Trigger Trends

Gerhard Raven
VU VRIJE UNIVERSITEIT AMSTERDAM NIKHEF LHCb

TIPP '14

International Conference on Technology and Instrumentation in Particle Physics

2 – 6 June 2014 / Amsterdam, The Netherlands

"Instrumentation as enabler of Science"

2007

Keynote

Some Quotes

- "Apple is already dead.”
    - Nathan Myhrvold, former Microsoft CTO, 1997

- “There is no reason anyone would want a computer in their home.”
    - Ken Olsen, Founder of DEC, 1977

- “Prediction is very difficult, especially about the future.”
    - Niels Bohr, Danish physicist (1885 - 1962)

Yet Another Prediction

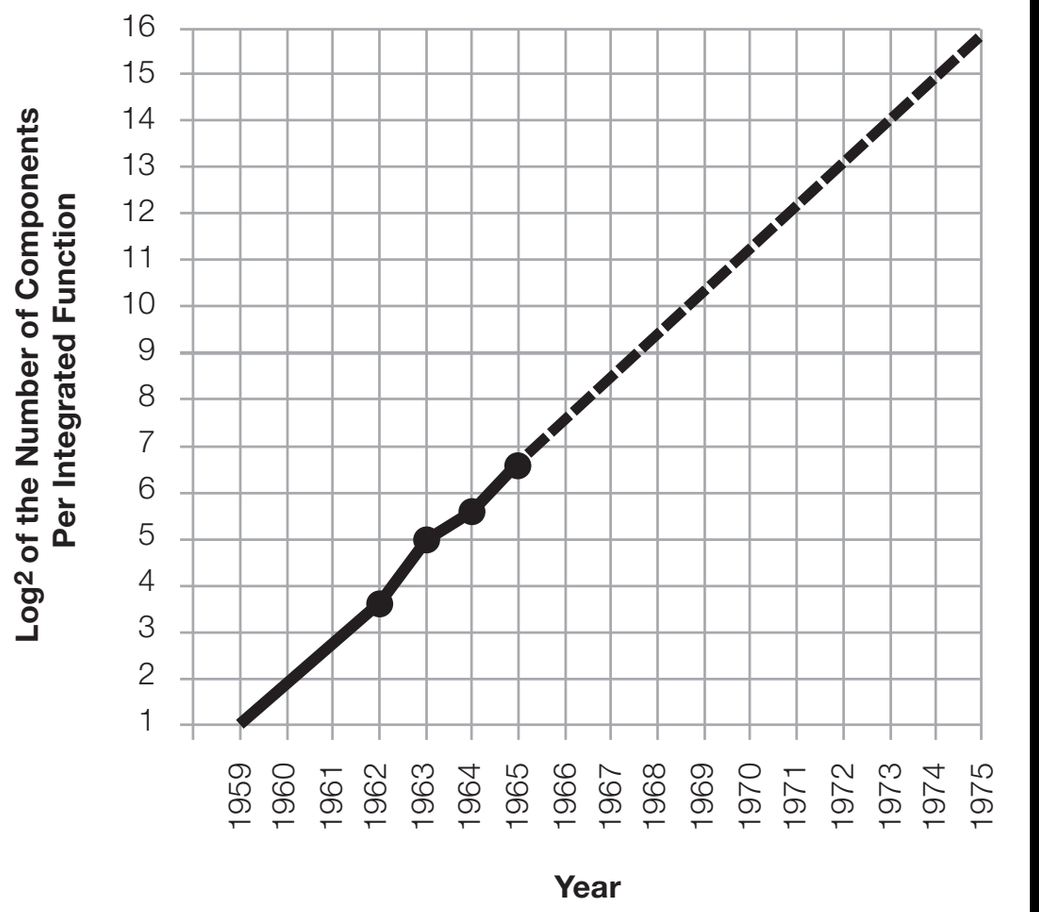
Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

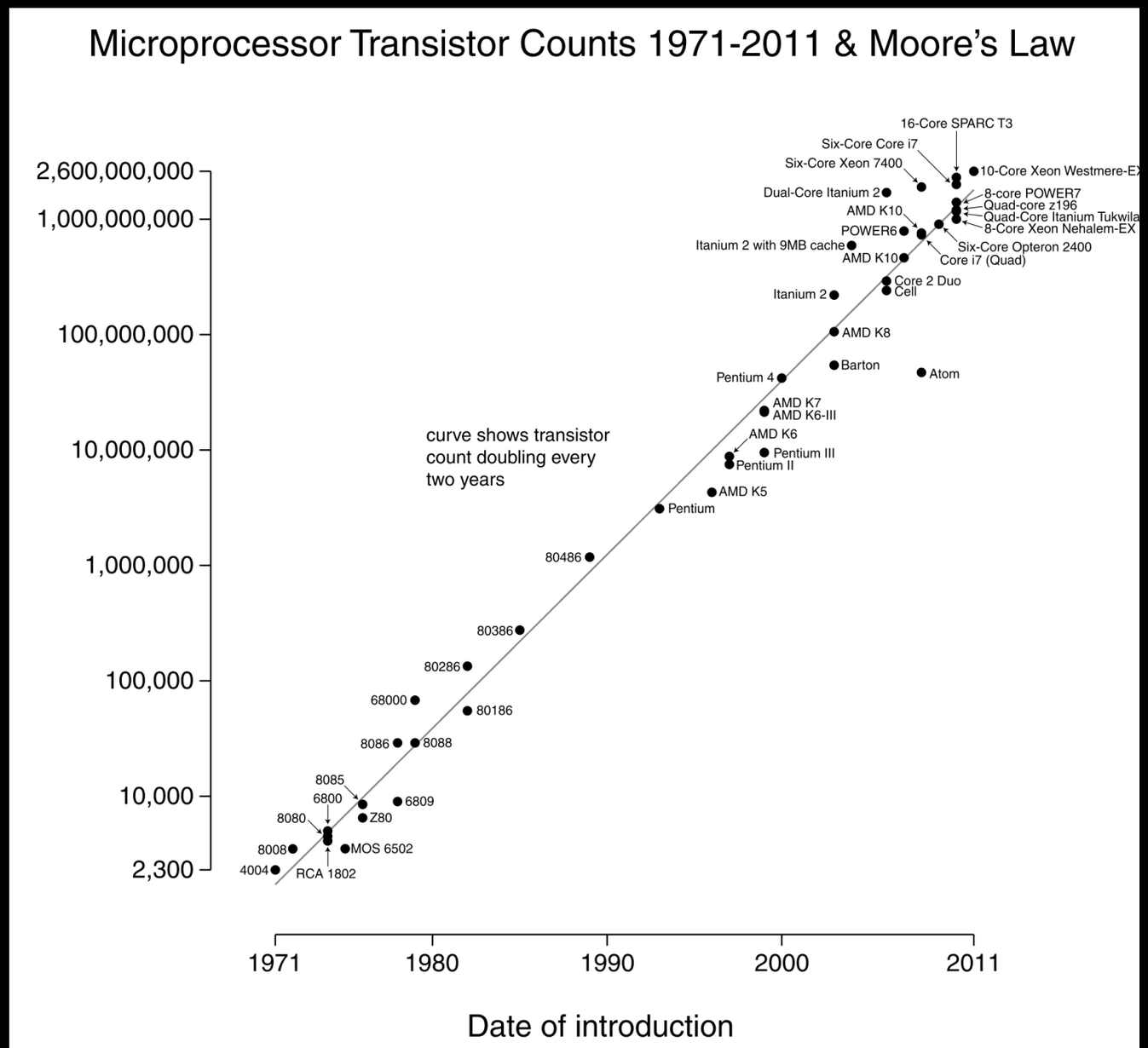
By Gordon E. Moore

Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

Electronics, Volume 38, Number 8, April 19, 1965



Yet Another Prediction



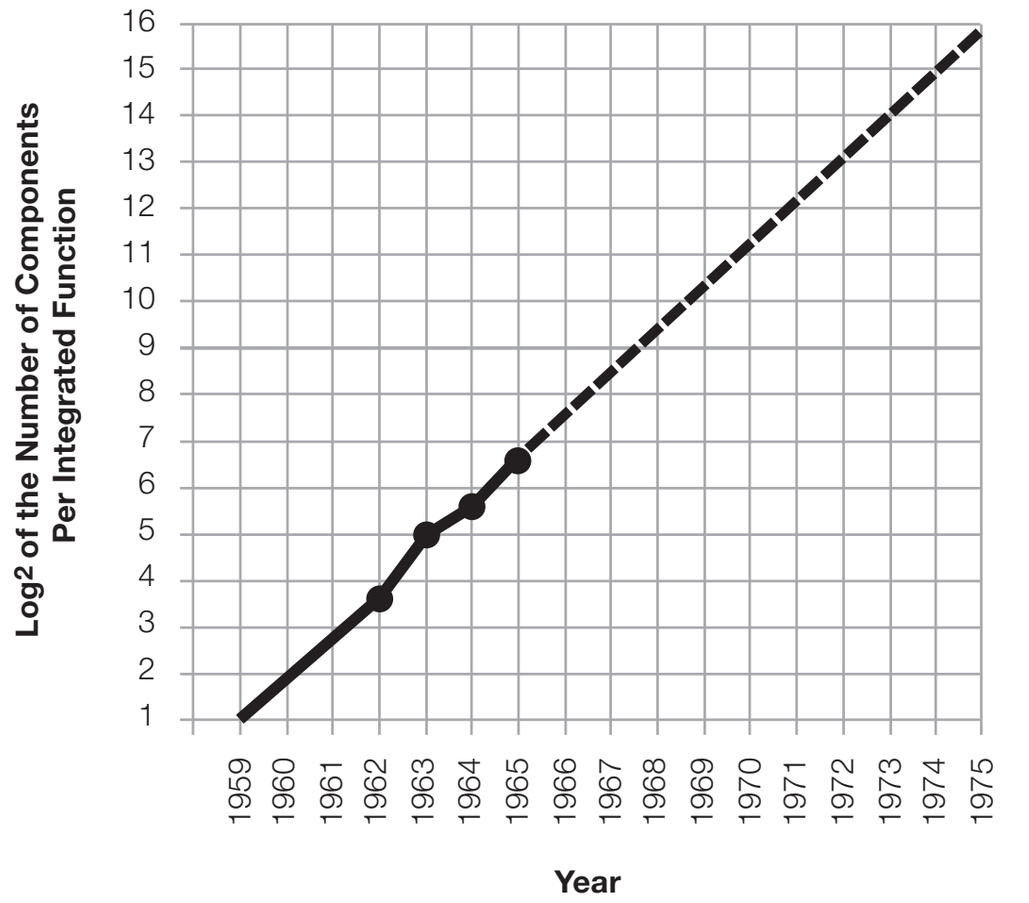
Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

By Gordon E. Moore

Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

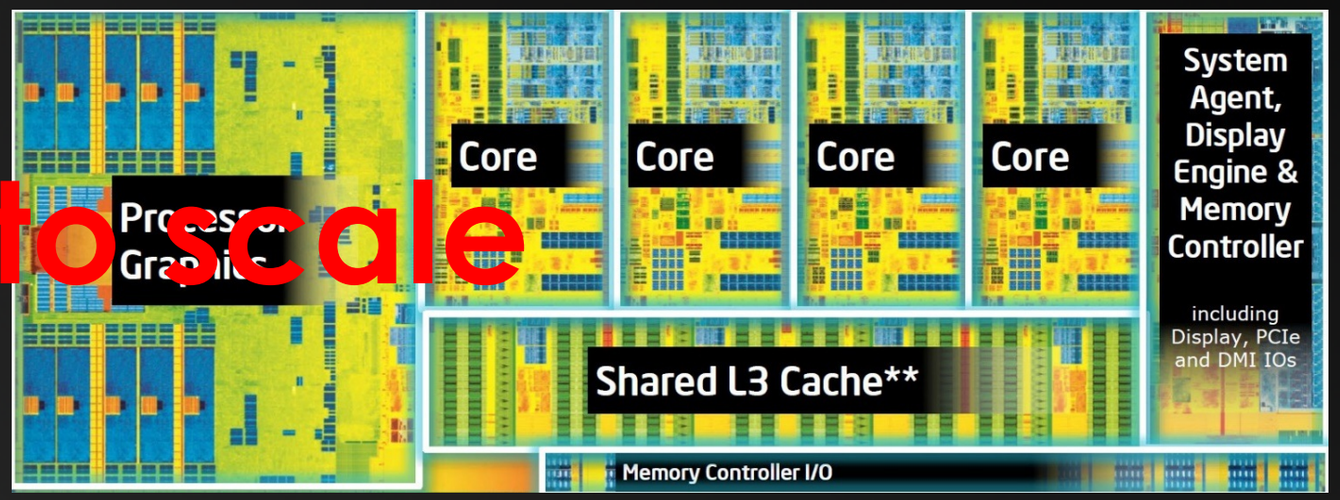
Electronics, Volume 38, Number 8, April 19, 1965



