Triggering at 100 TeV?

Elliot Lipeles

University of Pennsylvania

100 TeV Workshop

Apr 23, 2014

Machine Parameters

Machine Parameters are not very defined yet

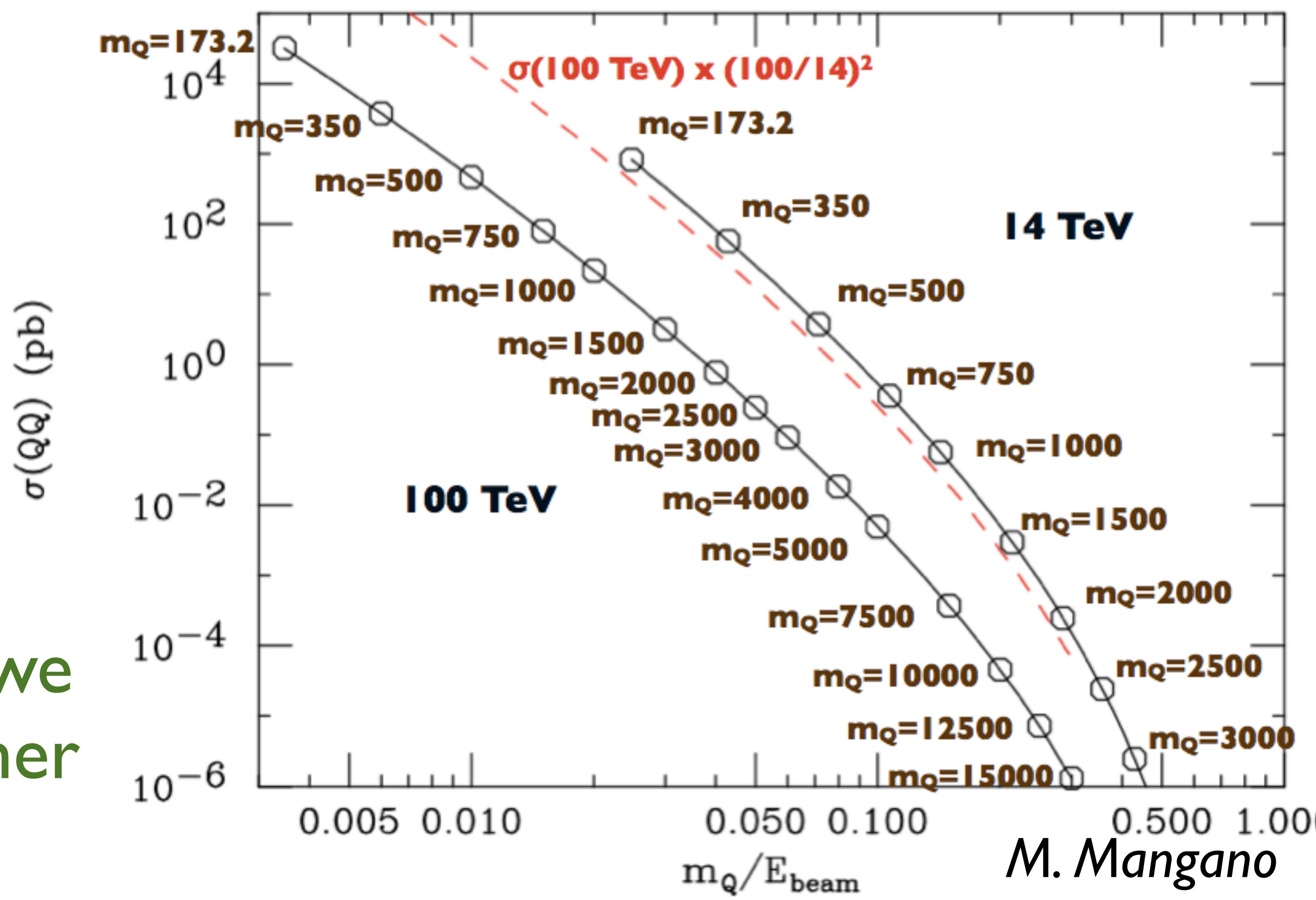
Consider two main cases:

Luminosity	Bunch Spacing	Pile-up
5×10^{34}	25 ns	170
5×10^{34}	5 ns	34

But need to pay attention to breaking points

Need $(100\text{ TeV}/14\text{ TeV})^2=49$ luminosity to reach the same sensitivity in new particle mass fraction of beam energy

If a way can be found, it is likely we will want to push luminosity higher



Triggering Purposes

Hardware = level-1 (L1)

Goal: Reduce data volume extracted from front-end chips

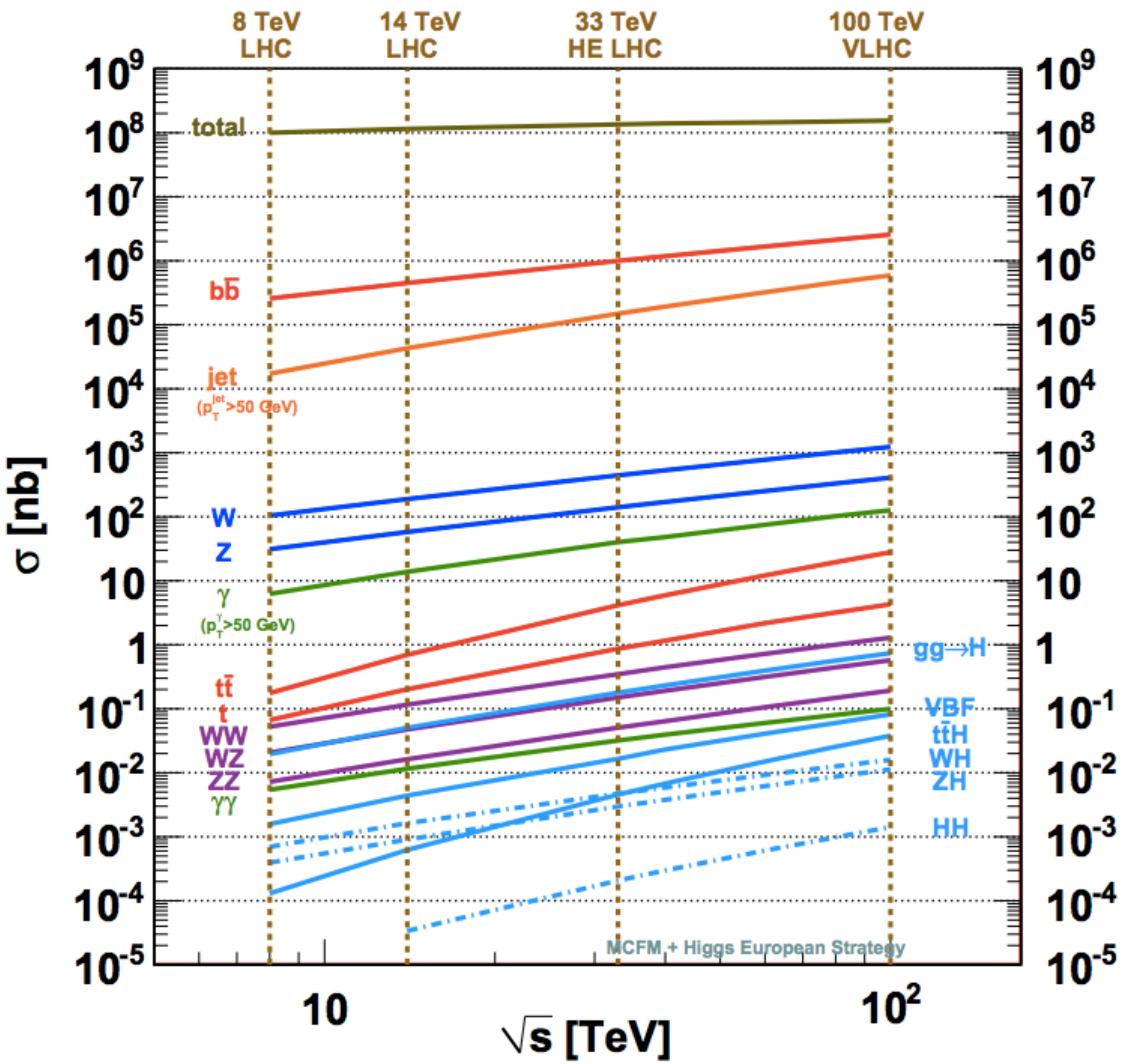
- Fundamental difference between inner detector (tracker) and outer detectors (calorimeter and muons)
- Tracker readout is necessarily in the tracker volume for a 4π detector
  - Contributions to tracker material
  - Tracker plays a big role in L1 discussion
- Outer detector readout is ~external to tracking volume
  - Assume full beam crossing rate readout and use in L1

Software = high-level trigger (HLT)

Goal: Reduce stored data volume

- PC-based with software
- Expect hardware to keep up with industry = Moore's Law?
- Expect affordable storage to scale with Moore's Law?

The Landscape



Minbias ~140 mb

~170 per crossing

Trigger Backgrounds

$b\bar{b}$ 150 MHz

~6 per crossing

~1 per crossing w/ lepton

Jets p_T > 50 GeV 25 MHz

~1 per crossing

Electroweak Physics

W+Z = 70 KHz!

Top ~30 mb

~1500 Hz

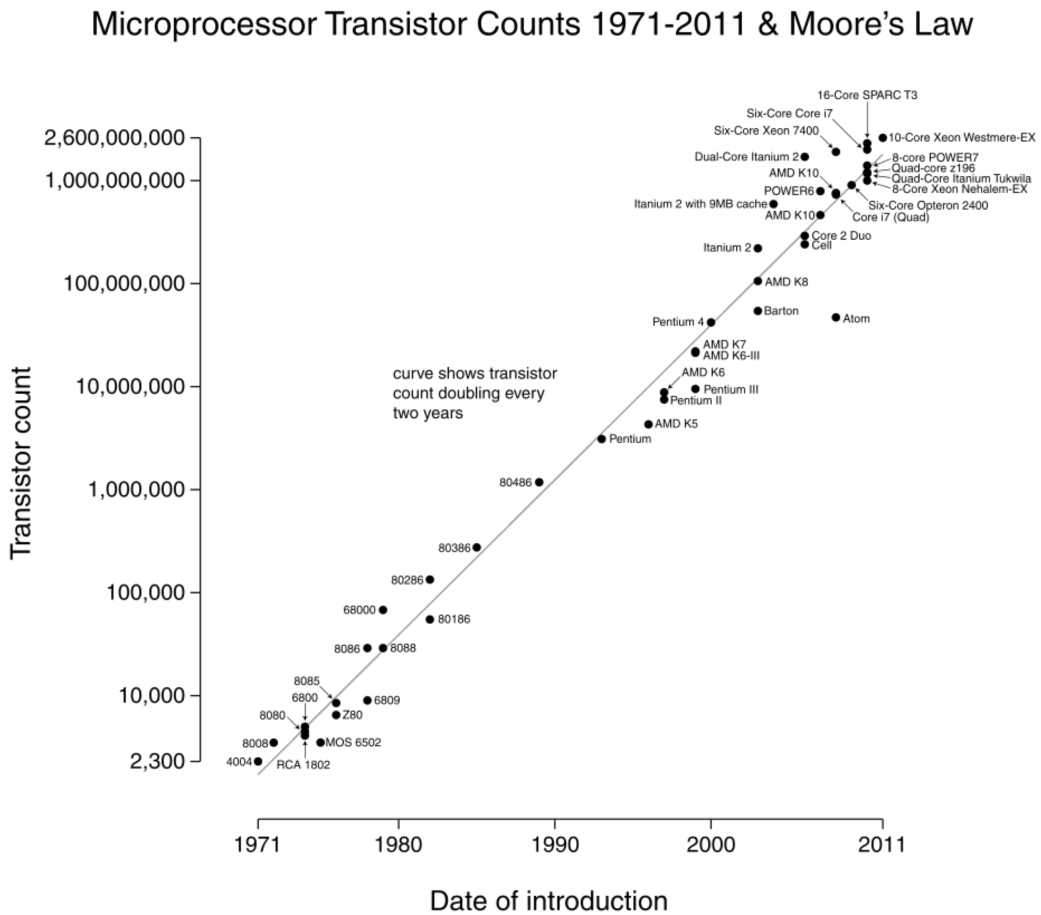
Moore's Law easily accommodates saving all the electroweak

Moore's Law

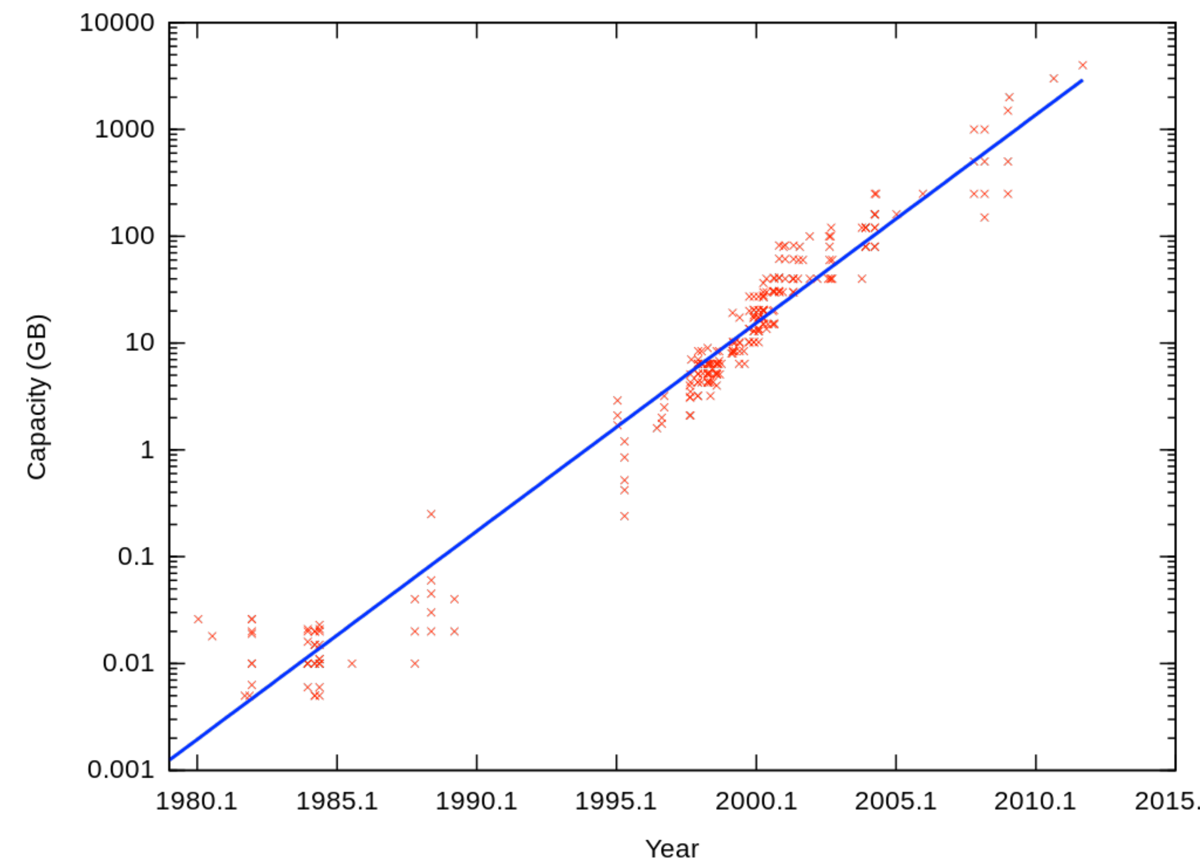
Emperical law (self-fulfulling?)

May not continue (physical limits, change in market factors,...)

Transistors in CPU



Storage



Doubles every ~2 years
Factor of ~8000 by 2040

Doubles every ~1.5 years
Factor of ~165,000?! by 2040

HLT and Moore's Law

UNIVERSITY of PENNSYLVANIA

HLT Output Rate (scales like storage)

- Run2(2015) HLT outputs expected to be 1 KHz
- 100+ KHz *probably* no problem in 2040
- Assume HLT output rate is not an issue

HLT Input Rate (scales like CPU+networking)

- HL-LHC rates expected to be ~200 KHz to 1 MHz
    - Detector readout limited anyway
- Input rates of order the beam-crossing rates probably achievable

Probably don't need to worry about HLT

What do we want to trigger on?

Easy stuff... core high p_T program

- Very high p_T leptons (incl τ), photons, jets, and met
- Hadronic SUSY with MET (with or without MET), Z', WW scattering, anomalous TGCs/QGCs, running of couplings

Pretty easy... single leptons

- precision/rare Higgs, top
- Many HH channels
- Much of electroweak SUSY

Challenges...

- HH to $bb\tau\tau$
- H to $Z\gamma$
- Monojets (+X)
- Exotics (monopoles, long lived hidden valley)
- Displaces Vertices

What do we want to trigger on?

