Upgrade of the ATLAS Level-1 Muon Trigger for the sLHC

R. Richter, J. Dubbert, S. Horvat, O. Kortner, H. Kroha

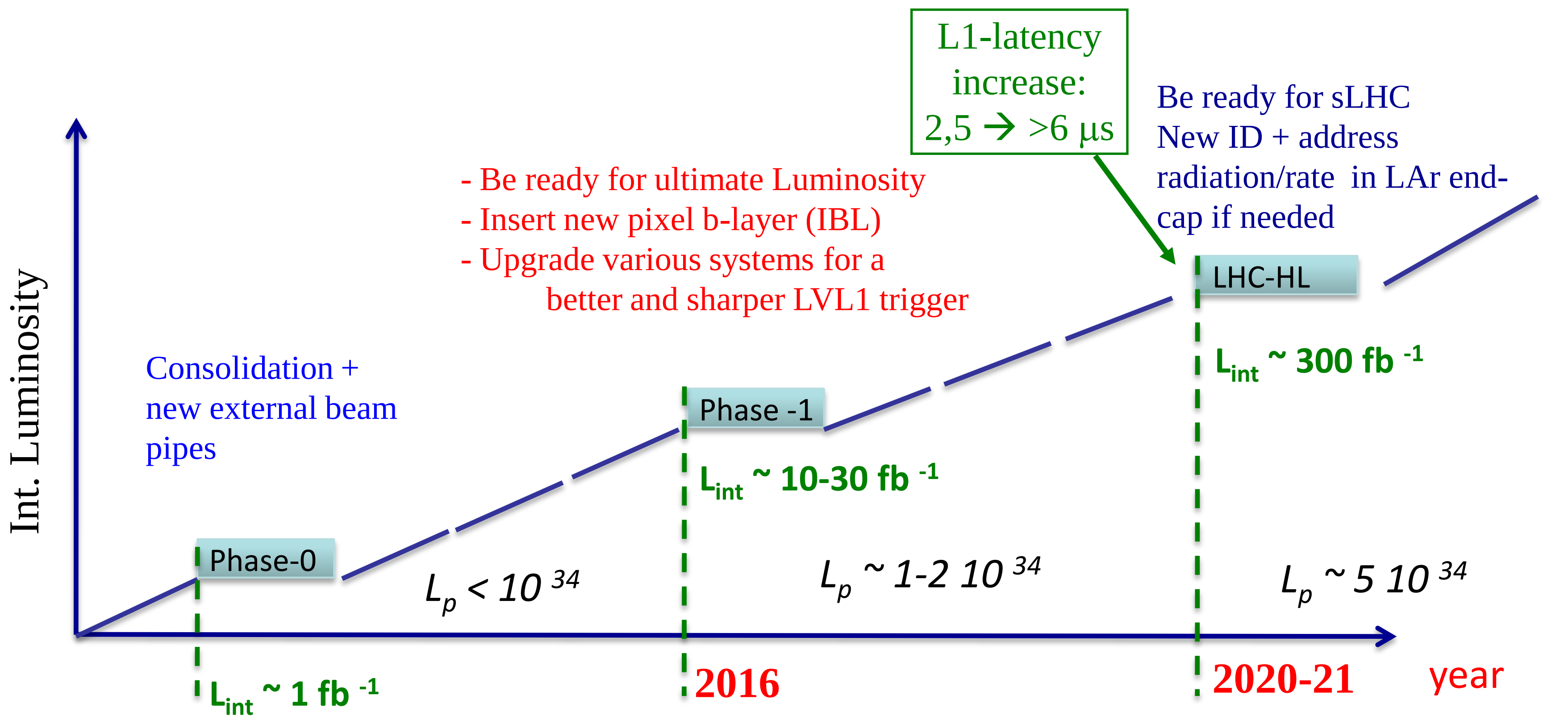
Max-Planck-Institute für Physik, München

Topical **W**orkshop on **E**lectronics in **P**article **P**hysics
Aachen, 20th-24th Sept. 2010

Outline

- The Muon trigger at the sLHC:
  → how to identify high-p_T tracks at the Level-1 ?
- A concept for improving the high-p_T selectivity
- Proposed technical realization
- Detailed estimate of the required Level-1 latency
- Robustness towards high background rates
- What has to be changed

Long term planning for the LHC
(M. Nessi, 19.08.2010)



Time for concept, prototyping, production, installation...

Need improved L1 trigger; otherwise miss new physics !

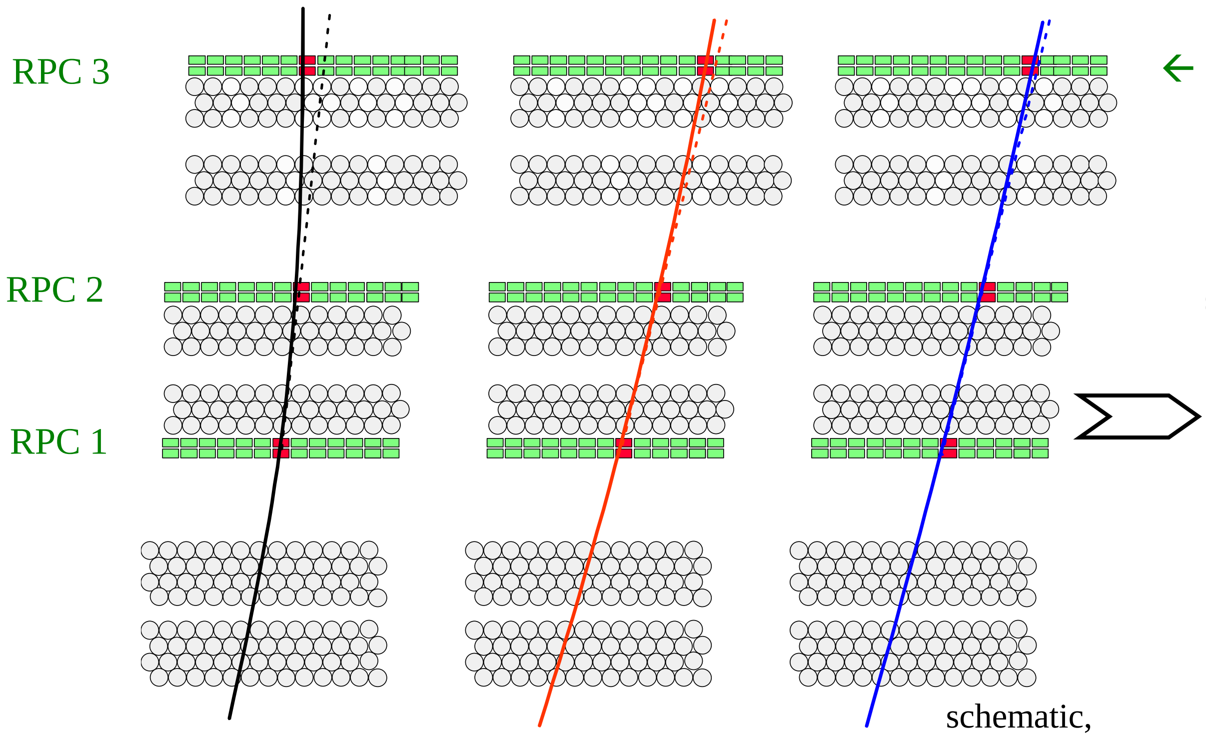
We are here!

The Level-1 Trigger for the MDT barrel: problems, solutions, history

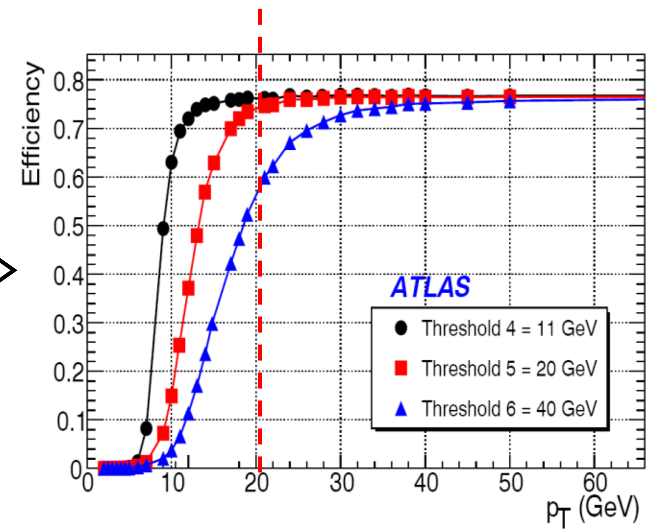
- With all trigger thresholds constant, the trigger rate would be ~ proportional to luminosity
  - However: even at sLHC the total L1 rate is limited to 100 kHz !
  - → The selectivity of L1 for „interesting“ physics has to be increased
- Raise p_T threshold for L1 muons to reach higher trigger selectivity
  - However: present L1-trigger can‘t select small deviations from straight tracks, due to limited spatial resolution of the trigger chambers
  - AND: L1 latency in the present system is limited to 2,5 μs !

- History: the present L1 muon trigger was hand-taylored to standard LHC operation (cost, time) → there are no reserves for improvement
- Questions:
  - Is there any alternative to building new chambers with better resolution ?
  - What improvement would be possible with a latency of > 6 μs ?

The problem of RPC granularity and single muon L1 rate



← RPC strip width ~30mm

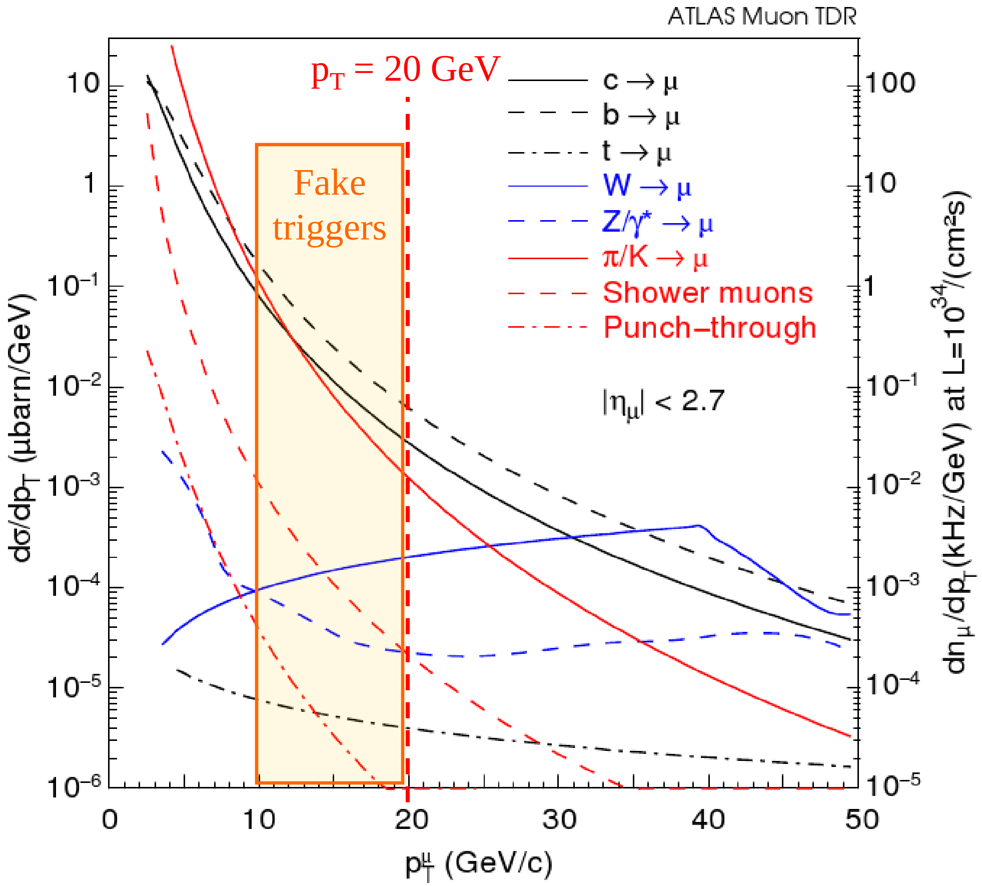


	p_T = 10 GeV	p_T = 20 GeV	p_T = 40 GeV
$\sigma_\mu > p_T$:	734 nb	47 nb	3 nb
actual trig. rate	110 kHz	24 kHz	11 kHz

High-p_T muons are a clear signature for interesting physics !

