Hardware Tracking for the Trigger (HTT) in ATLAS

Richard Brenner - On behalf of the ATLAS TDAQ Collaboration

design towards first prototypes

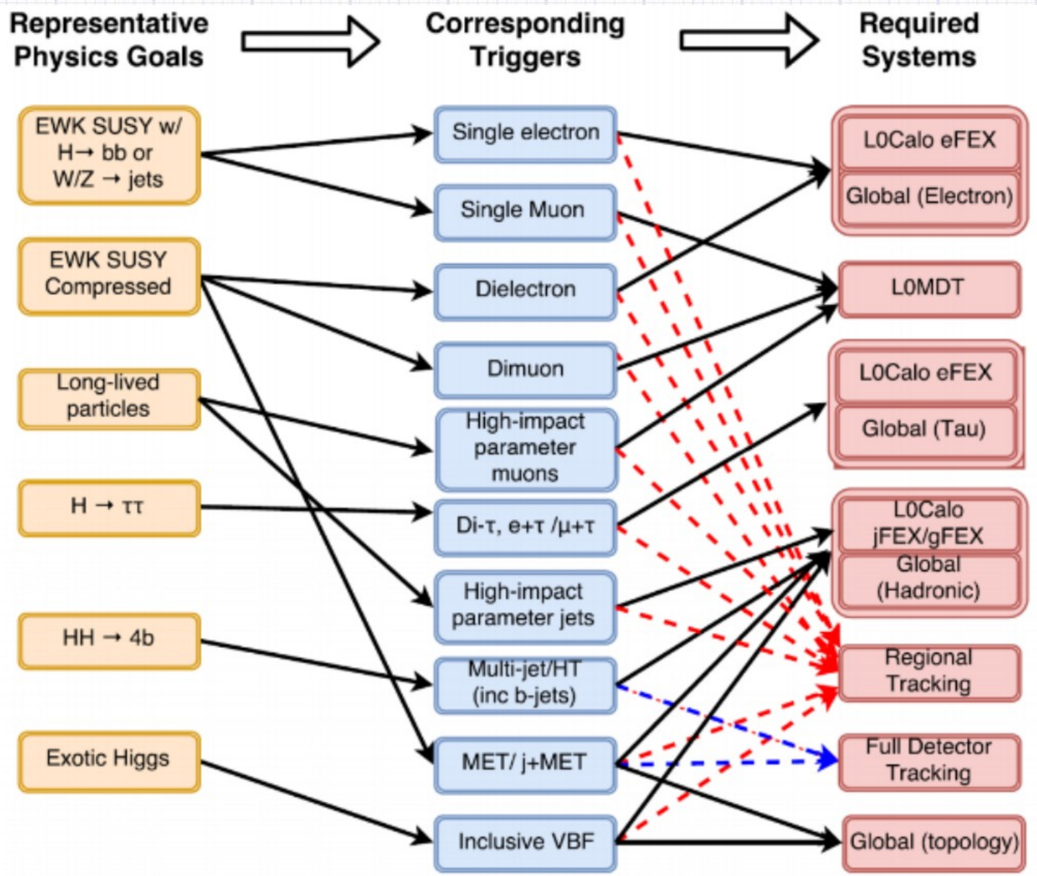
Material in this presentation has been reported in Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System

From LHC to HL-LHC

UPPSALA UNIVERSITET

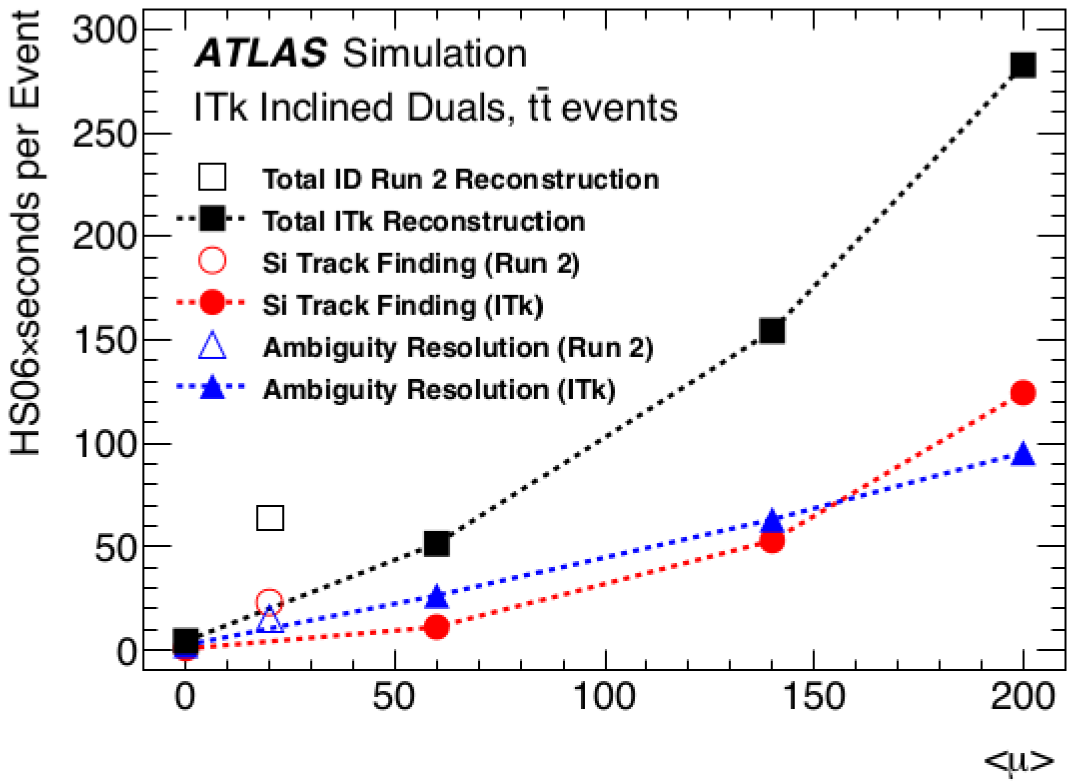
- ATLAS has pioneered tracking with hardware based computing:FastTracker (FTK).
- The Hardware Tracking for the trigger , HTT, is built on the legacy of the FTK system but several improvements have been made to make the implementation of HTT easier.
  - HTT runs in Event Filter which allows for commissioning without beam.
  - HTT system built with fewer unique components than FTK which will improve interfacing.
- An evolution system capable of higher trigger rates (L1Track) is foreseen.
- The firmware components of HTT is a direct evolution of FTK.
- Detailed studies of the architecture, requirements and specification of the HTT for HL-LHC has been done and will be presented in this talk.

ATLAS trigger menu target



- ATLAS is planning to run a single level trigger (L0) with a maximum rate of 1MHz (currently 100kHz).
- A regional Hardware Tracking for the Trigger (rHTT) will run as a co-processor in the Event Filter (EF) reducing the rate <350kHz.
- A second stage global HTT (gHTT) will run on full detector information up to 100kHz rate reducing it further.