Easy stuff... core high p_T program

- Very high p_T leptons (incl τ), photons, jets, and met
- Hadronic SUSY with MET (with or without MET), Z', WW scattering, anomalous TGCs/QGCs, running of couplings

Pretty easy... single leptons

- precision/rare Higgs, top
- Many HH channels
- Much of electroweak SUSY

Challenges...

- HH to bbττ
- H to Zγ
- Monojets (+X)

- Exotics (monopoles, long lived hidden valley)
- Displaces Vertices

Challenges: HH to bbττ

Theorist analysis (ignores fakes)

	HH	$b\bar{b}\tau\bar{\tau}$	$b\bar{b}\tau\bar{\tau}\nu_\tau\bar{\nu}_\tau$	ZH	S/B	$S/\sqrt{B}$
Cross section NLO [fb]	2.47	2.99×10^4	8.17×10^3	2.46×10^1	6.48×10^{-5}	6.93×10^{-1}
Reconstructed Higgs from τs	2.09×10^{-1}	8.35×10^1	1.58×10^2	5.70×10^{-1}	8.63×10^{-4}	7.36×10^{-1}
Reconstructed Higgs from bs	1.46×10^{-1}	6.34×10^{-1}	1.43×10^1	3.75×10^{-2}	9.75×10^{-3}	2.07
Cut on M_{HH}	1.30×10^{-1}	1.37×10^{-1}	1.74	1.26×10^{-2}	6.88×10^{-2}	5.18
Cut on $P_{T,H}$	1.10×10^{-1}	7.80×10^{-2}	7.17×10^{-1}	1.15×10^{-2}	1.36×10^{-1}	6.71
With 112.5 GeV $< M_{\tau\bar{\tau}} <$ 137.5 GeV	1.10×10^{-1}	3.41×10^{-2}	3.76×10^{-1}	3.15×10^{-3}	2.67×10^{-1}	9.37

Table 9: Cross section values of the of HH signal and the various backgrounds expected at the LHC at $\sqrt{s} = 14$ TeV, the signal to background ratio S/B and the significance $S/\sqrt{B}$ for $\int \mathcal{L} = 3000$ fb^{-1} in the $b\bar{b}\tau\bar{\tau}$ channel after applying the cuts discussed in the text.

J. Baglio, A. Djouadi, R. Grober, M.M. Muhlleitner, J. Quevillon and M. Spira, arXiv:1212.5581v2

Potential trigger objects

- 2 b-jets, pT > 30 GeV (probably not really possible offline)
- 2 τ, pT > 30 GeV

b and τ are both difficult without tracking

If you can't ensure that the objects came from the same vertex, this is just two dijets collisions

Challenges: Monojets

"Wino" Dark Matter

$\chi^{\pm}$
χ_0

Small $\Delta m(\chi^{\pm} - \chi_0)$

"Higgsino" Dark Matter

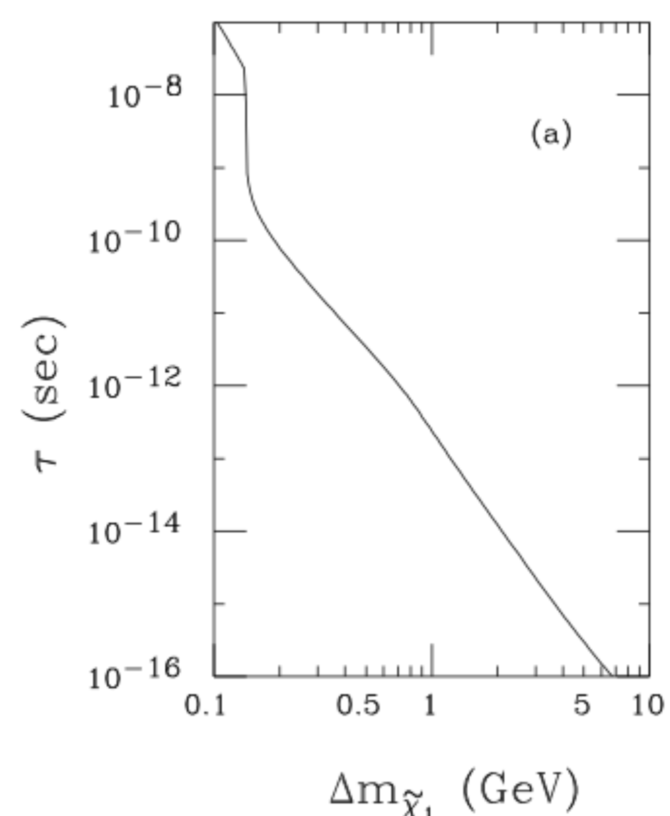
χ_2
$\chi^{\pm}$
χ_0

Small $\Delta m(\chi^{\pm} - \chi_0)$
Small $\Delta m(\chi_2 - \chi^{\pm})$

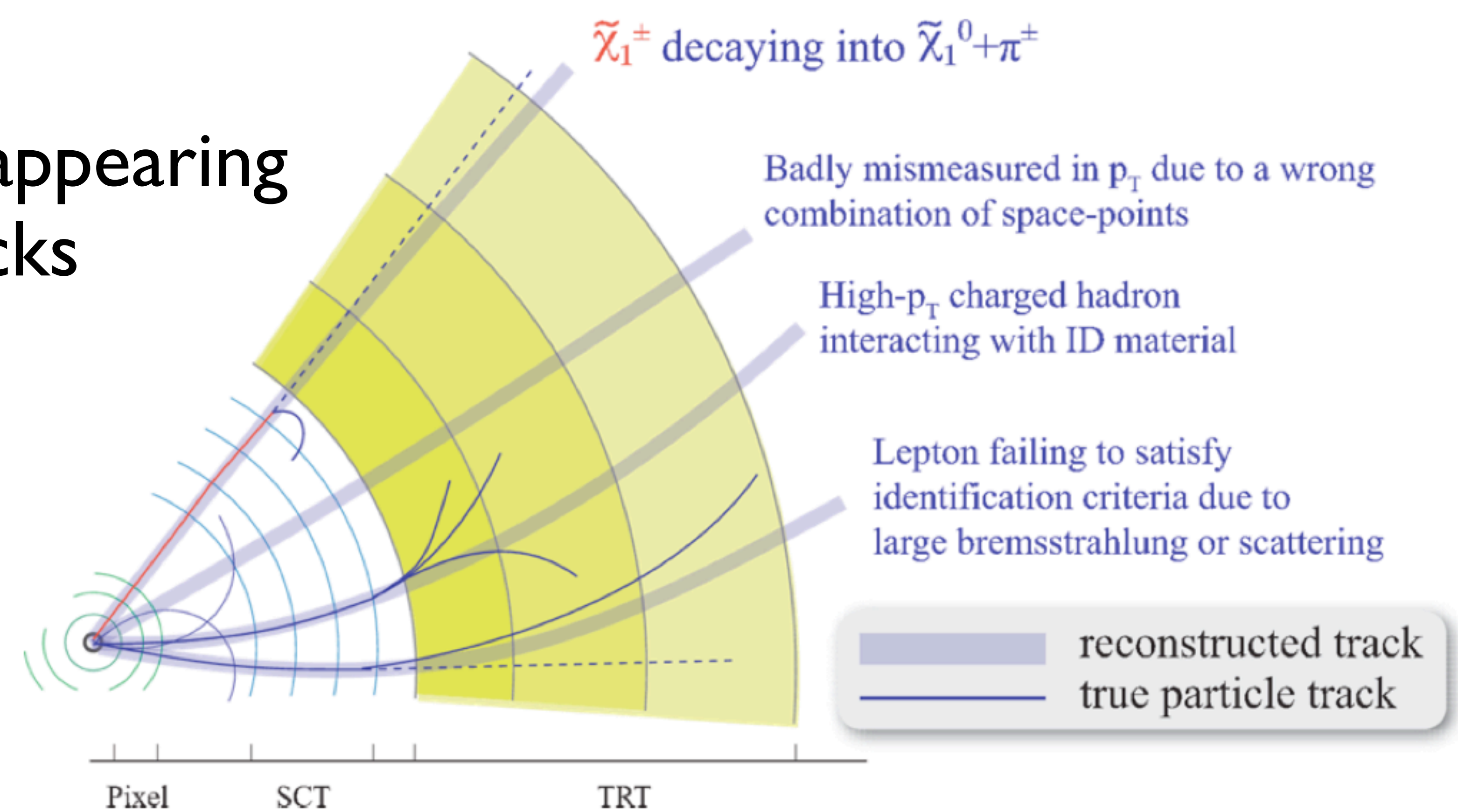
Possibly long-live

Disappearing Tracks

Soft Leptons



Chen, Drees and Gunion, hep-ph/9902309

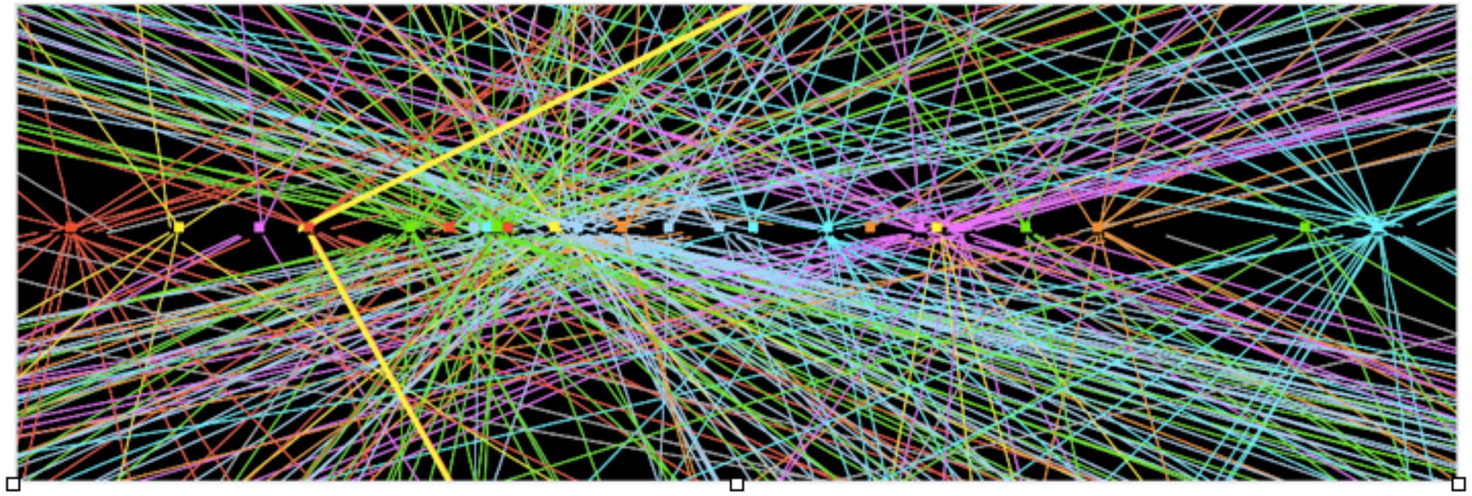


Distinctive offline signature; Need to trigger on moderate MET

High χ_0 mass could mean small cross-section

Pile-up scaling issues

Multiobject triggers scale badly with pile-up...



If p is the probability that a single collision produces object passing a given threshold, then the trigger rate for that object is

$$\text{Rate} = p\mu f$$

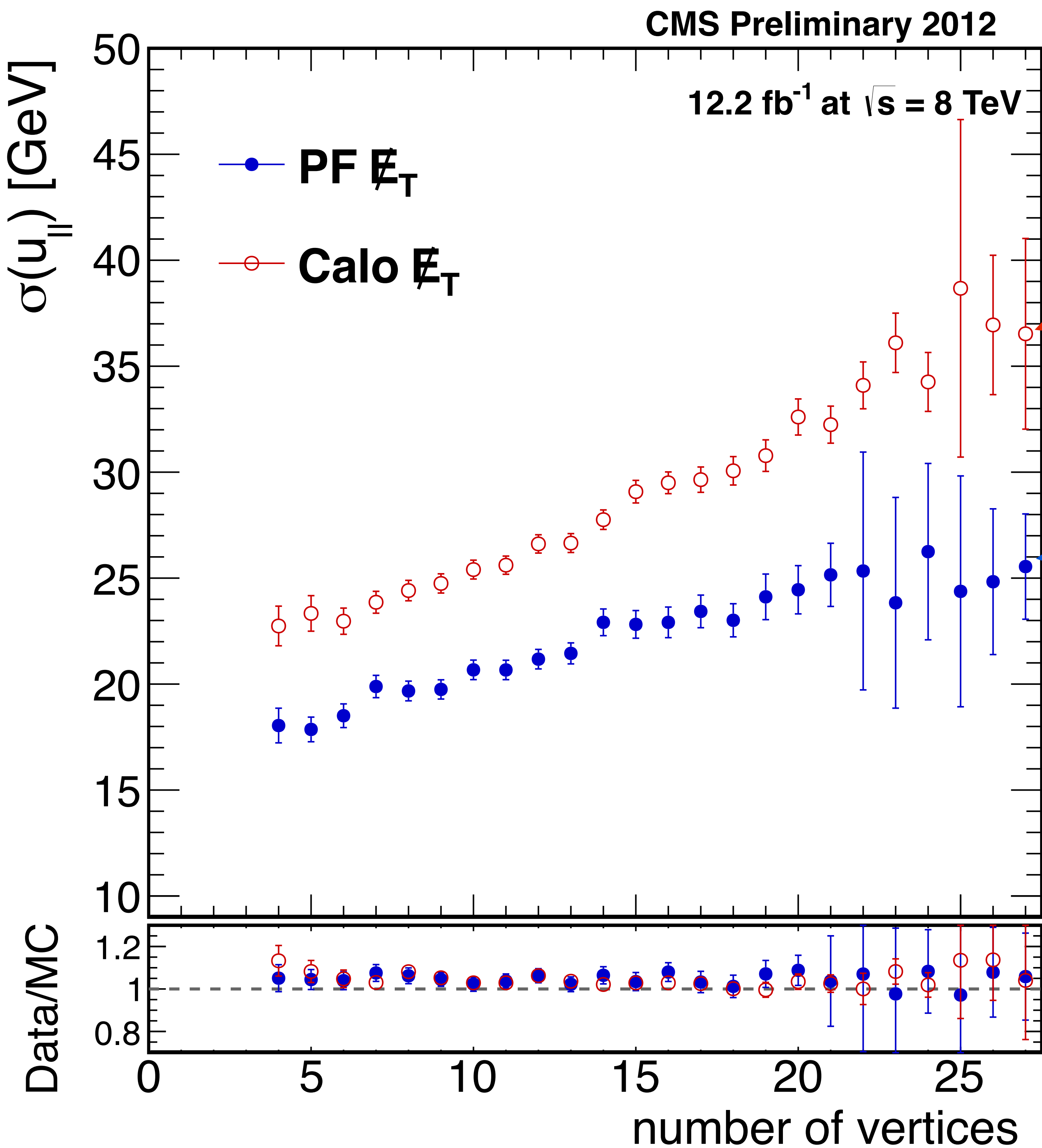
where f is the frequency of crossings and μ is the number of collisions per crossing

The rate for a coincidence of two such objects is approximately

$$\text{Rate} = \frac{1}{2}(p\mu)^2 f$$

I.e. it grows with the square of μ, and worse for more objects!!!

