Quarrie, Chris Bee, Livio Mapelli, Giacomo Polesello, Hans von der Schmitt

Streams: TDAQ side



Charge to the Study Group

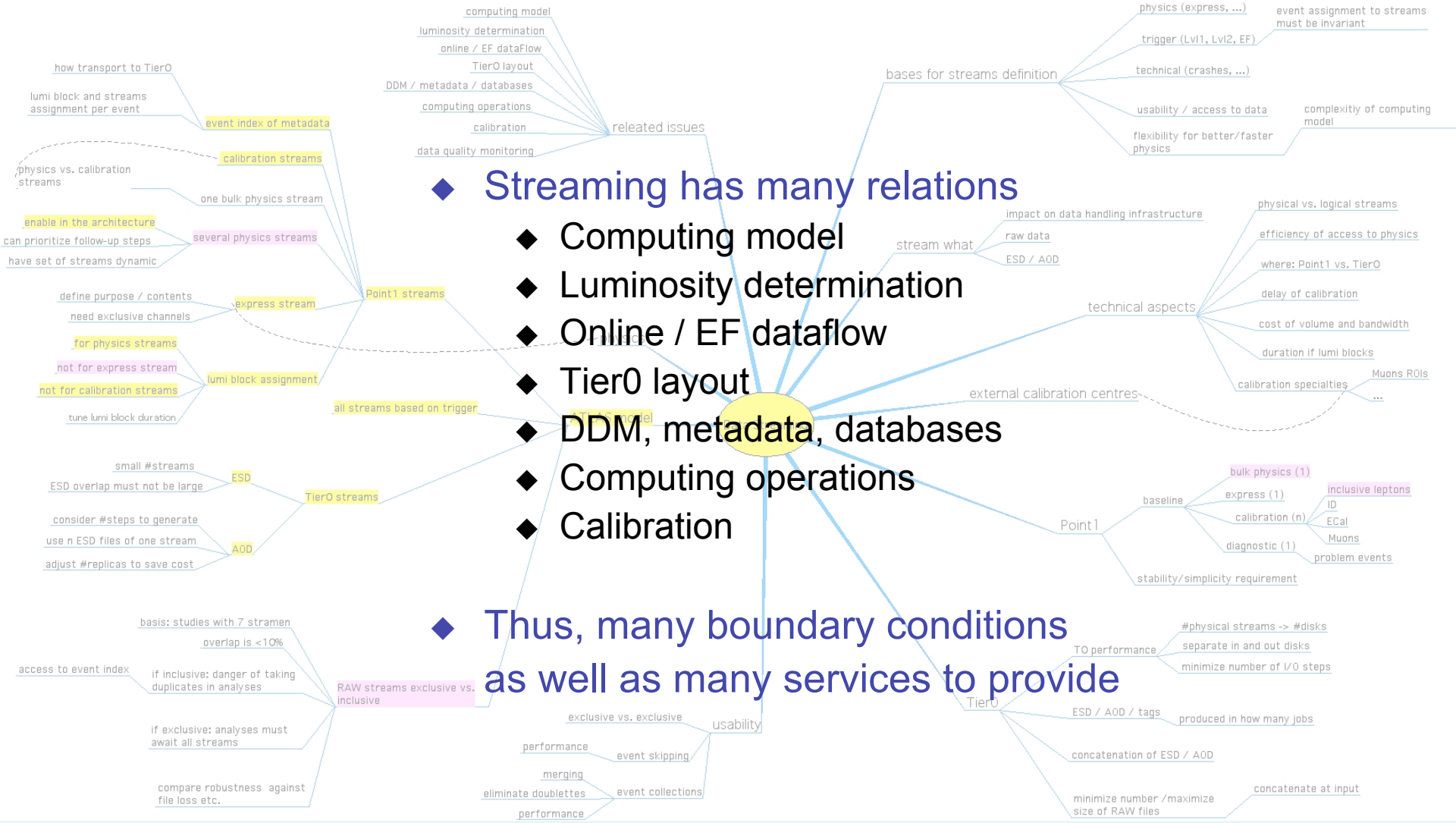
- Perform quantitative studies on questions including
  - Which streams generated at Point1 and at Tier0
  - Overlapping vs. Exclusive streams
  - Streams and merging at the ESD and AOD level
- and give recommendation to COB

- Underlying questions
  - Should we deviate from "baseline" and for which goals
  - Feasibility / implications for TDAQ and Data Management
  - Interplay with luminosity blocks

- Study group has been set up in December 2005
- Participation from
  - TDAQ, Software, Computing Model, Physics, and Luminosity

- More on the web:
  http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/COMMISSIONING/streaming.html
  use cases: https://twiki.cern.ch/twiki/bin/view/Atlas/DataStreaming