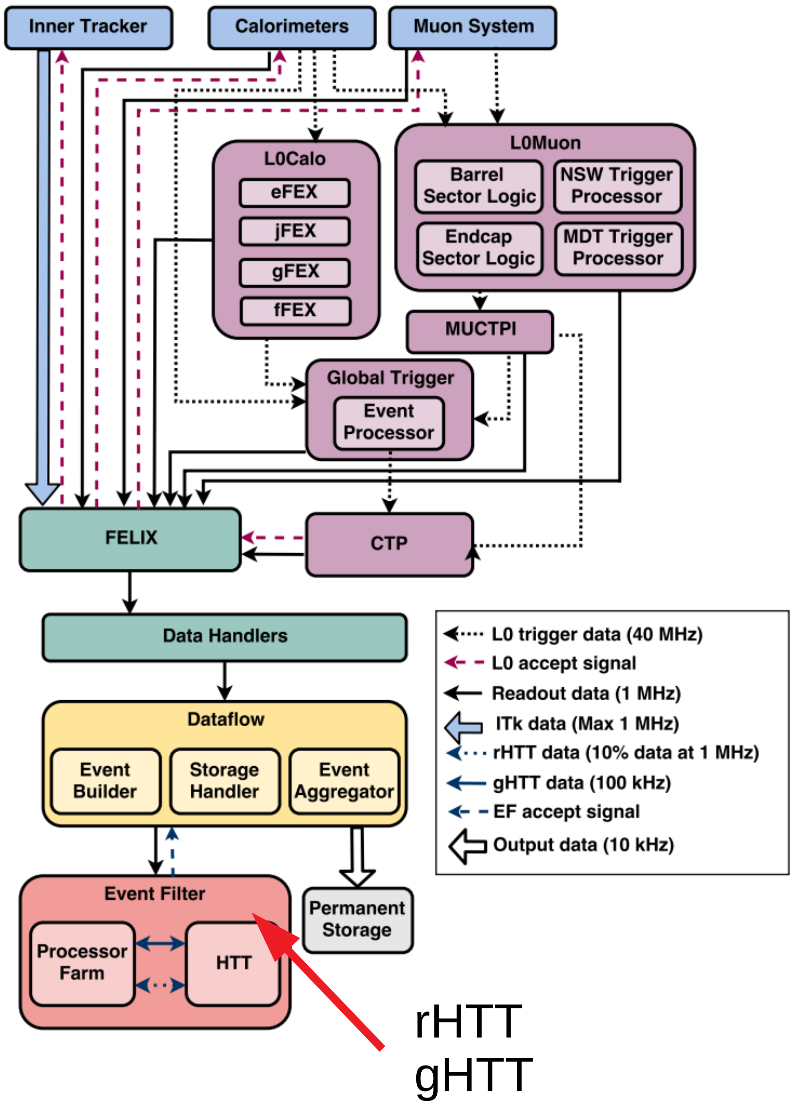
CPU resources for tracking run 2 vs. upgrade Phase II



- Hardware tracking is technologically more risky than CPU but there are advantages that lead to the selection of hardware based tracking.
  - efficient use of resources
  - low(er) power consumption
  - low latency

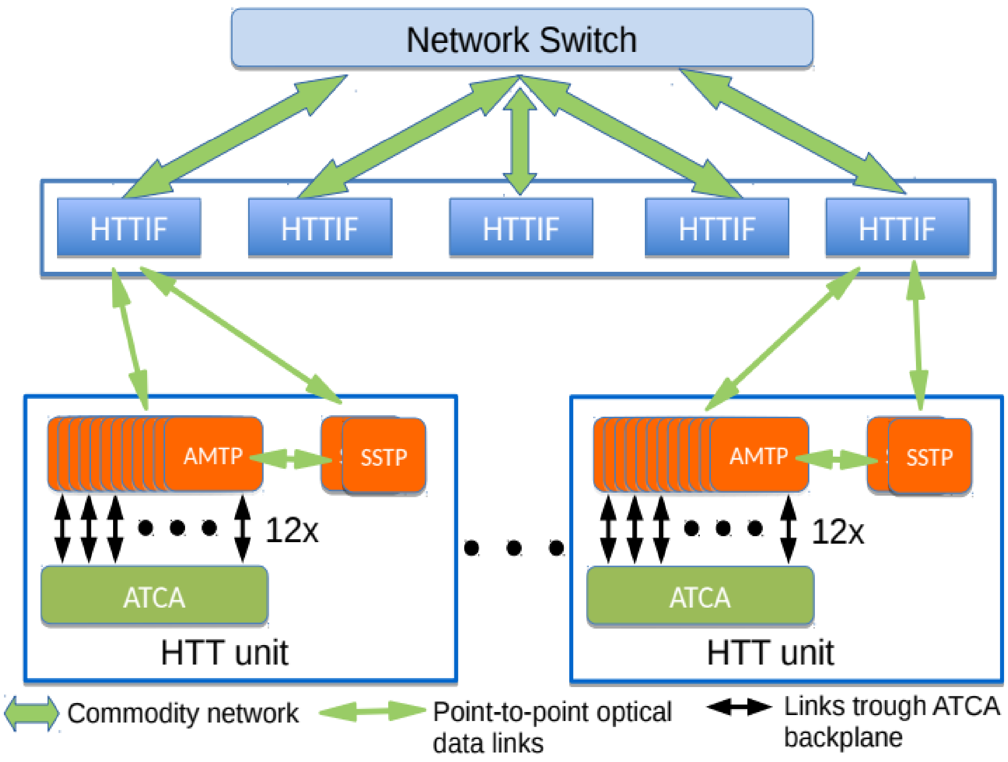
Baseline TDAQ architecture

UPPSALA UNIVERSITET



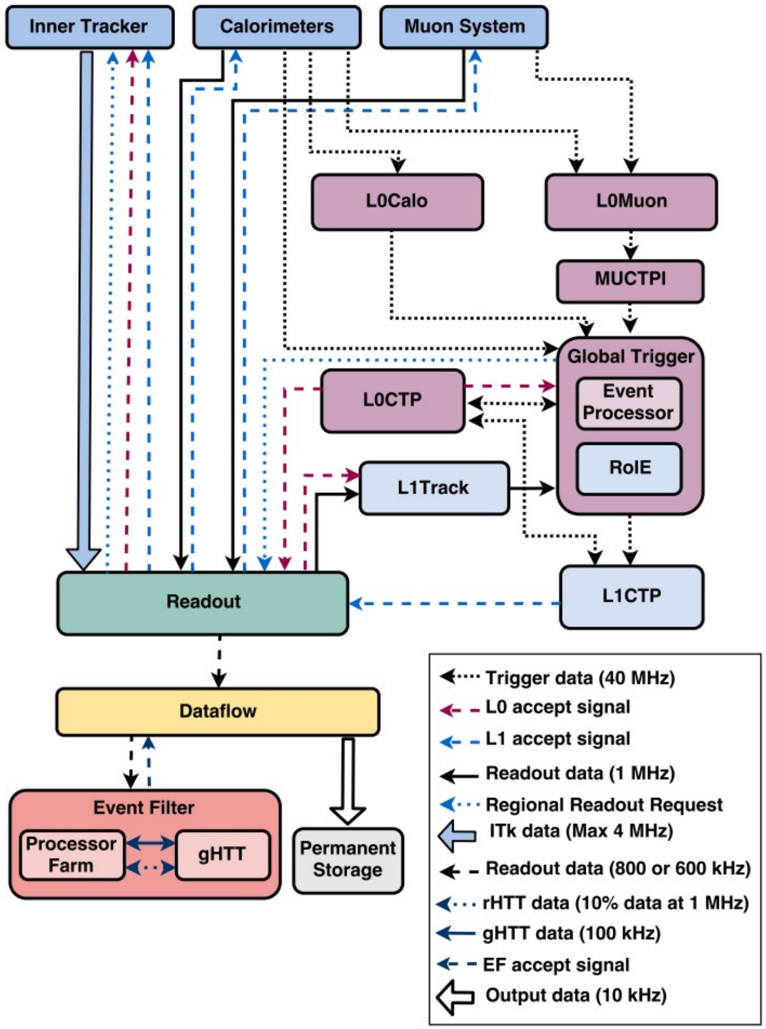
- The Event Filter is a CPU farm connected with high speed commodity network.
- HTT acts as a co-processor for the Event Filter.
  - **rHTT** can process events up to 1 MHz, the same as the EF (pT>2GeV).
  - **gHTT** can process 10% of events (100 kHz), when required by the trigger menu (pT>1GeV).
  - The EF processor unit taking care of sending data to (g/r)HTT.
  - No strict latency constraint.
- Lepton trigger threshold lower than in run 1,2 &3 (a full list of triggers, thresholds and rates in back-up)