HARDWARE LANDSCAPE

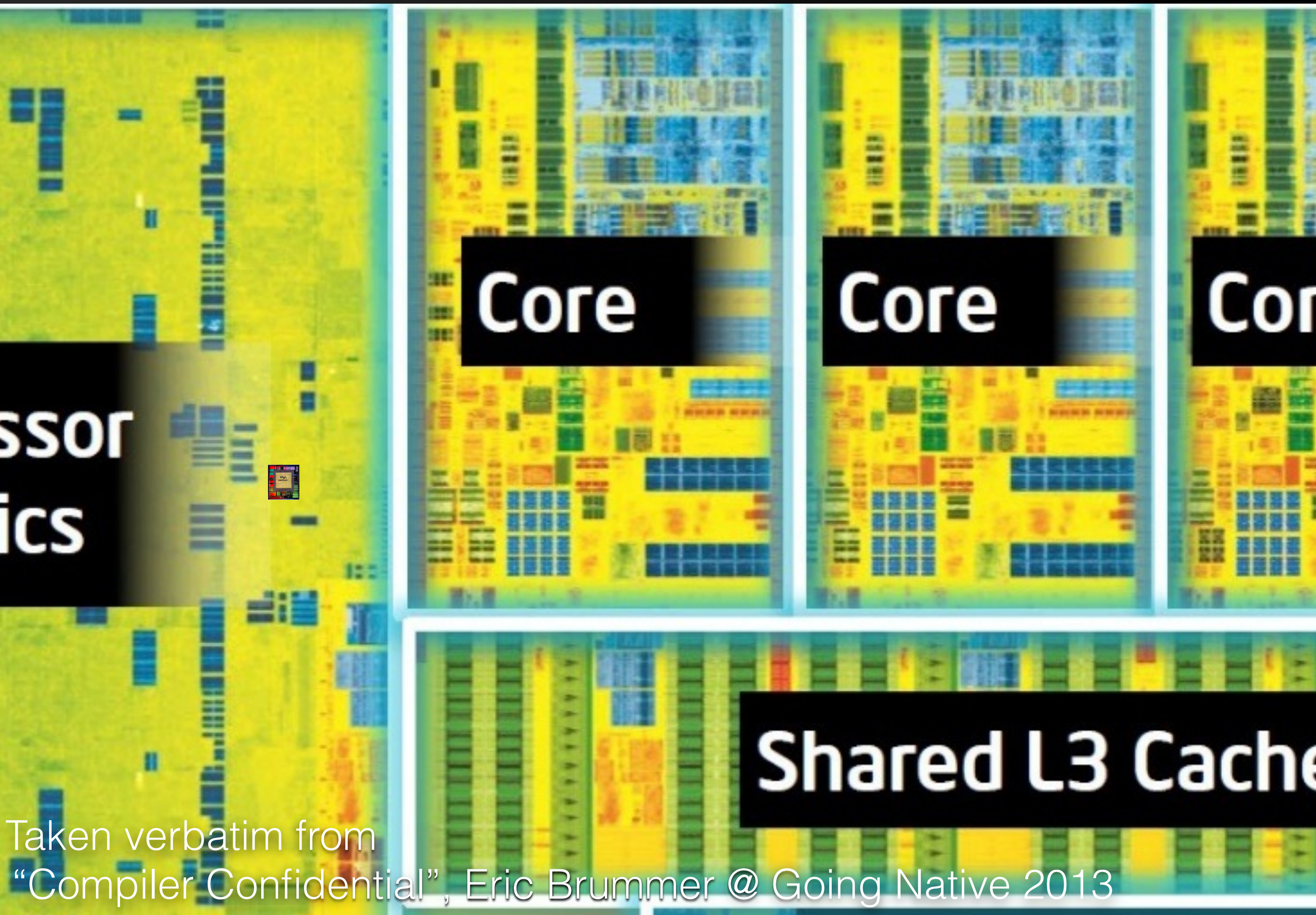
"Yesterday"
3.1 million transistors

Today
1.4 billion transistors



Not to scale

Taken verbatim from
"Compiler Confidential", Eric Brummer @ Going Native 2013



Taken verbatim from
"Compiler Confidential", Eric Brummer @ Going Native 2013

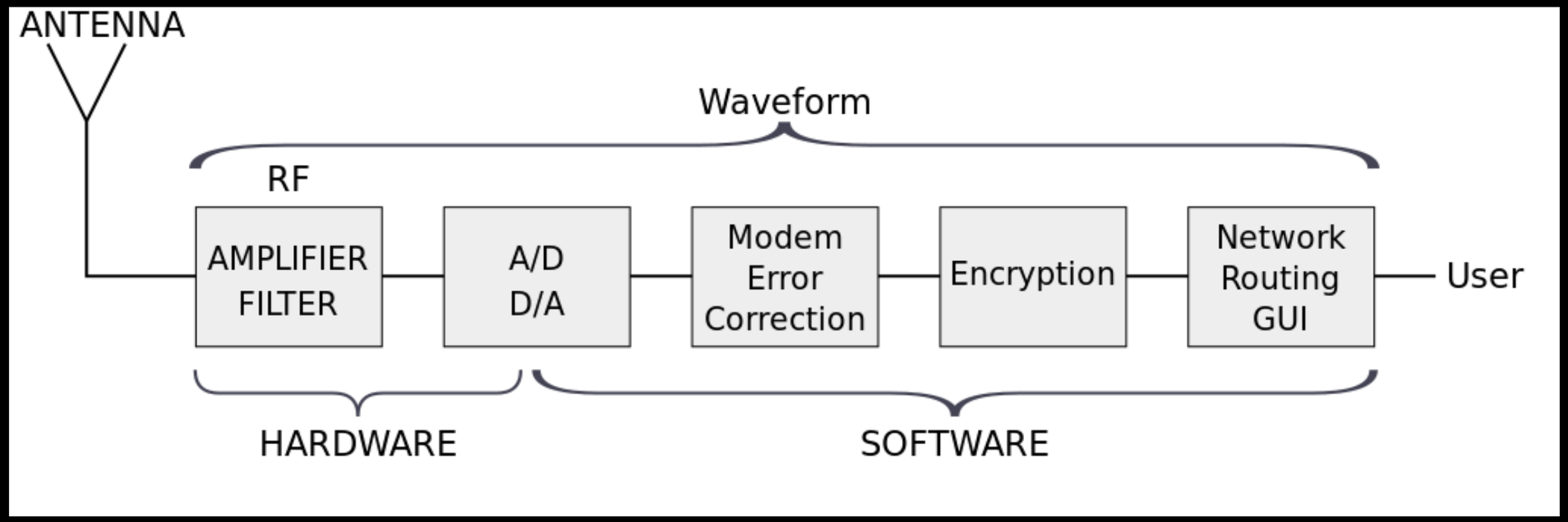

Moto Q
BlackBerry
Palm Treo
Nokia E62

iPhone premise:

Avoid single-purpose hardware *when software can do the job*

Footnote:
Apple is (now) big enough to *define* ‘commodity’ and optimize their ARM SoCs to ‘match’ hardware and software even better

Software Defined Radio

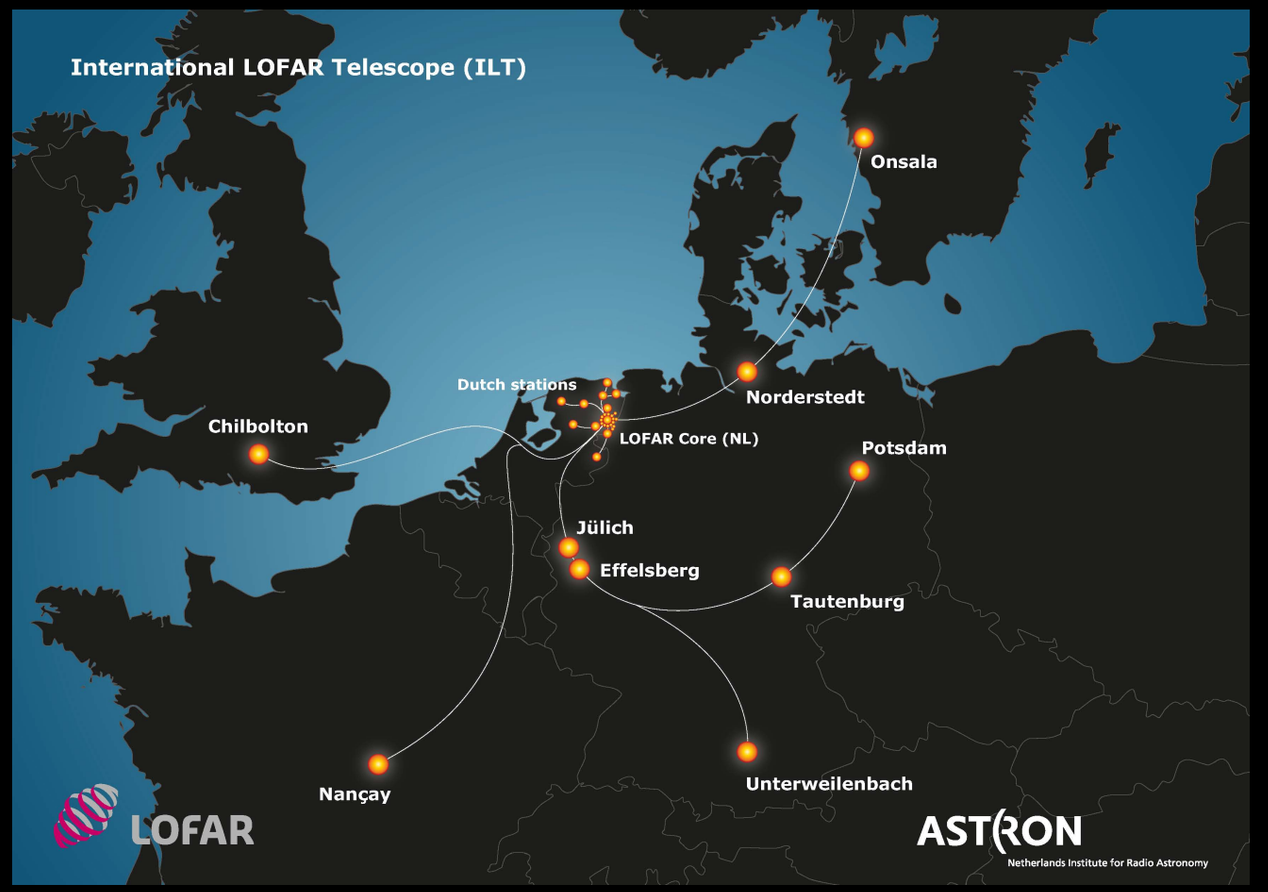


Wikipedia: ”The ideal receiver scheme would be to attach an analog-to-digital converter to an antenna”

“The ideal scheme is not completely realizable due to the limits of technology. The main problem is the difficulty of conversion between digital and analog domains at a high enough rate and accuracy at the same time.”

LOFAR

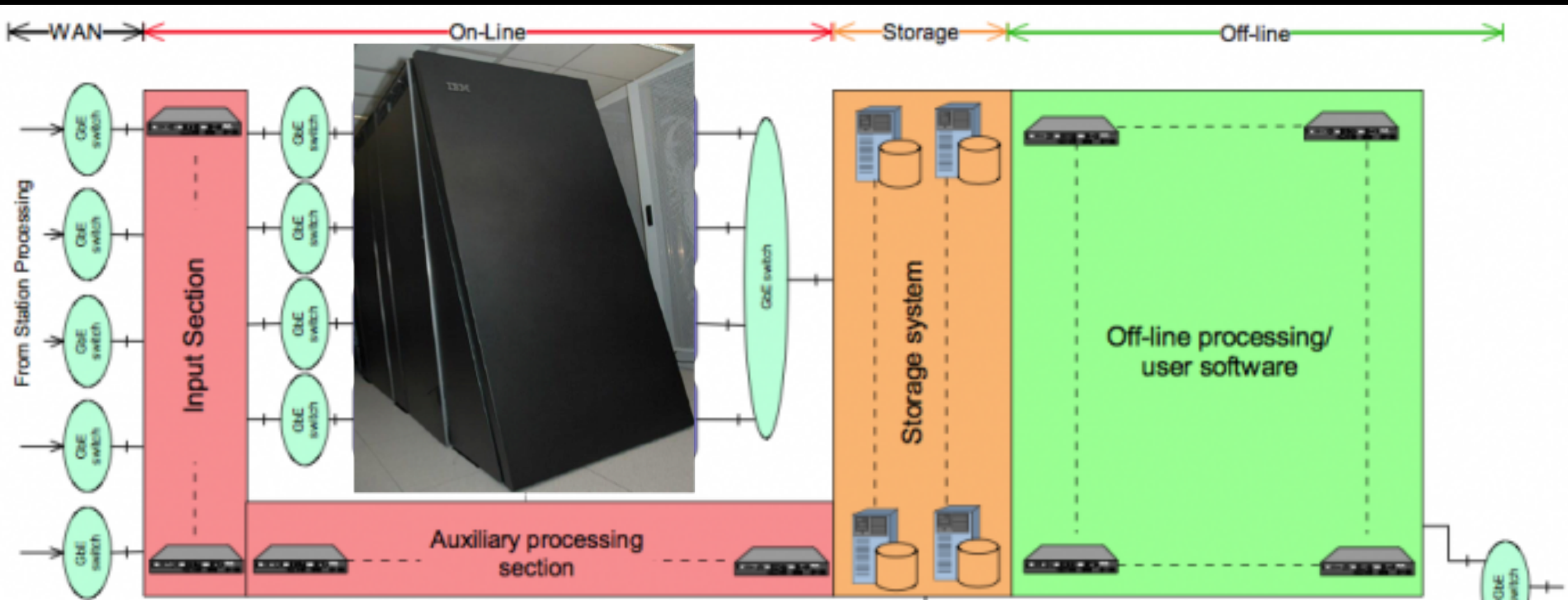
- Software Interferometer
- Fully operational since Dec 2012
- Interesting “connection” to last paragraph of Gordon Moore’s paper:
  - “The successful realization of such items as phased-array antennas, …, could completely revolutionize radar”