However: the present L1-trigger system has insufficient spatial resolution to tag muons above 10 GeV

Muon rates vs. p_T: the interesting physics is mainly at high p_T

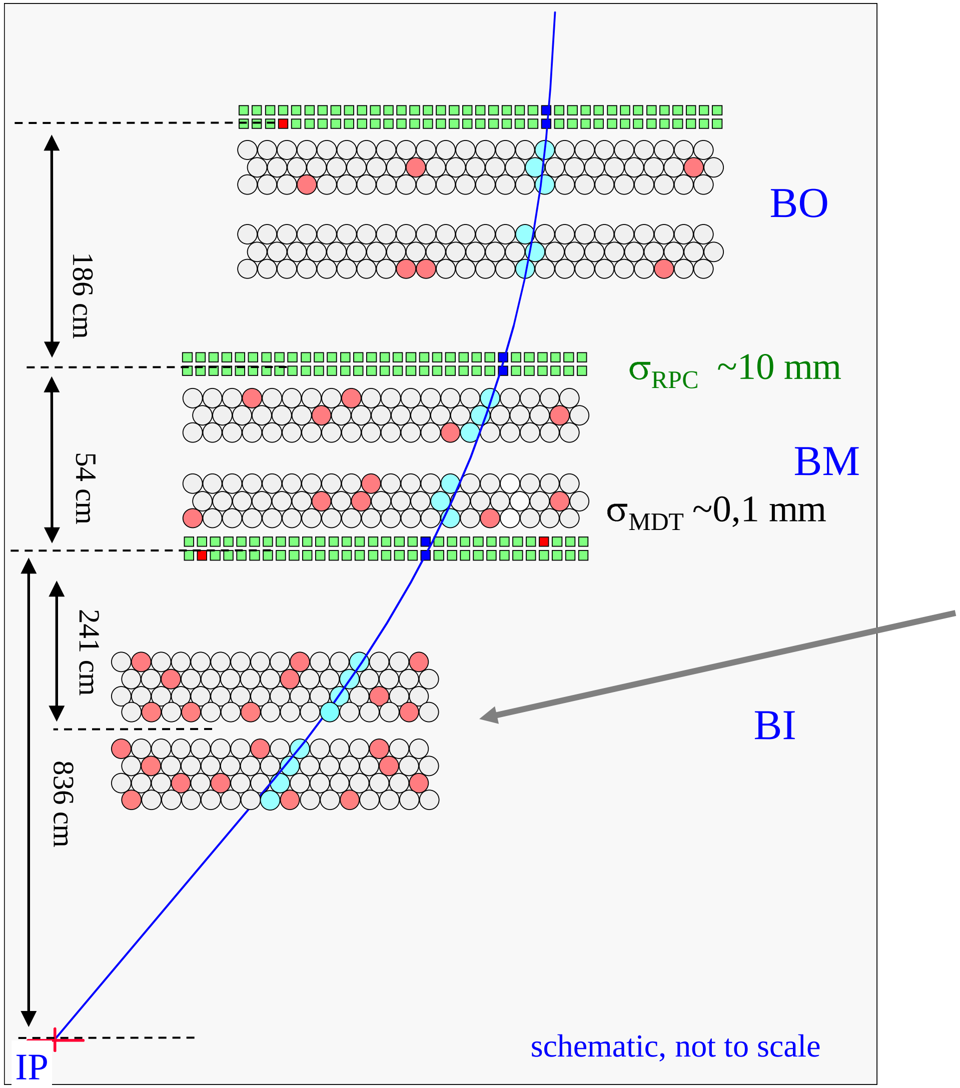


regular L1 triggers
p_T >20 GeV: ~47 nb

fake L1 triggers
p_T >10 GeV: ~400 nb

The steep slope of the pT spectrum combined with the width of the p_T resol. curve leads to high fake trigger rates.

What can the MDT do for the L1 trigger (example: barrel) ?



Strip width of RPC (2,6 – 3,5 cm) leads to a resolution of about σ = 10 mm, insufficient for **high** high-p_T thresholds > 20 GeV

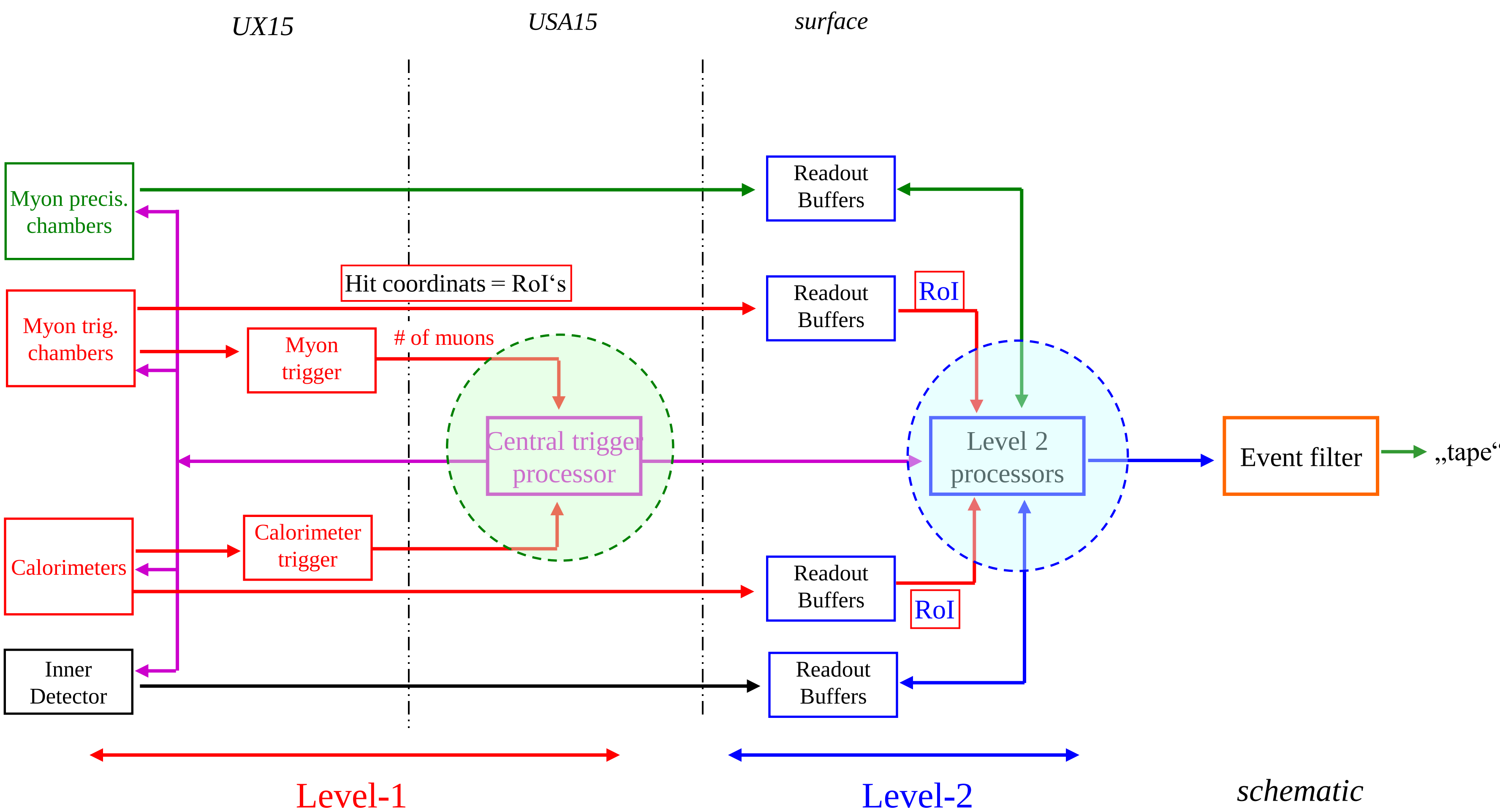
- MDT provides 100 x better resolution, but only factor ~ 10 needed !
  → can relax on drift time resolution (use only BX, ignore fine time)

In the present system the BI layer is not used for the L1 → 50% of the bending power dismissed for the trigger!

In the present system the high spatial resolution of the MDT is only used at Level 2 → reduces rate by 2 orders of magnitude.

Could we have part of this reduction already at Level 1?

Present system: MDT info only used at Level-2



In the muon system trigger and precision chambers **don't talk to each other** at Level 1

→ information only combined at Level-2

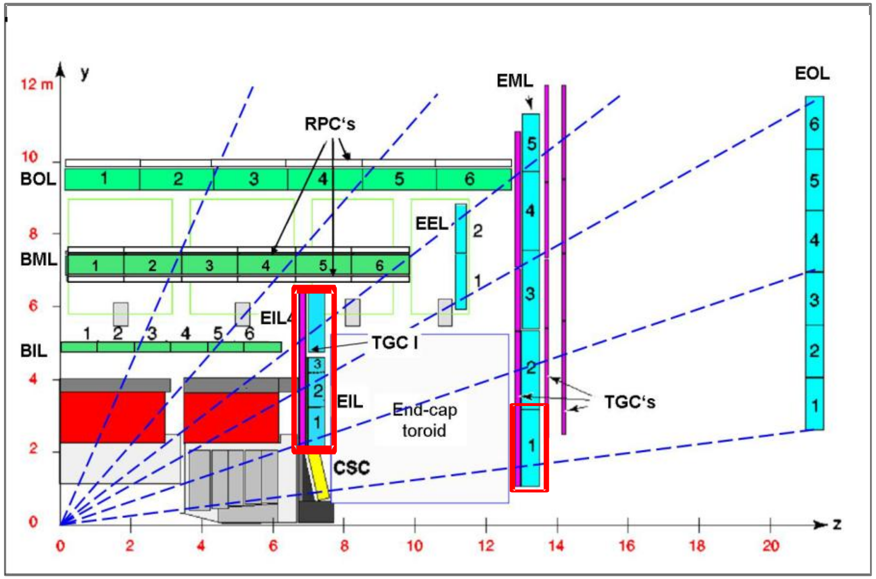
Big Q's

- Can the MDT information be merged into the L1 decision to sharpen the L1 trigger threshold?
- This certainly requires more data transfer and logics. → Can it be done inside the maximum allowed latency of 2,5 μs ?
- If NO: what latency would we need?