Baseline HTT architecture



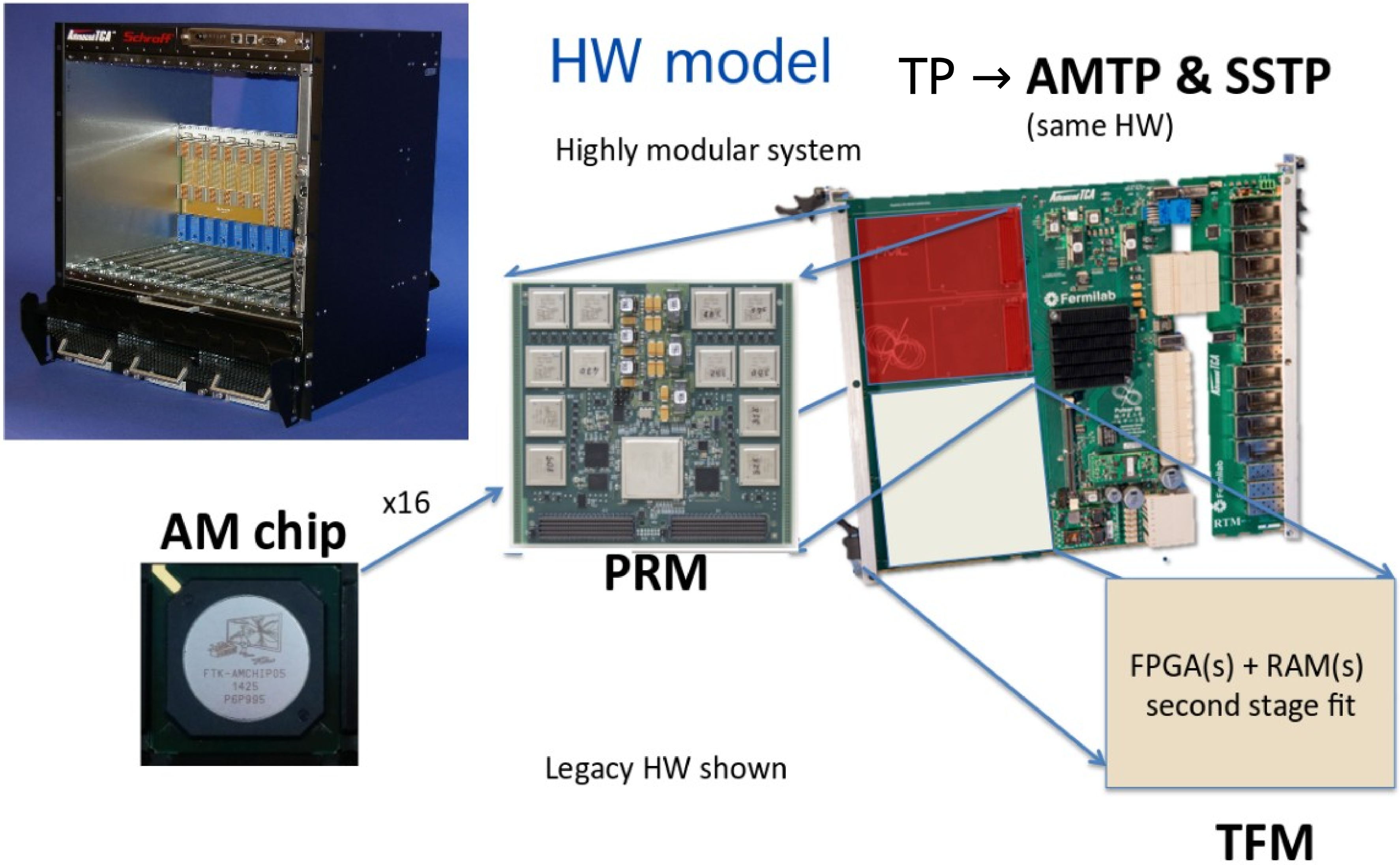
- The HTT units (both global gHTT and regional rHTT) are interfaced to the same commodity network via dedicated CPU servers (**HTTIF**).
- A HTT unit consists of:
  - 12 Associative Memory Tracking Processor (**AMTP**) cards performing pattern recognition
  - 2 Second Stage Trigger Processor (**SSTP**) performing track fitting.
- The AMTP and SSTP cards are located in separate **ATCA** crates.
- The AMTP and SSTP cards are built on a common Tracking Processor (**TP**) platform which are given its functionality by mezzanine cards.