Pile-up, MET, and tracking



Calorimeter only

Using tracking information

Both LHC experiments are using tracking extensively in missing energy calculations

Importance increases substantially with pile-up

Potential uses of tracking in trigger

Object Id

- Electrons: track-shower matching
- Muons: p_T measurement
- Taus: track counting and isolation
- Calo-isolation for highly boosted tops maybe difficult

Vertex confirmation: leptons (incl τ) , jets

- Key for multiobject triggers
  - Because of high coincidence rate at high pile-up

Missing Energy

- Calorimeter MET degraded with pile-up
- Need to match offline performance
- Association of MET to vertex

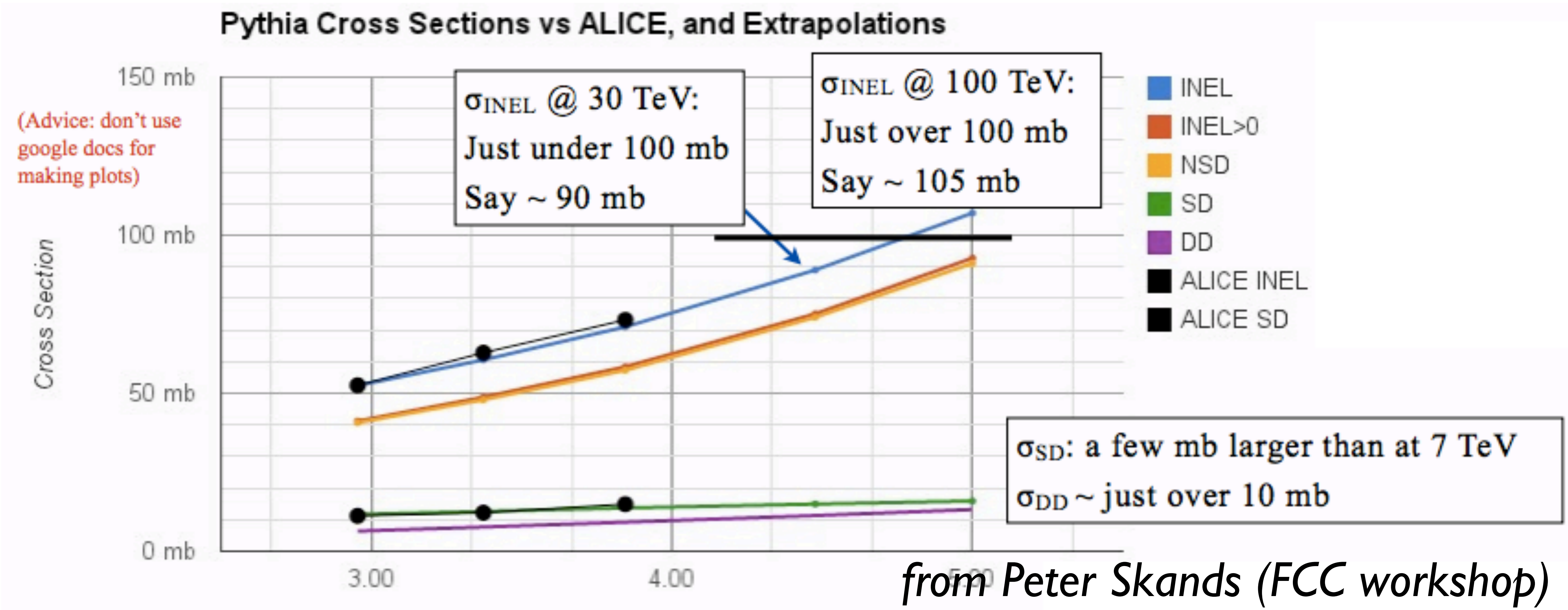
The big question is how much tracking is feasible to get into L1

This could also influence detector geometry and design

Data Volumes Possible: Tracker

Data volume scales by ~ (cross-section)x(multiplicity)x(rapidity coverage)

	100 TeV/ 14 TeV
Cross-section	100 mb/70 mb = 1.4
Multiplicity	
Rapidity Coverage	

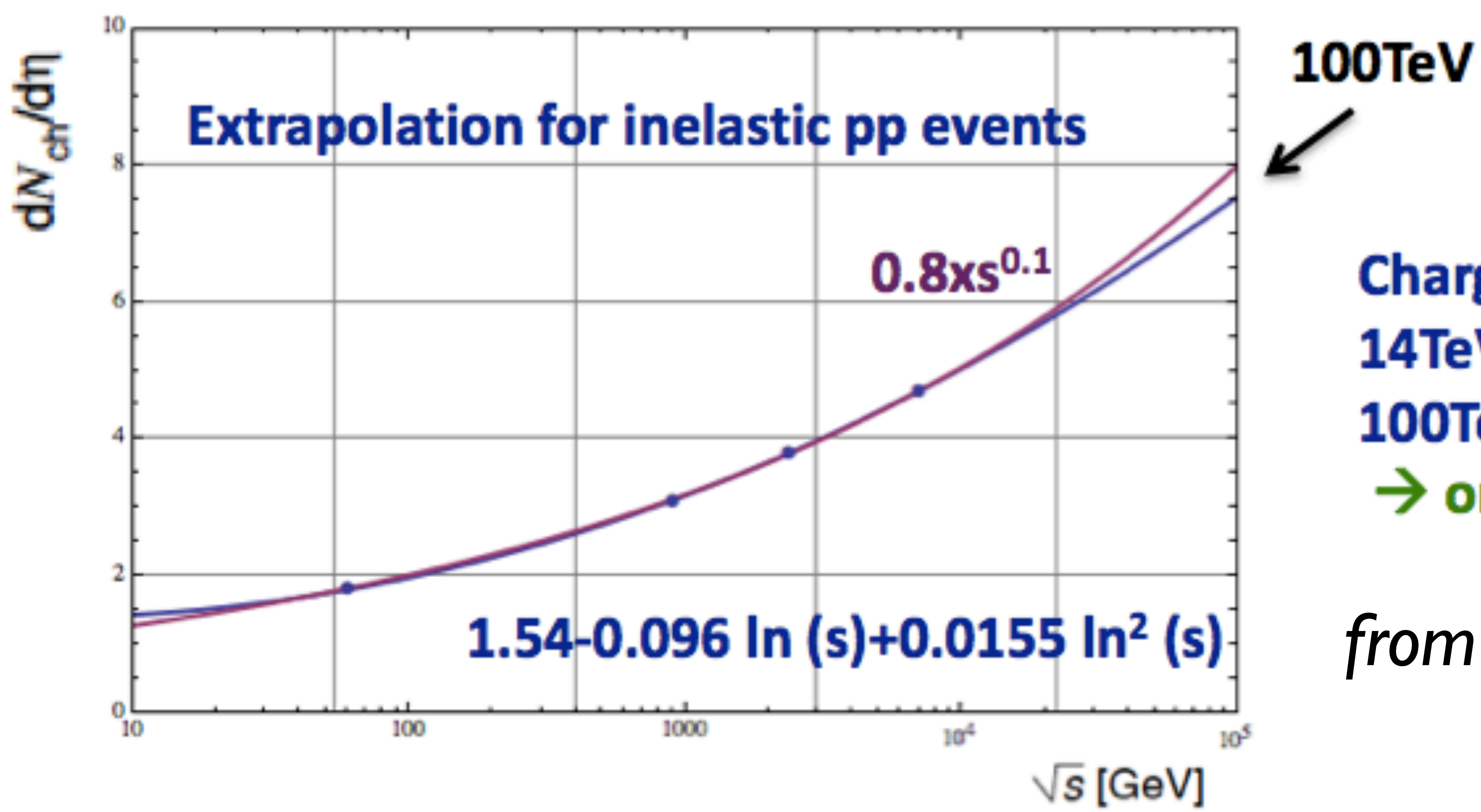


from Peter Skands (FCC workshop)

Data Volumes Possible: Tracker

Data volume scales by ~ (cross-section)x(multiplicity)x(rapidity coverage)

	100 TeV/ 14 TeV
Cross-section	100 mb/ 70 mb = 1.4
Multiplicity	1.5
Rapidity Coverage	



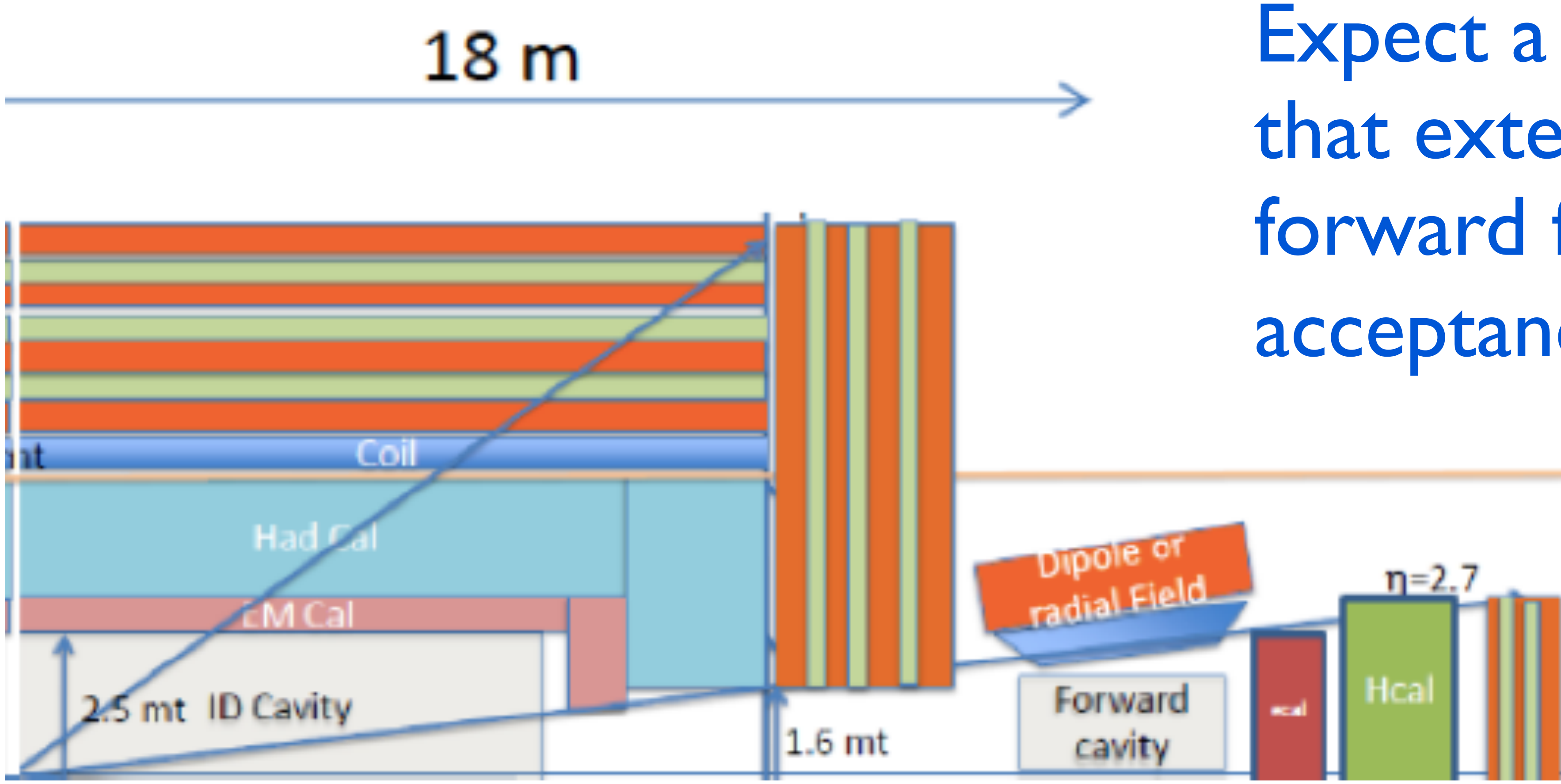
Charged particle multiplicity at
14TeV ≈ 5.4
100TeV ≈ 8
→ only about 1.5 times larger

from Werner Riegler (FCC workshop)

Data Volumes Possible: Tracker

Data volume scales by ~ (cross-section)x(multiplicity)x(rapidity coverage)

	100 TeV/ 14 TeV
Cross-section	100 mb/ 70 mb = 1.4
Multiplicity	1.5
Rapidity Coverage	5/2.5 = 2



Expect a geometry that extends farther forward for tracking acceptance ~ |η|<5

Data Volumes Possible: Tracker

Data volume scales by ~ (cross-section)x(multiplicity)x(rapidity coverage)

	100 TeV/ 14 TeV
Cross-section	100 mb/ 70 mb = 1.4
Multiplicity	1.5
Rapidity Coverage	5/2.5 = 2

HL-LHC plans

- ATLAS ~ 250 KHz readout (may be increased)
- CMS 0.5-1.0 MHz readout

Data Volume scales by ~2-4

- Not so clear eta coverage counts (just more links)
- Forward eta region will have a much higher track density, but also has better access for services

Reading out at 40 MHz (or 200 MHz) requires increased readout of

(Data Volume factor)x(Readout frequency ratio) = ~80-320 increase

... and of course we don't want more material!

Future links

Current (rad hard)

~ 10 Gbit/sec ("upgraded GBT", 65 nm)

Multilevel (should be rad hard)

~ scale by 2-4

Industry, Multiwave length (rad hard???)

~ scale by order 100+?

Unclear whether full tracker readout will be possible

Depends on rad hard link development

Options for tracking in the L1

No L1:

- Full tracker readout at 40 MHz or 200 MHz
- Requires large bandwidth ... already discussed

Tracking in the trigger

RoI-based = regional data filtering

- Needs larger latency (two-steps)

Self-seeded = p_T filtering

- Needs special tracker geometry

Buffering...

If you don't do full tracking readout you need to buffer to an extended time...

	Latency	Buffer Size
CMS HL-LHC	10-20 μs	400-800
ATLAS HL-LHC	~6 μs	~240
25 ns readout	Similar numbers to HL-LHC	
5 ns readout	HL-LHC x 5 ~= 2000-4000	

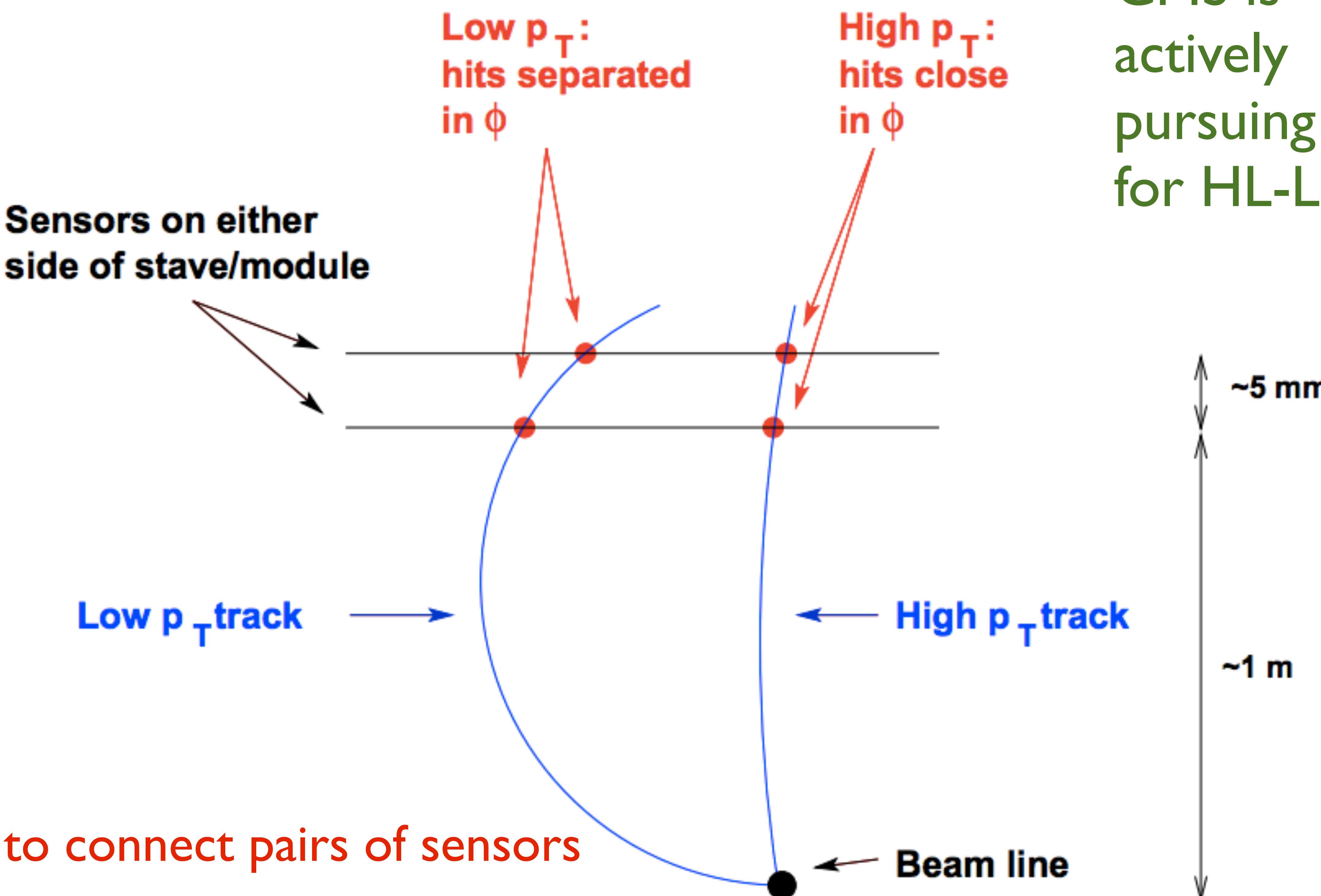
Significant latencies are not so far from what is achievable today even for 5 ns

- feature size will likely decrease, although small feature chip production can be expensive