Properties of the L1 trigger in the Muon barrel

There are a couple of things which help you!

- The trigger produced by the RPC is organized inside trigger towers: MDTs matching RPCs. There are about 200 trigger towers in the barrel (16 x 6 x 2).
- High p_T tracks, being ‘nearly’ straight, mostly travel inside one and the same tower
- The RPCs predict the location of the straight track with 1-tube-width precision! → defines search road for MDT hits

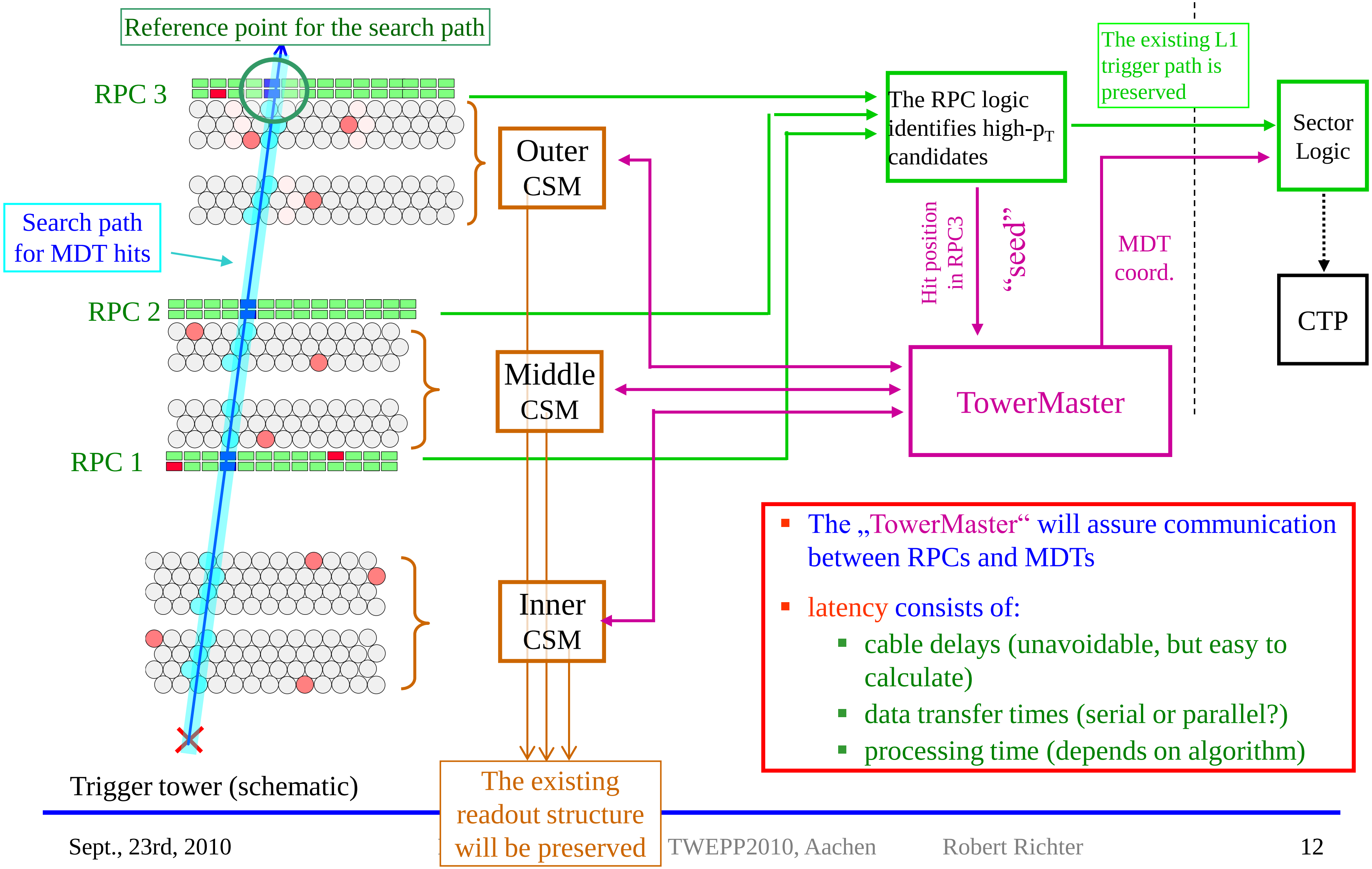


- The high-p_T RPC trigger is very selective and immune w.r.t. accidentals, even at sLHC
- The high-p_T trigger rate in any given tower is very low ~ 100 Hz, even at sLHC
- So: use the RPC trigger as a “seed”, don’t try “a stand alone” trigger with the MDT (my philosophy)

Strategies to keep L1-latency small

- Use RPC L1 trigger as "seed". The MDTs only verify p_T on request from the RPC! (No stand-alone trigger of the MDT.)
- Use the RPC hits to define a search road for the corresponding MDT hits
- Reduce time resolution from 0,78 ns to beam crossing frequency → simplify readout, save bits (i.e. data volume), but retain ~1 mm resolution!!
- Data have to be moved from the MDT frontend to a processing unit → keep cables short and at high bandwidth
- Transport of data costs time depending on cable/fiber BW → reduce word size, overheads and redundancy to the minimum
- Data recording (d.t. measurement) and data processing cost time → use "parallel processing" where possible
- Operate in pipeline mode: Request comes exactly 48 BX after particle passage → this way the absolute drift time becomes known and can be used to reject hits, corrupted by γ-conversions (see below)

Technical realisation: Implement communication between trigger- and precision chambers inside a trigger tower

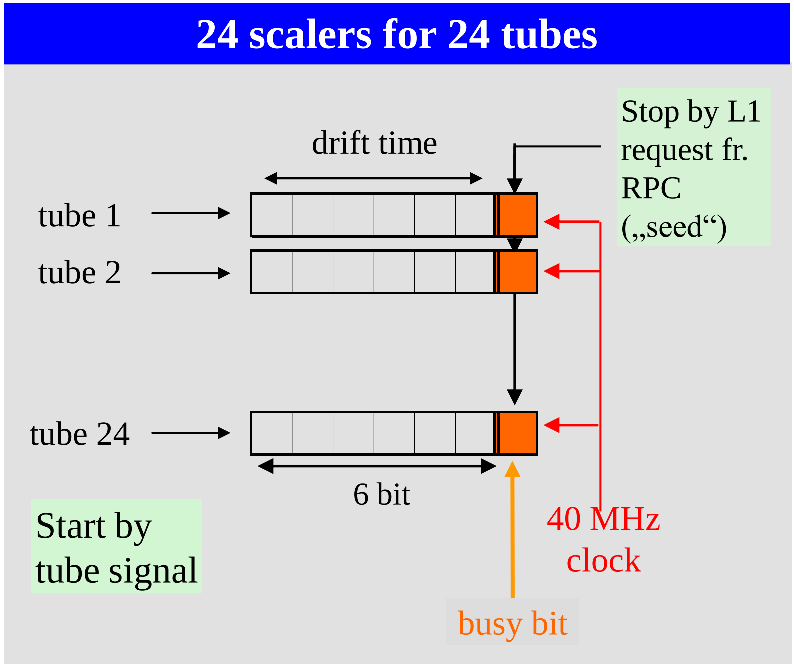


- The „TowerMaster“ will assure communication between RPCs and MDTs
- latency consists of:
  - cable delays (unavoidable, but easy to calculate)
  - data transfer times (serial or parallel?)
  - processing time (depends on algorithm)

Trigger tower (schematic)

Technical realisation:
measurement of drift times in the MDT tubes

Parallel drift time measurement on all tubes to save latency:



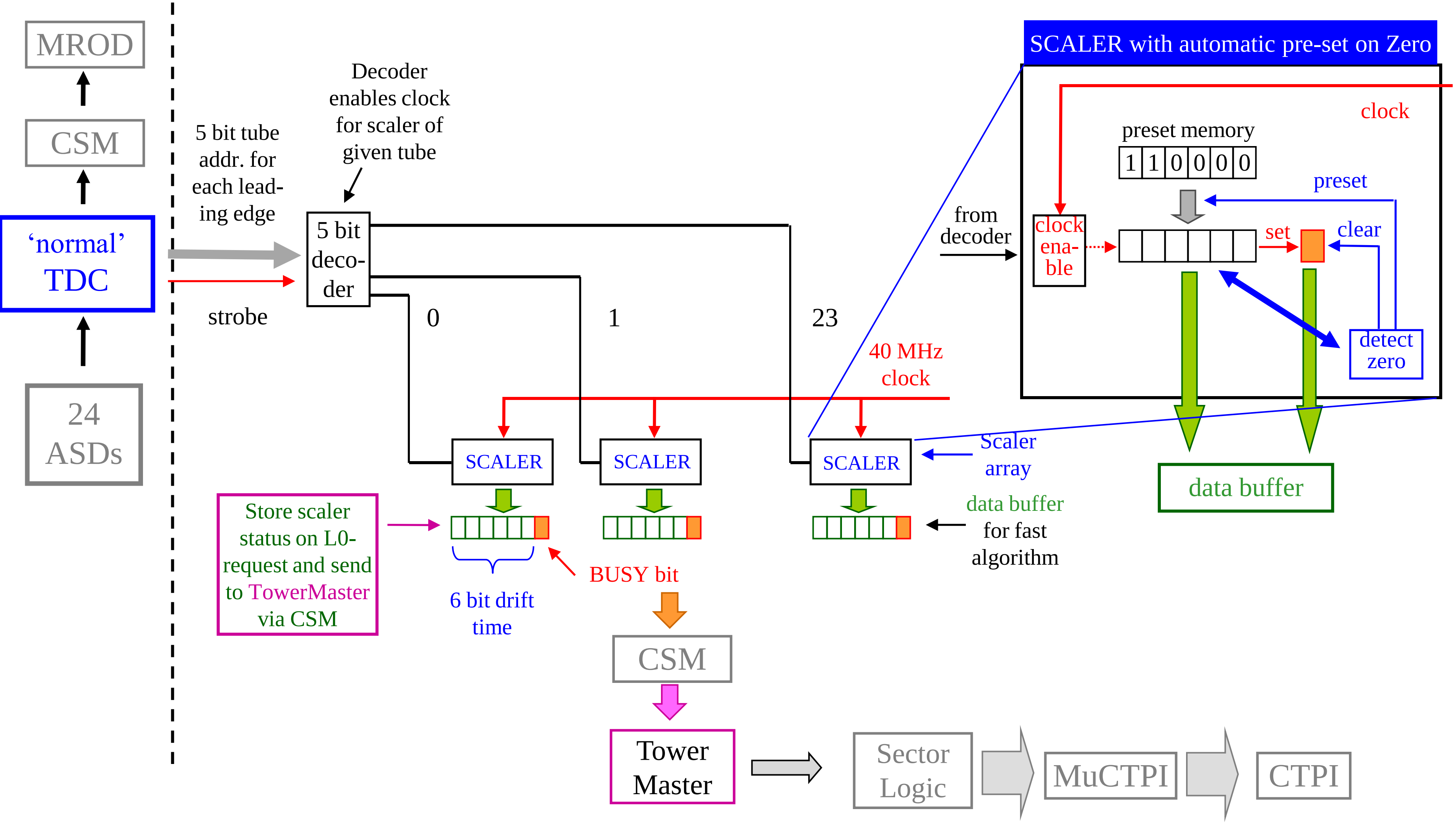
Fast readout of the drift times for the L1 trigger