Evolved system architecture

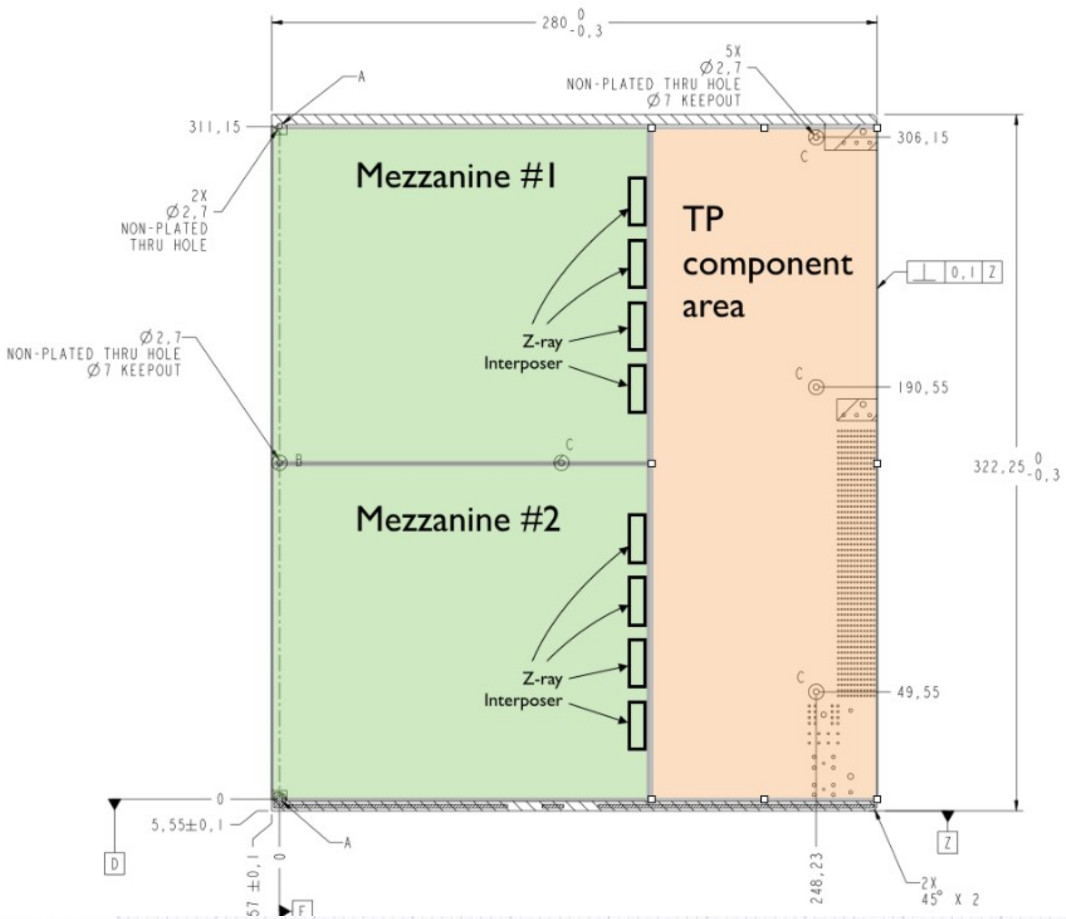


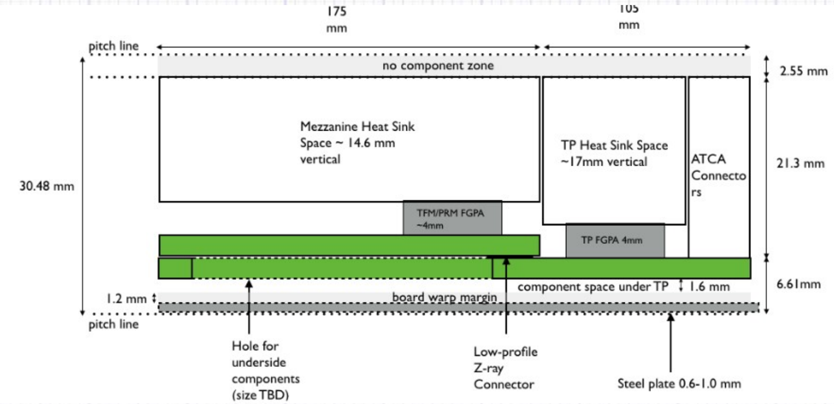
- Two level trigger architecture L0/L1 with L0 running up to 4MHz and L0<1MHz.
- HTT reconfigured to **L1Track** that run regional tracking on up to 10% of 8 tracking layers at a rate up to 4MHz (pT>4GeV).
- L1Track is run on data stream before EF.
- L1Track has a 6μs latency constraint → duplication of pattern banks.
- gHTT run as co-processor to EF.
- Primarily gain for hadronic triggers. (Details to be found in back-up)

Overview of Hardware



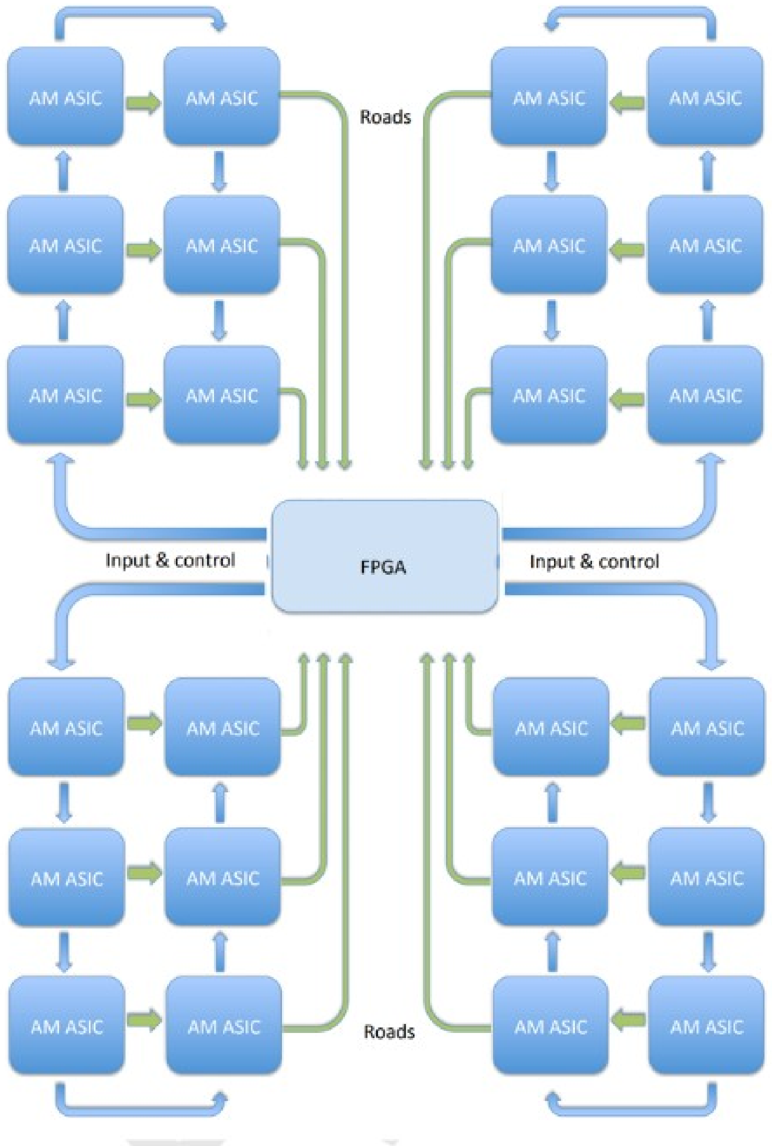
Tracking Processor board (TP)

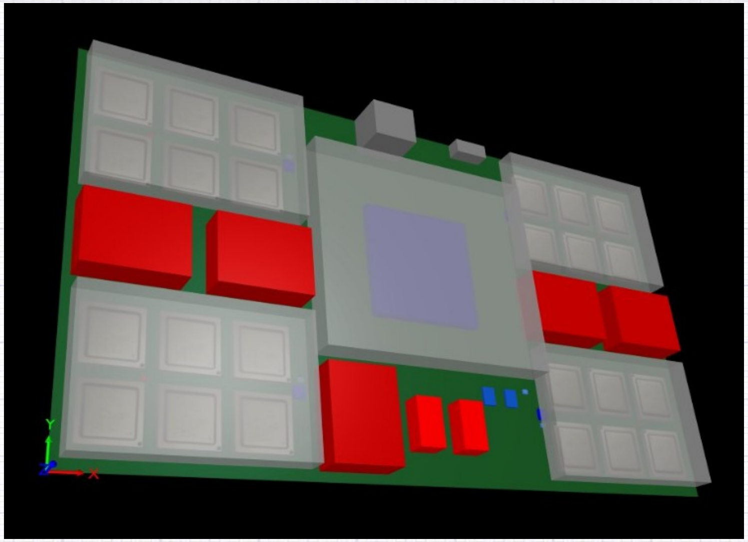




- The TP will handle the data communication and formatting to /from the mezzanine cards.
- Due to the high power dissipation the cooling of the electronics must be maximized
- We are investigating several low-profile connectors and selected Z-ray for our boards.