Reducing the data flow: Filtering on p_T

"Unseeded"/Doublet Method

CMS is actively pursuing this for HL-LHC



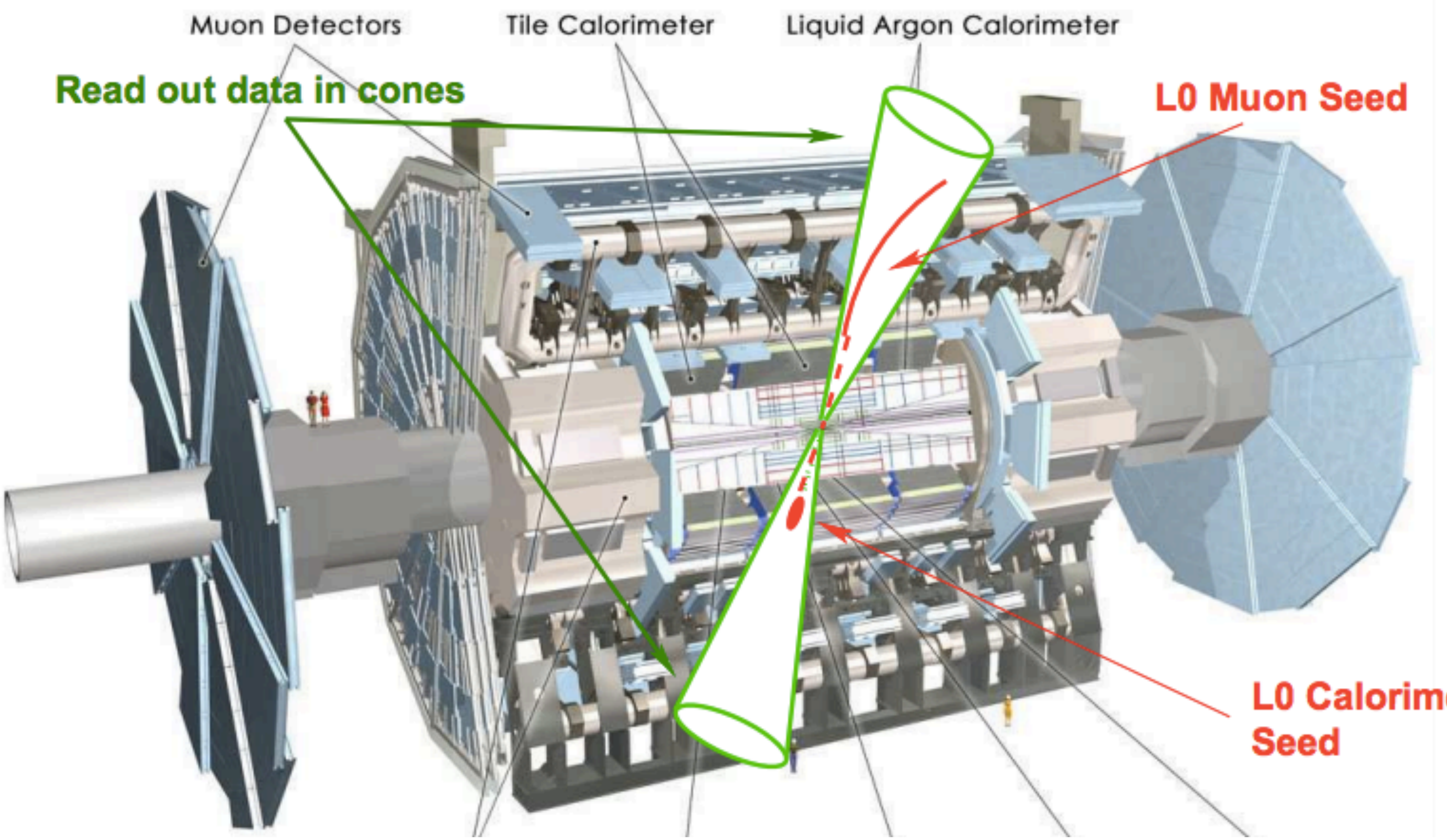
Need to connect pairs of sensors

Called "Intelligent" tracking

Reducing the data flow: Region method

Two-level trigger: L0 and L1

- **L0** uses calorimeter and muon system to define regions of interest (RoIs)
- **L1** extracts tracking for just RoIs from detector front-ends

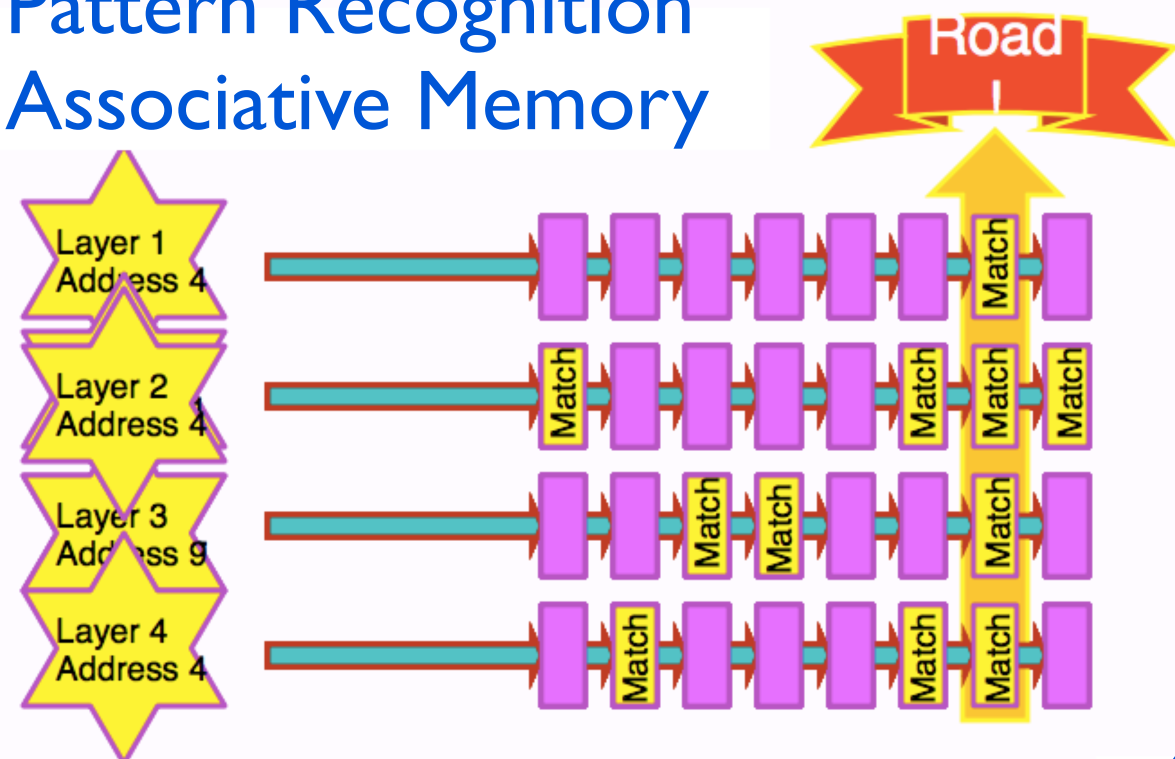


ATLAS is actively pursuing this for HL-LHC

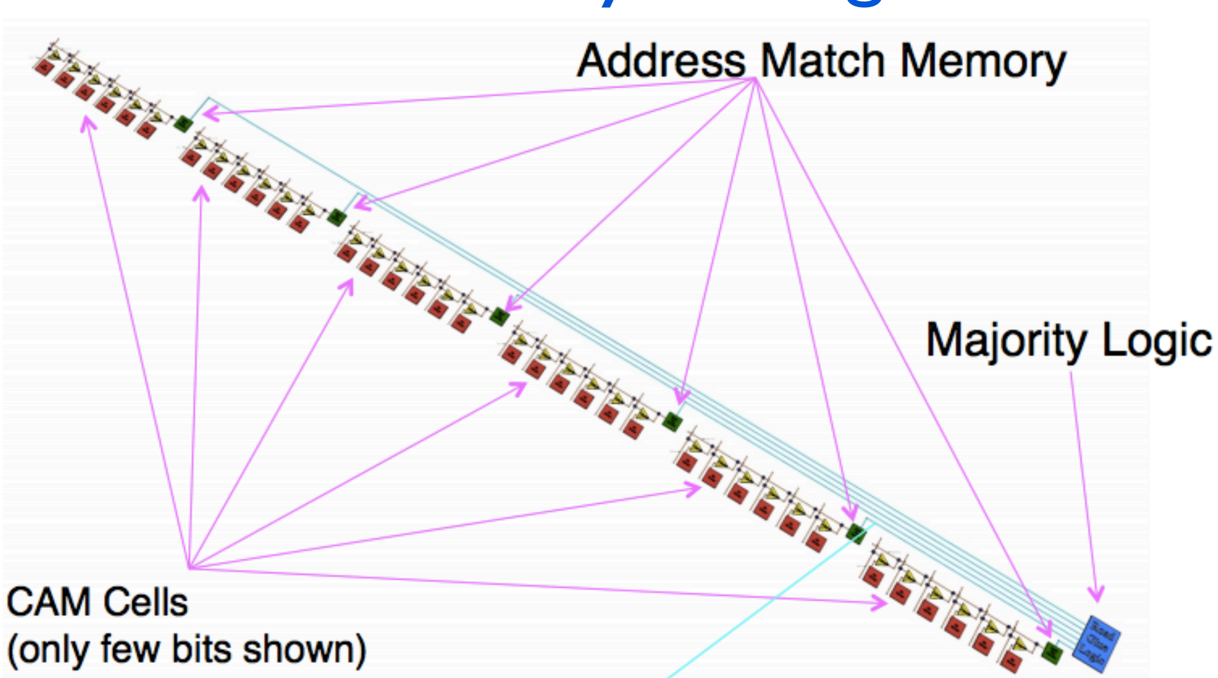
Pattern Recognition

Once we get the tracking data off the detector, we need to find tracks!

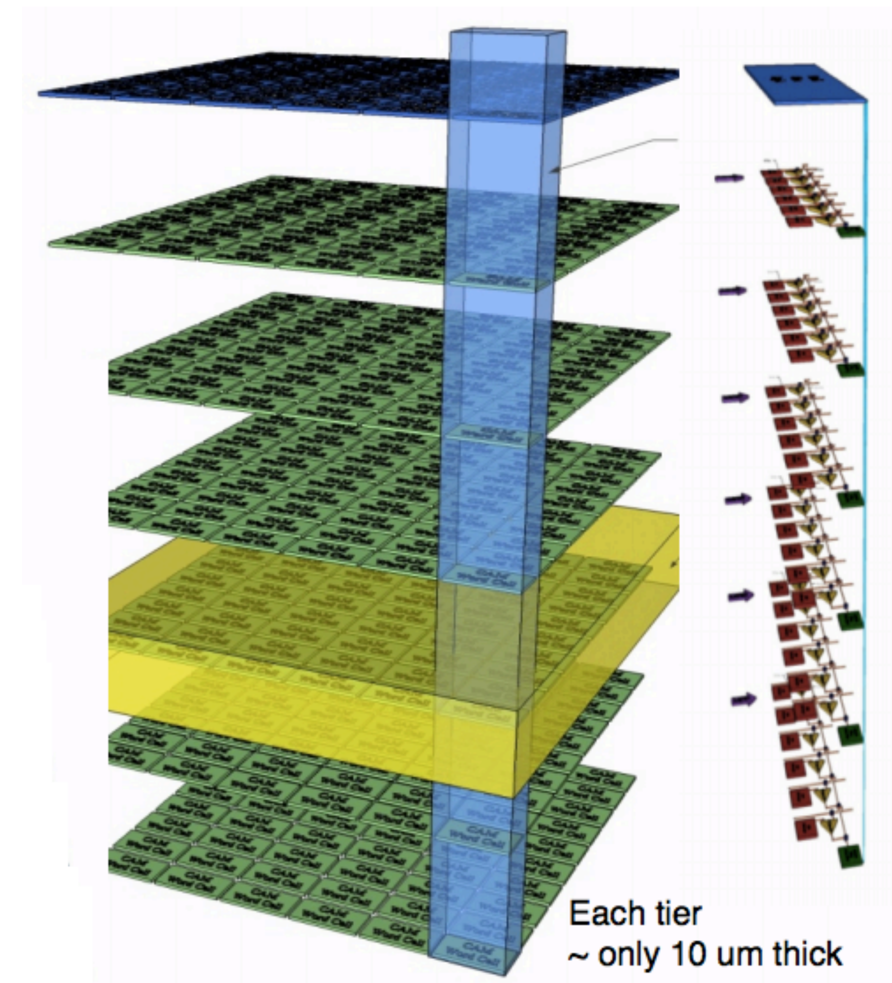
Pattern Recognition Associative Memory



Intrinsically 3D algorithm



Use a 3D chip!



Ted Liu, FNAL

Faster, more patterns

3d fab technology likely to advance

Need depends on L1 tracking goals

Comparing Methods

Doublet/"unseeded"/push method (p_T filtering)

- Delivers: Higher-pT tracks for all crossings
  - Good for lepton id, isolation
    - Vertex association maybe difficult p_T filtering in the inner tracker layers is harder ... high B-field helps
  - Good for missing energy?
    - What threshold in p_T corresponds to what data reduction?
    - What threshold in p_T is sufficient for MET calculation?

Region of Interest/"seeded"/pull method (regional filtering)

- Delivers: Regional tracking (in some or all crossings)
  - ATLAS HL-LHC version only looks in subset of events, but could do regions at 40(200) MHz
  - Good for lepton id, isolation, vertex association
  - Good for missing energy?
    - Can do jet-vertex association, but not "soft-term" for the unclustered energy

May even want to do both? What combined rejection can be acheived?

Summary

Reasonably Moore's law assumptions makes HLT probably CPU and storage probably not a big issue

Tracking is the core question for 100 TeV

- Difficult channels examples: HH to bb$\tau\tau$, Monojets
- RoIs or Self-seeded both good solutions for local high-p_T objects
- MET is probably the trickiest case

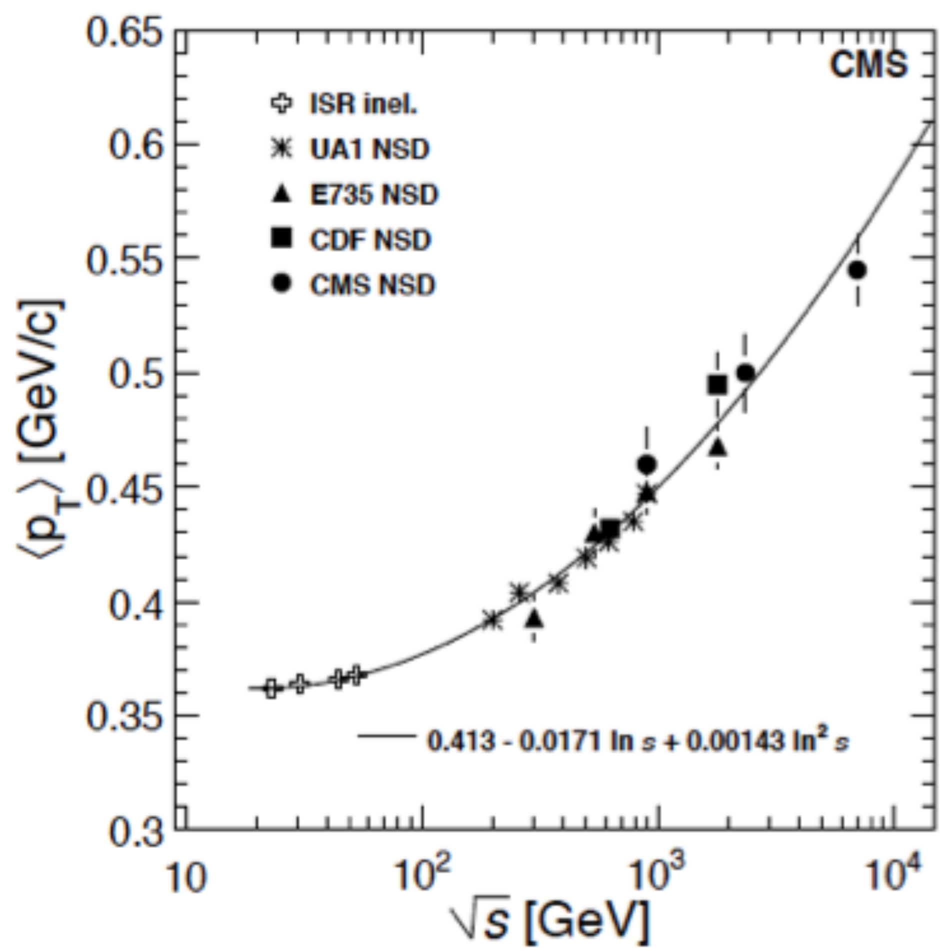
Full tracking at beam crossing rate technologies:

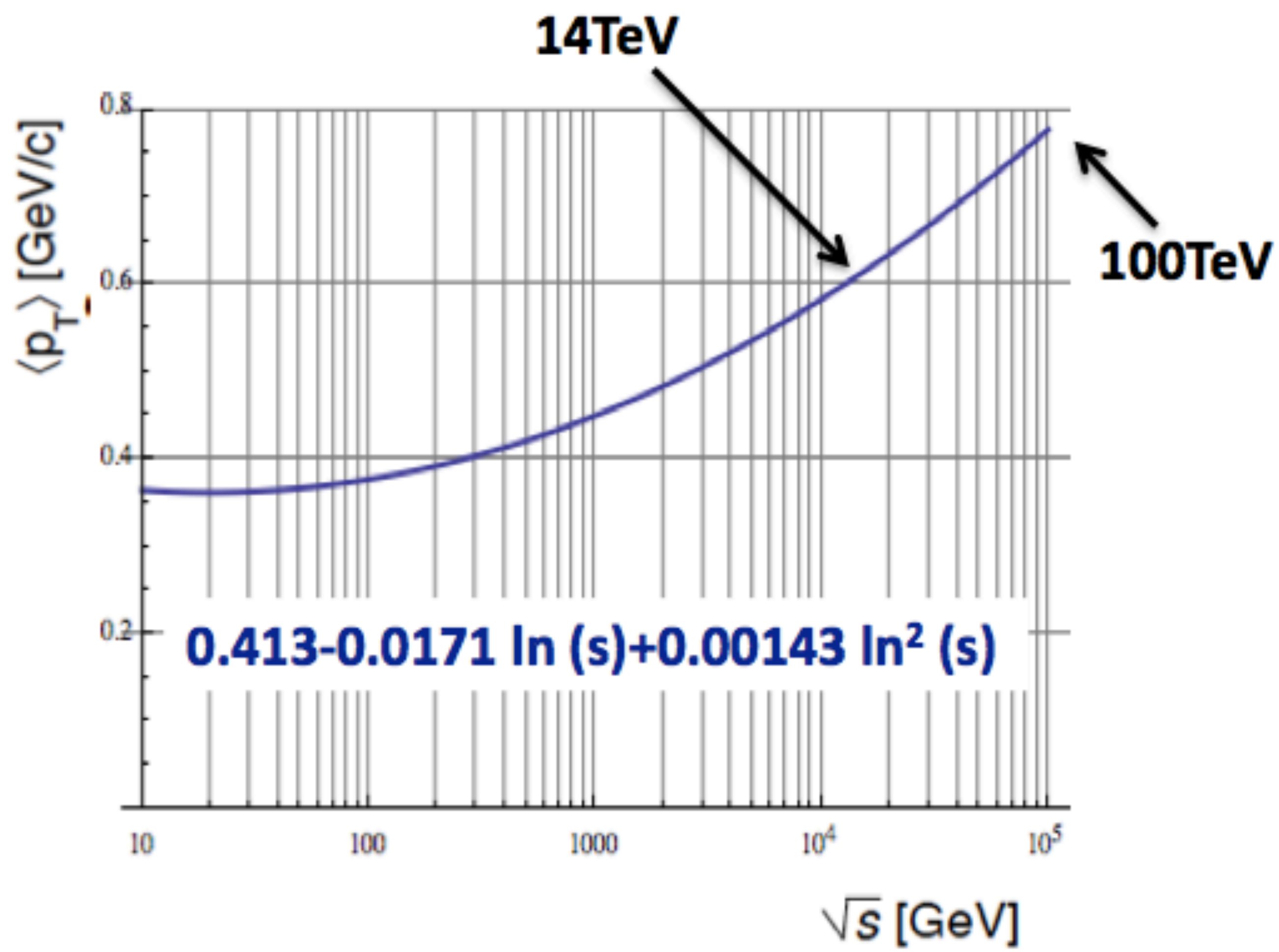
- Data extraction depends on rad hard link development
- Pattern recognition, new technologies look promising
- Correlating sensors at front-end = "Intelligent" tracking

Back up: Data Volume

From Werner Riegler

Average Particle Momentum





Average p_T approx. 0.6GeV/c for 14 TeV and 0.8GeV/c at 100TeV i.e. increase of 33%.

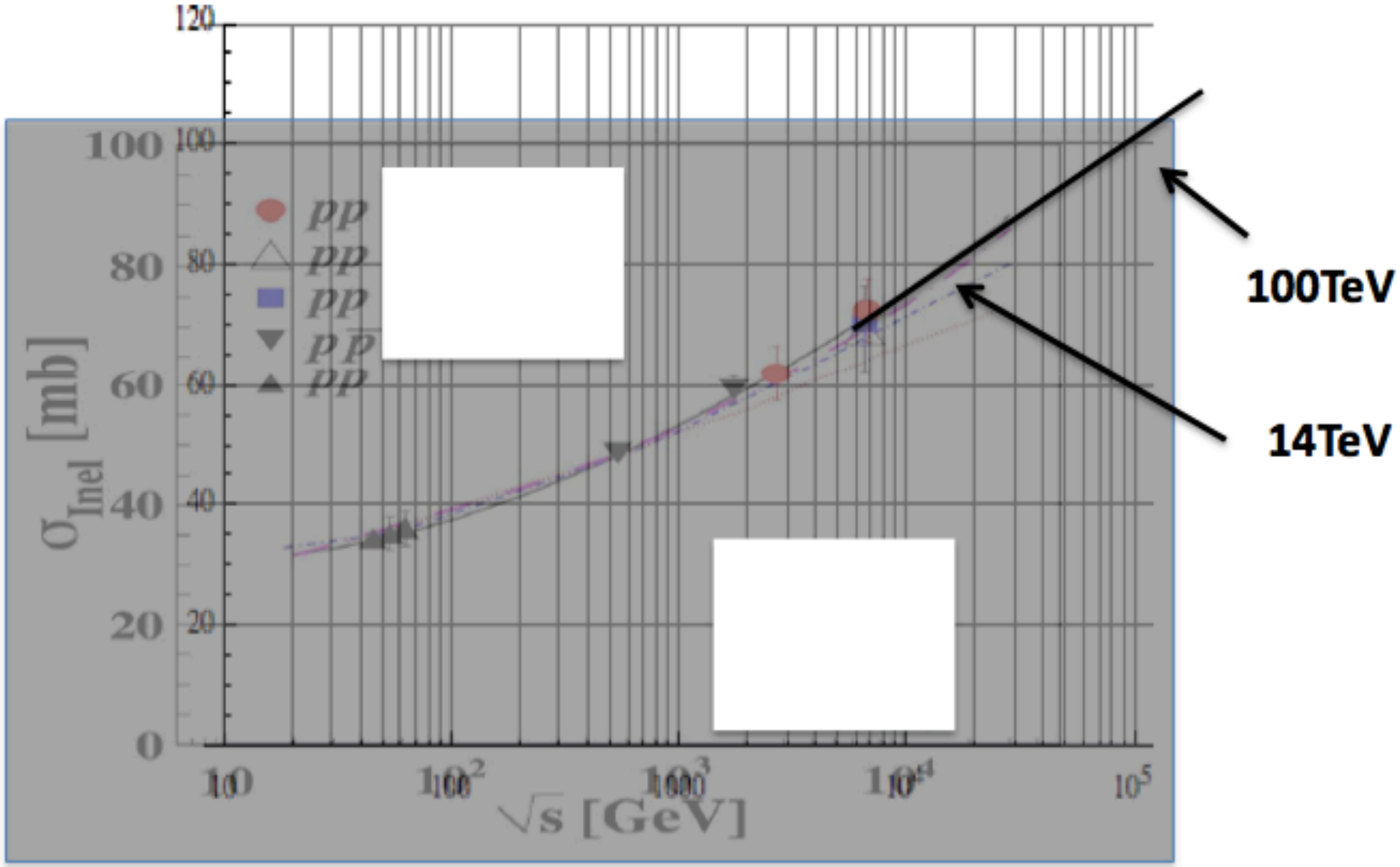
Bending in radius in 4T field:
R[m] = 3.33 * p_T[GeV/c] / B[T] = 3.33 * 0.8/4 = 0.67m

→ Average particle will curl with 1.33m diameter inside the ID.

14/02/2014 W. Riegler, FCC kickoff

From Werner Riegler

Inelastic pp crossection



Inelastic pp crossection, hand extrapolation from data up to 7 TeV:

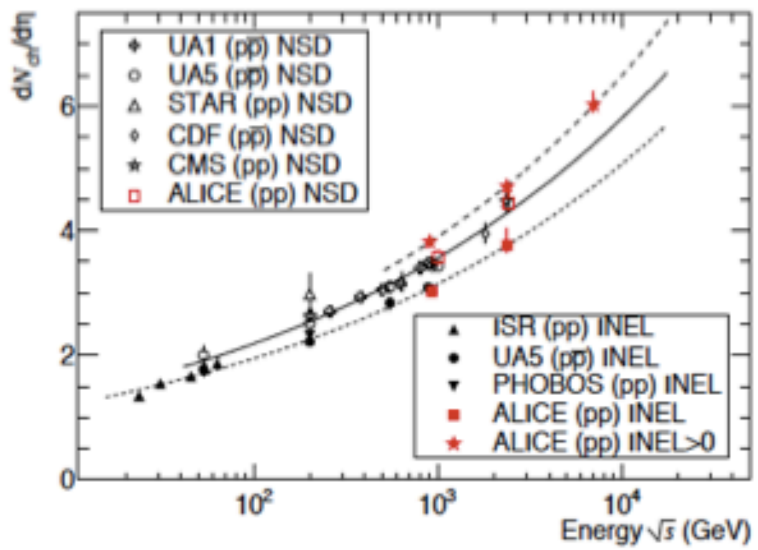
≈ 80mb at 14TeV

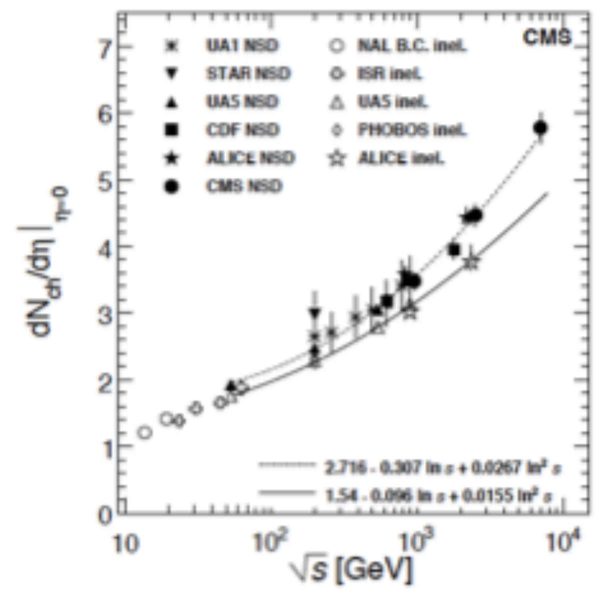
≈ 100mb at 100TeV

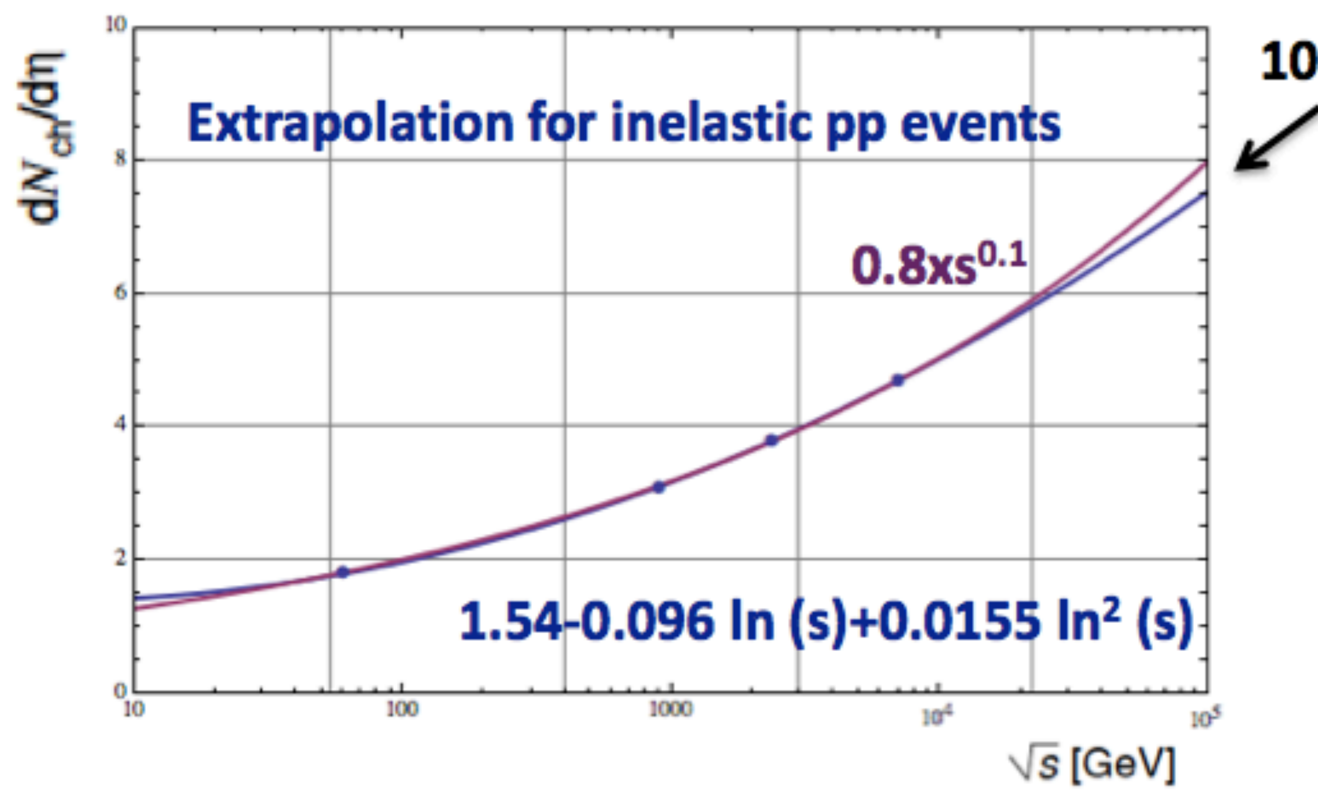
→ 25% increase

From Werner Riegler

Multiplicities







100TeV

Charged particle multiplicity at
14TeV ≈ 5.4
100TeV ≈ 8
→ only about 1.5 times larger

14/02/2014 W. Riegler, FCC kickoff

From Werner Riegler

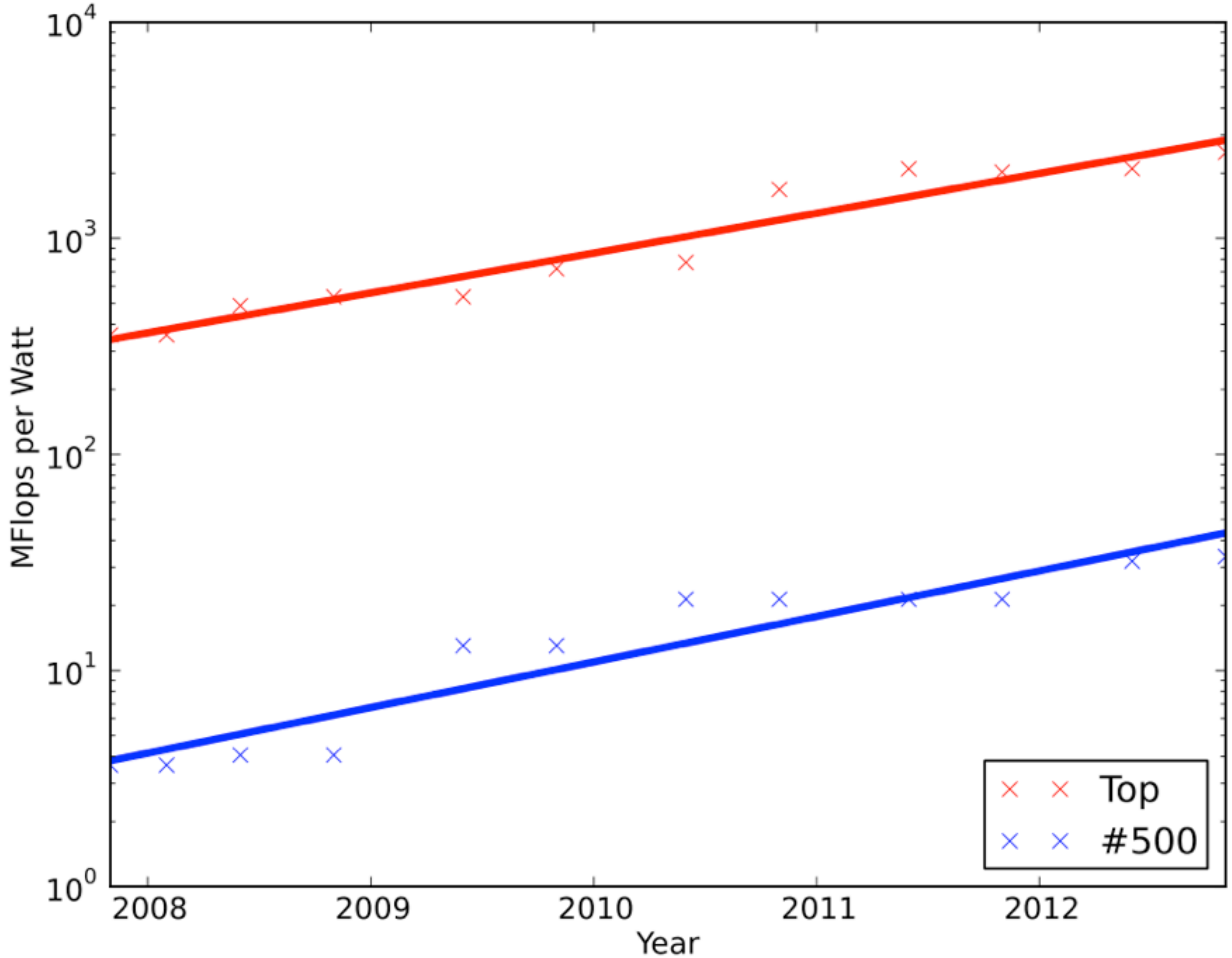
Assume a full pixel tracker:

- $L=5x10^{34}$ at 100TeV → $5x10^{9}$ pp collisions/second
- dN/dη = 8 i.e. 80 tracks inside η ±5
- Each track crosses 15 tracking stations
- In each station 5 pixels are fired.
- Each hit is encoded in 5 Bytes
- Factor 5 for background + curling etc.