The drift times of all 24 tubes of a mezz card are recorded by a bank of 24 scalers

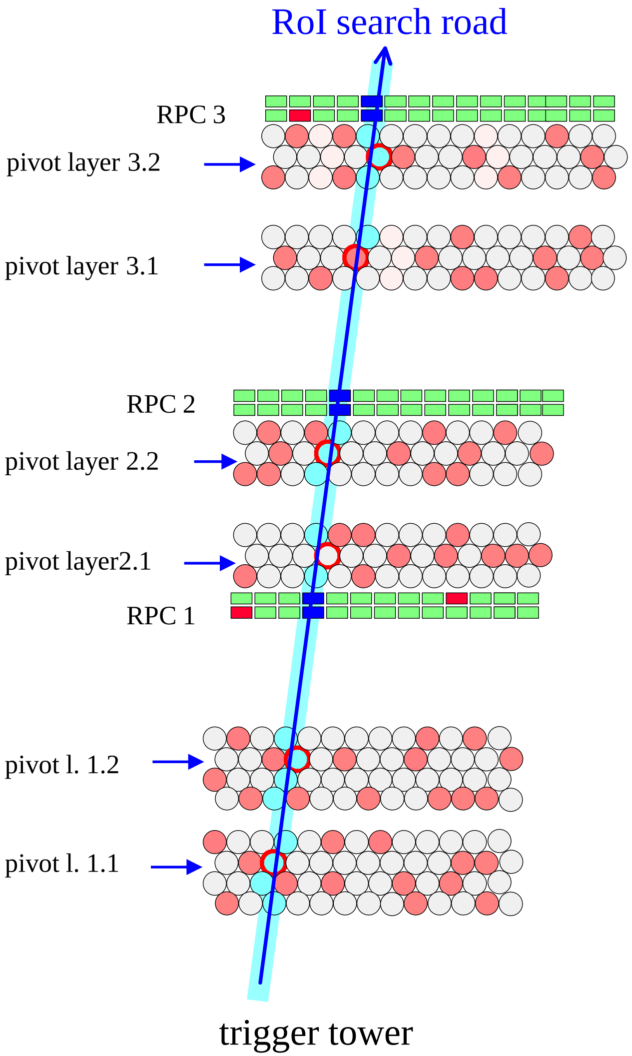
Scalers are

- started by the arrival of the ionisation at the wire
- stopped by a L1 request from the trigger chambers
- The trigger request comes a fixed number of BX after the particle passage; thus the absolute drift time and the distance from the wire are known
- The scalers only need a depth of 6 bit, corresponding to a maximum running time of 48 BX = 1,2 µs
- This provides a pos. resol. of σ = 0,5 mm

„Parallel processing“: a separate scaler per tube for fast drift time recording

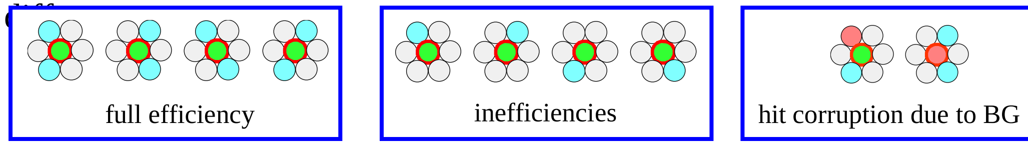


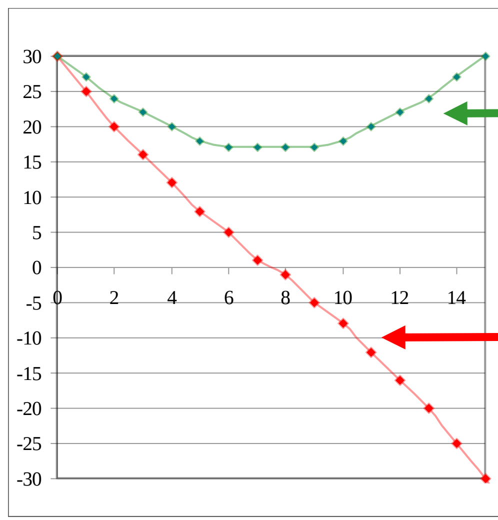
Determination of the drift time



Search strategy for drift time pairs

The RoI defines the (most likely) tube address in the pivot layer → „pivot tube". If this tube is empty, try the neighbour.
Once the pivot tube identified, several hit configurations are possible around the pivot tube.
→ define search rules for partner tubes via LUT to form drift time



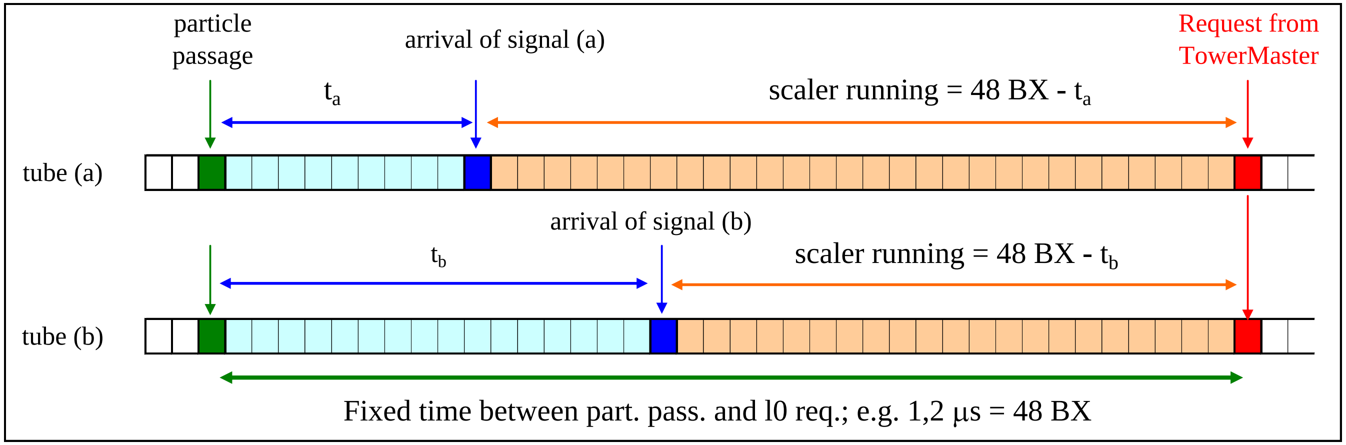


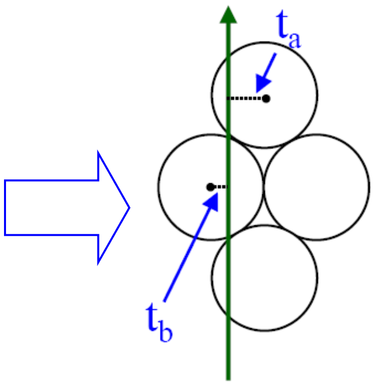
Sum and difference of the drift times of adjacent tubes vs. track position [units of BX]

The drift time SUM allows to identify corrupted hits. → Reduction: ~ factor 10

The drift time DIFFERENCE yields the position of the hit

Examples of typical timing signatures
(simple case: track at normal incidence)





The total drift time $t_a + t_b$ has to be inside predefined limits, otherwise the measurement is likely to be corrupted by a γ-conversion (→ a valuable quality criterion of the d.t. measurement)

If no request from RPC/tower master (no „seed“) the scaler runs to 48 BX and resets itself to zero, waiting to be started by the next hit from its tube.

Breakdown of latency

Transfer times on cables, due to

- cable length
- data volume
- processing time, decisions making

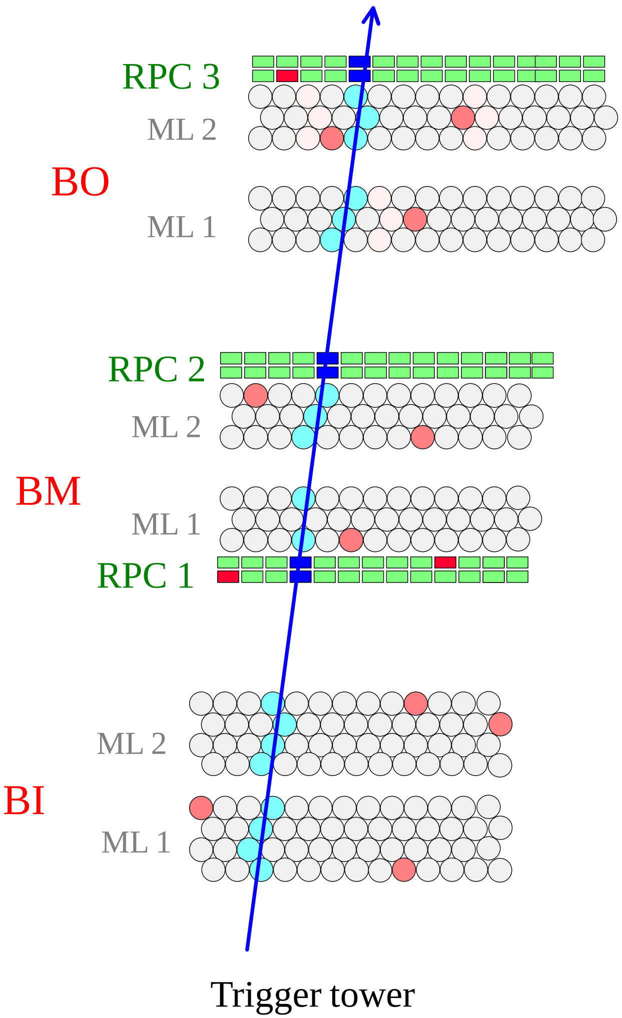
Two options:

- local processing at the frontend (e.g. coordinate finding in CSM, sagitta determ. in tower master)
- shift all raw data to rear-end and process there