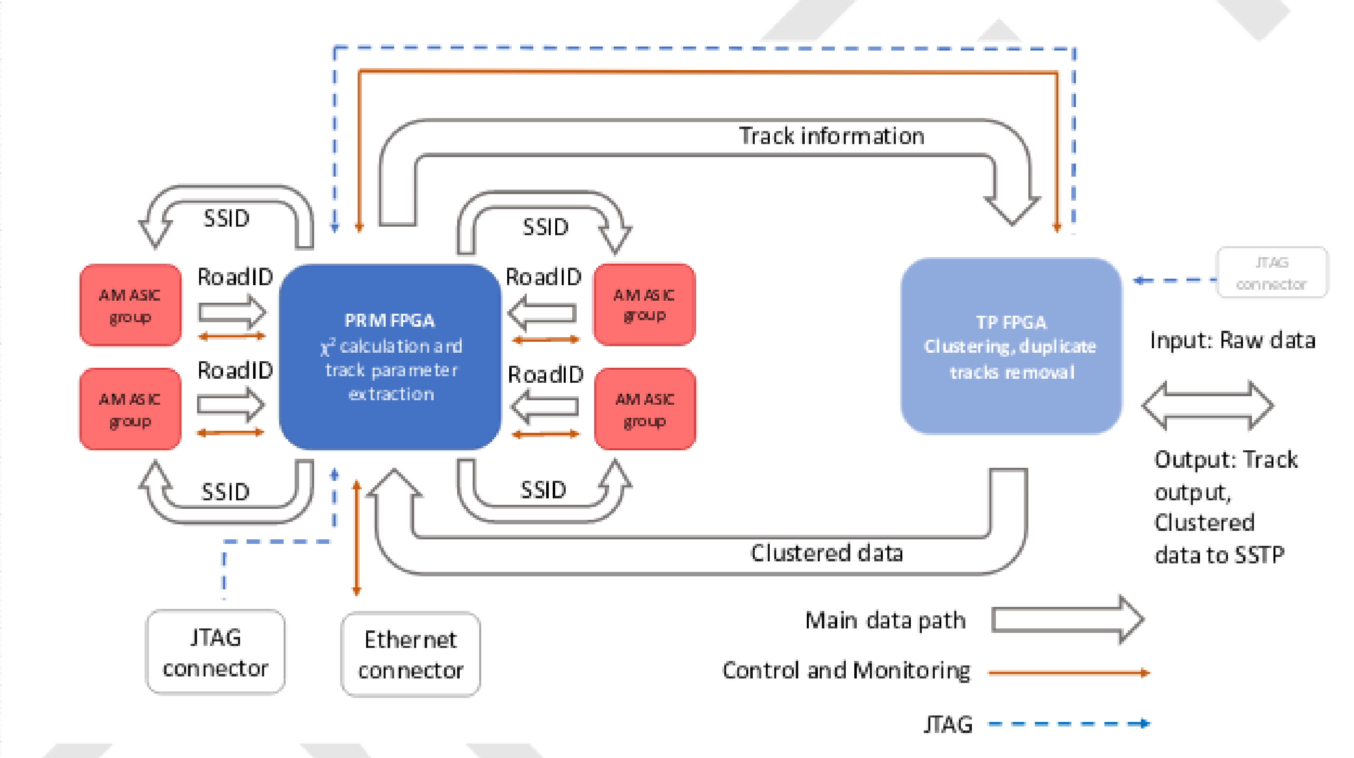
Pattern Recognition Mezzanine (PRM)





- PRM is a single board occupying full width of TP (1 mezzanine/AMTP)
  - Associative Memory (**AM**) ASICs/PRM: 24
  - Patterns/AM group: 2.3M
  - I/O bandwidth: 10 Gbps
  - Peak cluster rate/layer: 250MHz
  - Fit rate: 1GHz
- PRM will hold patterns for both rHTT and gHTT
- Performs first level fitting (8-layers)

PRM interaction

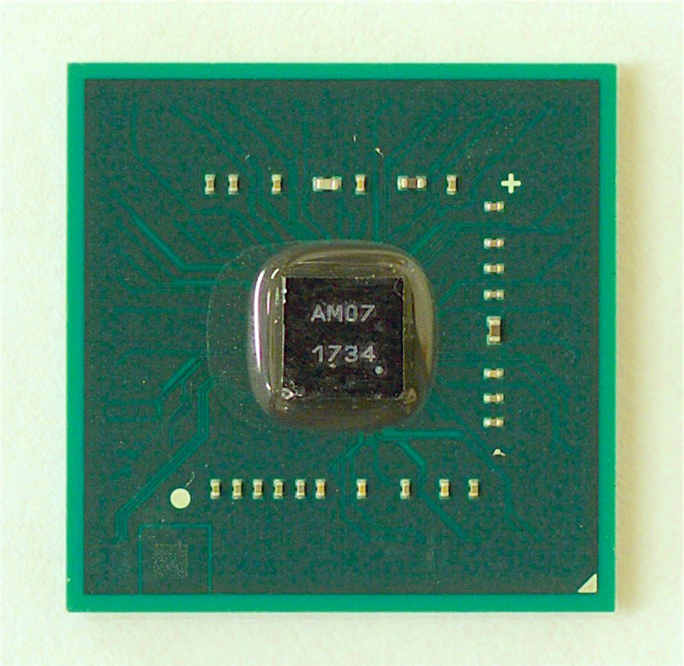


- PRM:
  - convert clusters into patterns of Super Strips
  - patterns are matched to patterns loaded in AM chips
  - Match patterns → roads. roads → cluster combinations
  - cluster combinations are fitted with first stage fitting
  - track information sent to AMTP
- AMTP:
  - receives data from EF and sends clusters to PRM
- AMTP:
  - **rHTT:** tracks and clusters returned to EF
  - **gHTT:** tracks and clusters sent to SSTP

AM09 ASIC

- **AM09:** production version used in HTT. A low energy usage per operation: to achieve this requirement a new CAM cell has been designed: the KOXORAM. Simulation of the new CAM cell predicts better energy efficiency: 0.30 fJ/comparison/bit compared to the 0.80 fJ/comparison/bit for the XORAM used before.
- 8 layers with an input bandwidth of 4 Gb/s per layer. 384 kpatterns/chip (3 times more in same area than previous production version **AM06**).
- Estimate of power consumption