→750 TByte/second into online system

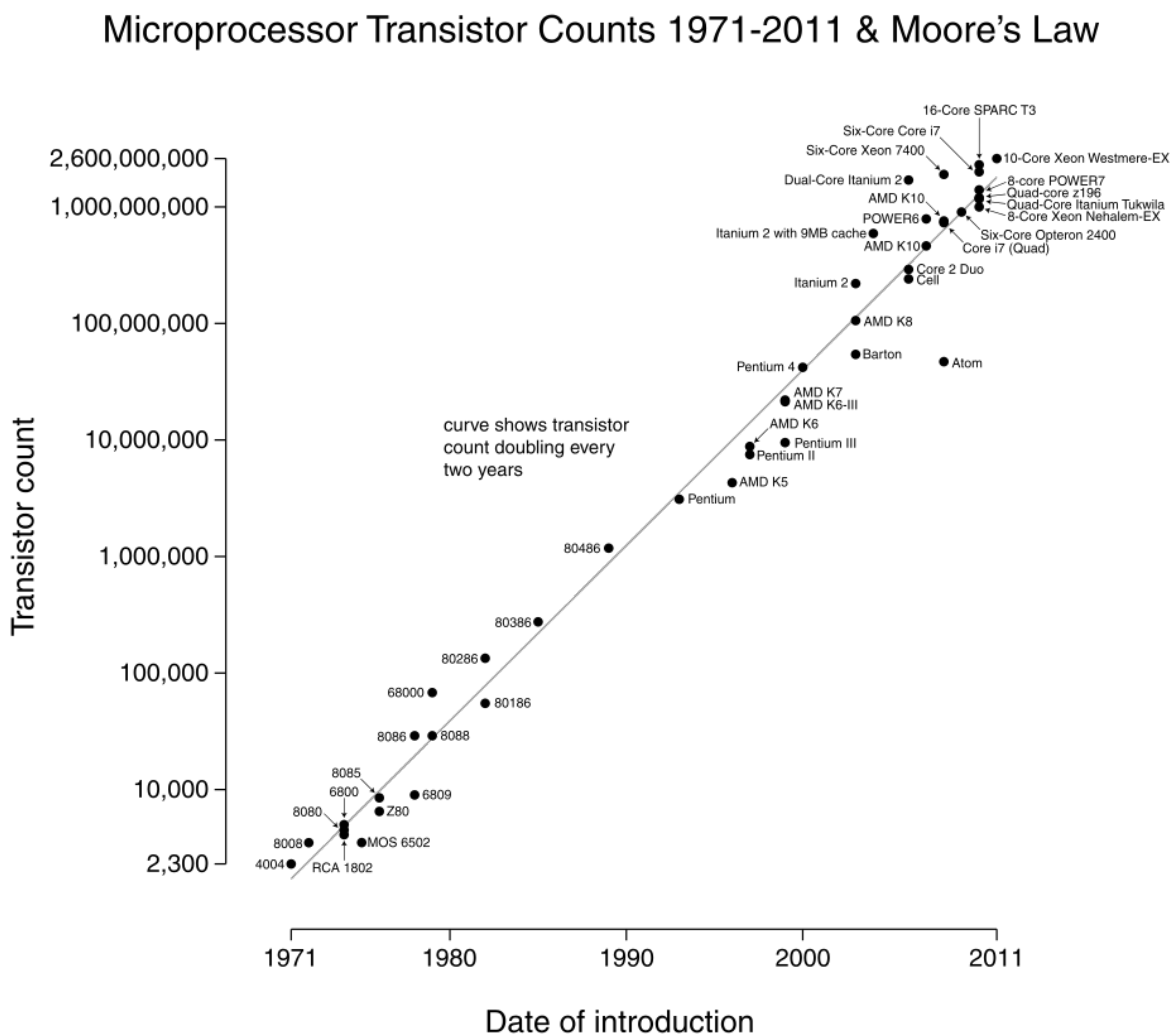
Back up: Moore's Law

Flops/Watt



Doubles every ~2.5 years?

Moore's Law: CPU



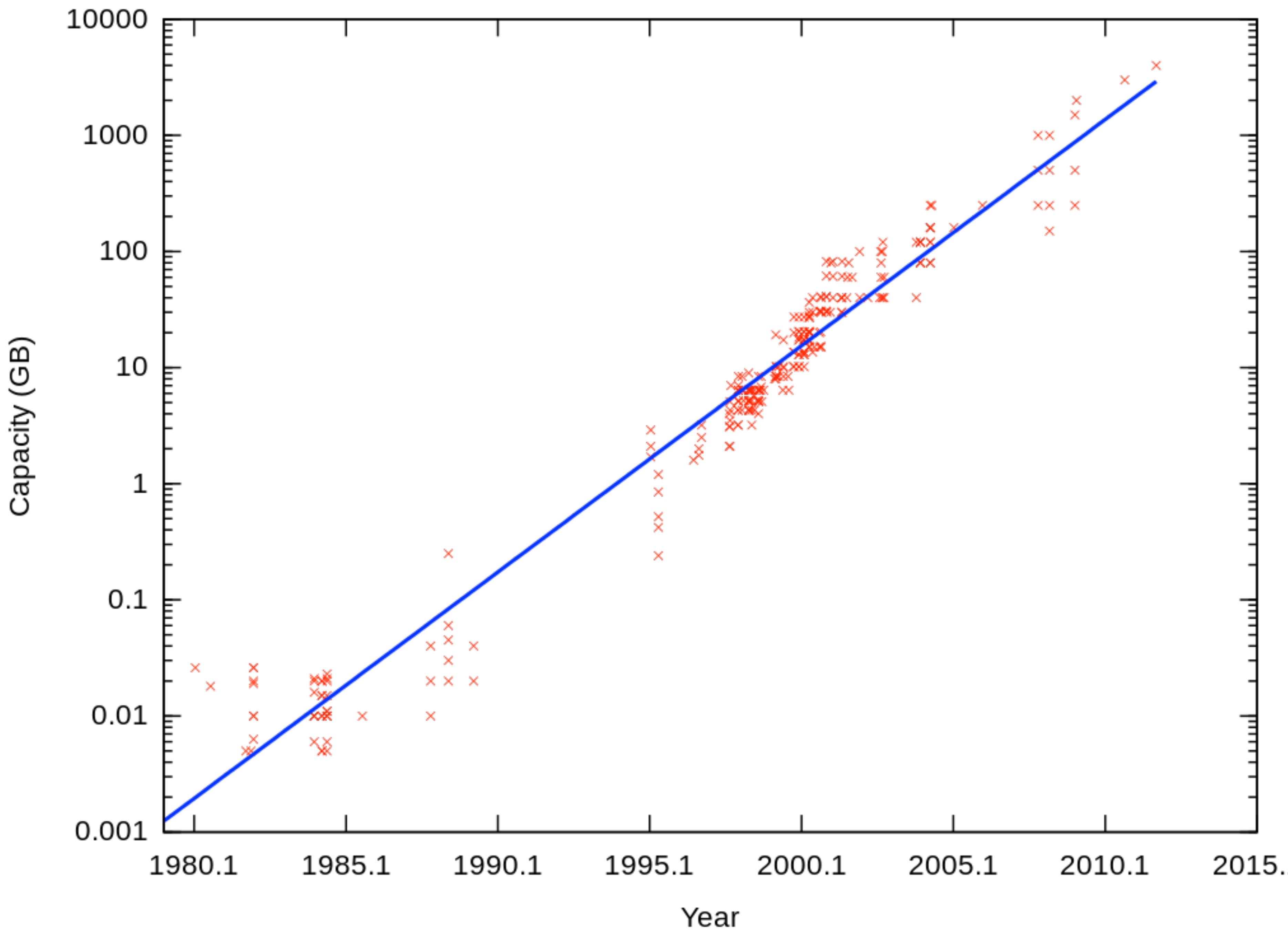
Doubles every ~2 years

Will it continue?

~factor of 8000 by 2040

Moore's Law Storage

Hard drive Capacity



Doubles every ~1.5 years

Will it continue?

~factor of 165,000 by 2040?

Back up: Track trigger filtering

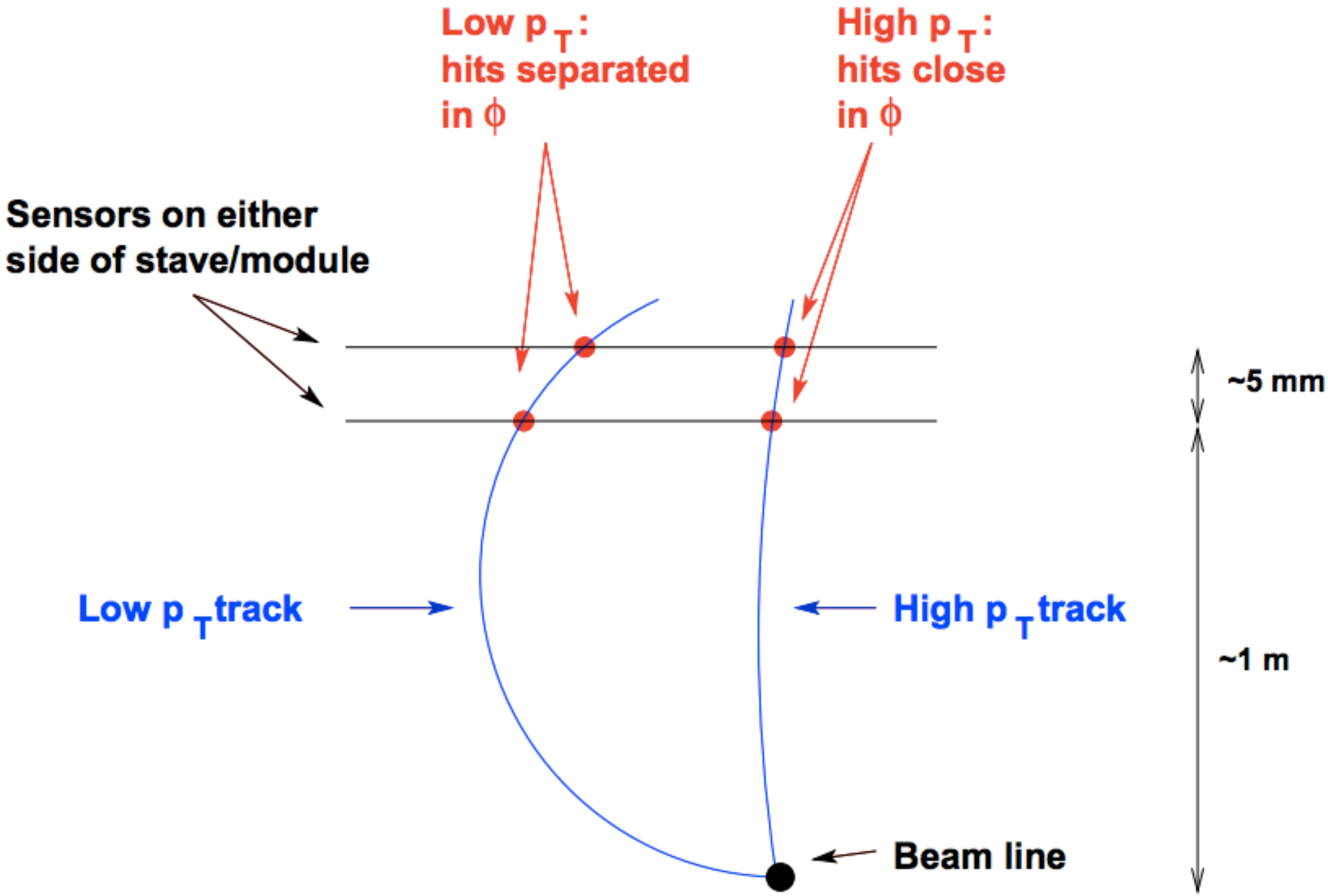
Filtering on p_T:

unseeded, doublet, push model

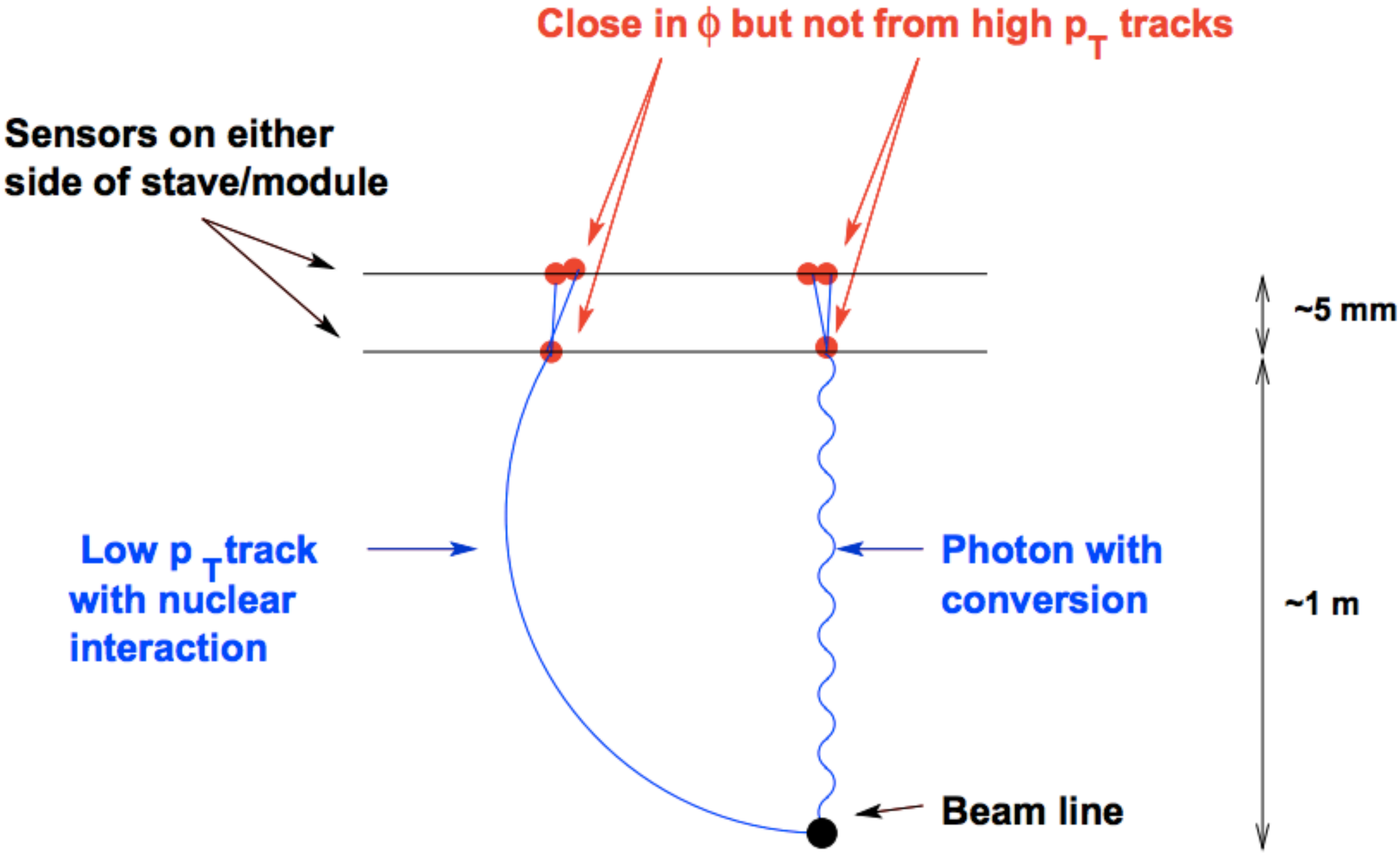
Reducing the data flow: Filtering on p_T