Low-band 30-80 MHz antennas

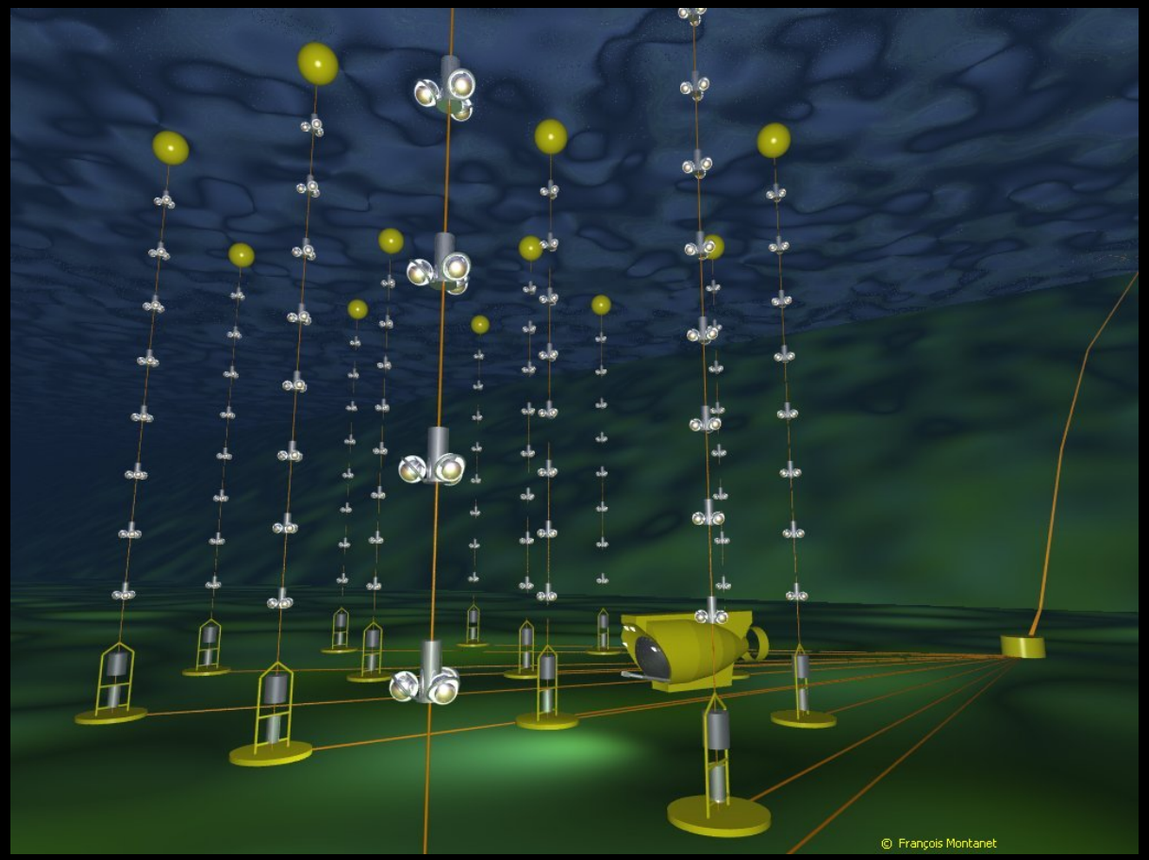
LOFAR

- Data rates: 2 - 20 Gbit/s dependent on antenna location.
- Total input rate: 2 Tbit/s
- WAN:10 Gigabit Ethernet (10 GbE) “This is the obvious choice for distances up to ~40 km with a single transmitter/receiver pair”.
- “A major advantage of the LOFAR Central Processor is that multiple applications and modes can be run concurrently”



Antares

- Neutrino telescope in the Mediterranean Sea
- Around 2000 decide to go triggerless — “all data to shore”
- Data rate: 0.3 - 0.5 Gbit/s



Panda: Triggerless DAQ

- Panda LOI (2004!):
  - “Key technologies … within the DAQ framework are high speed serial (10 GBit/s per link and beyond) and high-density FPGA …”
- 100 Gbit/s

International Conference on Computing in High Energy and Nuclear Physics 2012 (CHEP2012) IOP Publishing
Journal of Physics: Conference Series **396** (2012) 012027 doi:10.1088/1742-6596/396/1/012027

Modeling event building architecture for the triggerless data acquisition system for PANDA experiment at the HESR facility at FAIR/GSI

K Korcyl[1,2], I Konorov[3], W Kühn[4] and L Schmitt[5]

[1] Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
[2] Cracow University of Technology, Kraków, Poland
[3] Technische Universitaet Munchen, Physics Department, Garching, Germany
[4] Justus-Liebig-Universität, Giessen, Germany
[5] GSI, Darmstadt, Germany

E-mail: Krzysztof.Korcyl@ifj.edu.pl

Abstract. A novel architecture is being proposed for the data acquisition and trigger system of the PANDA experiment at the HESR facility at FAIR/GSI. The experiment will run without hardware trigger signal using timestamps to correlate detector data from a given time window. The broad physics program in combination with the high rate of $2 * 10^7$ interactions per second requires very selective filtering algorithms accessing information from many detectors. Therefore the effective filtering will happen later than in today's systems ie. after the event building. To assess that, the complete architecture will be built of two stages: the data concentrator stage providing event building and the rate reduction stage. For the former stage, which requires a throughput of 100 GB/s to perform event building, we propose two layers of ATCA crates filled with Compute Nodes - modules designed at IHEP and University of Giessen for trigger and data acquisition systems. Currently each board is equipped with 5 Virtex4 FX60 FPGAs and high bandwidth connectivity is provided by 8 front panel RocketIO ports and 12 backplane ports for the inter-module communication.

CERN/LHCC 2014-016
LHCb TDR 16
27th May 2014

LHCb Trigger and Online Upgrade Technical Design Report

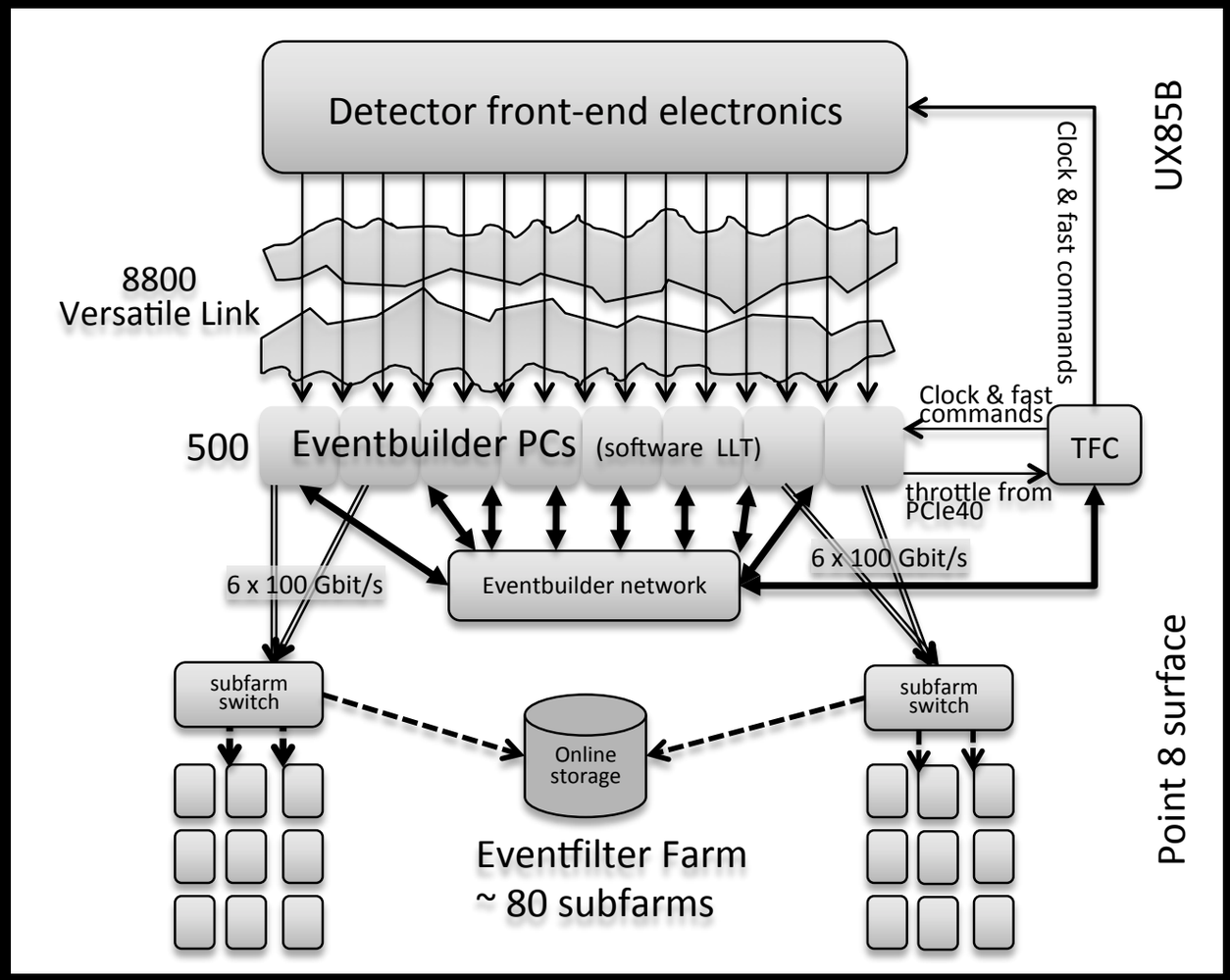
The LHCb collaboration

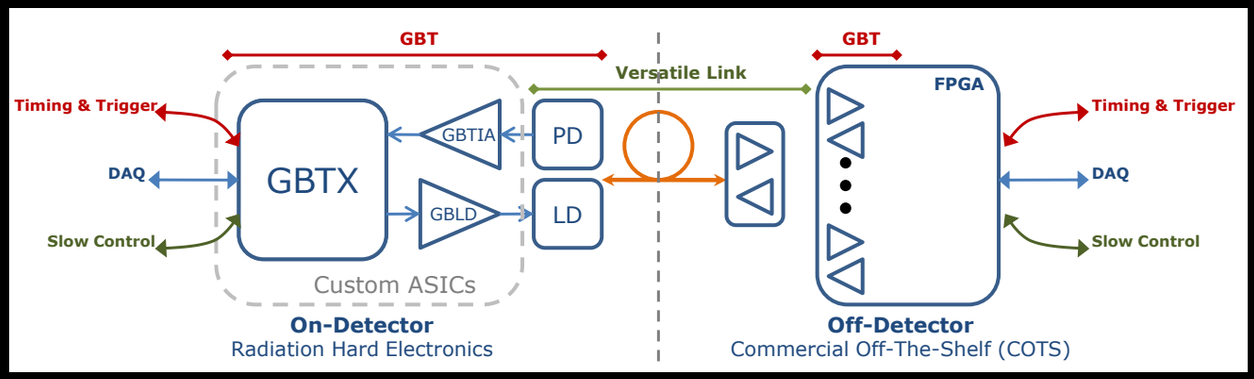
Abstract

The LHCb experiment will be upgraded between 2018 and 2019 in order to reach unprecedented precision on the main observable of the b and c-quark sectors. This Technical Design Report addresses the *trigger-less readout system* and the *full software trigger* features.