Context

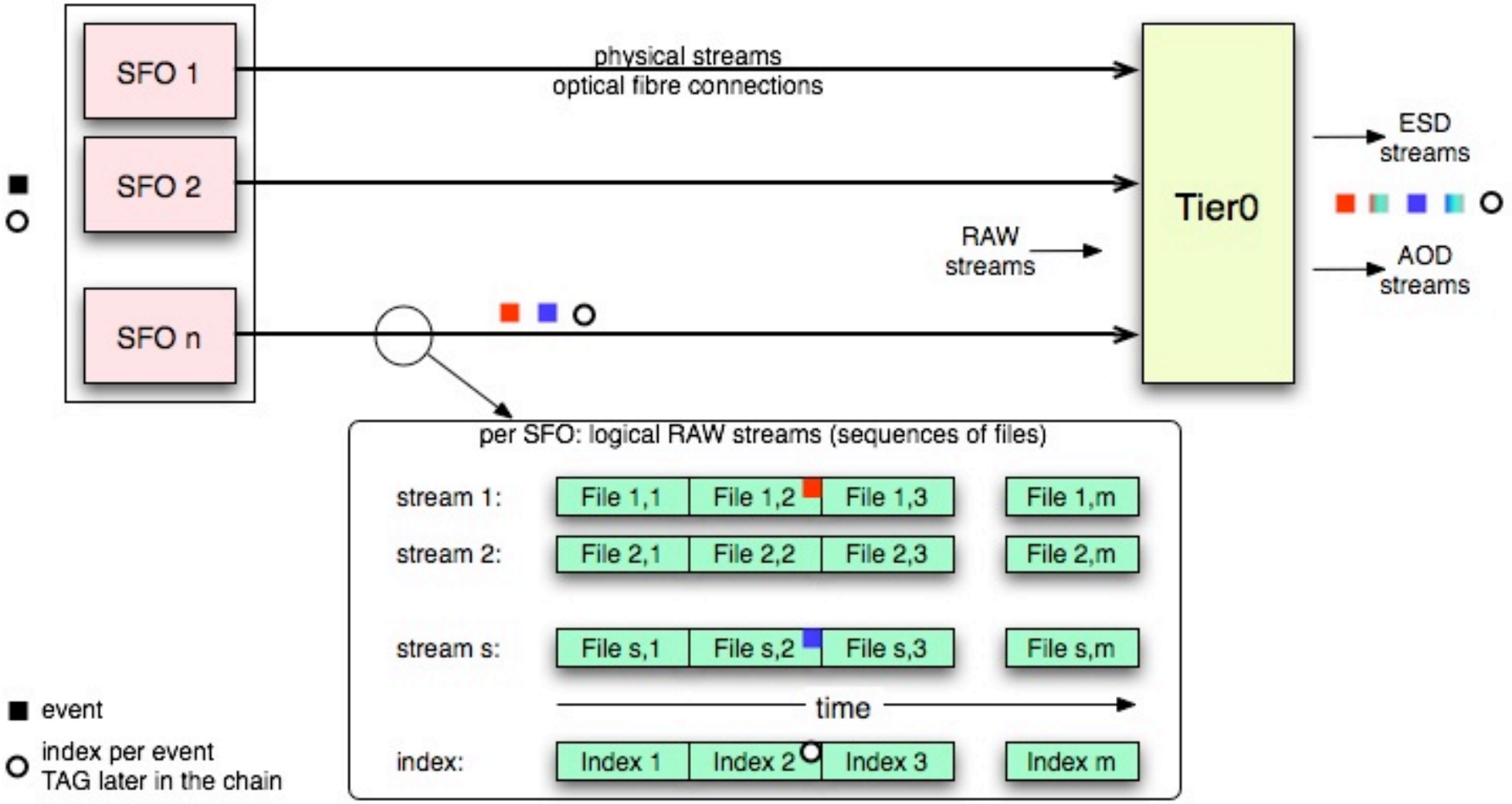


- Streaming has many relations
  - Computing model
  - Luminosity determination
  - Online / EF dataflow
  - Tier0 layout
  - DDM, metadata, databases
  - Computing operations
  - Calibration
- Thus, many boundary conditions as well as many services to provide