  P = 1W+ < inputrate > ∗0.05W/MHz
- Submission planned in 2020

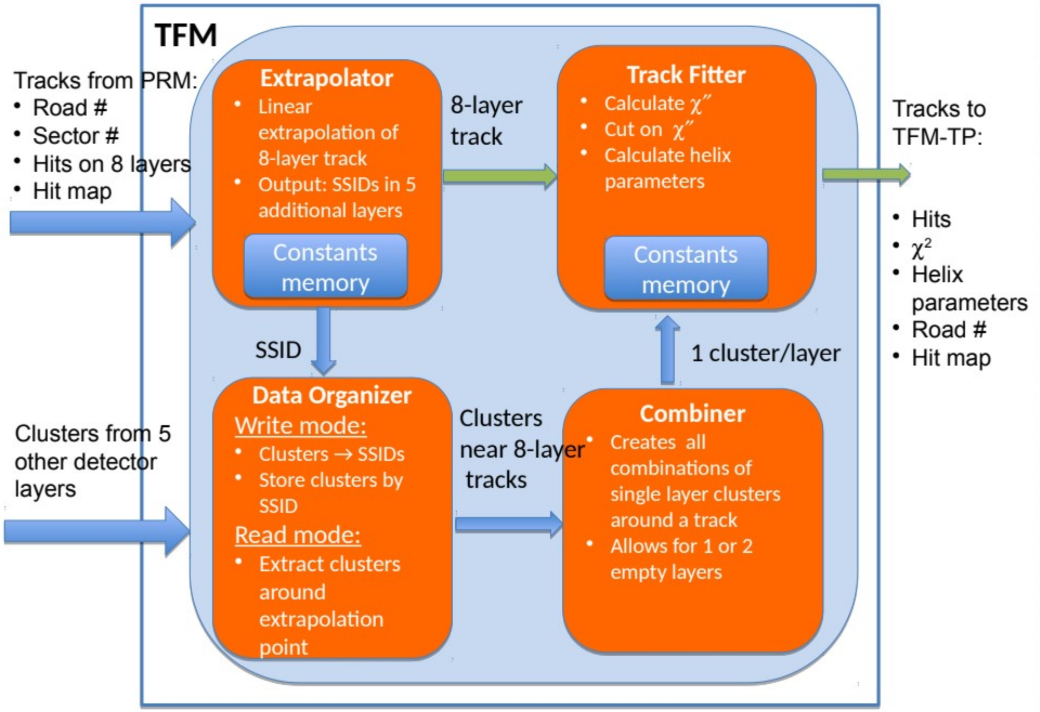


AM07 prototype with 16kpatterns

Chip name	Transistor count	Year	Brand	Technology	Area
Core 2 Duo Conroe	291,000,000	2006	Intel	65 nm	143 mm²
Itanium 2 Madison 6M	410,000,000	2003	Intel	130 nm	374 mm²
Core 2 Duo Wolfdale	411,000,000	2007	Intel	45 nm	107 mm²
AM06	**421,000,000**	**2014**	**AMteam**	**65 nm**	**168 mm²**
Itanium 2 with 9 MB cache	592,000,000	2004	Intel	130 nm	432 mm²
Core i7 (Quad)	731,000,000	2008	Intel	45 nm	263 mm²
Quad-core z196[20]	1,400,000,000	2010	IBM	45 nm	512 mm²
Quad-core + GPU Core i7 Ivy Bridge	1,400,000,000	2012	Intel	22 nm	160 mm²
Quad-core + GPU Core i7 Haswell	1,400,000,000	2014	Intel	22 nm	177 mm²
AM09	**1,684,000,000**	**2019**	**AMteam**	**28 nm**	**150 mm²**
Dual-core Itanium 2	1,700,000,000	2006	Intel	90 nm	596 mm²

Track Fitting Mezzanine (TFM)

- Receives tracks and clusters from first stage fitting of gHTT (rHTT do not use TFM)*
  → second stage fitting
- Extrapolates tracks to find cluster within search window (if a pixel hit is missing, the hit is guessed)
- Performs full fit of hit combinations
- Output hits and track parameters to SSTP



*) First stage fitting is performed on PRM with hits from the 8 layers used for pattern matching.

System performance (PRM)

UPPSALA UNIVERSITET

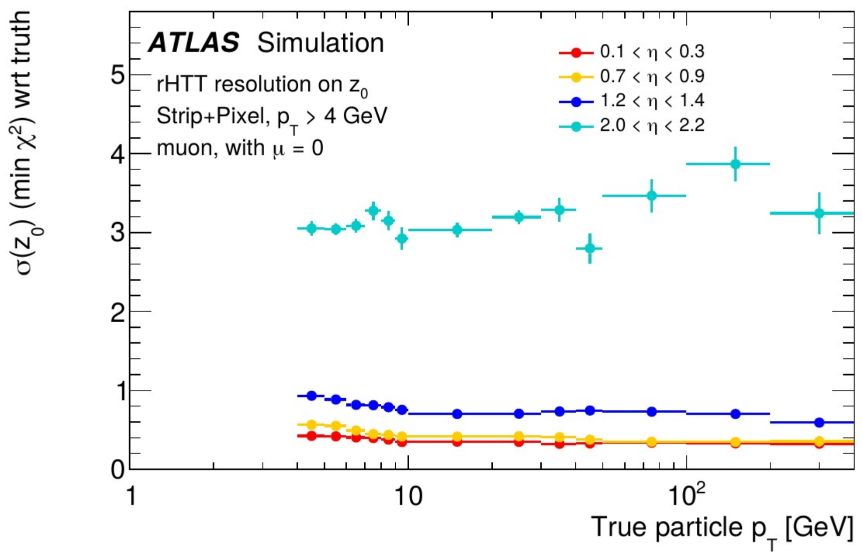
- The system has been simulated in 4 eta regions with full simulation with pT>4
- One region has been studied for pT > 1,2 and 4 GeV

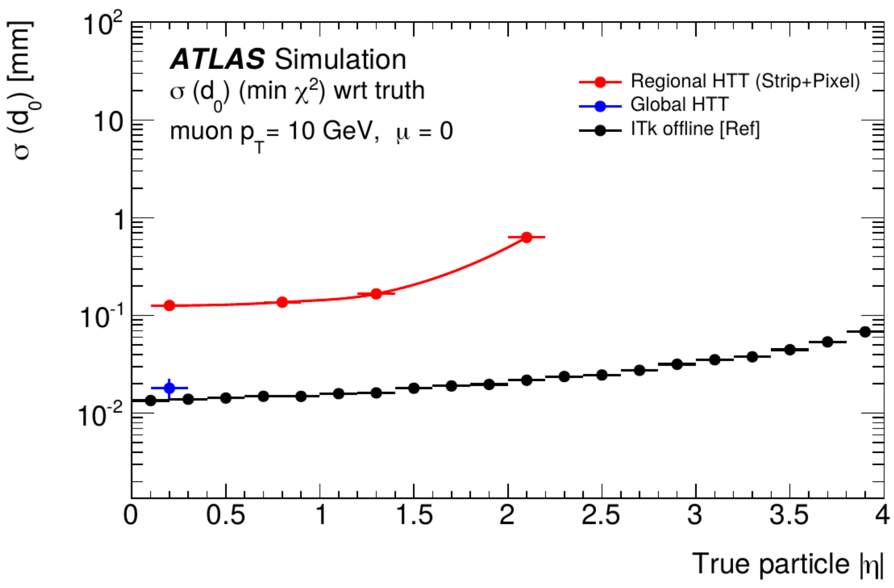
η range	muon eff.	mean matches pile-up	99% interv matches i pile-up
$0.1 < \eta < 0.3$	99.1%	31	151
$0.7 < \eta < 0.9$	99.2%	21	93
$1.2 < \eta < 1.4$	98.8%	42	159
$2.0 < \eta < 2.2$	98.7%	10	56

particle	min p_T	Eff. (%)	# roads	# fits	# tracks $\chi^2 < 40$	# tracks HitWarrior	# fit constants
muon	1 GeV	99.5	144	1115	55	4.6	73
muon	2 GeV	99.1	79	586	23	1.9	40
muon	4 GeV	99.2	48	313	16	1.2	23
jets	1 GeV		195	1519	77	6.2	97
jets	2 GeV		104	804	29	2.4	52
jets	4 GeV		51	344	13	1.1	26
min-bias	1 GeV		110	842	38	3.6	58
min-bias	2 GeV		48	359	6	0.8	27
min-bias	4 GeV		21	133	1	0.2	12

HitWarrior: Duplicate removal run in TP. Identifies tracks that share more than a given number of hits. The threshold of hits can be tuned to give high efficiency and low number of duplicates.