Event Building @ 40 MHz

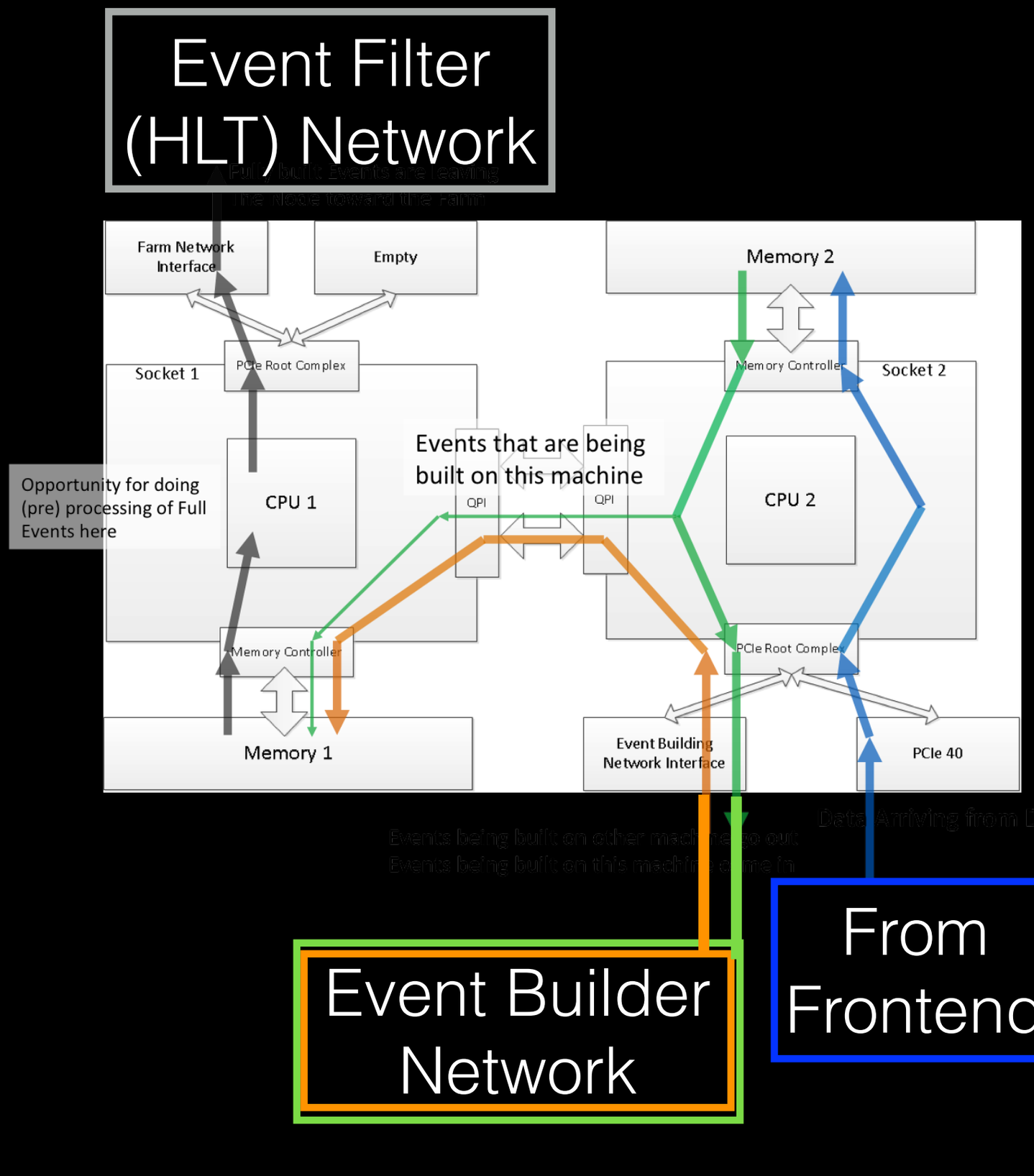
- 32 Tbit/s
- “All data to the surface”
- Decouple front-end electronics from event builder network
  - Frontend→GBT→PCIe
  - GBT: Rad-hard, integrated into front-end, so no commodity solution possible…
- Buffering in PC memory





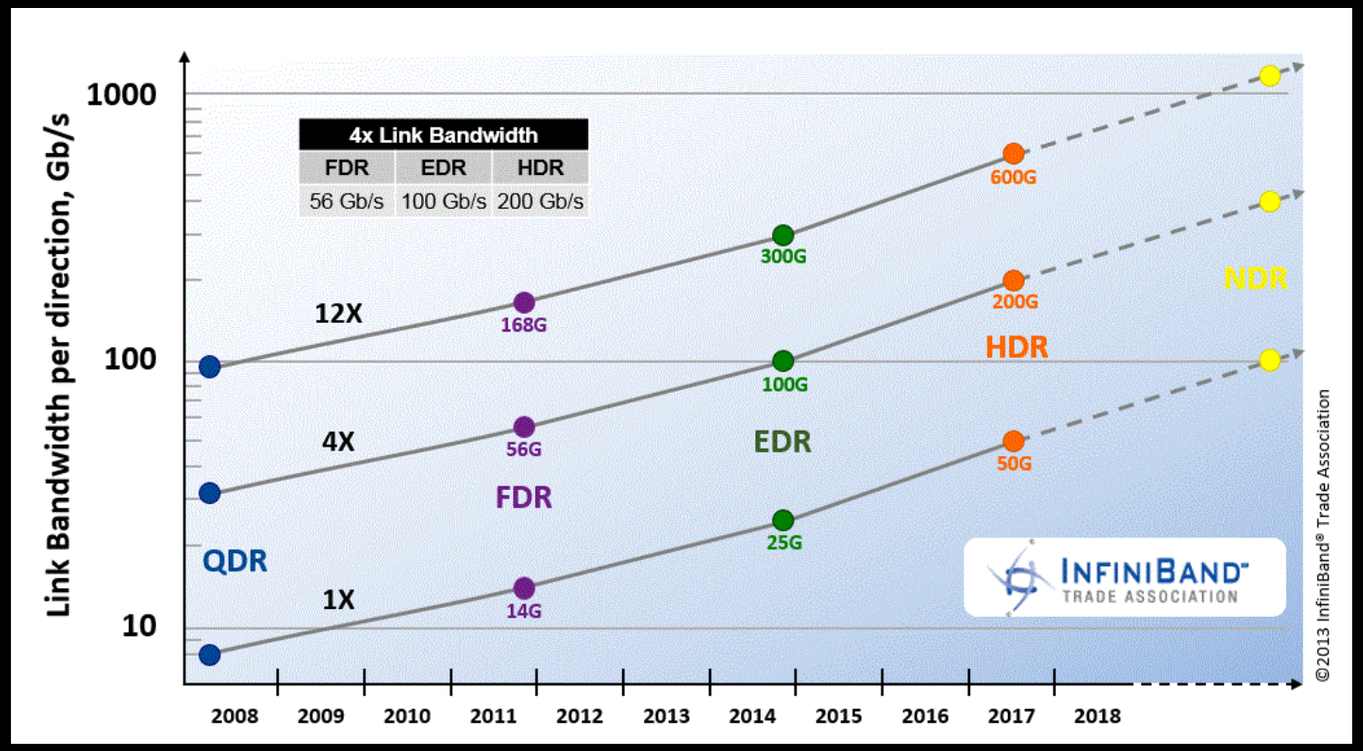
Event Building @ 40 MHz

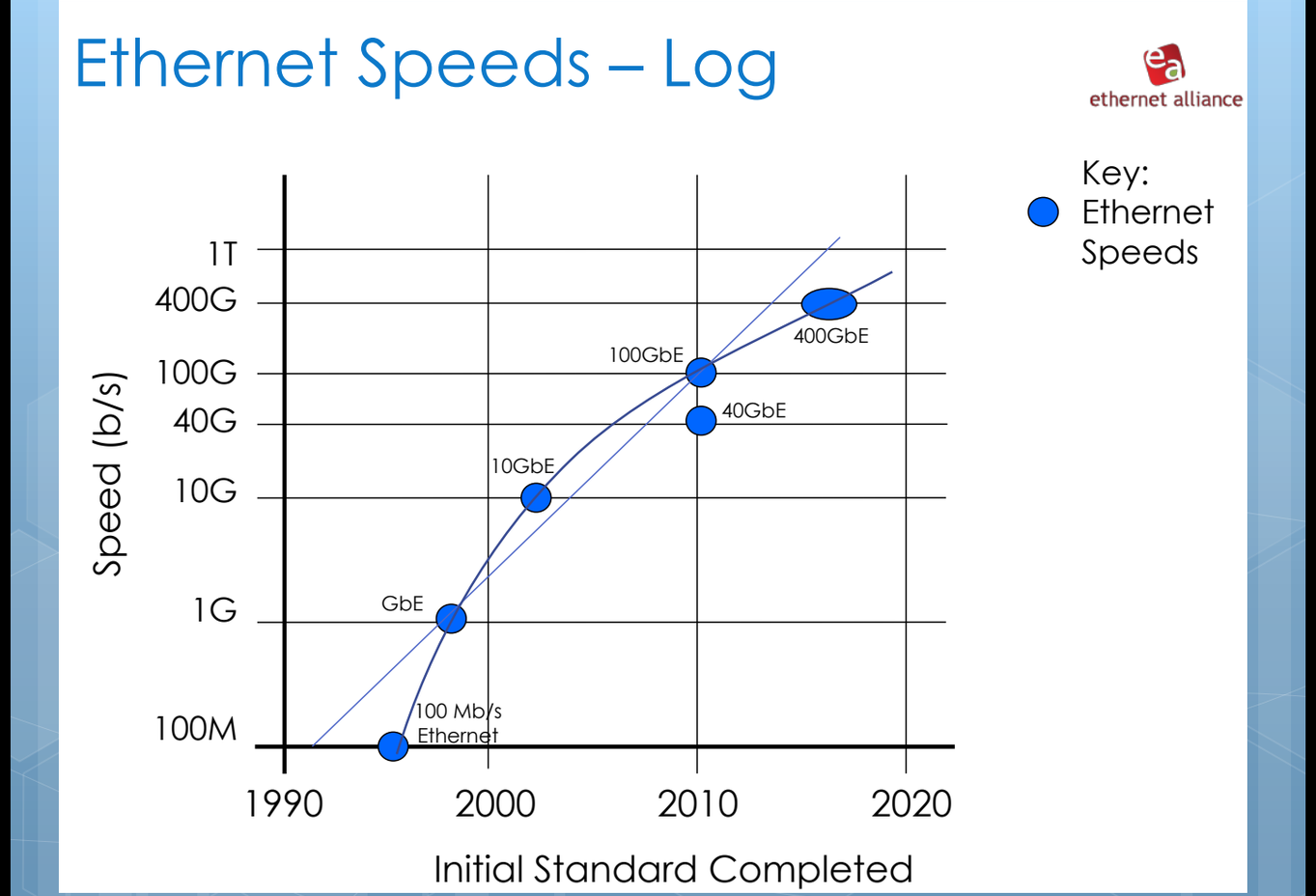
- "COTS" as soon as possible
- O(500) servers for event building
- "Data Center" ("thin" switch, Infiniband/Ethernet/ …) instead of "Telecom" (ATCA, "fat" switch)
- Event Filter: O(1000) servers



Event Building @ 40 MHz

- 32 Tbit/s
- O(500) servers for event building
- “Data Center” (“thin” switch, Infiniband/ Ethernet/…) instead of “Telecom” (ATCA, “fat” switch)
- Event Filter: O(1000) servers





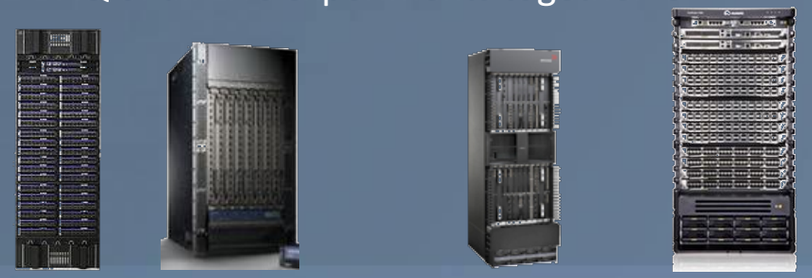
Event Building @ 40 MHz

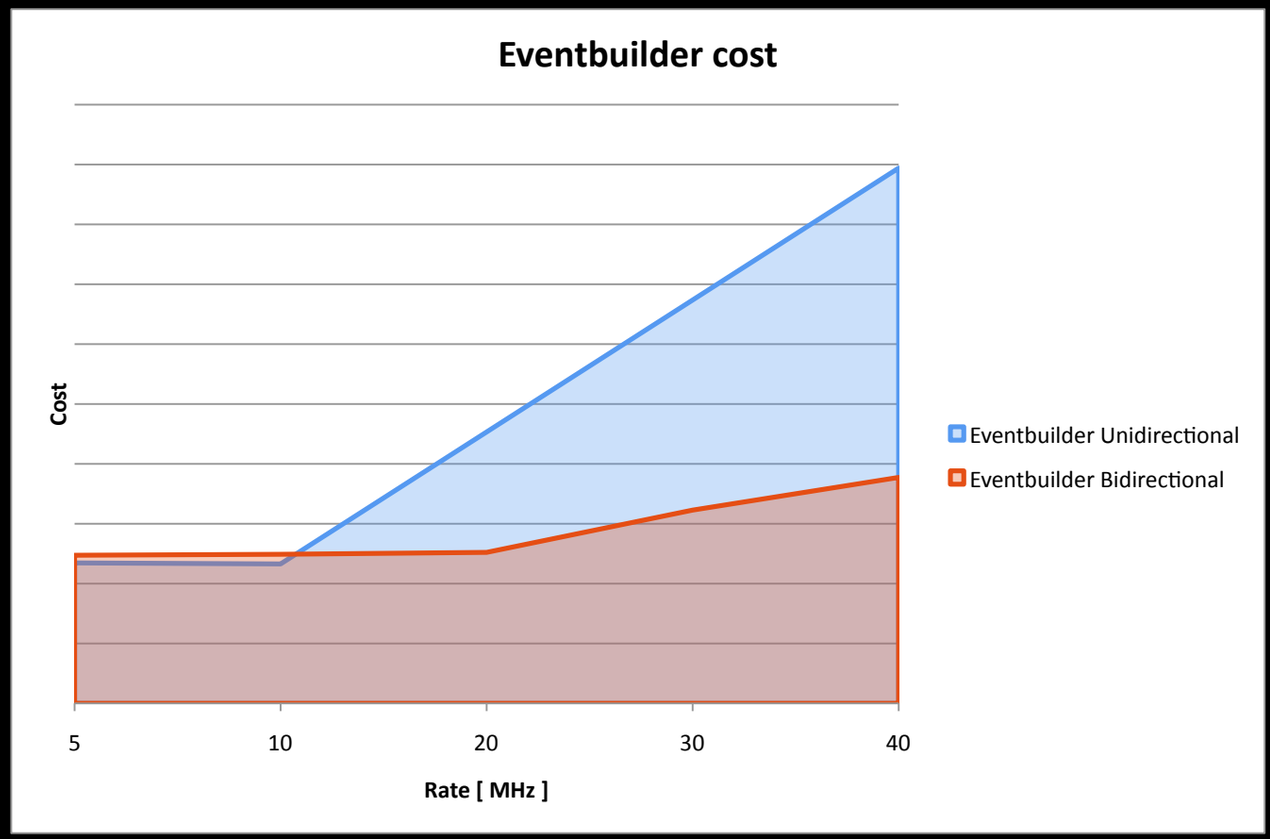
- 32 Tbit/s
- O(500) servers for event building
- “Data Center” (“thin” switch, Infiniband/ Ethernet/…) instead of “Telecom” (ATCA, “fat” switch)
- Event Filter: O(1000) servers