Penn
UNIVERSITY *of* PENNSYLVANIA

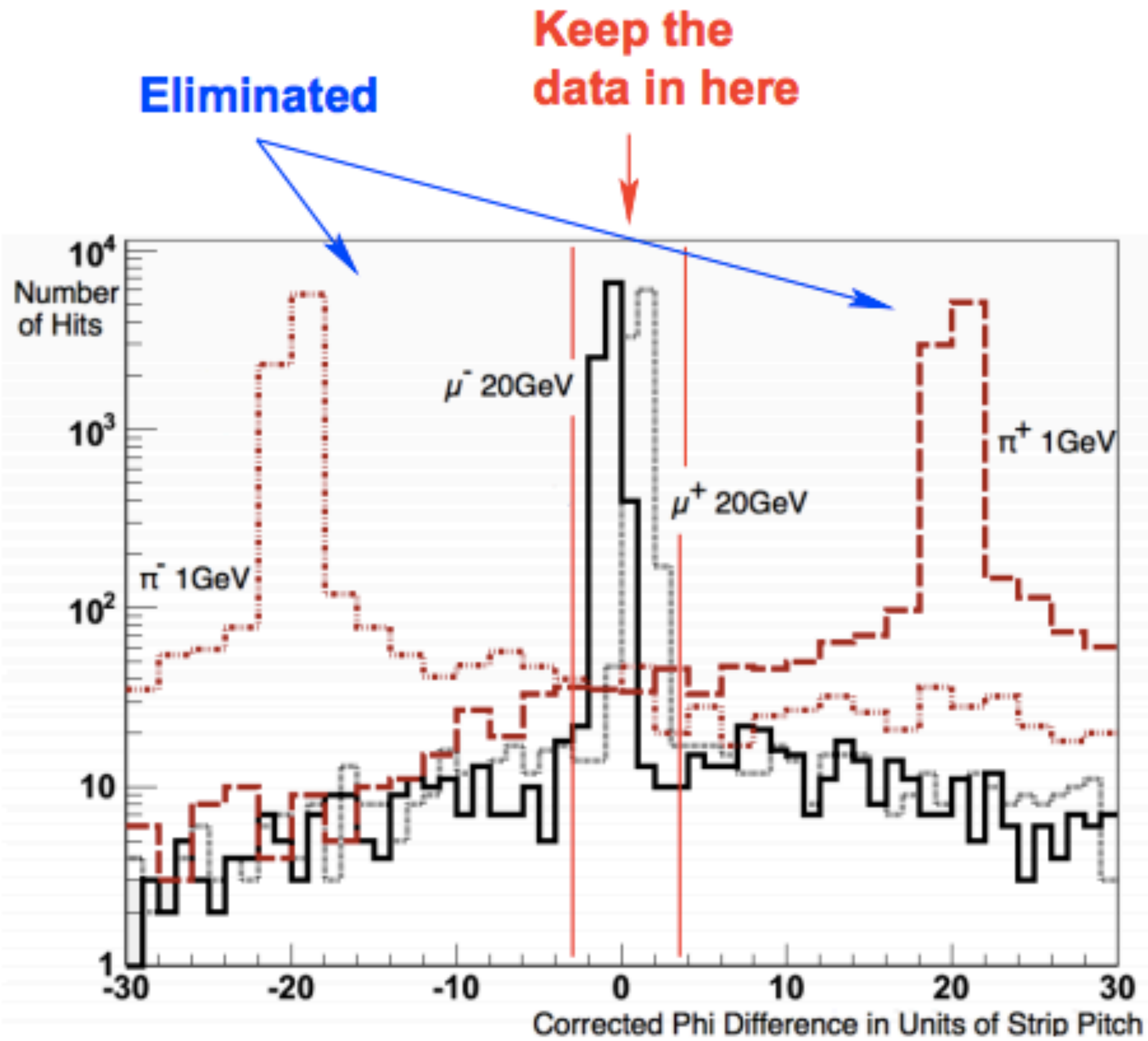
"Unseeded"/Doublet Method

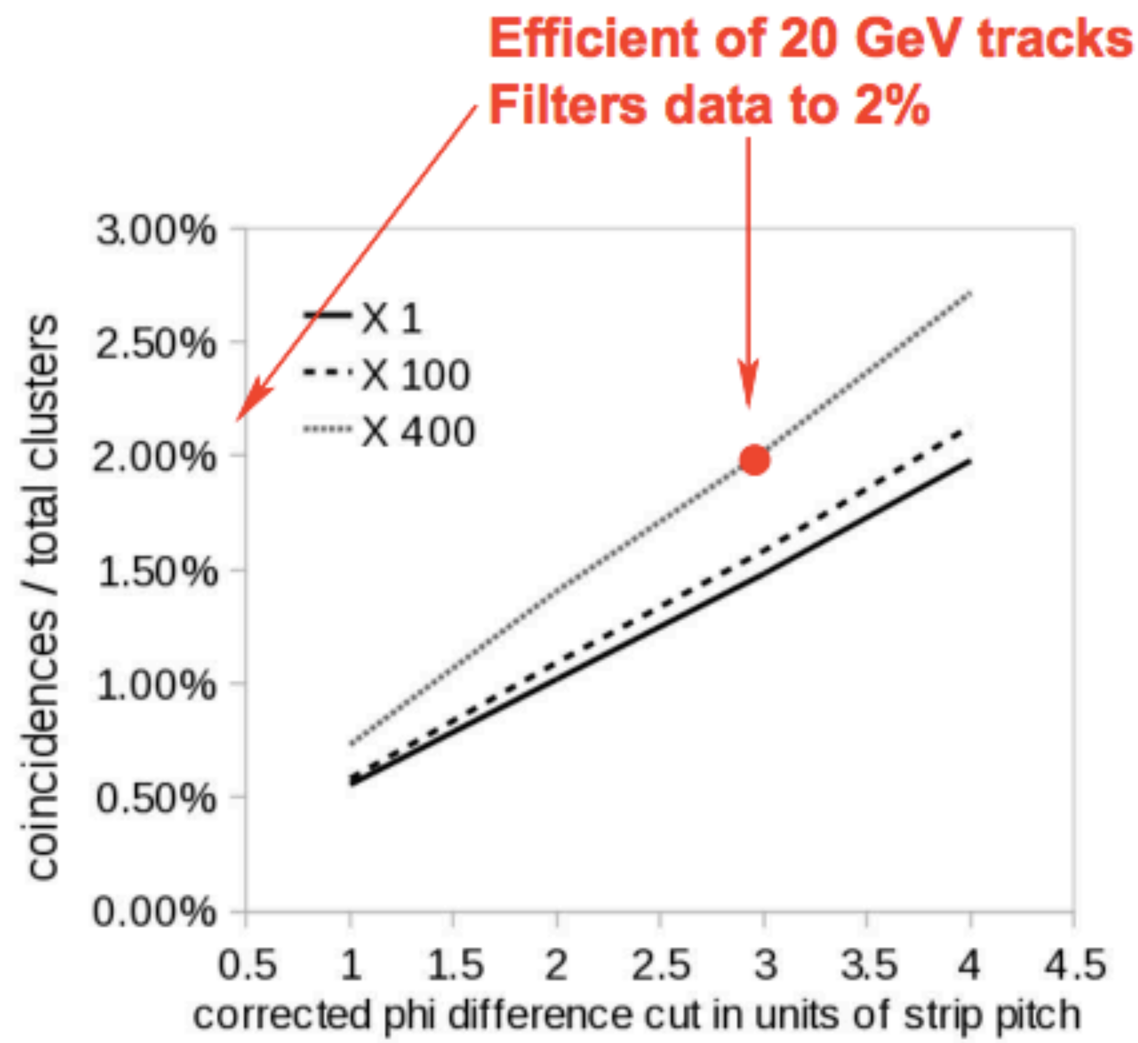


Other sources of doublet coincidences



Doublets: The data reduction

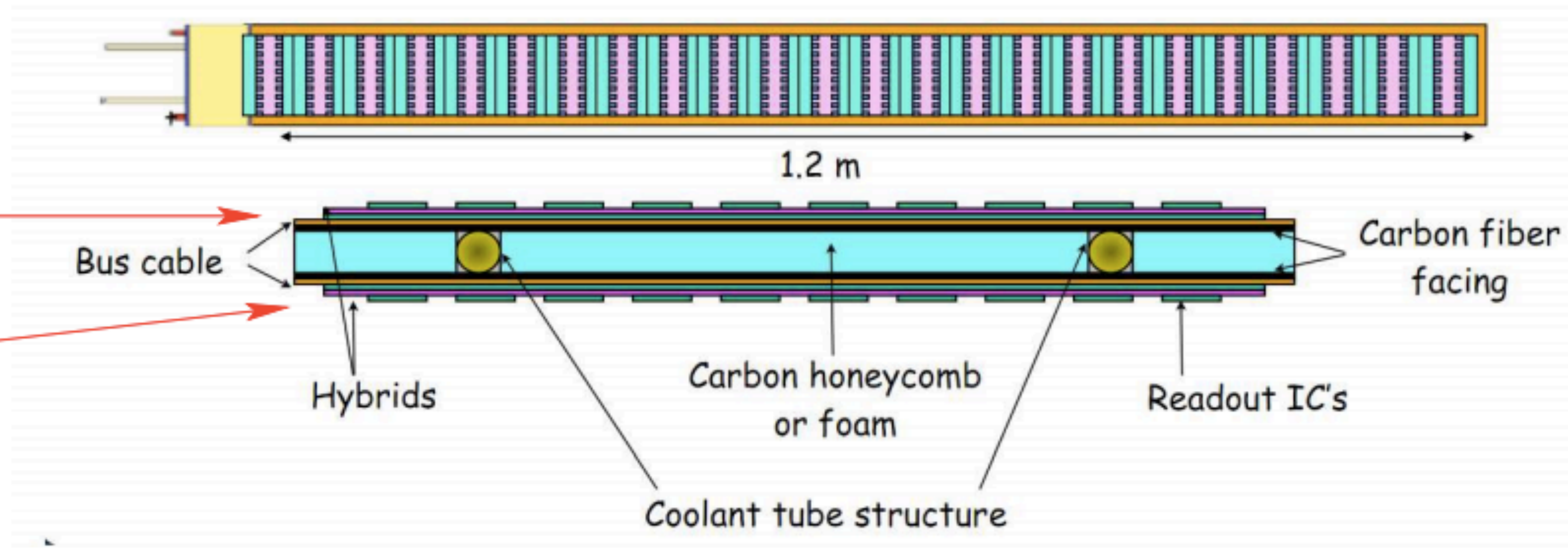




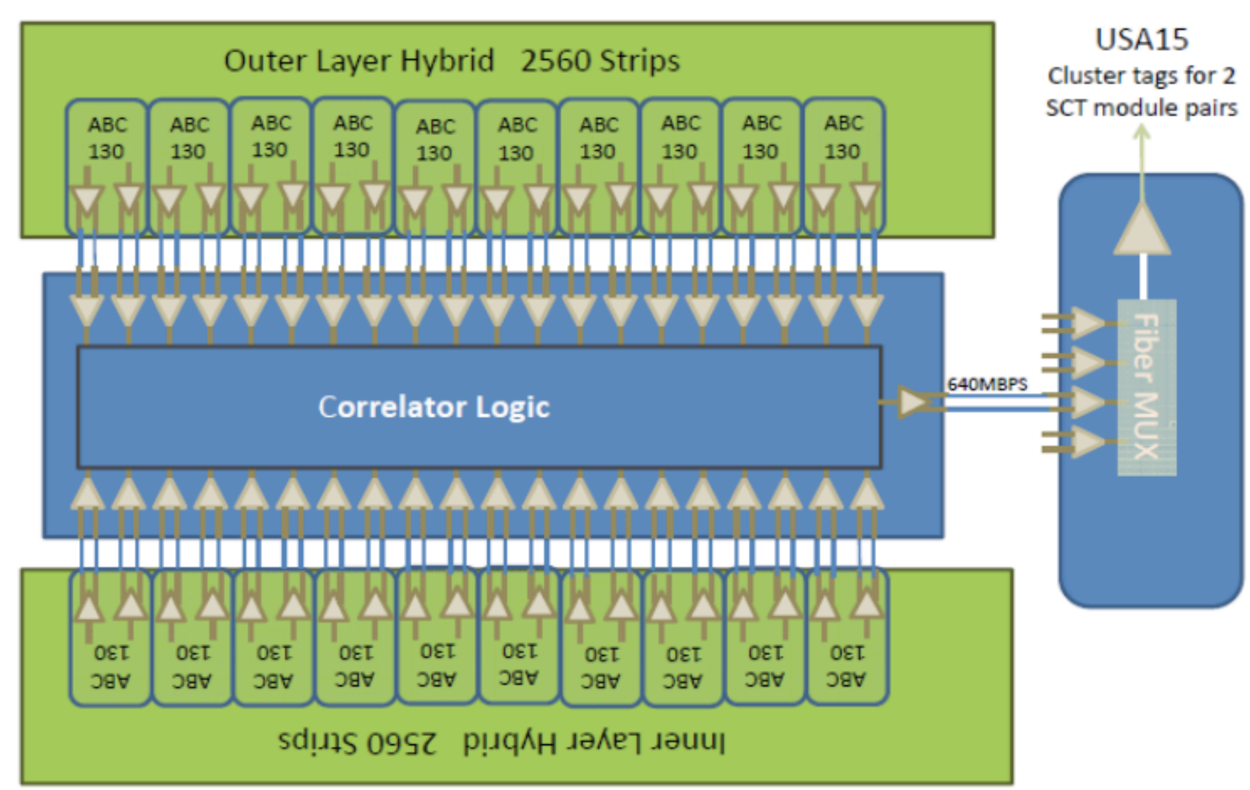
- Two-trigger layers at 0.8 m and 1.0 m have roughly double the readout rate as an offline only design
- Total bandwidth for outer layer with doublet readout is comparable to an inner layer without
- Must eliminate stereo angle for outer layers (impact not that serious)

Communication between the two-sides

Need to get from here to here



- Add wrap around cable with a high-speed serial interconnect for each 128 channels
- Add correlator chip for each ~10 cm module



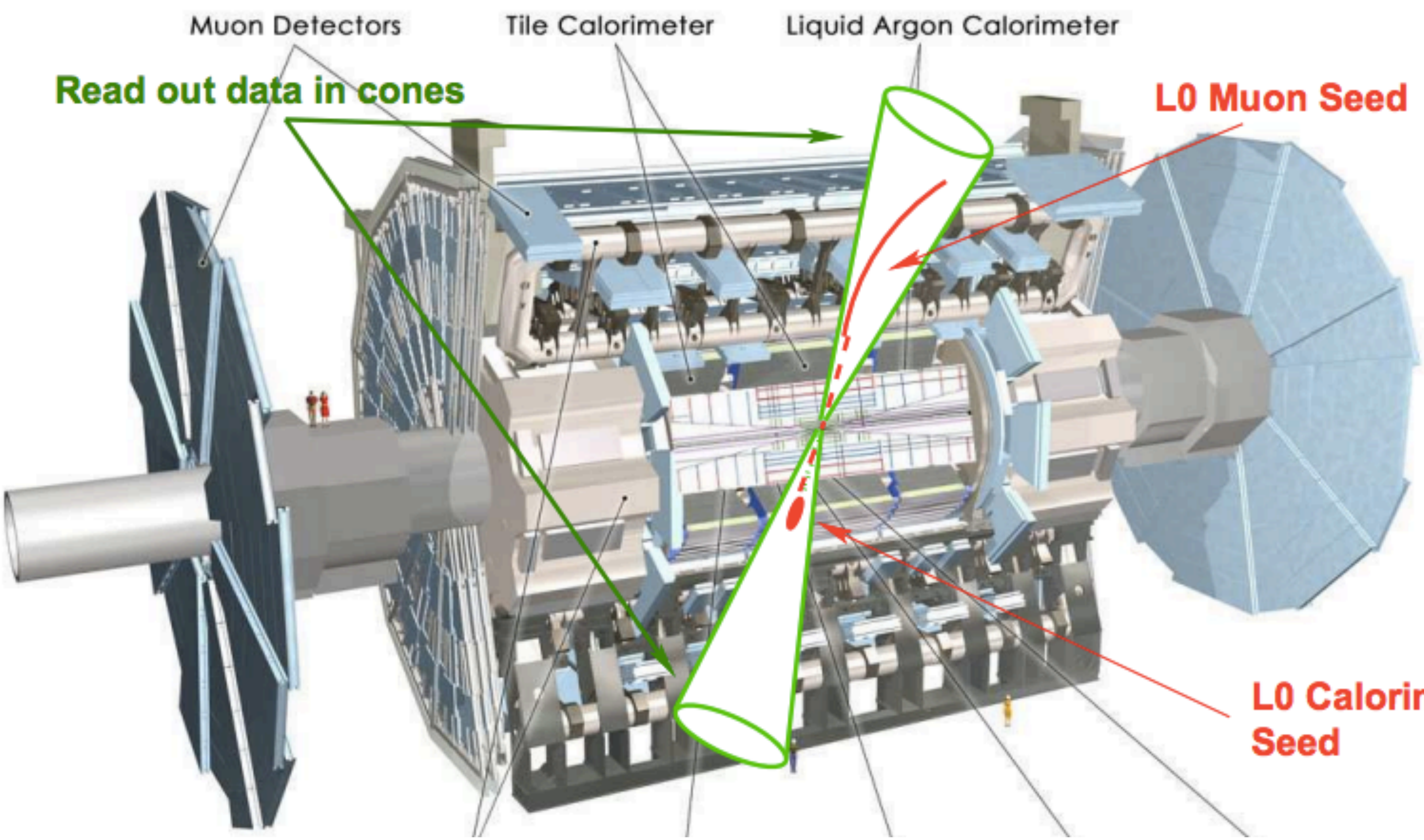
Filing on Region:

Two-level trigger, Pull method

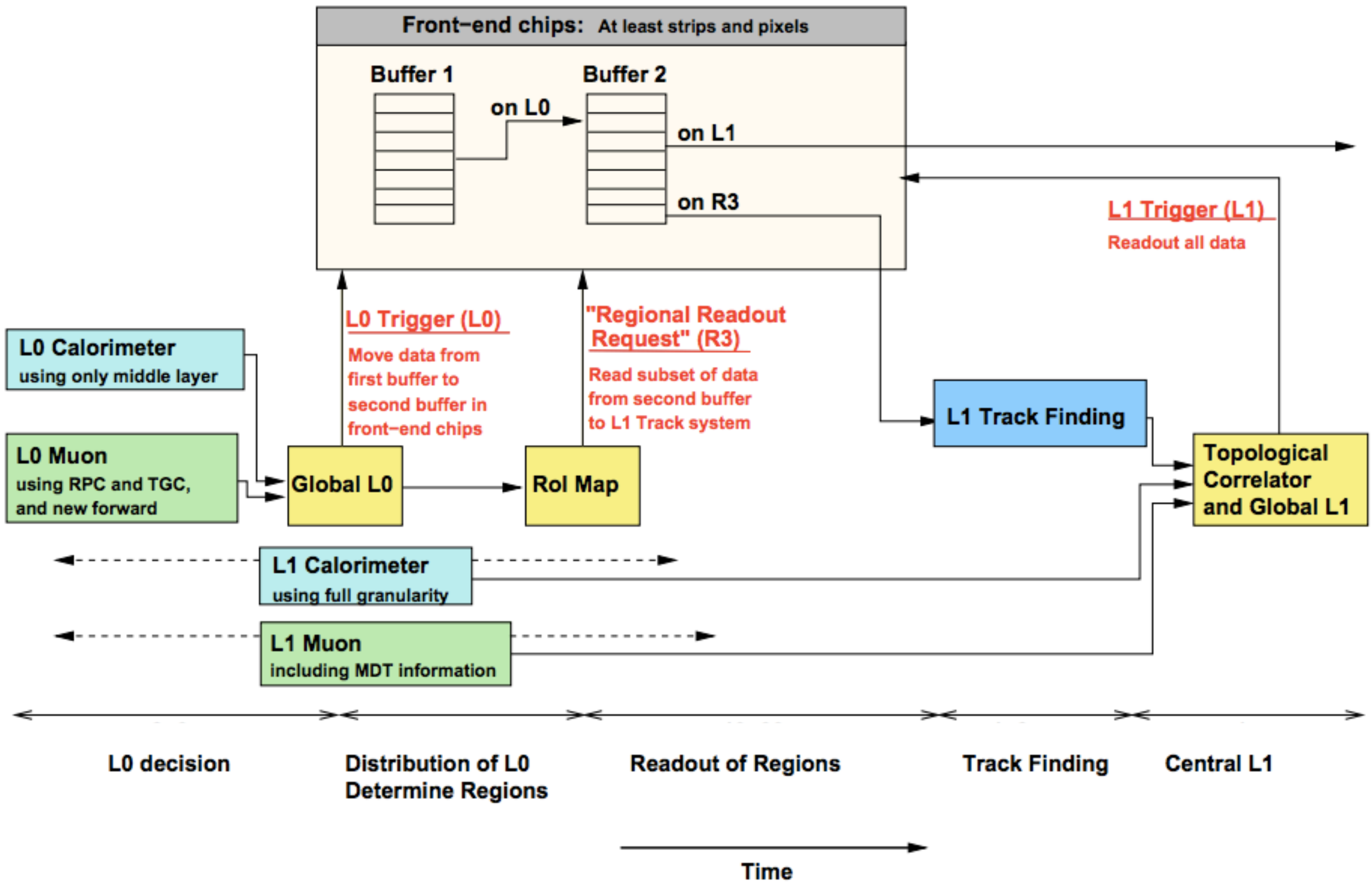
Reducing the data flow: Region method

Two-level trigger: L0 and L1

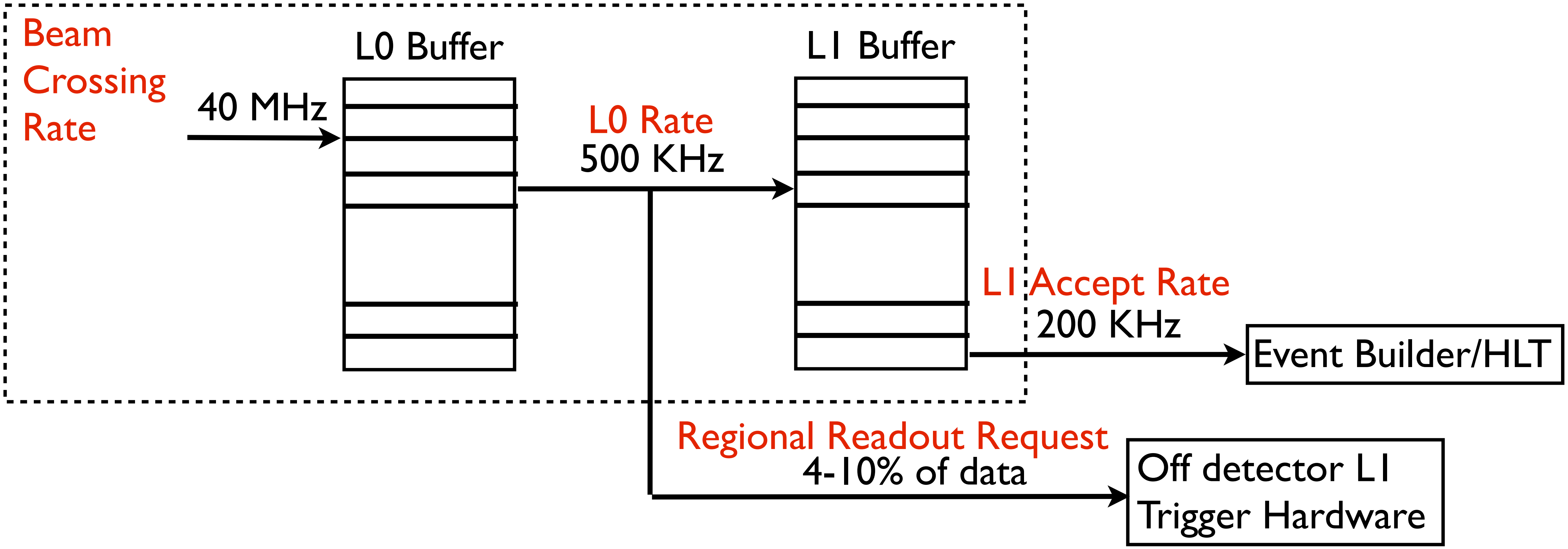
- **L0** uses calorimeter and muon system to define regions of interest (RoIs)
- **L1** extracts tracking for just RoIs from detector front-ends



Two-buffer scheme



Two-buffer scheme



Bandwidth = L1 Rate + L0 Rate × fraction of data in RoI

Nominal parameters:

L0 Rate = 500 KHz, L1 Rate = 200 KHz, RoI fraction = 10%

$$\text{L0 Latency} = \frac{\text{L0 Buffer length (in events)}}{\text{Beam Crossing rate}} \approx \frac{128}{40\text{MHz}} \approx 3.2\mu s$$

$$\text{L1 Latency} = \frac{\text{L1 Buffer length (in events)}}{\text{L0 Rate}} \approx \frac{128}{500\text{KHz}} \approx 256\mu s$$

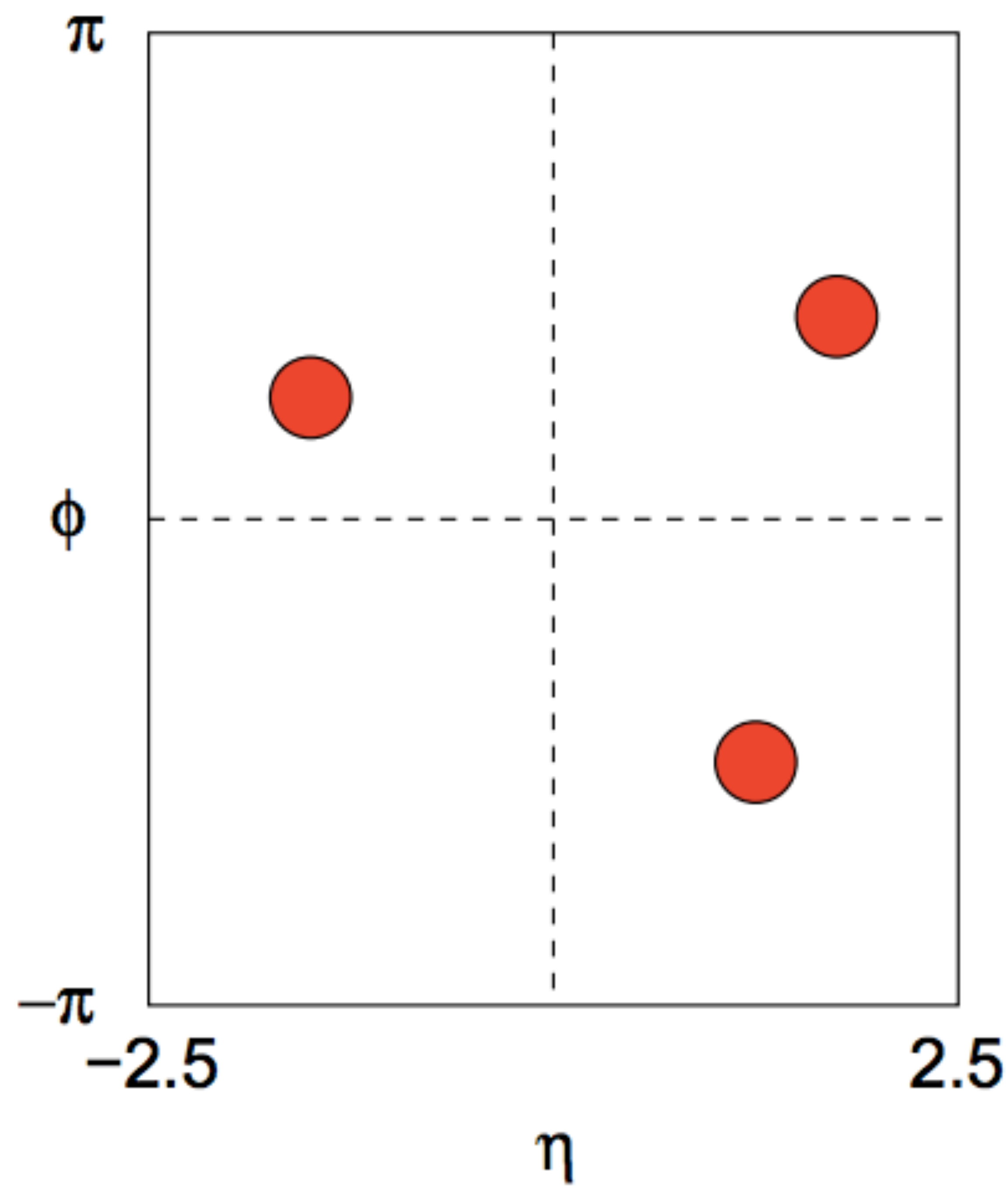
Data Reduction from Regions

Consider cones in $\eta - \phi$ space

- Typical cones size used for isolation are $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.2 - 0.4$
- Fractions of tracking volume in a cone of $\Delta R < r$ is

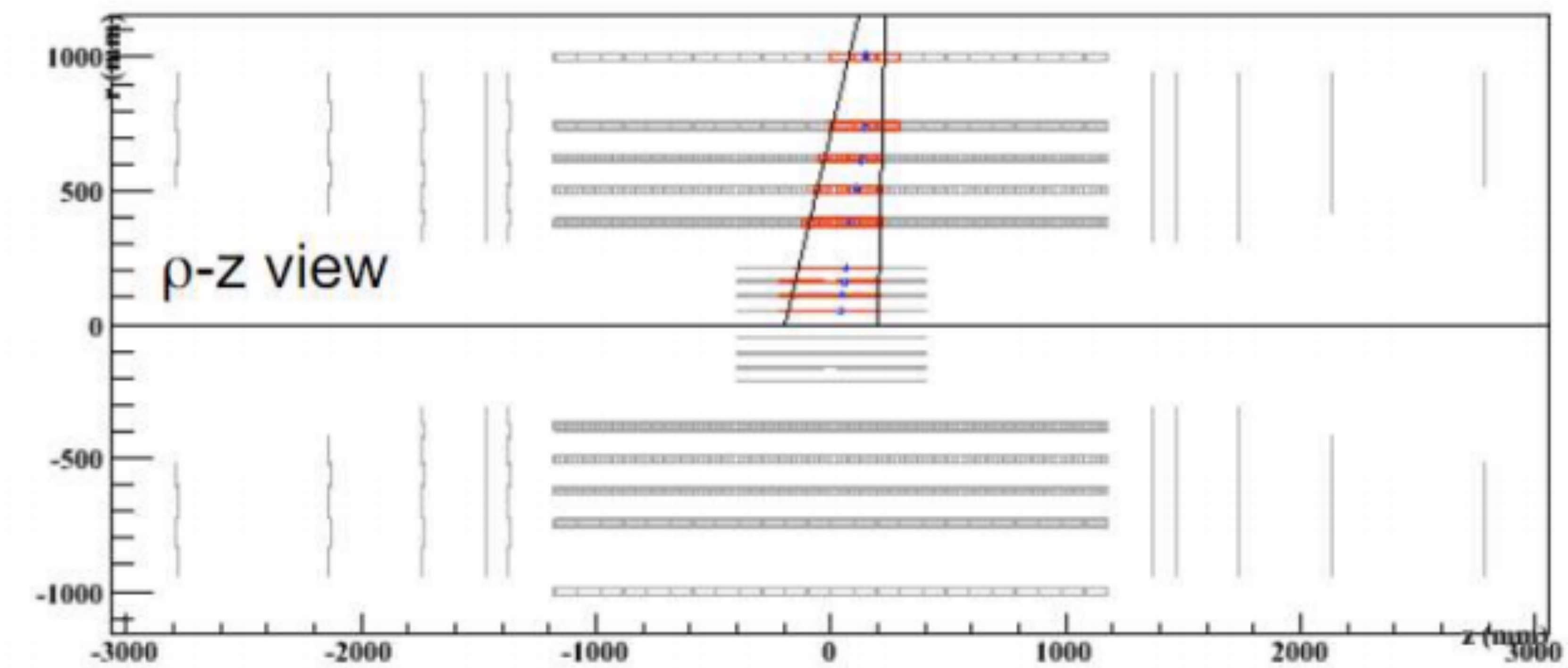
$$\frac{\pi r^2}{(\eta \text{ range}) \times (\phi \text{ range})}$$

- For a cone of $\Delta R < 0.2$ this is 0.4%
- This allows for a large number of RoIs and a safety margin to fit in 10% RoI request fraction



Data Reduction from Regions

A tricky challenge



•Because of beam spot spread, RoI need to be elongated along beam direction

•Large request rate for central wafers in inner pixel layers