Track fitting performance PRM and TFM

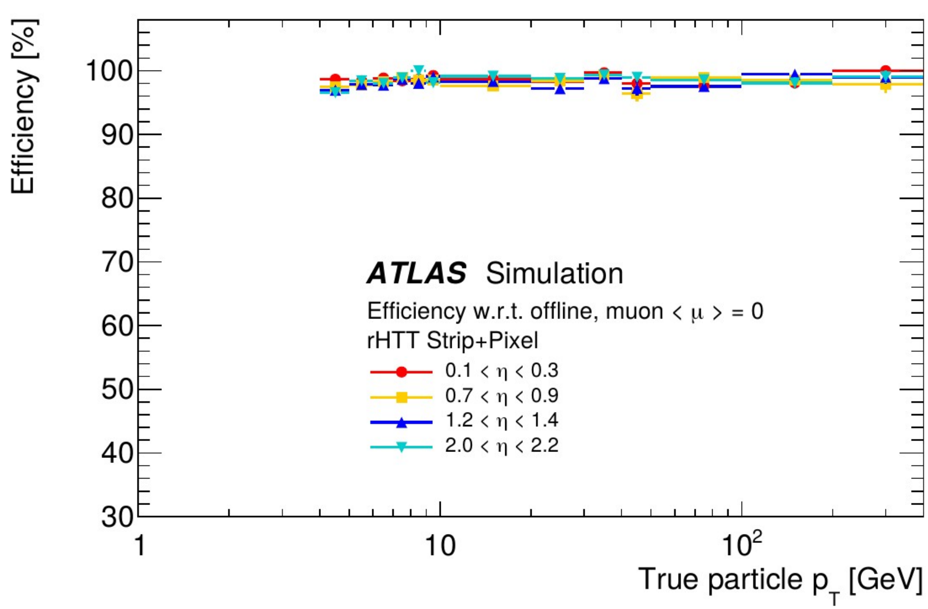


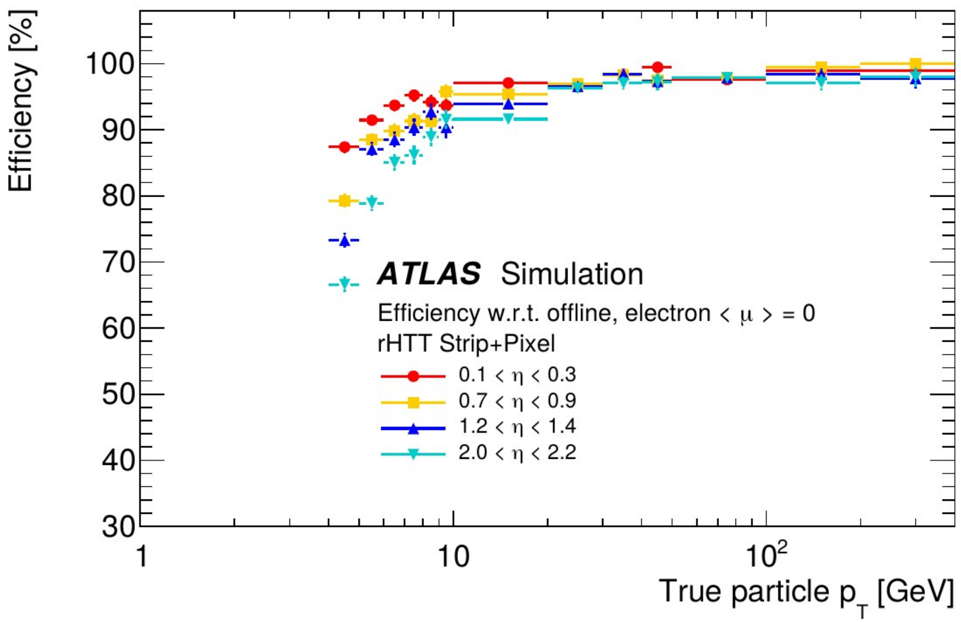


- 1^{st} stage track fitting done in the PRM: z_0 and d_0 performances limited by the short lever arm and distance to impact point.
- 2^{nd} stage track fitting in the TFM: uses all ITk layers and give near off-line quality

η range	η	ϕ	q/P_t [GeV^{-1}]	d_0 [mm]	z_0 [mm]
$0.1 < \eta < 0.3$	0.004	0.003	0.021	0.42	2.9
$0.7 < \eta < 0.9$	0.004	0.003	0.031	0.52	4.5
$1.2 < \eta < 1.4$	0.011	0.013	0.048	0.87	19.3
$2.0 < \eta < 2.2$	0.014	0.012	0.059	1.03	22.1

Tracking efficiency





- Tracking efficiency is flat over the full pT range for muons (and pions) while electron efficiency drops at low pT because of Bremsstrahlung
- The efficiency is for muons and pions similar in all studied regions while electrons are effected by the increase in material at high eta.

System size and power

UPPSALA UNIVERSITET

Size

Item	Number
Number of HTTIF PCs	24
Number of ATCA shelves for AMTP	48
Number of AMTP blades per shelf	12
Number of AMTP blades per HTTIF	24
Total number of AMTP	576
Number of PRM per AMTP	1
Total number of PRM	576
Number of AM ASIC per PRM	24
Total number of AM ASIC	13824
Number of ATCA shelves for SSTP	8
Number of SSTP blades per shelf	12
Number of SSTP blades per HTTIF	4
Total number of SSTP	96
Number of TFM per SSTP	2
Total number of TFM	192
Number of ConMon PCs per ATCA shelf	1
Total number of ConMon PCs	56

Power(preliminary)

AMTP main card (including DC/DCs)	100 W
PRM FPGA	30 W
PRM 12 AM ASIC	50 W
PRM others (RAMs, IO fanout, DC/DC)	25 W
Total/AMTP	310 W
SSTP main card (including DC/DCs)	100 W
TFM 2 x FPGA	75 W
TFM others (RAMs, IO fanout, DC/DC)	25 W
Total/SSTP	300 W

Dataflow summary

UPPSALA UNIVERSITET

1st stage fitting

	rHTT/event	gHTT/event	HTT rate	available
# Cluster /PRM (layer average)	200	260	46 MHz	60 MHz
# Roads/PRM	170	270	45 MHz	400 MHz
# Constants read/PRM	90	140	23 MHz	30 MHz
Fits/PRM	1500	2250	400 MHz	1 GHz
Tracks after χ^2/PRM	80	280	36 MHz	
Tracks after HitWarrior/AMTP	10	35	4.5 MHz	
rHTT output bandwidth /AMTP	640 Mb/s			
<Tracks after HitWarrior>/AMTP	7	20	2.6 MHz	
Total output bandwidth	250 Gb/s	750 Gb/s	1 Tb/s	
Processed event size	30kB	900kB		
Average event size	30kB	250kB		

2nd stage fitting

	Needed per event	Capability per event
# of clusters/TFM (max layer)	580	5000
# of clusters/TFM (average)	380	5000
Cluster rate/TFM (max layer)	58 MHz	500 MHz
Cluster rate/TFM (average)	38 MHz	500 MHz
# of roads/TFM	48	120
#of constant sets read/TFM		
Extrapolator	192	500
Fitter	100	230
Fits/TFM	260	600
<Tracks after HitWarrior>/SSTP	30	
Total output bandwidth	200 Gb/s	
Average event size	250kB	

Based on:
Xilinx KU085

Summary

UPPSALA UNIVERSITET

- A Hardware Tracking for the trigger has been designed to meet the physics goals of ATLAS at HL-LHC.
- The baseline HTT system runs as a co-processor in the Event Filter which will help the commissioning of the system.
- The system is divided between a regional and global HTT running at 1MHz and 100kHz L0 trigger rates.
- If required the baseline system can be reconfigured to an evolution system L1Track running on data stream with regional tracking at 4MHz but with higher threshold than rHTT.
- The system is more homogeneous than its predecessor FTK
- The system delivers high tracking efficiency and near off-line quality track parameters.
- The power budget is a challenge.