Baseline

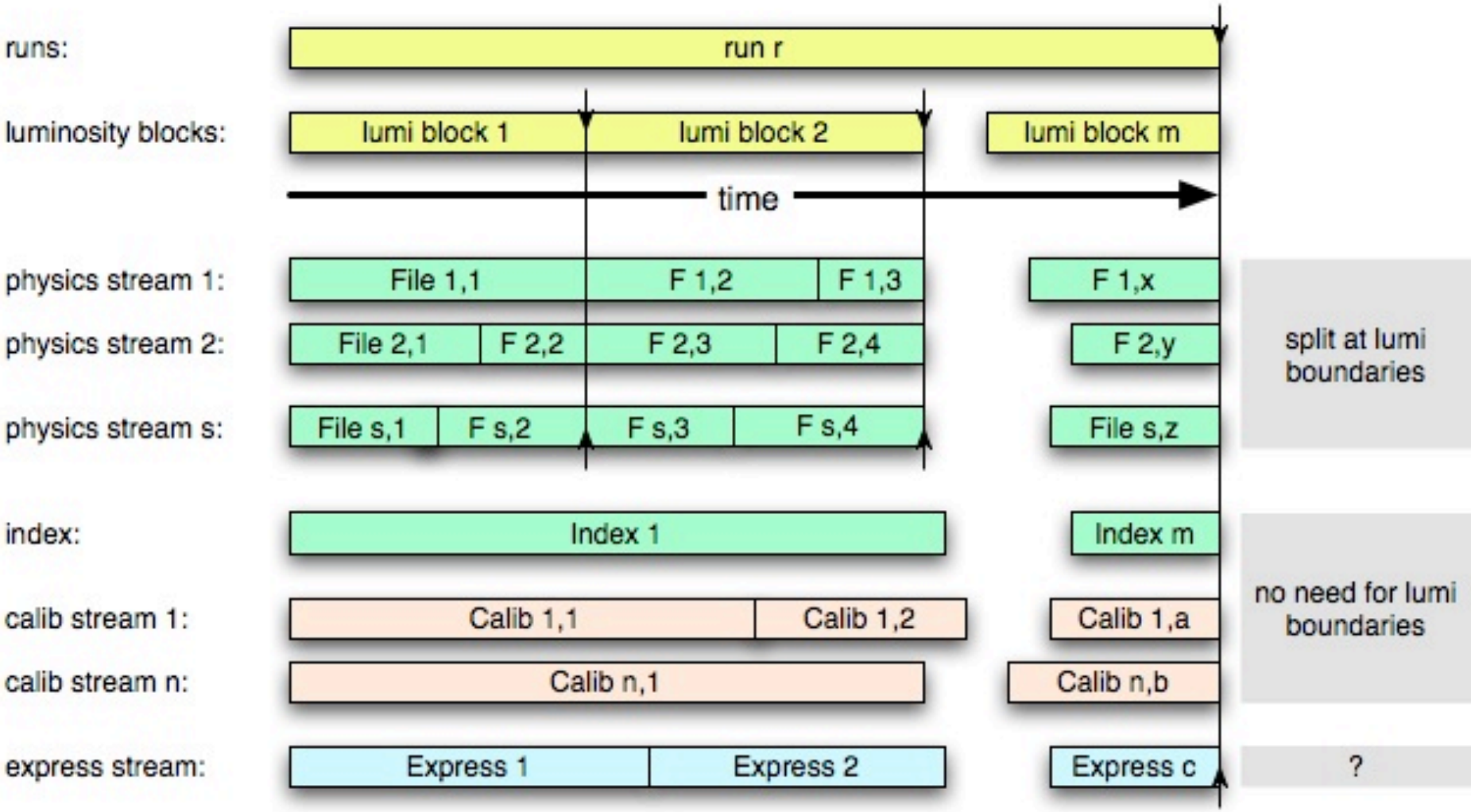
- Relevant documents include:
  - Online-Tier0 Taskforce (December 2004)
  - ATLAS Detector Calibration Model (February 2005)
  - Computing TDR (June 2005)
- Streams generated at Point1 as foreseen in the baseline:
  - One bulk physics stream
    - Important part of work is to study subdivision of this stream
    - If more than one, overlap between streams becomes an issue
  - One express physics stream
  - Several calibration streams
    - Ranging from partial event data (e.g. L1 Regions of Interest)
    - To full events which are used in physics analyses
  - One debug / diagnostic stream
- Guiding principles: efficiency, flexibility
  - Efficiency in the HLT has other criteria than offline, etc.

Overview of streams, files



- Hierarchy of (logical) streams: streams can be divided into narrower ones - but streams are never mixed