The evolution of switches

Date of release

- Brocade MLX: 768 10-GigE
- Juniper QFabric: up to 6144 10-GigE
- Mellanox SX6536: 648 x 56 Gb (IB) / 40 GbE ports
- Huawei CE12800: 288 x 40 GbE / 1152 x 10 GbE
- Each with sufficient bandwidth to run the entire Run #2 DAQ of all LHC experiments together ☺

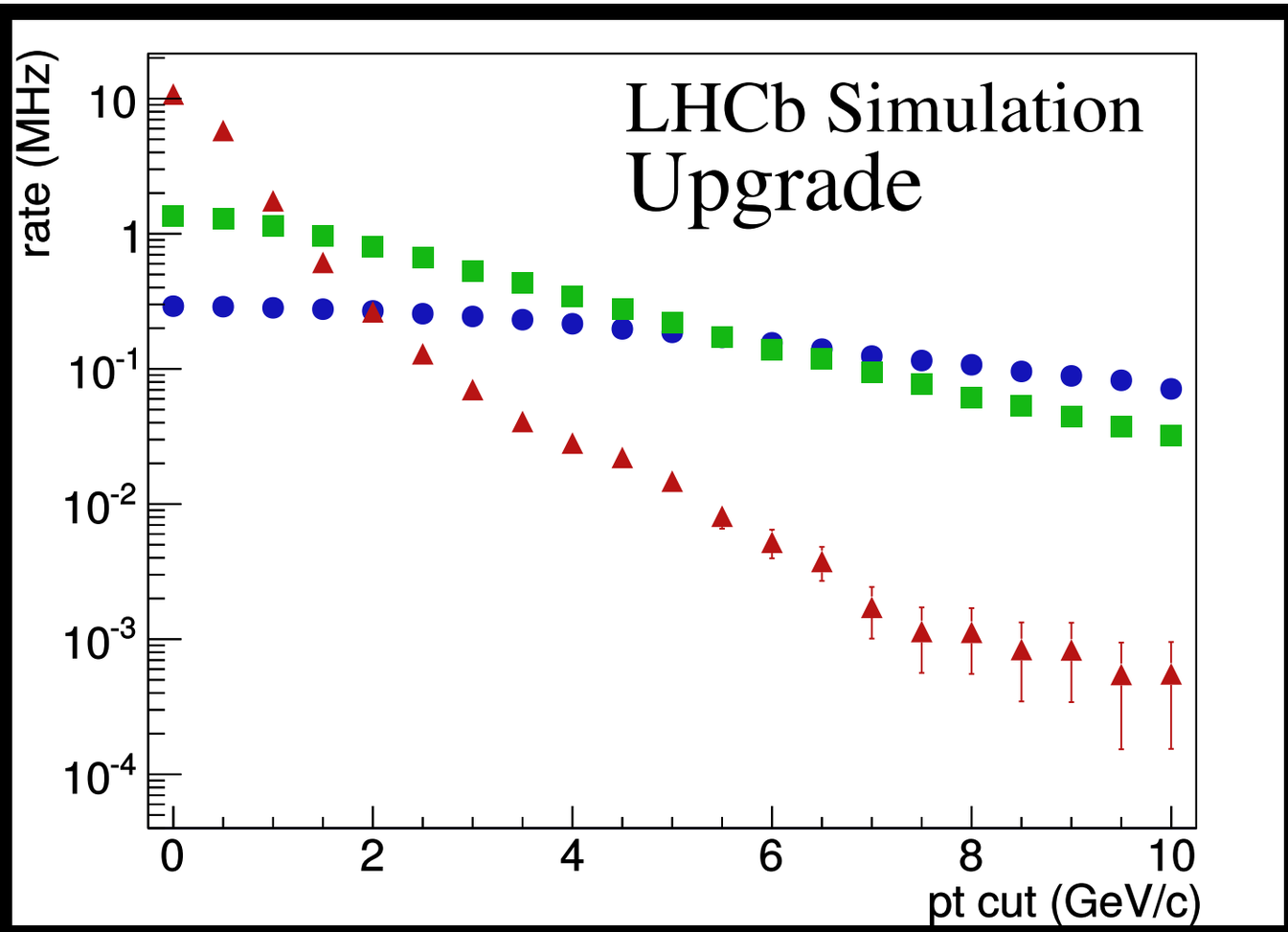




Why go “triggerless"?

- Both Panda and LHCb upgrade share a “challenge”:
  - High “signal” rate
  - *By construction*: “high-precision” experiments
- LHCb upgrade: 1 / 5 beam crossings contains “reconstructible charm”
- The problem is “classification of signal” not “rejection of background”

Run I	Per event	with vertex in VELO	Rate [GB/s]
b-hadrons	0.0258 ± 0.0004	0.0029 ± 0.0001	0.9
c-hadrons	0.297 ± 0.001	0.0422 ± 0.0005	3.3
light, long-lived hadrons	8.04 ± 0.01	0.511 ± 0.002	1.1
Upgrade	**Per event**	**with vertex in VELO**	**Rate [GB/s]**
b-hadrons	0.1572 ± 0.0004	0.01874 ± 0.0001	27
c-hadrons	1.422 ± 0.001	0.2138 ± 0.0005	80
light, long-lived hadrons	33.291 ± 0.006	2.084 ± 0.001	26



Why go “triggerless"?

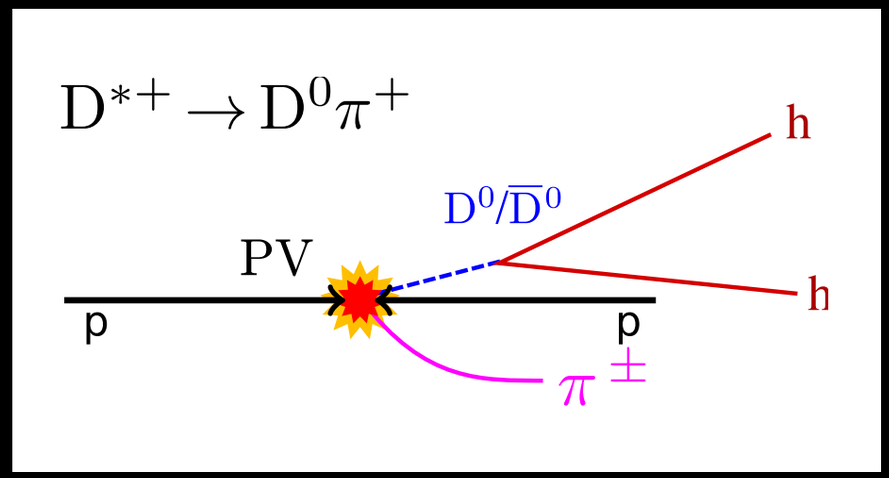
- Both Panda and LHCb upgrade share a “challenge”:
  - High “signal” rate
  - *By construction*: “high-precision” experiments

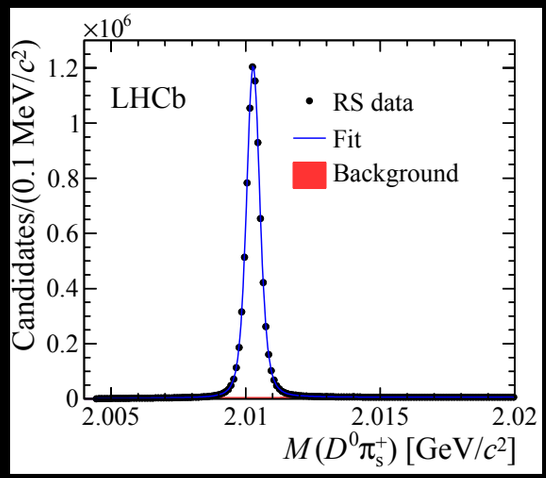
Run I	Per event	with vertex in VELO	Rate [GB/s]
b-hadrons	0.0258 ± 0.0004	0.0029 ± 0.0001	0.9
c-hadrons	0.297 ± 0.001	0.0422 ± 0.0005	3.3
light, long-lived hadrons	8.04 ± 0.01	0.511 ± 0.002	1.1
Upgrade	**Per event**	**with vertex in VELO**	**Rate [GB/s]**
b-hadrons	0.1572 ± 0.0004	0.01874 ± 0.0001	27
c-hadrons	1.422 ± 0.001	0.2138 ± 0.0005	80
light, long-lived hadrons	33.291 ± 0.006	2.084 ± 0.001	26

- LHCb upgrade: 1 / 5 beam crossings contains “reconstructible charm”
- The problem is “classification of signal” not “rejection of background”

boring, large rate:
$D^{*+} \rightarrow D^0(K^-\pi^+)\pi^+$

interesting, “doubly Cabibbo” suppressed by factor ~250:
$D^{*+} \rightarrow D^0(\pi^-K^+)\pi^+$





Why go “triggerless"?

- Both Panda and LHCb upgrade share a “challenge”:
  - High “signal” rate
  - *By construction*: “high-precision” experiments
- LHCb upgrade: 1 / 5 beam crossings contains “reconstructible charm”
- The problem is “classification of signal” not “rejection of background”

boring, large rate:
$D^{*+} \rightarrow D^0(K^-\pi^+)\pi^+$

interesting, “doubly Cabibbo” suppressed by factor ~250:
$D^{*+} \rightarrow D^0(\pi^-K^+)\pi^+$

Run I	Per event	with vertex in VELO	Rate [GB/s]
b-hadrons	0.0258 ± 0.0004	0.0029 ± 0.0001	0.9
c-hadrons	0.297 ± 0.001	0.0422 ± 0.0005	3.3
light, long-lived hadrons	8.04 ± 0.01	0.511 ± 0.002	1.1
Upgrade	**Per event**	**with vertex in VELO**	**Rate [GB/s]**
b-hadrons	0.1572 ± 0.0004	0.01874 ± 0.0001	27
c-hadrons	1.422 ± 0.001	0.2138 ± 0.0005	80
light, long-lived hadrons	33.291 ± 0.006	2.084 ± 0.001	26