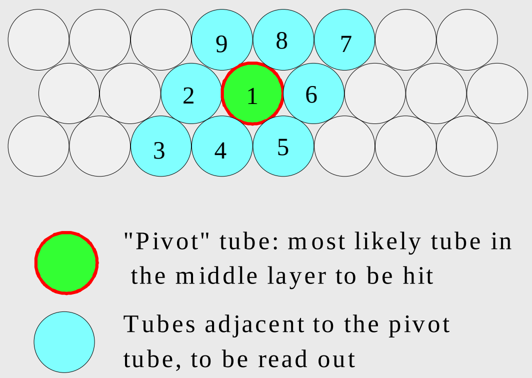
Trigger decision at the rearend:

estimation of data volume and transfer times

Transfer rate on fiber: 0,8 ns/bit
Transfer time to Sector Logic: **0,4 µs**
(+0,5 µs fiber delay)



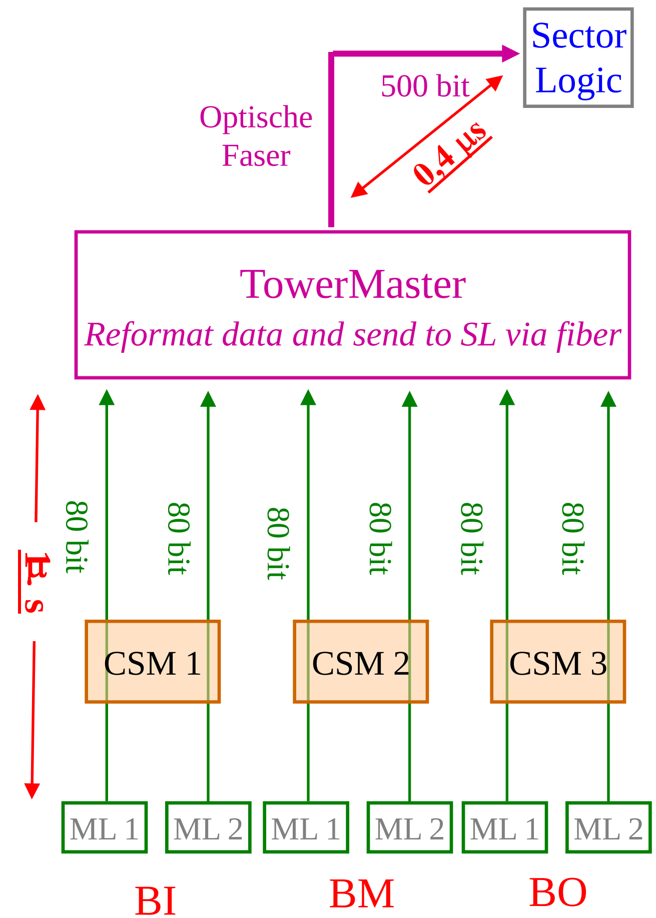
The pivot + 8 surrounding tubes are read out and transfered to the SL (no tube addr. needed)



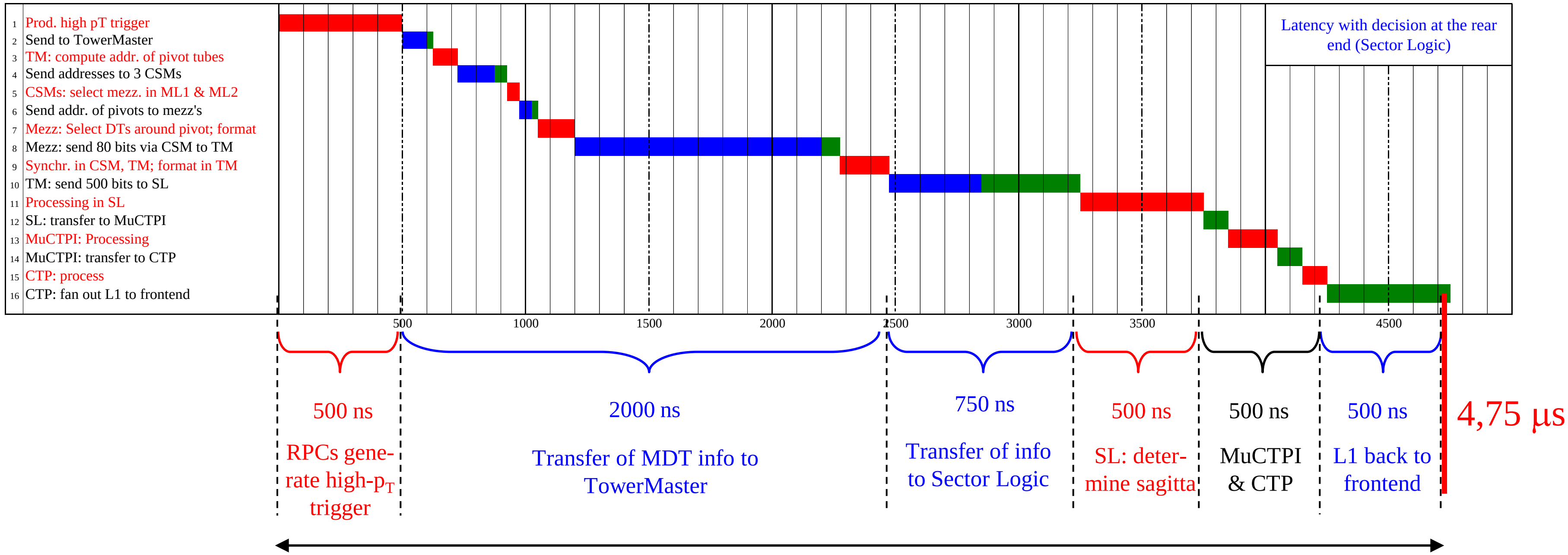
Data volume to be read out:

9 x Scaler content (7 bits) = 63
Parity bits etc. ~ 15
~ 80 bits

Trf. rate on cable: 12,5 ns/bit
→ Trf. time to TowerMaster: **1 µs**



Latency estimate



→ Total latency from particle passage to L1 at the front-end: 4,75 µs

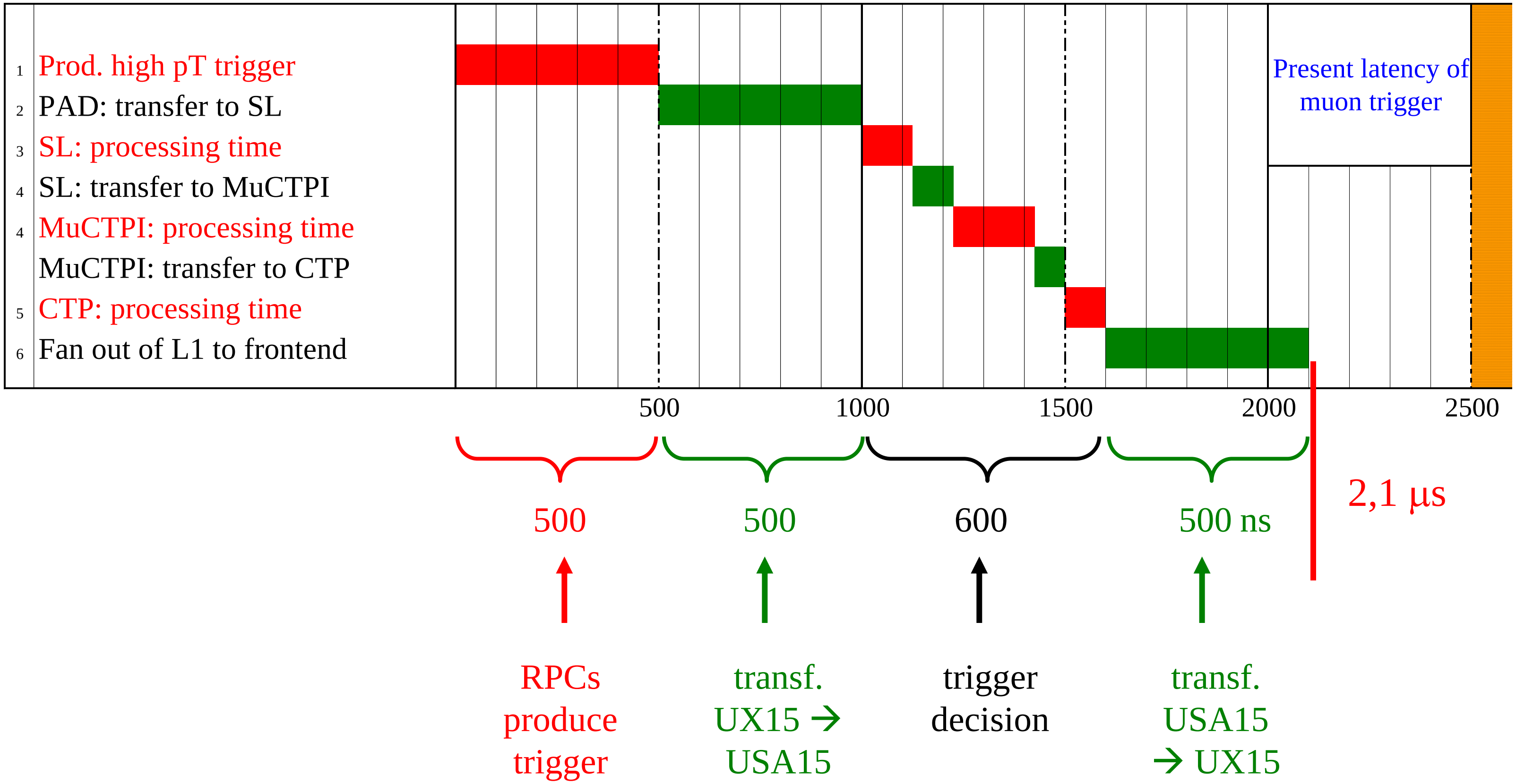
- → Estimates are generous, but more work needed on algorithms, data formats and processing times.
- → Do MuCTPI and CPT need extra latency out of the 6 µs budget?
- → We need an agreed-on latency budget for the muon trigger!