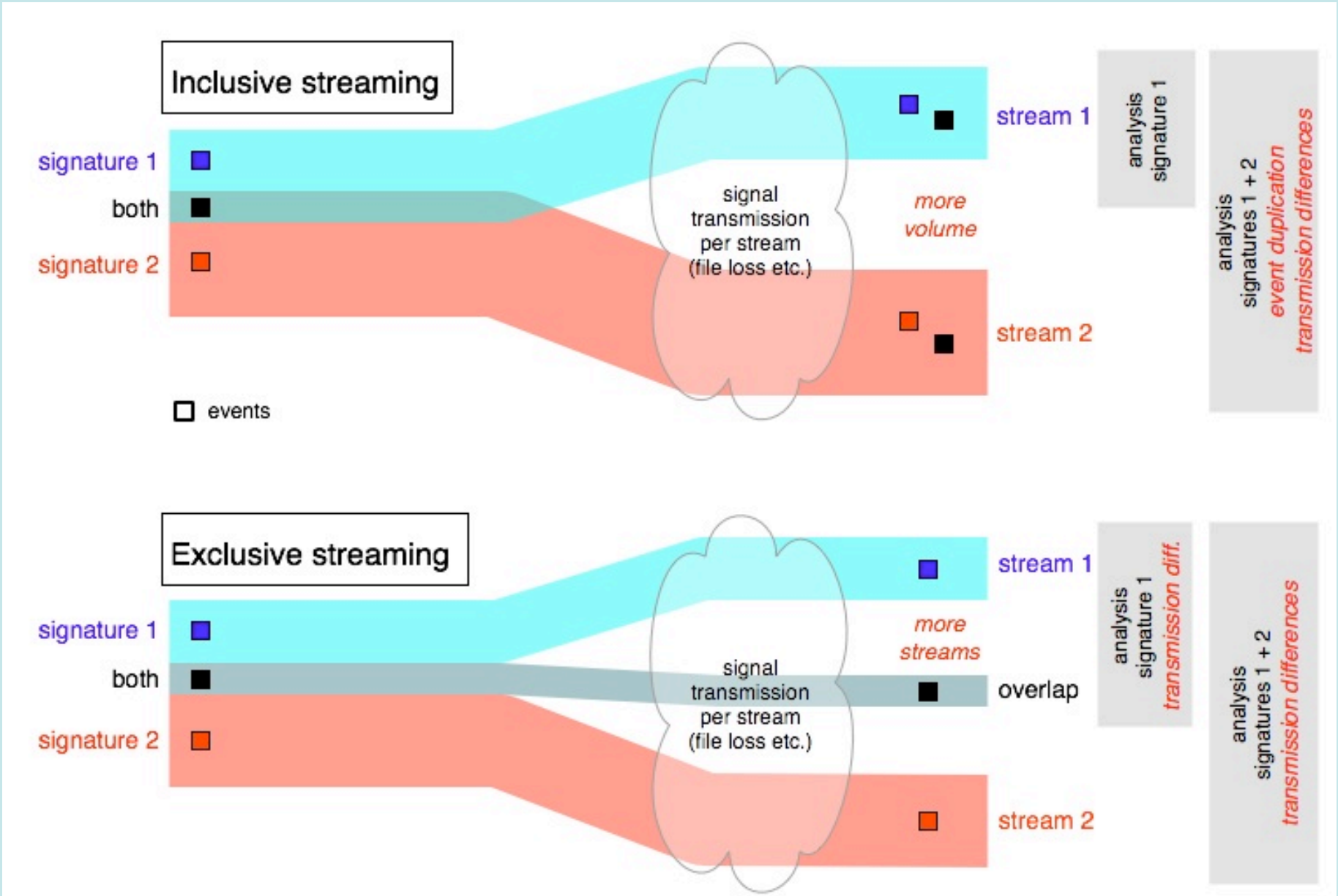
Overview of streams, files, lumi blocks



- That much we have per SFO
- More on event index: https://twiki.cern.ch/twiki/bin/view/Atlas/EventIndex

Overlapping (nclusive) vs. Disjoint (exclusive) streams

(possible problems in red italic)



Physics streams under study at present - "strawmen"

(J.-F. Arguin, LBL)

- **Express stream**
- **Stream A: electrons**
  - e25i, 2e15i, e15imu10 + prescaled triggers
- **Stream B: muons**
  - mu20i, 2mu10 + prescaled triggers
- **Stream C: jets**
  - j400, 2j350, 3j165, 4j110, jetSumEt1000, SumEt1000 +prescaled triggers
- **Stream D: photons**
  - gam60i, 2gam20i + prescaled trigger
- **Stream E: missing Et and taus**
  - tau60i, tau35i+xE45, xE200, j70+xE70 + prescaled trigger
- **Stream F: B physics**
  - 2mu6+mass, 1mu6, etc.
- **Stream G: Luminosity, zero-bias**
  - min. bias, zero bias, roman pots, etc.

- Each stream contains similar physics:
  - similar users and reprocessing needs
  - Minimize overlaps
- Based on triggers information (TDR)
- Overlaps allowed between streams but not within a stream (see example next page)
- Note: separate e and mu to avoid excessive access to a "super" single lepton stream

- Scenarios: 1-2*10^33, also initial trigger menu for 1*10^29
- Maybe combine A with D

Rates seen per stream (Rome and CSC samples)

(J.-F. Arguin, LBL)

Event Rates I (Hz) ($\int \mathcal{L}dt = 2 \times 10^{33} cm^{-2} s^{-1}$)

	Stream A			Stream B		all Stream C	
Processes	e25i	2el 5i	el5imu10	mu 20i	2mu10	j400	2j 350
Dijet (17-35 GeV)	0±0	0±0	0±0	0±0	0±0	0±0	0±0
Dijet (35-70 GeV)	42±5.7	0.79±0.79	0.79±0.79	3.1±1.6	13±3.1	1.6±1.1	0±0
Dijet (70-140 GeV)	4.4±0.79	0.14±0.14	0.28±0.2	0.28±0.2	2.8±0.63	0.28±0.2	0.28±0.2
Dijet (140-280 GeV)	0.41±0.092	0±1.1e-24	0.041±0.029	0.062±0.036	0.97±0.14	0.41±0.092	0.041±0.029
Dijet (280-560 GeV)	0.0023±0.00055	0±7e-27	0.00027±0.00019	0.00027±0.00019	0.029±0.0019	0.31±0.006	0.27±0.0056
Dijet (560-1120 GeV)	0.00042±0.00024	0.00014±0.00014	0±0	0±0	0.031±0.002	0.7±0.0016	0.64±0.0031
Dijet (1120-2240 GeV)	9.1e-06±3.7e-06	0±2.8e-28	0±0	6.1e-06±3e-06	0.0014±4.3e-05	0.011±1.5e-06	0.011±5.2e-06
Dijet (>2240 GeV)	3.9e-07±1.5e-07	0±1.8e-29	5.6e-08±5.6e-08	7.2e-07±2e-07	0.00013±2.3e-06	0.00048±0	0.00048±5.6e-08

Event Rates II (Hz) ($\int \mathcal{L}dt = 2 \times 10^{33} cm^{-2} s^{-1}$)