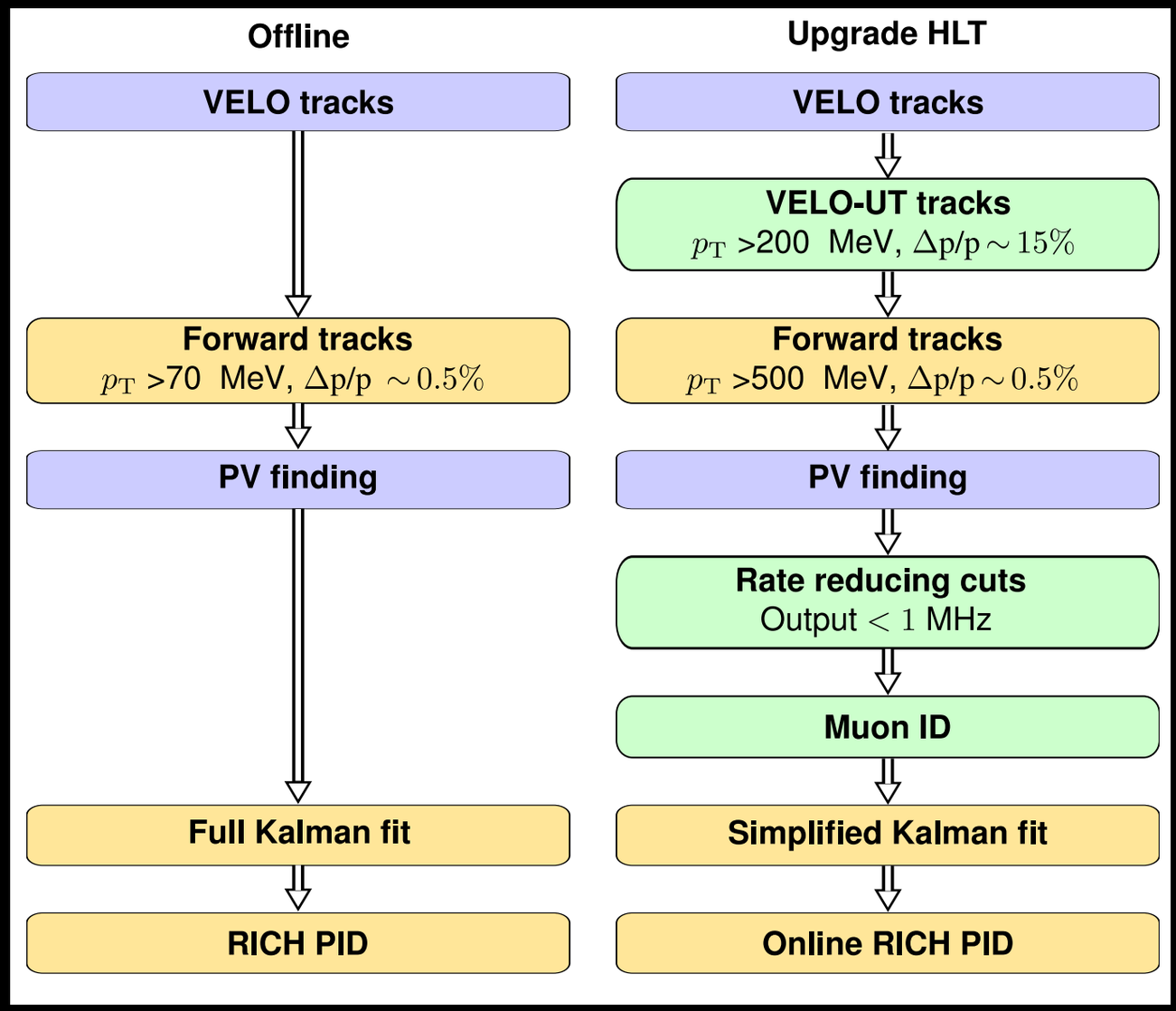
“Offline” Tracking @ 30 MHz

- Run “Software Tracking” at full 30 MHz *interaction rate*!

- Need O(1000) servers to receive the data…

- …and tracking is estimated to use O(40%) of those CPU resources (in 2020)

  - extrapolation based upon past in-situ experience & Moore’s law



“Offline” Tracking @ 30 MHz

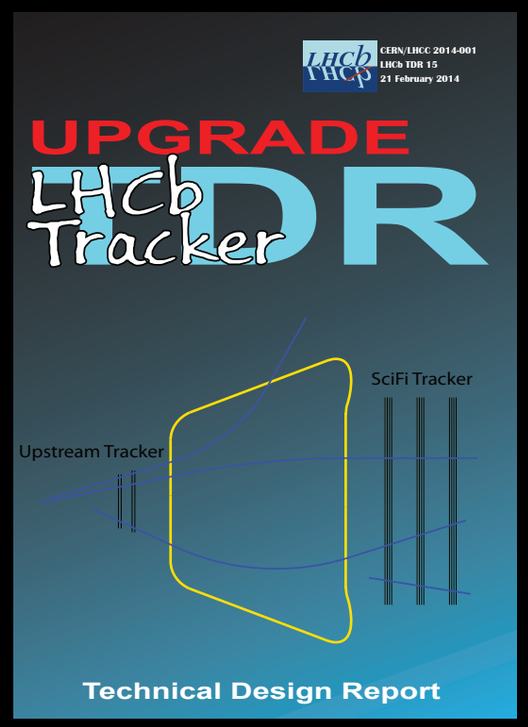
- Run “Software Tracking” at full 30 MHz *interaction rate*!

- Need O(1000) servers to receive the data…

- …and tracking is estimated to use O(40%) of those CPU resources (in 2020)

  - extrapolation based upon past in-situ experience & Moore’s law

- Thanks to new tracking system, software runs >3x faster @ 5x luminosity!



“Offline” Tracking @ 30 MHz

- Run “Software Tracking” at full 30 MHz *interaction rate*!

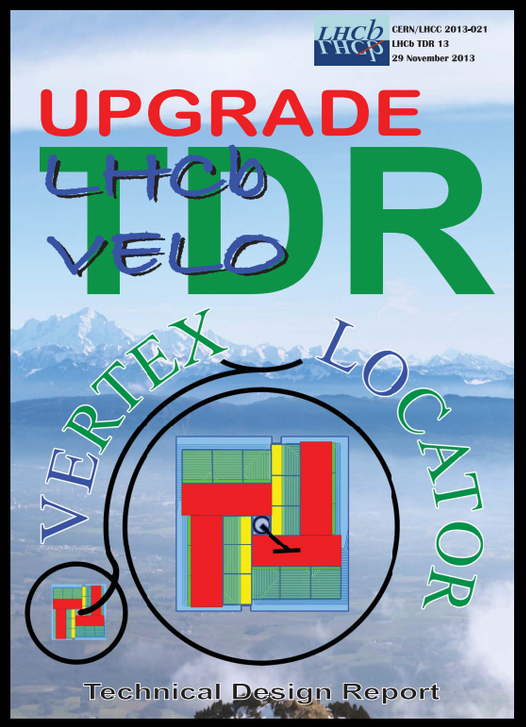
- Need O(1000) servers to receive the data…

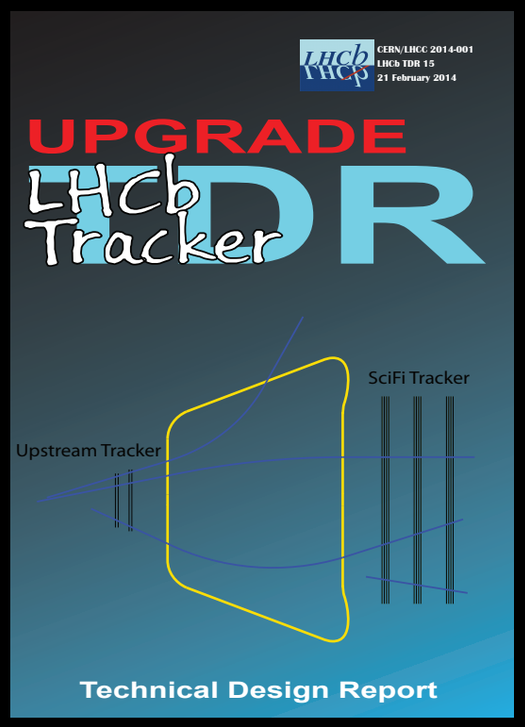
- …and tracking is estimated to use O(40%) of those CPU resources (in 2020)

  - extrapolation based upon past in-situ experience & Moore’s law

- Thanks to new tracking system, software runs >3x faster @ 5x luminosity!

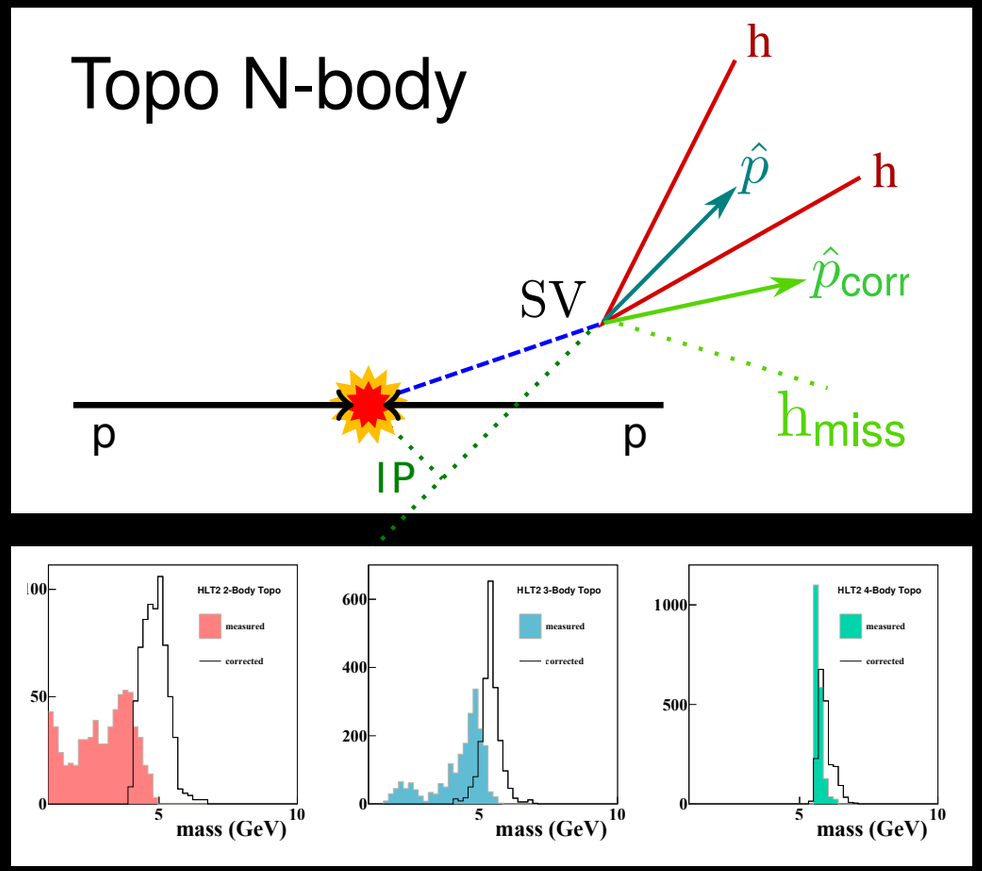
- Decided to pursue *a full-software trigger system*

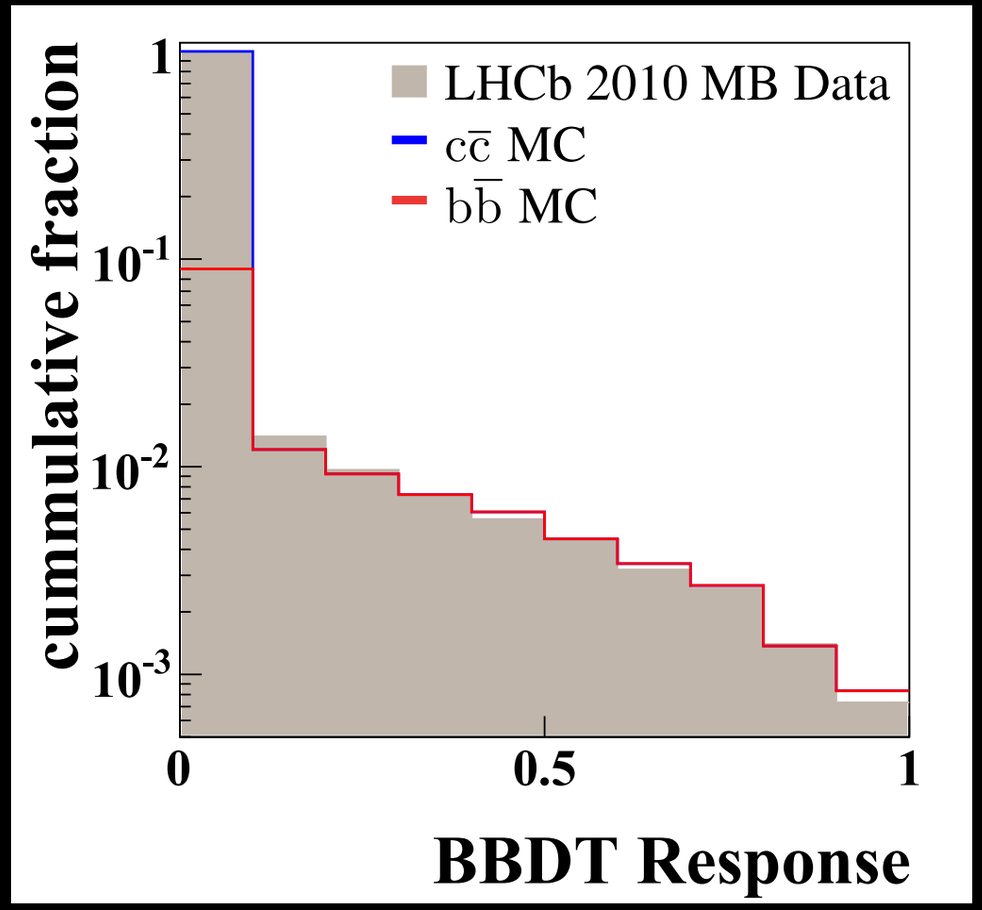




Trigger Selection

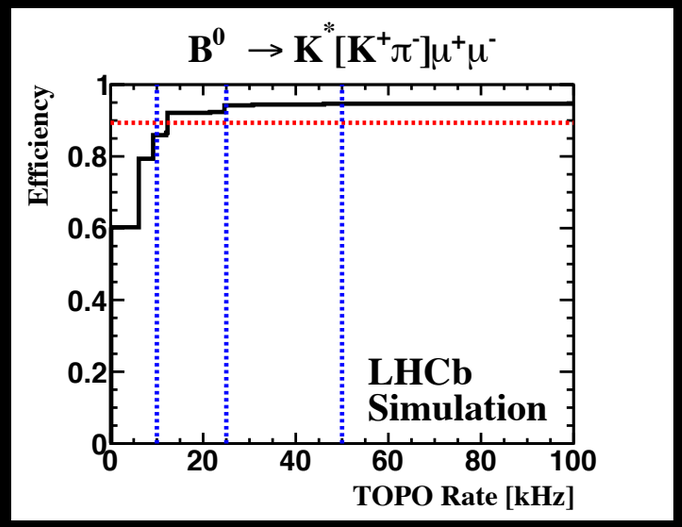
- Main B-physics trigger: upgrade version of the 2011/2012 *inclusive* 'Topological Trigger'
- Combines displaced tracks, invariant mass & "Boosted Bonsai Decision Tree"
- Sufficiently efficient to fill entire output bandwidth with 'pure B'