processing

delay due to data volume

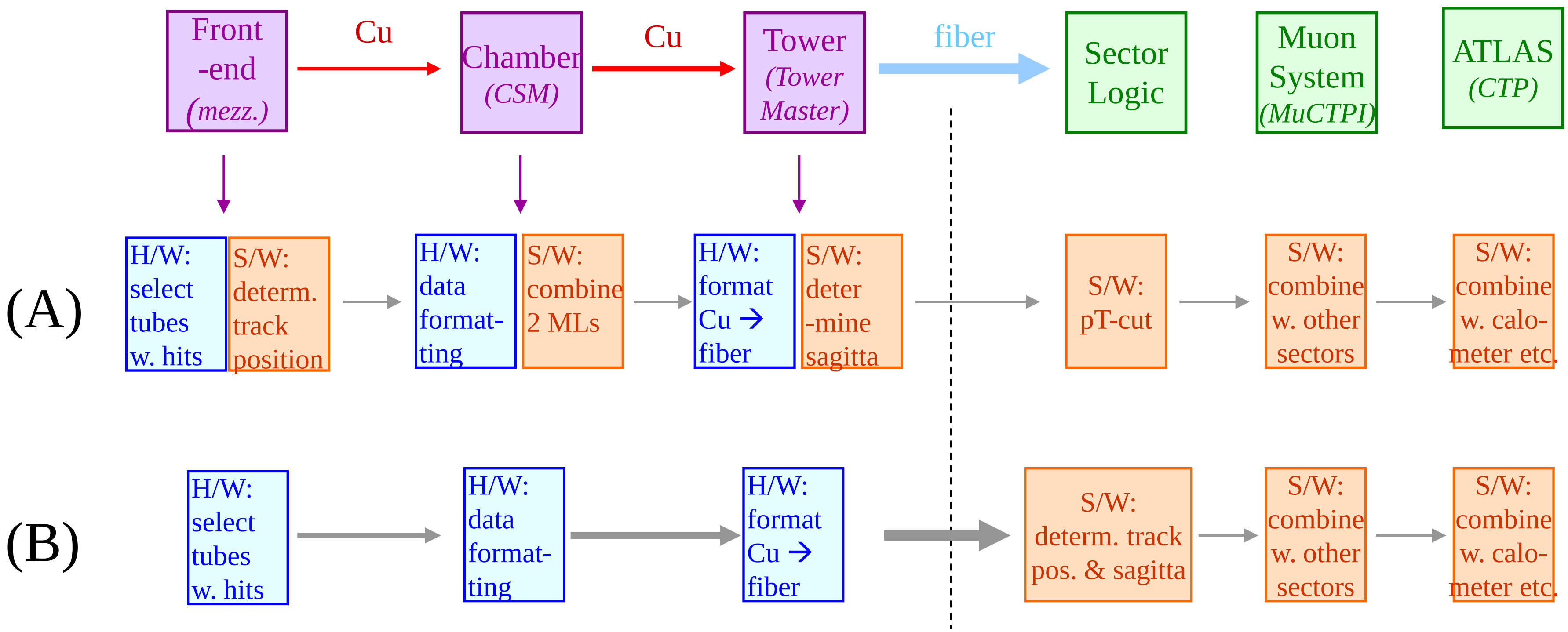
delay due to cable length

Present latency in the Muon system



Historical reason for very tight latency:
cost limitation, in particular for subdetectors with analog storage

Two philosophies: where to do the L1 decision ?



A: Decisions at the Frontend (on-chamber):

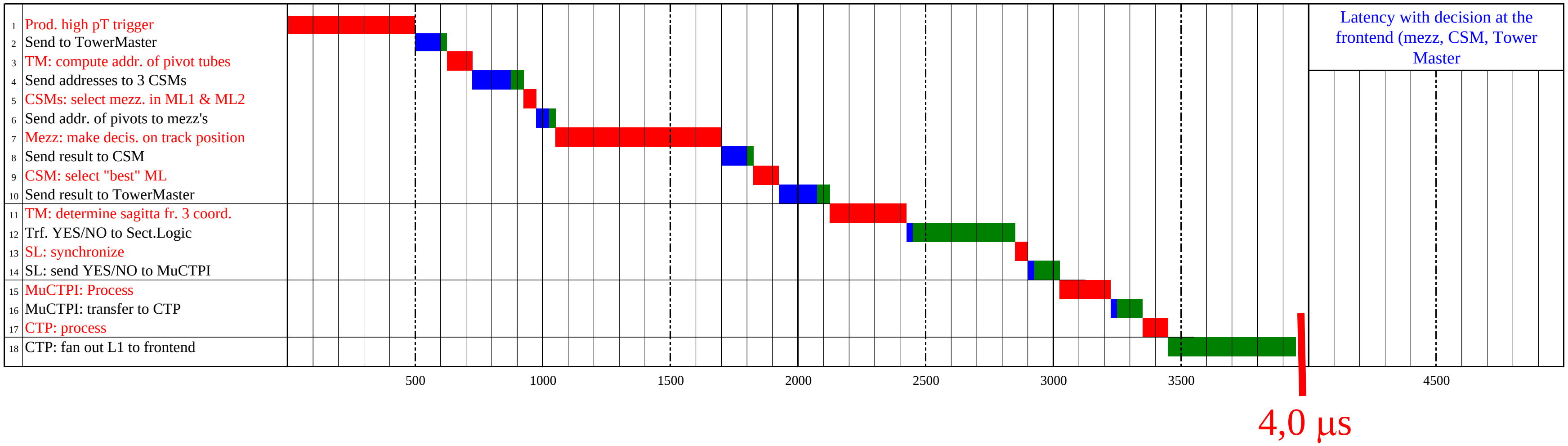
PRO: small data volume to be transferred
CON: programmable device in hot area
risk of SEUs
problem to maintain code in many devices
need 2nd R/O path to keep original data

B: Decisions at the Rearend (USA15):

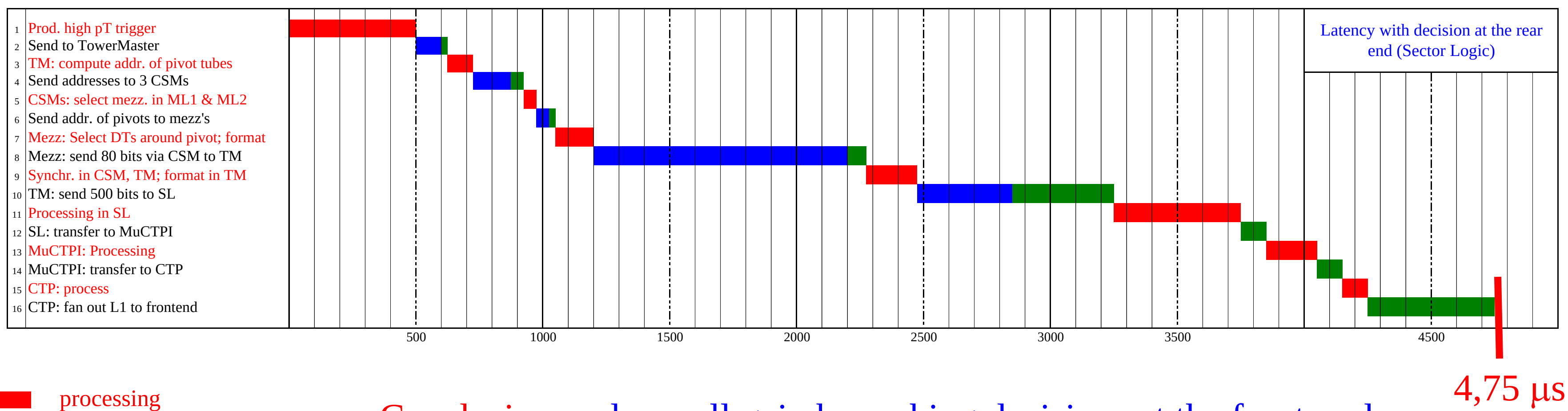
PRO: work on original data (can store a safty copy)
easy access for s/w updates
programmable devices only in shielded area
no need to maintain s/w in the frondend
CON: more data transfer → more latency

Latency comparision: decision at front/rear end ?

FE:



RE:



processing

delay due to data volume

delay due to cable length

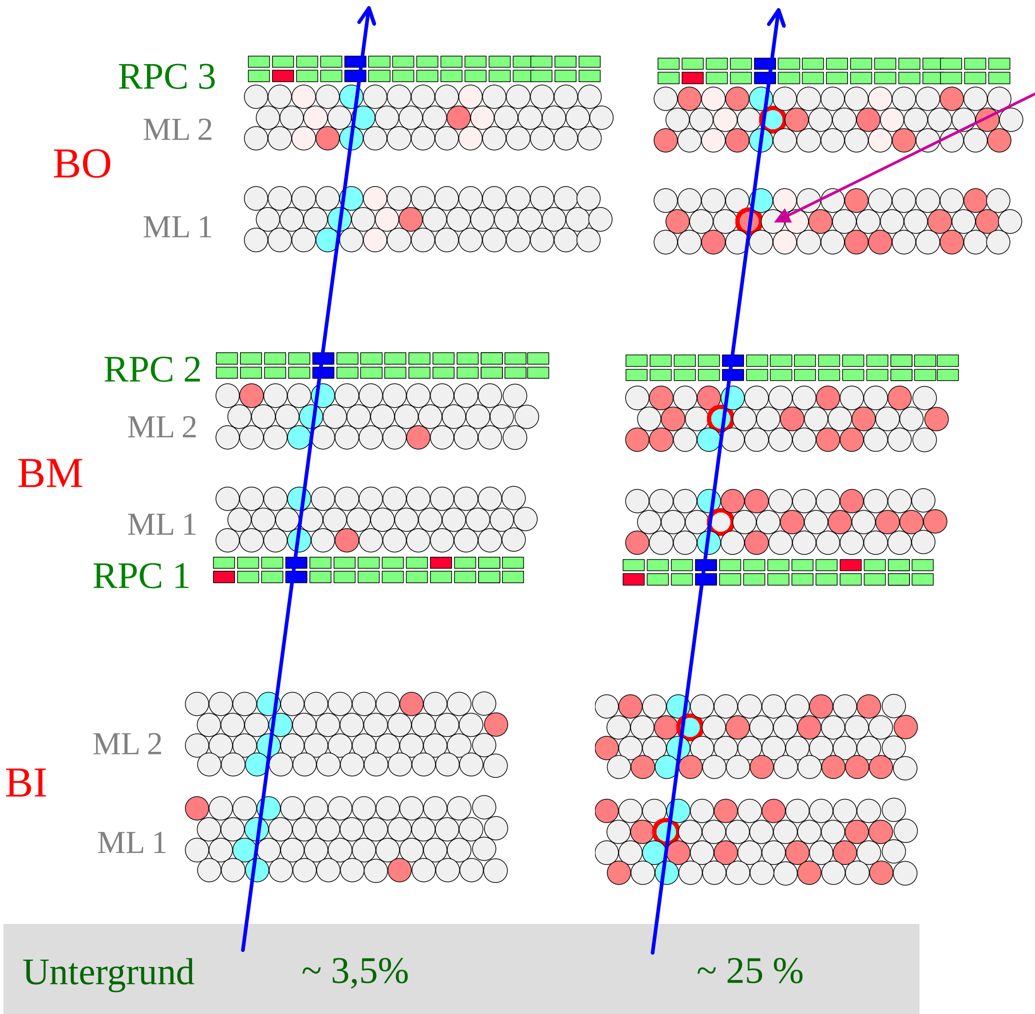
Conclusion: only small gain by making decisions at the front-end

Parameters for latency estimates

Latency under various assumptions about processing time and bit transfer rates					
	Assumptions about processing time	conservative (100%)		optimistic (50%)	
	BW on cable (Mbps)	80	160	80	160
Trigger decision	on chamber	4,0 μs	3,6 μs	3,3 μs	**3,0 μs**
	sector logic	**4,7 μs**	4,1 μs	4,3 μs	3,4 μs

Conclusion: no way to reach 2,5 μs → need more latency → wait for 2020 !