	all Stream C		Stream D		Stream E		Stream F
Processes	3j165	4j110	gam60i	2gam20i	tau35ixE45	j70xE70	2mu6cut
Dijet (17-35 GeV)	0±0	0±0	0±0	0±0	0±0	0±0	0±0
Dijet (35-70 GeV)	0.79±0.79	2.4±1.4	17±3.7	2.4±1.4	0±0	10±2.8	1.6±1.1
Dijet (70-140 GeV)	0.43±0.25	0.14±0.14	17±1.6	0.43±0.25	0.28±0.2	3.1±0.66	0.14±0.14
Dijet (140-280 GeV)	1.9±0.2	1.2±0.16	1.9±0.2	0.041±0.029	0.21±0.065	3±0.25	0.041±0.029
Dijet (280-560 GeV)	0.14±0.0042	0.055±0.0027	0.015±0.0014	0.00066±0.0003	0.002±0.00051	0.074±0.0031	0.00027±0.00019
Dijet (560-1120 GeV)	0.21±0.0046	0.085±0.0033	0.0072±0.001	0.00098±0.00037	0.0017±0.00049	0.12±0.0037	0.00084±0.00034
Dijet (1120-2240 GeV)	0.005±6.5e-05	0.0023±5.3e-05	0.0002±1.7e-05	9.1e-06±3.7e-06	2.4e-05±6e-06	0.0042±6.3e-05	3.8e-05±7.6e-06
Dijet (>2240 GeV)	0.00023±2.6e-06	0.00011±2.2e-06	9.8e-06±7.3e-07	3.3e-07±1.4e-07	1.1e-06±2.5e-07	0.00028±2.6e-06	4.1e-06±4.7e-07
γ+jet	0.0076±0.0054	0±0	8.8±0.18	0.45±0.041	0.038±0.0038	0.12±0.021	0±0
$W \to e\nu$	0±0	0±0	0.064±0.012	0.0021±0.0021	1.3±0.051	0.16±0.018	0±0
$W \to \mu\nu$	0±0	0±0	0.0031±0.0015	0.0015±0.0011	0.0077±0.0024	0.36±0.016	0.00077±0.00077
$Z \to ee$	0.00018±0.0001	0.00018±0.0001	0.017±0.00099	0.0029±0.00042	0.0034±0.00045	0.00048±0.00017	0±0
$Z \to \mu\mu$	0±0	0±0	0.00035±0.00011	3.2e-05±3.2e-05	0.00038±0.00011	0.039±0.0011	0±0
$Z \to \tau\tau$ (loose)	2.7e-05±2.7e-05	0±0	0.0013±0.00019	0.00083±0.00015	0.0029±0.00028	0.007±0.00043	0±0
γ/Z (30<M<81 GeV)	0±0	0±0	0.00061±0.00012	0.00037±9.4e-05	0.0028±0.00026	0.005±0.00034	0±0
γ/Z (M>100 GeV)	2.8e-05±7.5e-06	1.6e-05±5.7e-06	0.0035±8.3e-05	0.00035±2.6e-05	0.016±0.00018	0.011±0.00014	2e-06±2e-06
$\gamma\gamma$	2.2e-06±2.2e-06	0±0	0.017±0.00018	0.06±0.00029	8.9e-06±4.5e-06	3.4e-05±8.7e-06	0±0
$ZZ \to 4l$	2.3e-08±5.7e-09	1.9e-08±5.2e-09	1.8e-06±5e-08	3.3e-07±2.2e-08	1.4e-05±1.3e-07	8.1e-06±1e-07	8.6e-09±3.5e-09
ttbar ($\geq$ 1l)	0.0056±0.00017	0.0066±0.00018	0.0081±0.0002	0.00064±5.8e-05	0.099±0.00068	0.32±0.001	0.00018±3.1e-05
Single-top (Wg fusion)	8.5e-05±2.6e-05	1.5e-05±1.1e-05	0.00083±8e-05	4.6e-05±1.9e-05	0.0096±0.00026	0.021±0.00037	0±0
Single-top (Wt)	2.9e-06±1.7e-06	2.9e-06±1.7e-06	0.00029±1.7e-05	2e-05±4.4e-06	0.003±5.1e-05	0.0081±7.6e-05	2.9e-06±1.7e-06
Total	3.5±0.85	3.8±1.4	46±4	3.4±1.4	1.9±0.22	18±2.9	1.8±1.1
Predictions	~ 25 (Stream C)		~ 40		~ 5	~ 20	~ 10

Also prescaled triggers taken into account

(J.-F. Arguin, LBL)

- 53 prescaled triggers, inspired from HLT TDR
- Scenario: 1*10^33
- Remaining bandwidth (200-115 = 85Hz) shared between prescaled triggers

Rate for prescaled triggers

Streams	Triggers	Prescale	Output rate
A	e25	231	0.988±0.566
	e20i	102	2.4±0.894
	e15i	38.4	
	e10i	151	
	e7i	775	
	2e5i	1.77e+03	
B	mu20	60.5	
	mu15i	23.9	
	mu10i	94.9	
C	j300	7.74	
	j200	69.5	
	j150	277	
	j100	1.48e+03	
	j70	5.68e+03	
	j50	1.95e+04	
	j20	3.24e+05	
	2j200	17	
	2j150	81.9	
	2j100	473	
	2j70	1.72e+03	
	2j50	5.55e+03	
	2j20	9.9e+04	
	3j150	2.03	
	3j100	23.6	
	3j70	136	
	3j50	600	
	3j20	1.94e+04	
	4j100	2.22	
	4j70	18.6	
	4j50	118	
	4j20	5.61e+03	

Rate for prescaled triggers

Streams	Triggers	Prescale	Output rate
D	gam40i	87.3	1.19±0.571
	gam25i	509	0.912±0.561
	gam20i	1.03e+03	0.193±0.101
	gam15i	2.72e+03	0.413±0.394
	gam10i	1.01e+04	0.0756±0.0709
	gam7i	2.02e+04	0.397±0.394
	2gam7i	601	3.79±3.79
E	tau60i	21.6	1.75±0.704
	tau40i	35.1	4.51±3.8
	tau35i	38.4	0.688±0.401
	tau25i	53.9	0.833±0.413
	2tau35i	1.07	1.62±0.107
	2tau25i	1.81	1.55±0.0386
	xE200	1.03	1.62±0.396
	xE150	1.1	1.64±0.397
	xE100	3.15	1.61±0.572
	xE70	20.6	0.787±0.413
	xE45	237	0.192±0.101
F	mu6i	281	0.0282±0.0031
	mu5i	384	3.82±3.79
	2mu6	28	1.42±0.69
	2mu5	51.5	0.553±0.405
Total		–	85.8