Back-up

Preliminary trigger objects in baseline L0-only system

UPPSALA UNIVERSITET

Trigger Selection	Run 1 Offline p_T Threshold [GeV]	Run 2 (2017) Offline p_T Threshold [GeV]	Planned HL-LHC Offline p_T Threshold [GeV]	L0 Rate [kHz]	After regional tracking cuts [kHz]	Event Filter Rate [kHz]
isolated single e	25	27	22	200	40	1.5
isolated single μ	25	27	20	45	45	1.5
single γ	120	145	120	5	5	0.3
forward e			35	40	8	0.2
di-γ	25	25	25,25	20	20	0.2
di-e	15	18	10,10	40	10	0.2
di-μ	15	15	10,10	10	5	0.2
$e-\mu$	17,6	8,25 / 18,15	10,10	45	10	0.2
single τ	100	170	150	3	3	0.35
di-τ	40,30	40,30	40,30	200	40	0.5†††
single b-jet	200	235	180	25	25	0.35†††
single jet	370	460	400			0.25
large-R jet	470	500	300	40	40	0.5
four-jet (w/ b-tags)		45†(1-tag)	65(2-tags)	100	20	0.1
four-jet	85	125	100			0.2
H_T	700	700	375	50	10	0.2†††
E_T^{miss}	150	200	210	60	5	0.4
VBF inclusive			2x75 w/ ($\eta > 2.5$ & $\phi < 2.5$)	33	5	0.5†††
B-physics††				50	10	0.5
Supporting Trigs				100	40	2
Total				1066	341	10.4

† In Run 2, the 4-jet b-tag trigger operates below the efficiency plateau of the Level-1 trigger.

†† This is a place-holder for selections to be defined.

††† Assumes additional analysis specific requires at the Event Filter level

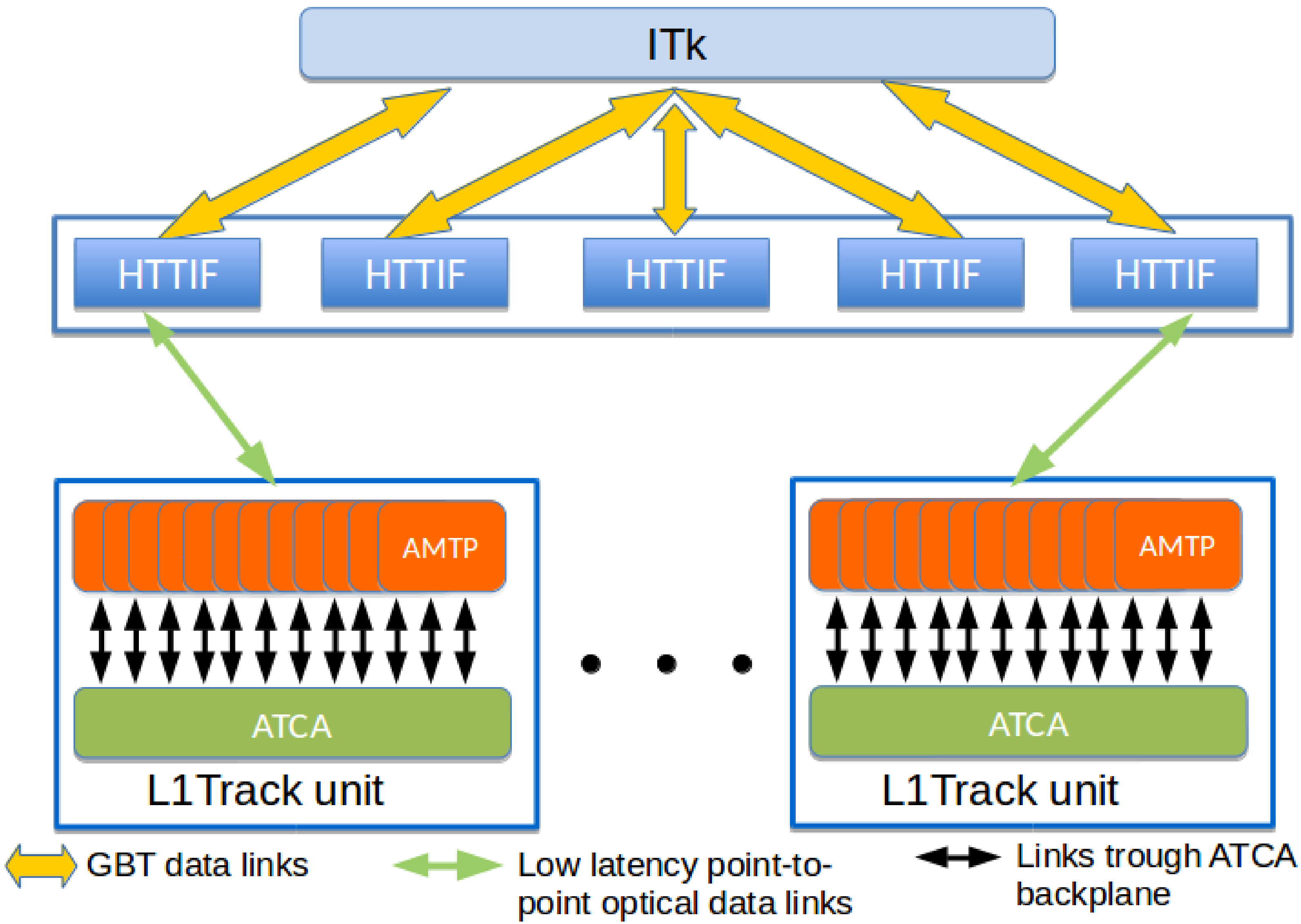
Additional triggers in evolved L0/L1 system

UPPSALA UNIVERSITET

Signature	Baseline Threshold	Evolved Threshold	Level-0 (kHz)	Level-1 (kHz)	EF before analysis specific cuts (kHz)	Gain
$E_{\mathrm{T}}^{\mathrm{miss}}$	210 GeV	160 GeV	800	80	3	2× acceptance for compressed SUSY model and 2.4× for $ZH \rightarrow \nu\nu bb$
di-τ	40, 30 GeV	30, 20 GeV	800	80	2.2	increased acceptance from 30% to 55% for VBF $H \rightarrow \tau\tau$ and 32% to 54% for $HH \rightarrow bb\tau\tau$
4 jet w/ 2-btags	65 GeV	55 GeV	800	100	0.4	improved limit in $HH \rightarrow 4b$ from 1.85 to 1.65 $\sigma/\sigma_{\mathrm{SM}}$
VBF Higgs	75 GeV + topological	60 GeV + topological	280	40	40	increased acceptance from 6.6% to 10% for inclusive VBF Higgs production
Total			2680	300	-	

L1Track architecture

UPPSALA UNIVERSITET



ATLAs trigger system in Run 2