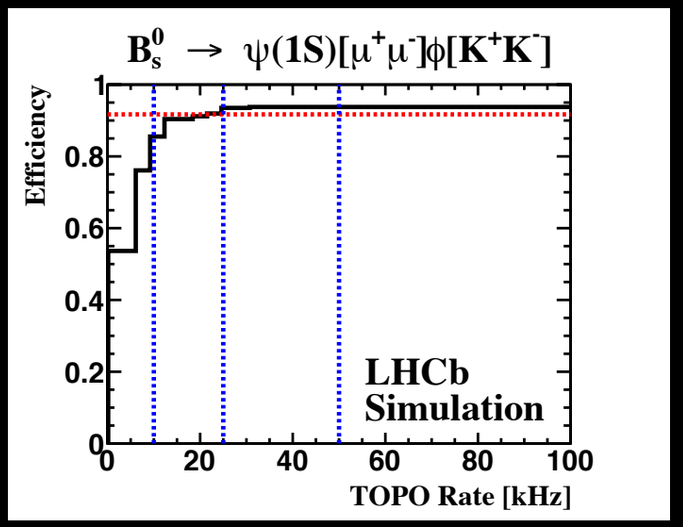


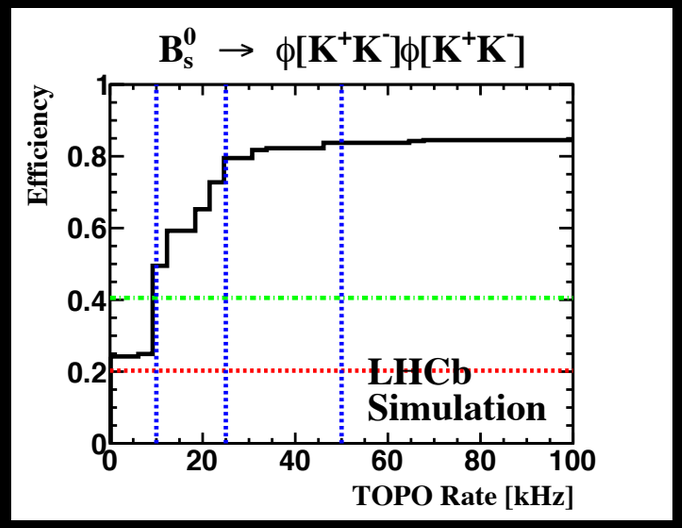


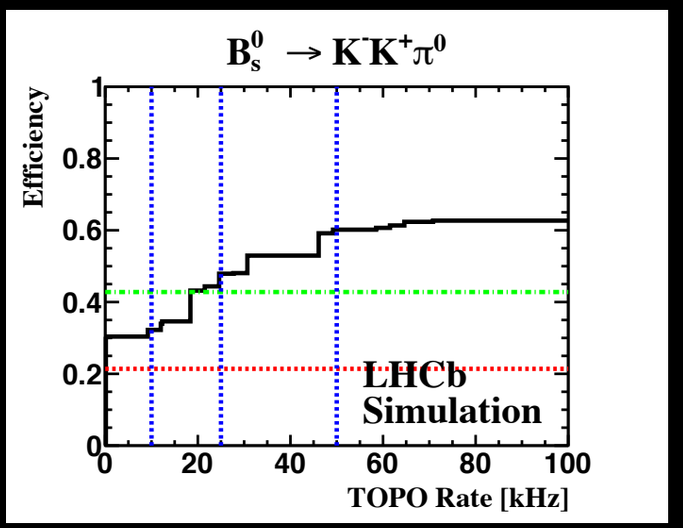
Trigger Selection

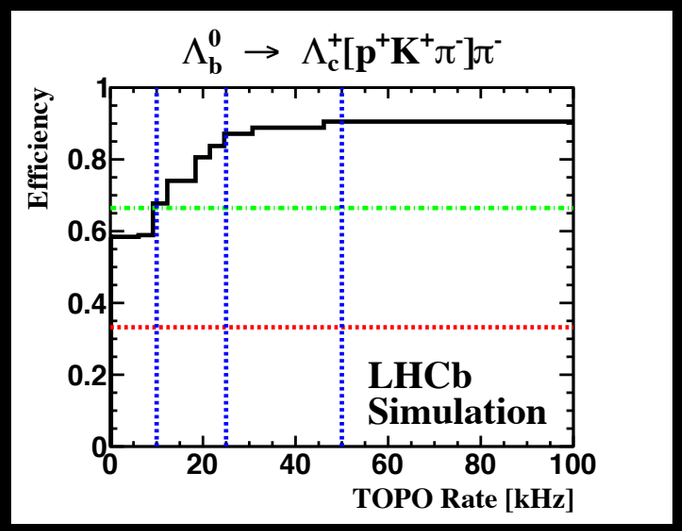
- Main B-physics trigger: upgrade version of the 2011/2012 *inclusive* 'Topological Trigger'
- Combines displaced tracks, invariant mass & "Boosted Bonsai Decision Tree"
- Sufficiently efficient to fill entire output bandwidth with 'pure B'
- At 25-50 kHz output rate, more than 2x efficient (at 5x luminosity, 2x cross-section) cmp. to 2012
- Output rate limited by allotted offline storage capacity….

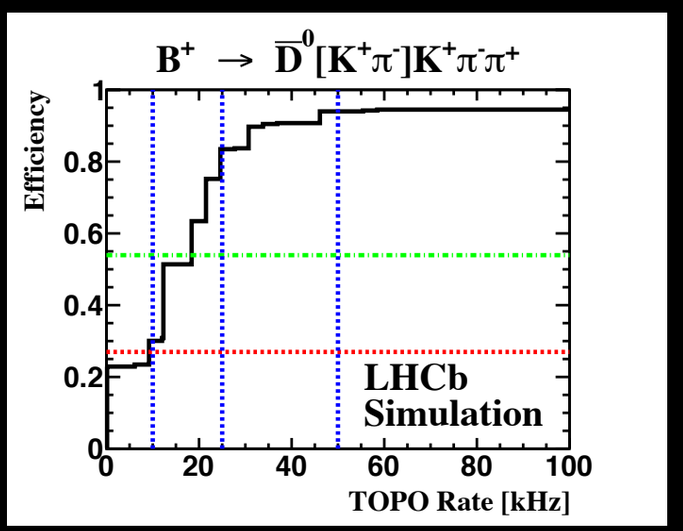












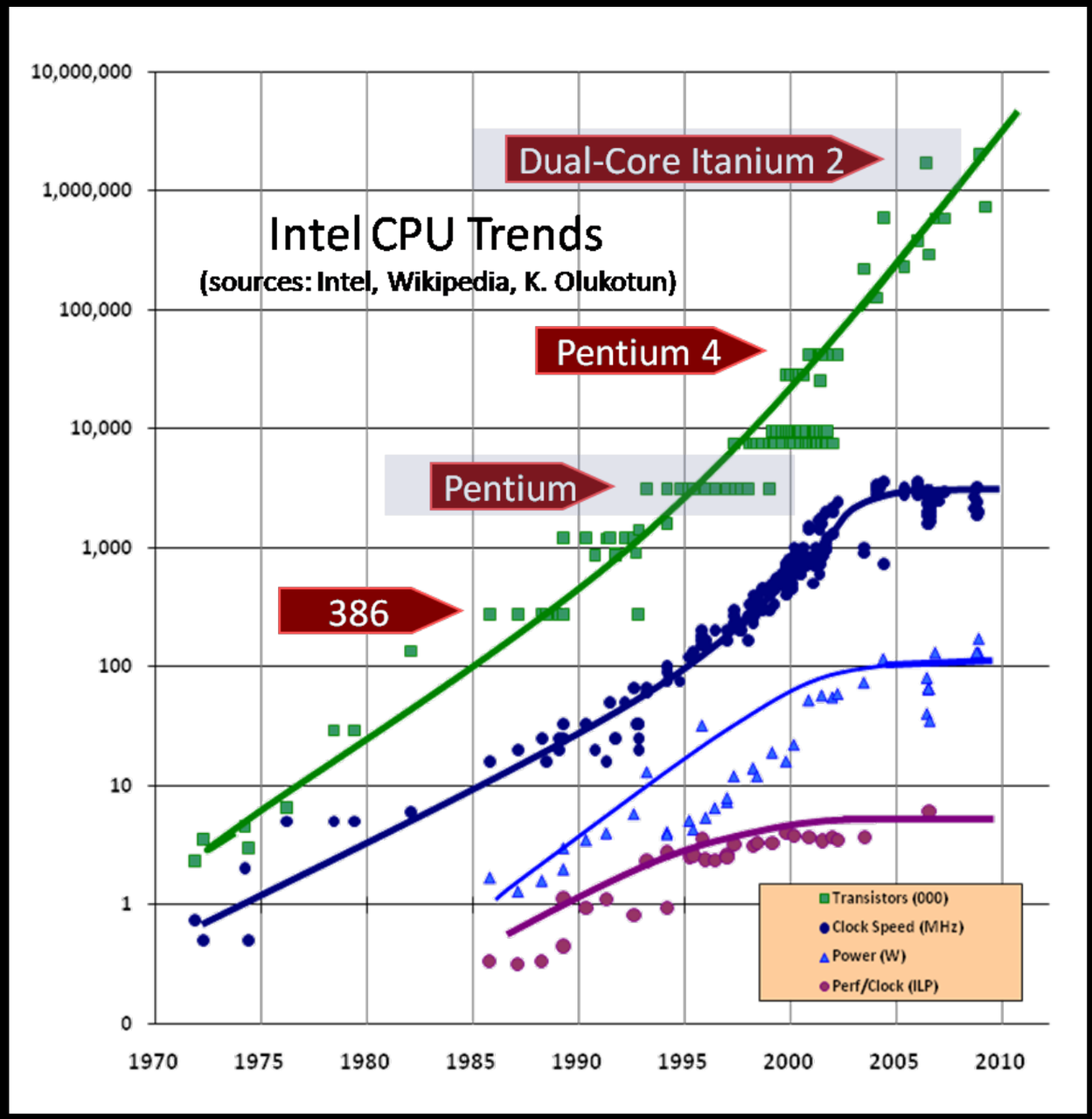
..... 2012 efficiency x 2

..... 2012 efficiency

“There ain’t no such thing as a free lunch”

R. A. Heinlein, *The Moon Is a Harsh Mistress*

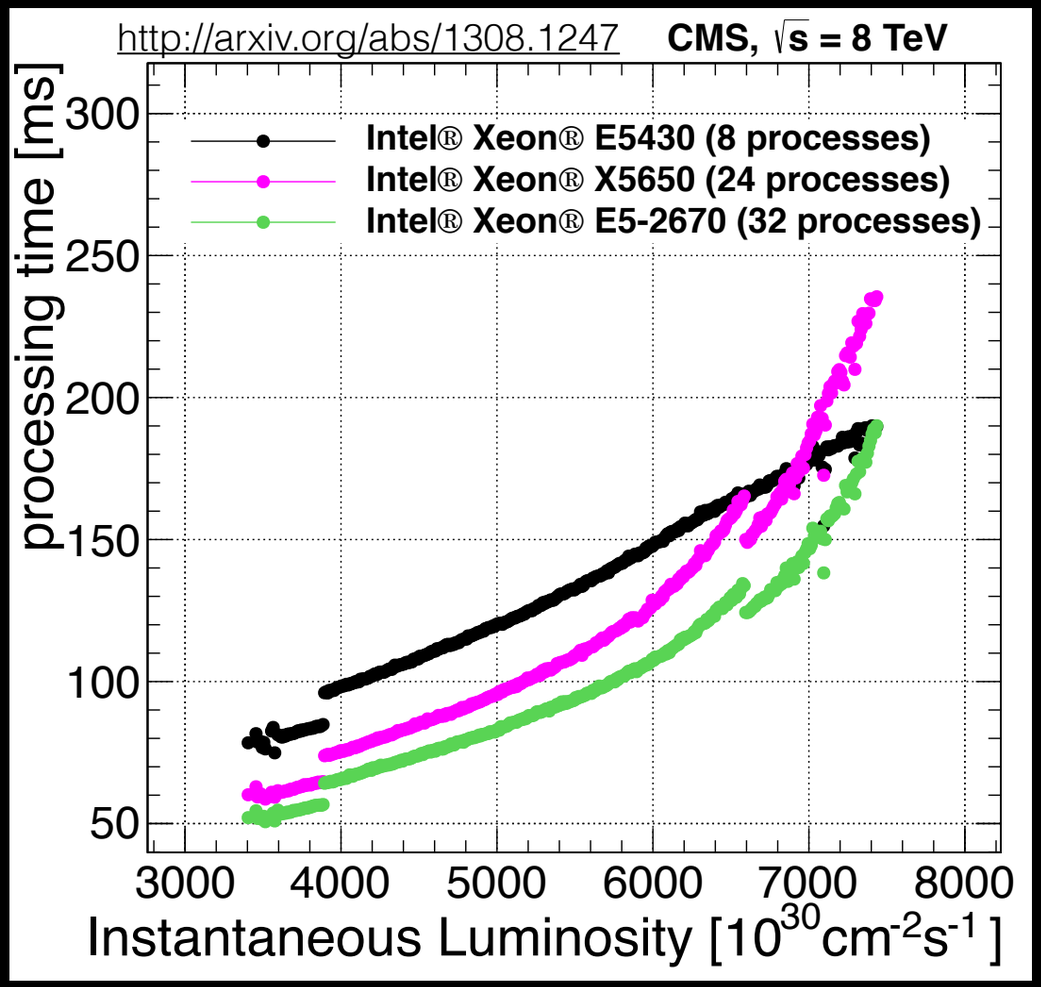
- Already in 2005 (*) it was realized that “things are about to change”
- First 2 GHz Intel CPU: August 2001
- Days of ever faster “sequential processing” are long gone…