Results for overlaps (unprescaled + prescaled triggers)

(J.-F. Arguin, LBL)

Final Overlap Table (rate in Hz)

Streams	electrons (A)	muons (B)	jets (C)	photons (D)	xE and taus (E)	B physics (F)
electrons (A)	38±8.8	0.1±0.0057	0.0039±0.0015	0.0077±0.0022	3.1±0.4	8.6e-06±4.5e-06
muons (B)	–	40±10	0.022±0.015	0.0034±0.002	0.3±0.075	1±0.57
jets (C)	–	–	54±8.8	0.11±0.03	0.7±0.4	0±0
photons (D)	–	–	–	40±7	0.055±0.013	0±0
xE and taus (E)	–	–	–	–	26±6.8	3.9e-06±3.9e-06
B physics (F)	–	–	–	–	–	1.6±0.7

Note: table contains only events passing 1 or 2 streams
Rate for passing 3 or more streams is 0.075 ± 0.019 Hz

Total overlap = 2.726 ± 0.698%

- Diagonal elements: rates for passing one and only one stream
- Off-diagonals: rates for passing two streams

➔ Overlap (in inclusive streaming) is not a problem of data volumes

Physics use cases with the strawman streams
(S. Binet, LBL)

- Top analyses:
  - mass (lepton+jet, dilepton)
  - x-section
- W analyses:
  - mass
  - x-section
- Higgs mass
  - various mass ranges
  - w/ w/o VBF
- Calibration
  - ECAL
  - Spectro.
  - Inter-calibration
  - JES:
    - Z/Gamma+jet
    - E/p
    - W -> jj

S.Binet

- SUSY:
  - SUGRA signatures
  - SUGRA @ large tanBeta
  - GMSB
  - R-parity breaking models
- Extra-dimensions
- Technicolor
- Drell-Yan
  - dilepton mass spectrum
  - gamma gamma prod.
- QCD
- B-Physics
- Minimum bias

Frequency of usage of streams (preliminary)

(S. Binet, LBL)

Analysis Name	Stream A	Stream B	Stream C	Stream D	Stream E	Stream F	Stream G	# Analyses
Top mass (dilepton, l+jet), cross section	3	3						3
W cross section, mass	2	2						2
Higgs mass analyses	11	10	1	2	1			12
SUSY	12	12	1	1	14			24
Technicolor	3	3	1					4
Extra-dimensions	2	2	1		2			5
Drell-Yan		1		1				2
QCD			1	1				2
B-physics						10		10
Calibration	4	5		1	2			7
mini. bias, zero bias							1	1
TOTAL	37	38	5	6	19	10	1	72
Ratio	~ 50%	~ 50%	~ 7%	~ 8%	~ 25%	~ 15%	~ 2%	

Analysis Name	A + B	A + D	C+D	C+E	A+B+E	A+B+D
Top mass (dilepton, l+jet), cross section	3					
W cross section, mass	2					
Higgs mass analyses	10	1	1		1	
SUSY	12	1		1		1
Technicolor	3					
Extra-dimensions	2					
Drell-Yan						
QCD						
B-physics						
Calibration	3			2	2	
mini. bias, zero bias						
TOTAL	35	2	1	3	3	1
Ratio	~ 50%	~ 3%	~ 2%	~ 5%	~ 5%	~ 2%

- e and mu streams are dominating, also in combination e+mu

File sizes from SFOs
(S. Gadomski, Bern)

Trigger menu and rated from TDR, table 13-11							
Luminosity 2x10^33			4	SFOs		1.6	MB/event
HLT signature	**Rate [Hz]**		**Stream**	**Rate [Hz]**	**[Hz/SFO]**	**ev/min/SFO**	**MB/min/SFO**
e25i	40						
2e15i	1						
gamma60i	25						
2gamma20i	2	A+D	egamma	68	17	1020	1632
mu20i	40						
2mu10	10	B	mu	50	13	750	1200
j400	10						
3j165	10						
4j110	10						
j70+xE70	20						
tau35i+xE45	5	C+E	jtauetmiss	55	14	825	1320
2mu6 with vertex etc	10	F	Bphysics	10	3	150	240
Others (incl calibration)	20	G++	others	20	5	300	480
Total	203			203	51	3045	4872

- Number and size of files are a potential worry at Tier0 (Castor I/O, book-keeping)
- Assumption: 1 minute lumi blocks, files closed at lumi block boundaries
- Note: fewer SFOs - matches network bandwidth; improved disk I/O rate used

➔ All physics file sizes are reasonable - but beware: calibration is subdivided further

File sizes from SFOs - exclusive streaming

(S. Gadomski, Bern)