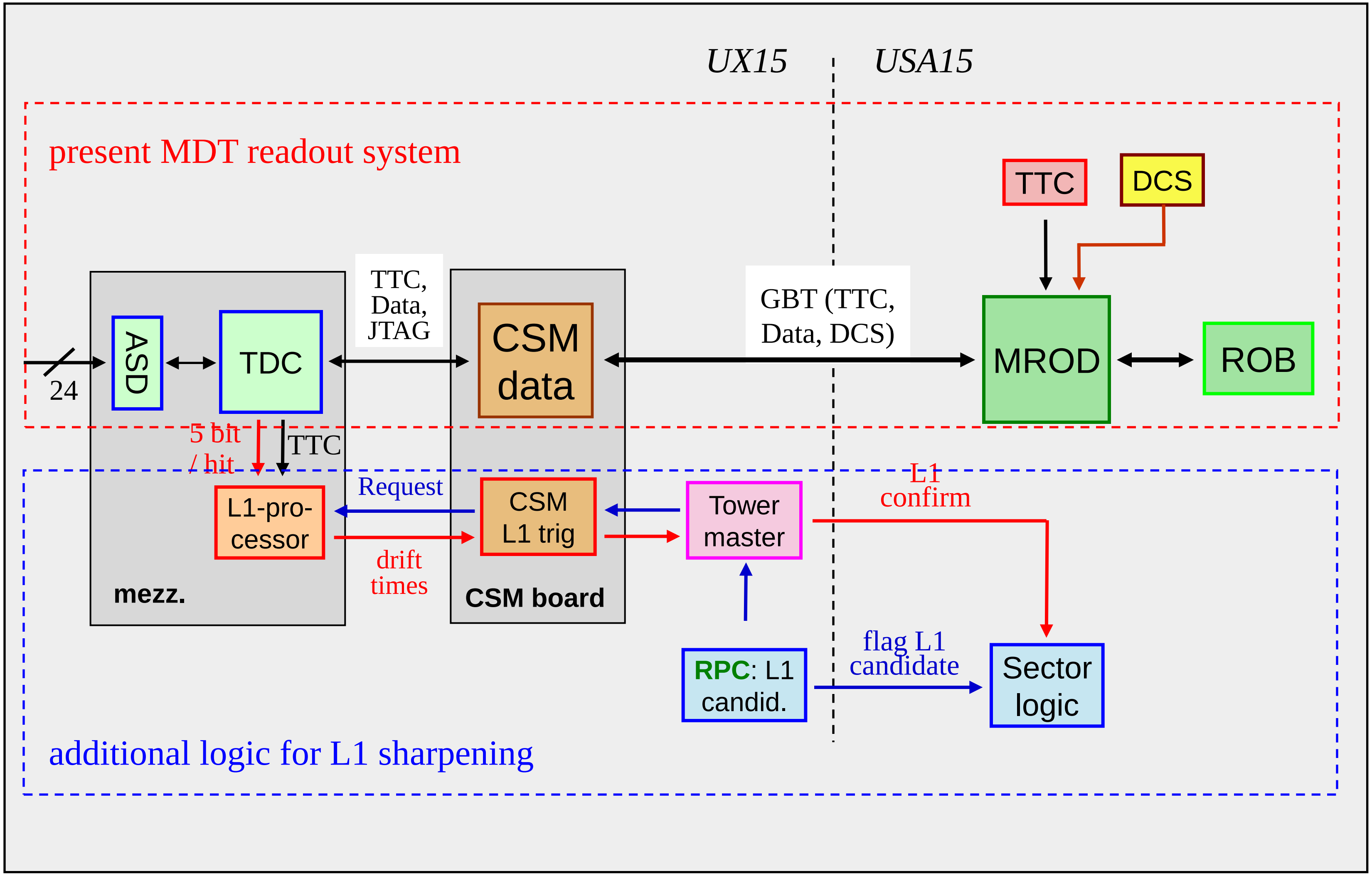
How does it work at high BG rates??



Example of a background hit masking the drift time signal of a MDT tube → coordinate unusable; however, there is some redundancy: ML2 may deliver the correct coordinate

Typical BG rates in the barrel are < 300 kHz per MDT-tube ($4 * 10^{34}$)
→ corresponds to a probability of ~10% of a BG-Hit masking a track hit (BG-hit arriving before the track hit)
→ probability of BOTH ML's being masked: ~ 1 - 2 %
→ redundandancy of 2 ML's improves immunity vs. BG-hits
→ for BG-rates > 30% the method becomes inefficient
→ wait for the real BG-rates in the experiment (effect of beryllium beam pipe ?)

A possible architecture of the new MDT elx



How about the end-cap region (TGC trigger) ?

- a similar approach is possible using the TGC info as a "seed" for the MDT
- due to different magn. field configuration, however, the pos. resolution provided by the MDT must be about a factor 3-4 higher
- this can be done, but requires a more sophisticated processing of the MDT info

Summary

- MDT precision can be used for L1 sharpening
- Need only extra latency of ~ 2 μs
- Benefits:
  - No additional trigger chambers required in the Barrel !
  - No interference with „normal“ readout
- Hardware consequences: concept needs
  - renewal of the MDT elx
  - modification of parts of the RPC elx (PADs, Sector Logic).
- Requires development of new chips and boards
  - new frontend board (mezzanine)
  - new CSM
  - architecture of „TowerMaster“
  - interface to RPC readout
- It is a big job and requires a long-term effort of the muon community (trigger and precision) and considerable resources (~ 10 MCHF ++)

Spares

Trigger thresholds and rates

	LHC, low lumi		LHC, high lumi		SLHC	
years	2007 - 2009		2009 - 2015		2016 - 2025 ?	
Luminosity [$cm^{-2}s^{-1}$]	10^{33}		10^{34}		10^{35}	
Trigger	thresh.	rate [kHz]	thresh.	rate [kHz]	thresh.	rate [kHz]
Single muon	p_T > 6 GeV	23	p_T > 20 GeV	4	p_T > 30 GeV	25
Pair of muons			p_T > 6 GeV	1	p_T > 20 GeV	few
Single, isolated EM cluster	E_T > 20 GeV	11	E_T > 30 GeV	22	E_T > 55 GeV	20 *)
Pair of isolated EM clusters	E_T > 15 GeV	2	E_T > 20 GeV	5	E_T > 30 GeV	5
Single jet	E_T > 180 GeV	0.2	E_T > 290 GeV	0.2	E_T > 350 GeV	1
jet + missing E_T	50 + 50	0.4	100 + 100	0.5	150 + 80	1 - 2
		36		32		52
	(TDR, 1998)		(TDR, 1998)		(A. Lankford, 2005)	

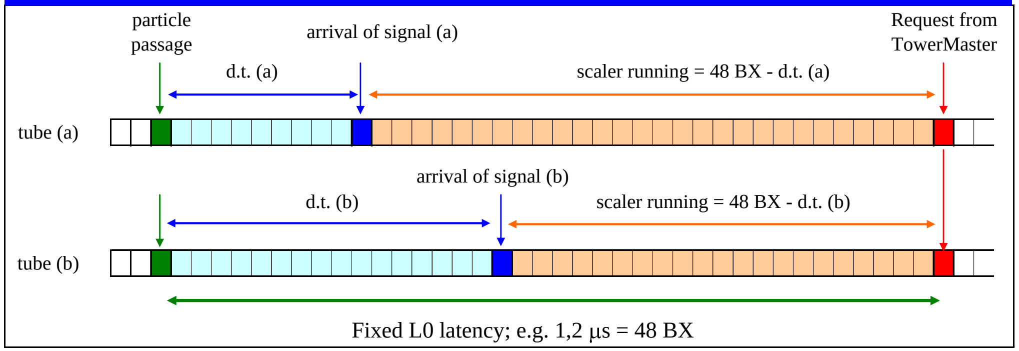
*) added degradation from pile-up not included

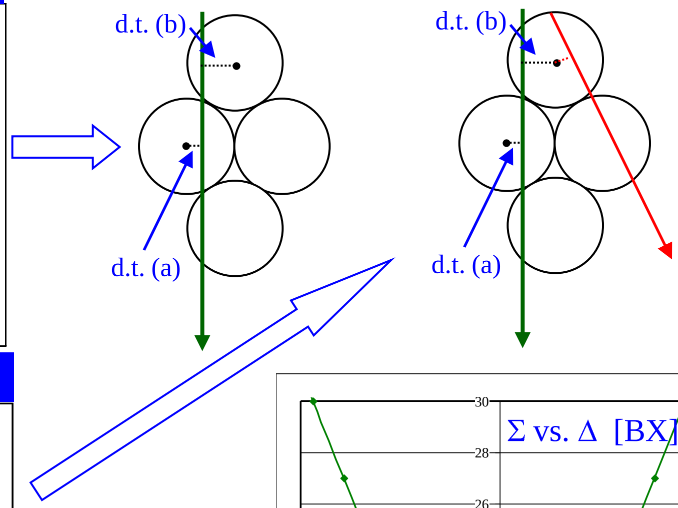
Presently not possible

The μ-trigger rate is shared between ~ 400 towers
→ small absolute rates in any given tower: < 100 Hz
→ probability of > 1 track per tower negligible

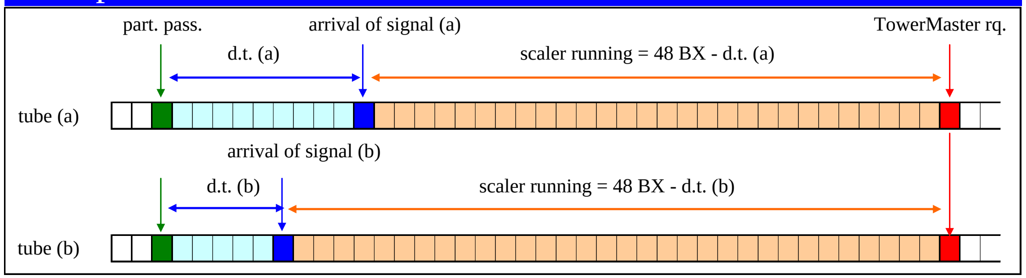
Examples of timing signature

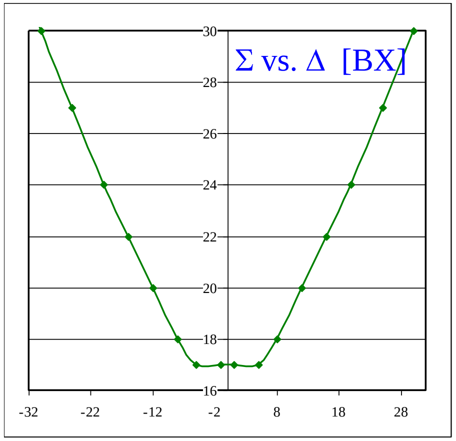
Good track measurement: drift time sum OK