Scenario 1
L=2e33, unprescaled only

parameters
SFO 10
LB [min] 5
MB/event 1,6

Stream	Rate [Hz]	ev/LB/SFO	MB/LB/SFO
electrons (A)	58,00	1740	2784
muons (B)	67,00	2010	3216
jets (C)	8,90	267	427
photons (D)	48,00	1440	2304
xE and taus (E)	34,00	1020	1632
B physics (F)	1,70	51	82
Overlap AE	1,30	39	62
Overlap CE	1,20	36	58
Other overlaps (13*XY+XYZ+…)	0,50	15	24
Total	220,60	6618	10589

- The additional overlap streams generate small, additional files
- Also, their choice depends on lumi (and background) conditions
- From online point of view, can afford exclusive streams only with some simplifications
- Prefer inclusive streams

Scenario 2
L=1e33, including prescaled

parameters
SFO 10
LB [min] 5
MB/event 1,6

Stream	Rate [Hz]	ev/LB/SFO	MB/LB/SFO
electrons (A)	38,00		
muons (B)	40,00		
jets (C)	54,00		
photons (D)	40,00		
xE and taus (E)	26,00		
B physics (F)	1,60		
Overlap AE	3,10		
Overlap BF	1,00		
Other overlaps (13*XY+XYZ+…)	1,38		
Total	205,08		

Scenario 2
L=1e33, including prescaled

parameters
SFO 10
LB [min] 5
MB/event 1,6

Stream	Rate [Hz]	ev/LB/SFO	MB/LB/SFO
electrons (A)	38,00	1140	1824
muons (B+F)	42,60	1278	2045
photons (D)	40,00	1200	1920
jet+tau+xE (C+E)	80,70	2421	3874
Overlaps	3,78	113	181
Total	205,08	6152	9844

Calibration streams to Tier0
(from Detector Calibration model, Feb. 2005)

System	Stream	Readout	Rate×size Hz × kB	Source	Comments/data type
ID	Generic tracks p_T>2 GeV	ID ROI	100 × 40	EF all	Custom trackfit+hit data
LAr	Electrons p_T>20 GeV	EM ROI	50 × 50	EF all	5-sample EM RAW data
Muon	Muon p_T> 6 to 20 GeV LVL1	MDT/CSC /RPC/TGC ROI	1000 x 1	LVL2 all	Custom trackfit using trigger and precision hits
MDT	High p_T muons in large/ small chamber overlap	MDT/RPC ROI	6 x 5	EF special	MDT/RPC hits only in large / small chamber overlap regions
HAD calos	isolated hadron p_T > ~ 20 GeV	ROI (here 0.4 x 0.4)	~ 5 x 400 ?	All	Single-prong tau trigger (needs more study)
All	Inclusive e/μ p_T>20 GeV	Full event	20 x 1600	EF all	Duplicate interesting events
All	di-leptons (Z→ ll)	Full event	1x1600	EF all	Duplicate Z→ ll events
All	Prescaled minimum bias	Full event	1x1600?	EF special	Duplicate minimum bias events

- Some streams have clean separation from physics
- Duplication calibration/physics (even triplication with express) to be avoided where possible - although calib streams are "cheaper" without lumi assignment
- Need to be revisited if more than one physics stream written
- Also need extra/updated input from some detectors

File size/number calculations at Tier0
(L. Goossens, CERN)

- Current model (one bulk physics stream from SFOs)
  - One sequence of RAW files per SFO, each file 1000 events (no lumi blocks) = 1.6GB, 17k RAW files/day
  - Reconstruction job inputs one RAW file, 1 ESD and 10 AOD files per RAW, 17k ESD + 170k (temporary) AOD files/day
  - Merging of temporary AOD files belonging to one run (~100) into one AOD, 1.7k AOD files/day
- New model (assume 7 streams, 7 SFOs, 1min lumi blocks: worst case)
  - 7 sequences of RAW files per SFO, 200Hz*1min/49 = 250 events per RAW file, 68k RAW files/day
  - Reconstruction job can input 4 RAW files, produces 2 AOD streams per RAW stream, 17k ESD + 34k AOD files/day
  - ➔ 4* number of RAW, but only 1/5 number of AODtmp due to streamed RAW
  - ➔ Absorb the factor 4 e.g. in 5min lumi block duration
- The small calibration files need additional attention
  - Proposed not to chop calibration streams into lumi blocks
  - Study merging of files from different SFOs for low-volume streams

Status

- Settled: physics streams definition is based on trigger signatures
- Settled: ability (and usefulness) to produce several RAW physics streams at Point1
  - react flexibly to different calibration needs
  - more flexible re-reconstruction - esp. during startup
  - some full-event calibrations streams will be physics streams instead - e.g. inclusive electrons, muons
- Luminosity conditions have little effect on overlaps
  - even for initial conditions (10^29 /cm2/sec)
- Inclusive vs. exclusive streams under further discussion
  - so far, inclusive streams preferred by most
- Need detailed contents of express stream
- Need updated list of detector calibration streams

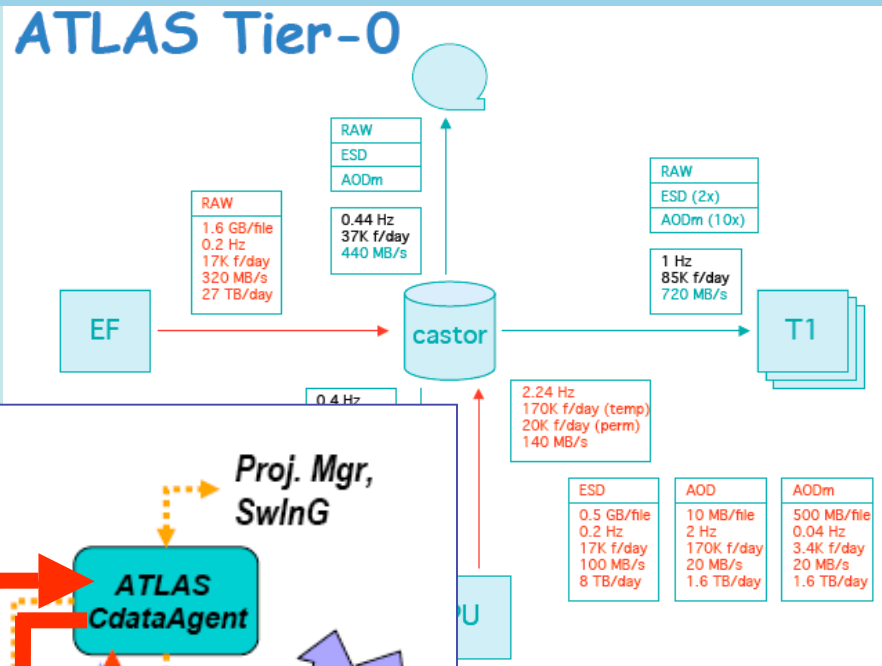
Extra slides

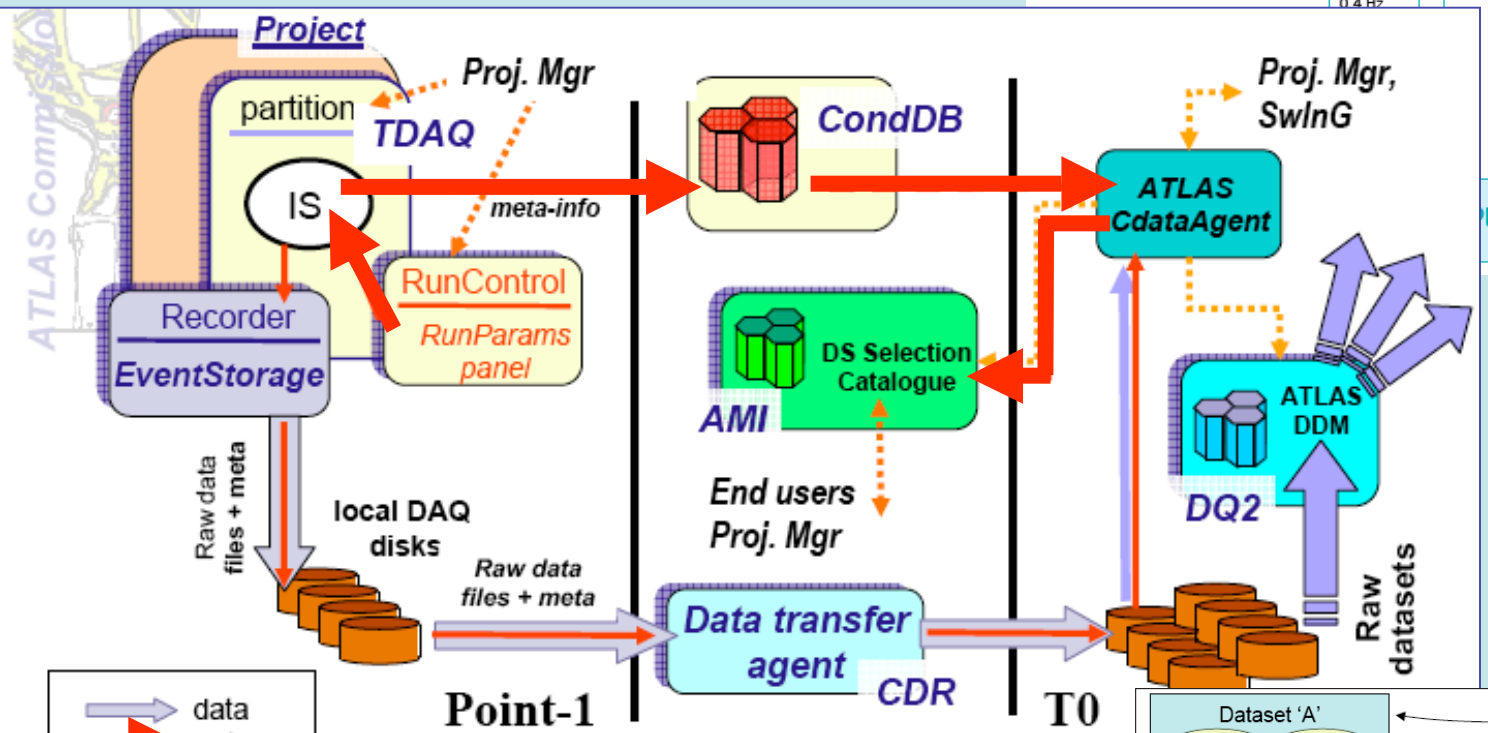
Distributed data management

Flow of Data and Metadata from the Cavern via Tier0 worldwide: **Distributed Data Management**

Primary metadata originate in RunControl and proceed via COOL to cataloging in DQ2 (AMI: metadata DB)

DQ2 replicates sets of data files (datasets) to the remote sites which can subscribe to datasets





Options under study:

CDR or DQ2 for data transfer from Point1 to Tier0

AMI or alternative (review underway)