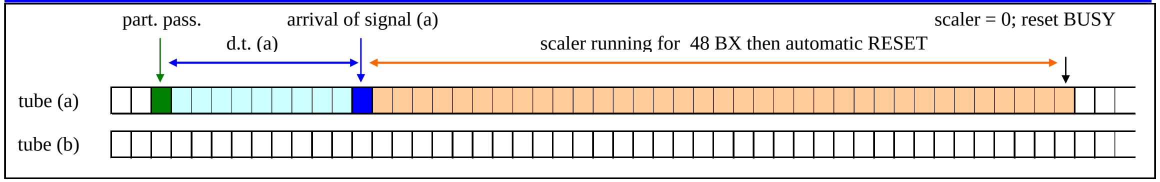


Corrupted track measurement: drift time sum too small





Most frequent signature, may mask a good track: n/γ conversion



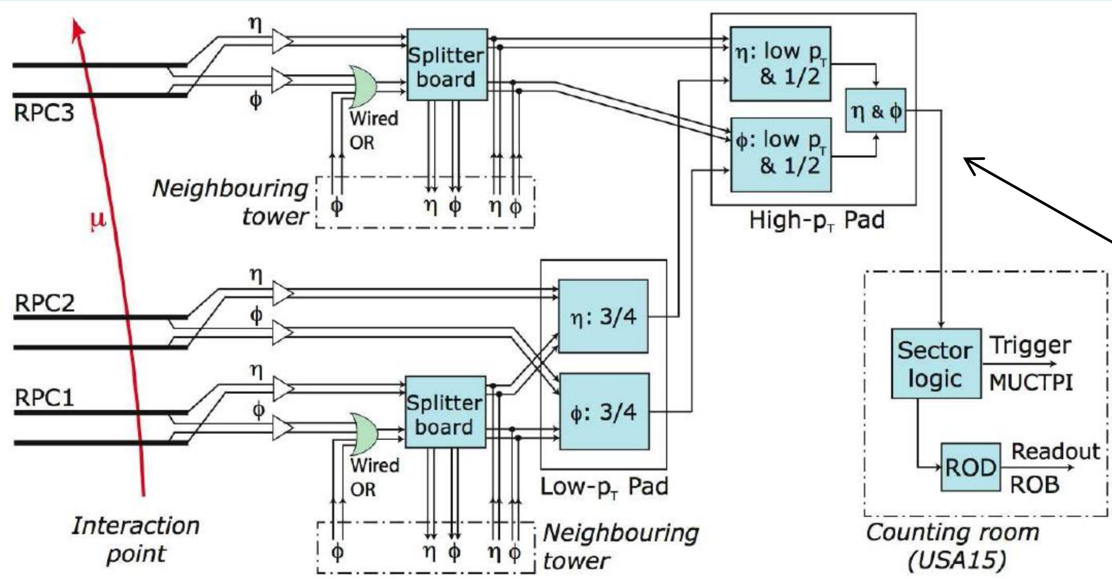
Due to the fixed delay of TriggerMaster request w.r.t. particle passage (e.g. 48 BX), the d.t. sum can be used as a quality criterion (this is not possible in the „standard“ R/O scheme, as the „start time“ is not known).

$\Delta = t_a - t_b = sc_b - sc_a = 48\,BX - t_b - (48\,BX - t_a)$

$\Sigma = t_a + t_b = sc_b + sc_a - 96\,BX$

→ Check on Δ – Σ correlation, using a LUT

Present/future trigger latency
(example of barrel)



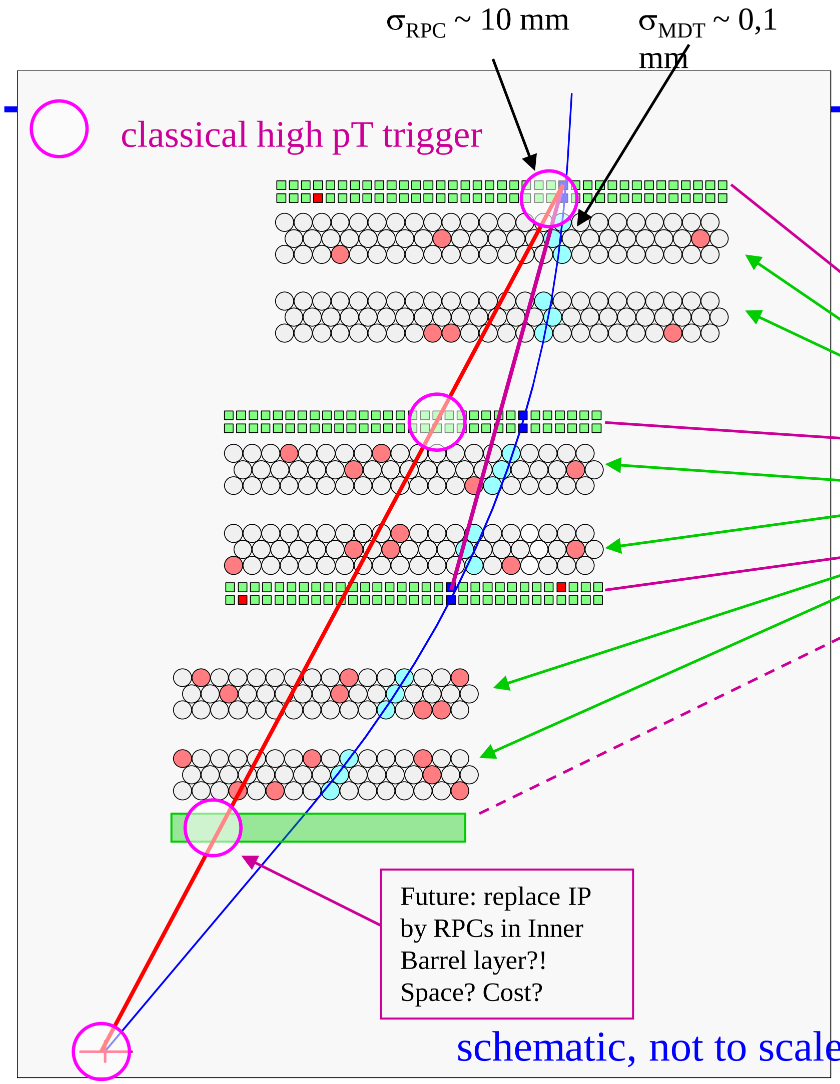
The RPC trigger & readout scheme

The RPC high-p_T delivers a clean, noise-immune signature and a precise prediction of the search road.
→ MDT only needs η-strip number in RPC3 and BX time to deliver "precise" sagitta

Present:
- Latency limit is 2,5, muon trigger using 2,1 μs
- RPC high p_T Pad ~ 0,5 μs
- Sect.Logic+ MuCTPI+ CTP use ~ 0,5 μs
- Fiber delay UX15 ←→ USA15 ~ 1 μs

We count latency from particle passage to arrival of L1 trigger at the frontend

Future:
- an ATLAS wide latency increase to 6,4 μs would give ~ 4 μs extra latency for L1 refinement
- → need fast, simple algorithms: addition, substraction, LUTs (no multiply etc.)
- transmit minimum info → small word size → fast transfers (serial?, parallel?)
- work in strict synchronicity with BX (pipeline)



Tower-wise modularity would be simplest:

→ high mom. tracks are nearly straight: sagitta @ 20 GeV ~ 25 mm, @ 50 GeV ~ 10 mm

→ boundary logic too slow (accept loss of these tracks)

Tower Master

→ determines L1 trigger road

→ combines results from MDTs to form L0 (tower)

MCTP

→ combines results from all towers

→ confirms/rejects L0A (muon)

CTP

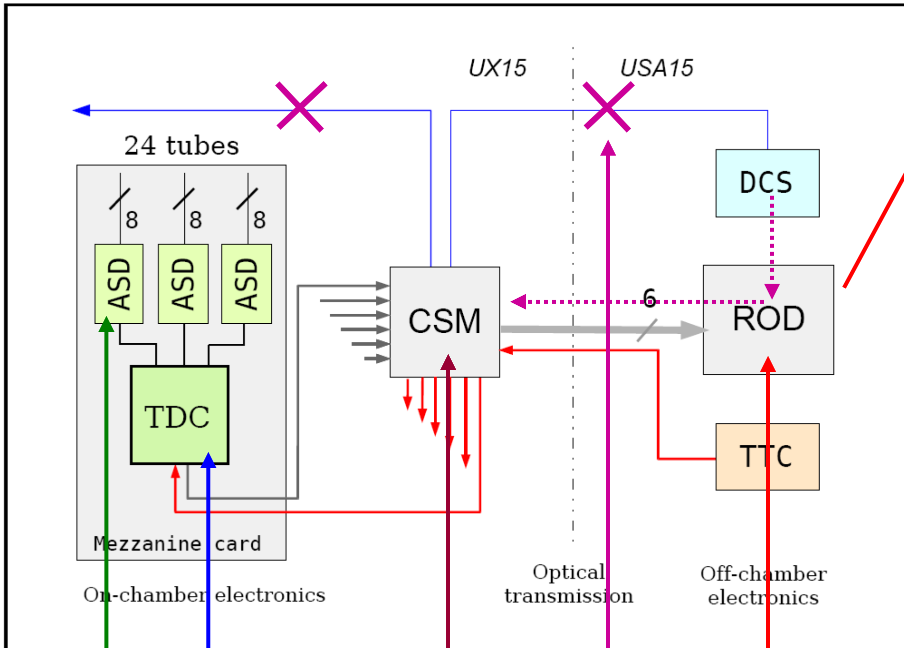
→ combines results from MCTP with calo info etc.

→ generates full L1

RoI info RPC to TowerMaster

Coordinate info TowerMaster to MDT & d.t. info back to TowerMaster

Summary: Modifications needed for MDT electronics in SmWh and BgWh due to high BG



cannot presently handle 6 CSMs at full. rate

obsol. tech-nology → new device

limit: 300 kHz/tb → new device

GOL chip no more available → use GBT

Get rid of ELMB, CANbus & cables → use GBT !

Adapt ROD to GBT, increase proc. power

Upgrade requirements for 4 x nom.

- new ASDs (old technology obsolete)
- new TDC with high throughput and fixed PAIR mode (& rad-tol)
- modified CSMs for GBT (& rad-tol)
- modify MRODs for GBT and higher data throuput

This upgrade for 128 MDTs requires

- ~ 128 new CSMs, GBTs
- ~ 25 new MRODs
- ~ 2500 new TDCs & mezz boards
- ~ 7500 new octal ASDs

A path to installation of new elx for improved BW and L1 trigger

Combining upgrade for increase BW <u>and</u> L1 trigger

- stage I: define architecture and interface lines with trigger chambers
- stage II: detailed definition of the new MDT readout system
- stage III: simulate operation of crucial components (ASICs, FPGAs) for timing and latency
- stage IV: produce prototypes of chips, test in lab and under realistic conditions
- stage V: decide on fine-tuning of system and make production prototypes
- stage VI: certify system with production prototypes and place volume orders
- stage VII: install new elx on the MDT