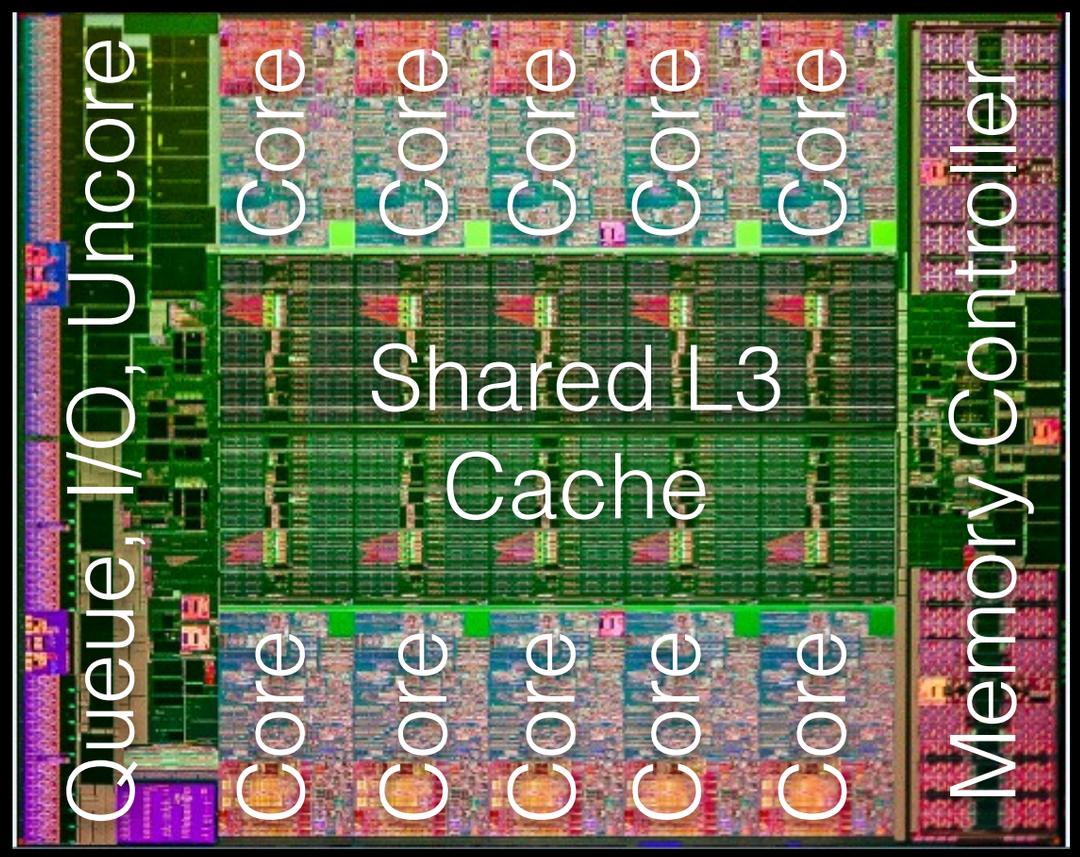


(*) Herb Sutter, Dr. Dobb's Journal, 30(3), March 2005.

Scaling…

- Days of ever faster “sequential processing” are long gone…
- To take advantage of Moore’s law, must ‘go wide’
  - vectorization, concurrency
- Event processing in HEP is “embarrassingly parallel” — that helps, but isn’t enough anymore…

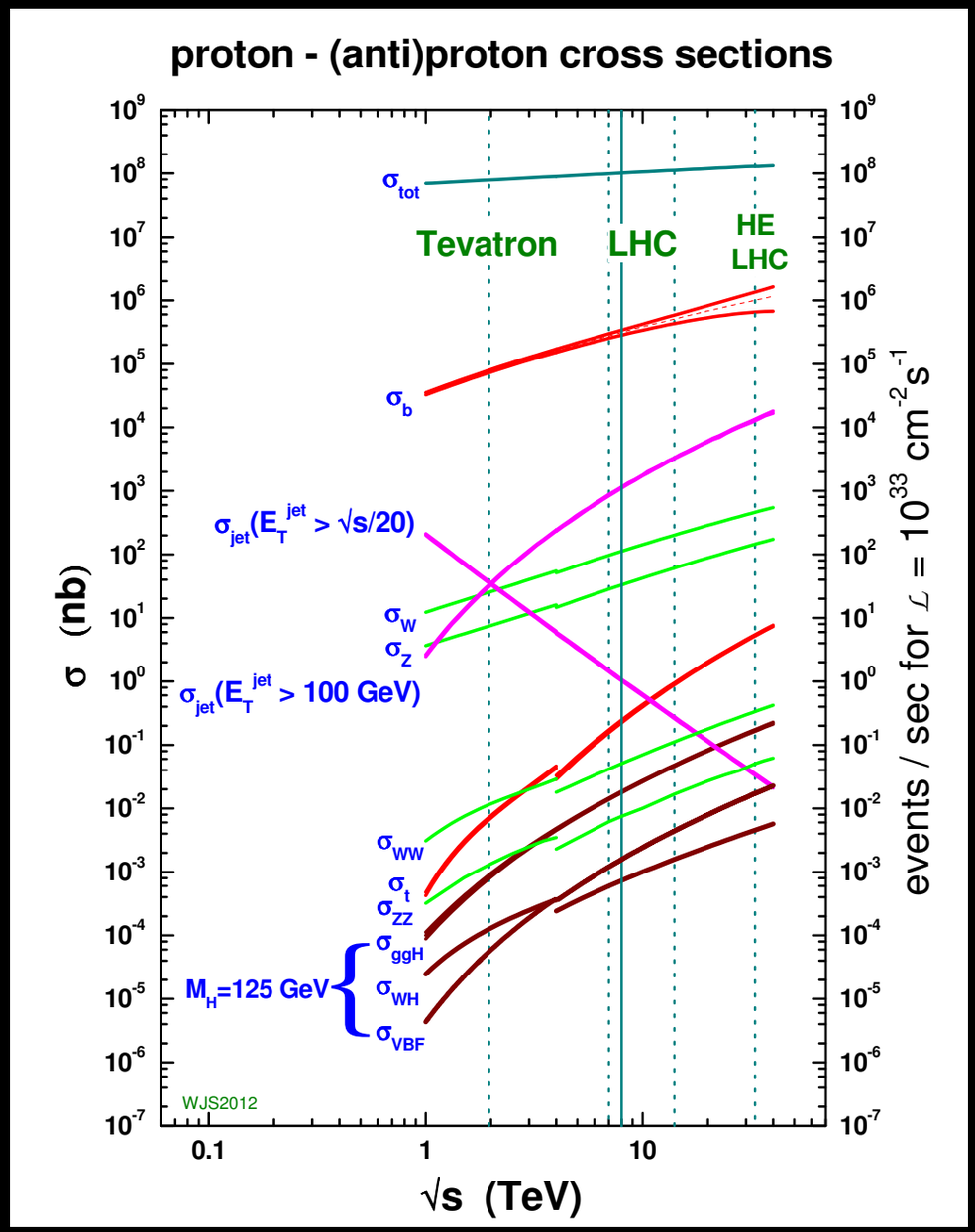




22 nm Xeon E5-2600v2

Energy Frontier

- Requires highest possible luminosity one can afford
- Initial data reduction relies on ‘simple’, ‘isolated’ signatures
  - High-P_T leptons, jets



W. Stirling, private communication

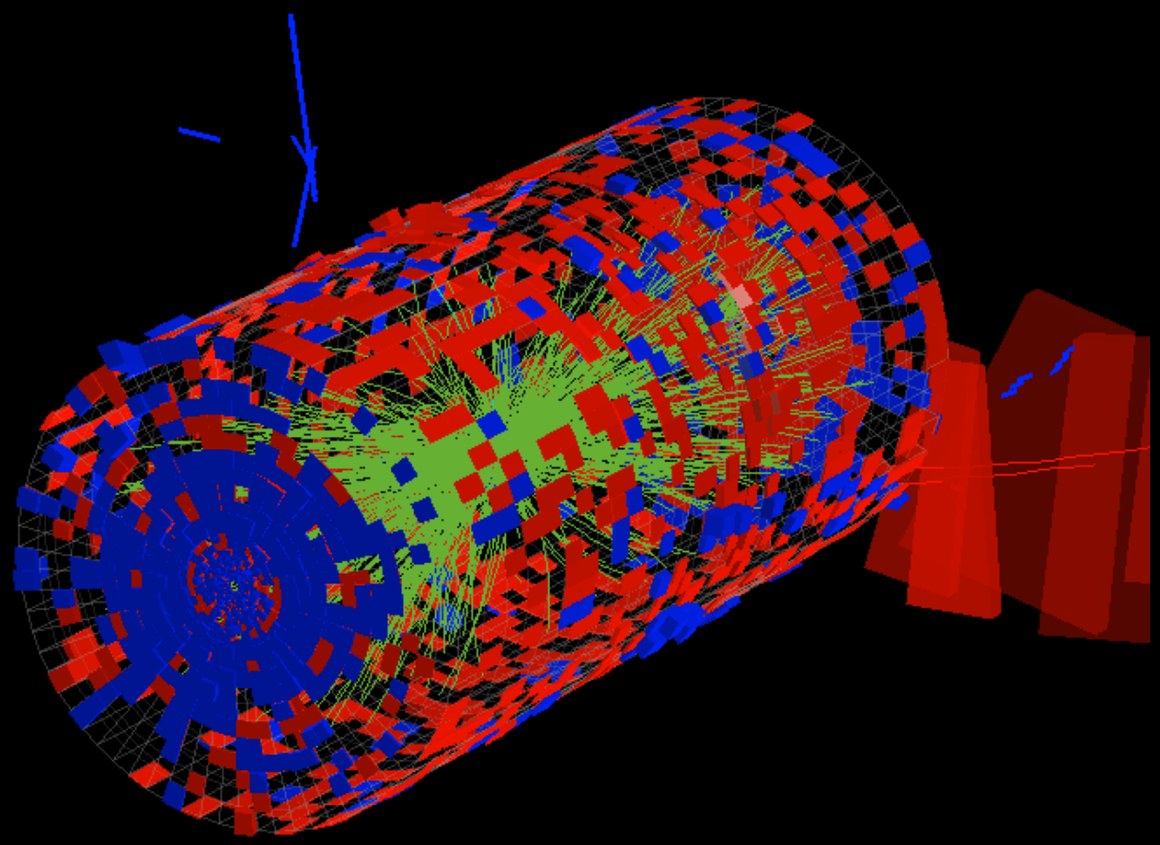
Energy Frontier

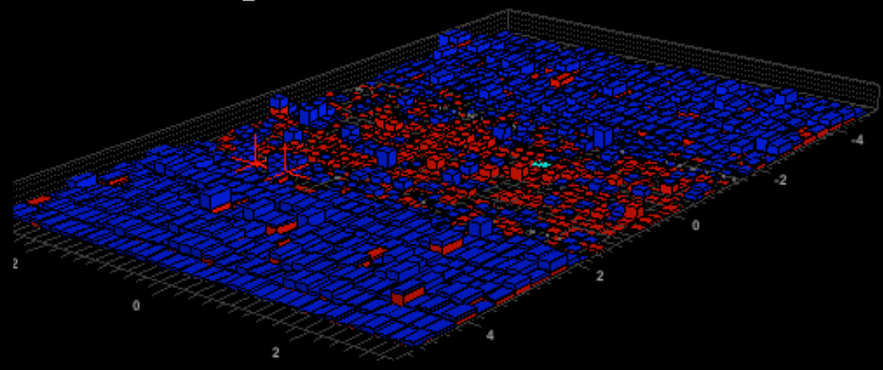
- Requires highest possible luminosity one can afford
- Initial data reduction relies on ‘simple’, ‘isolated’ signatures
  - High-P_T leptons, jets

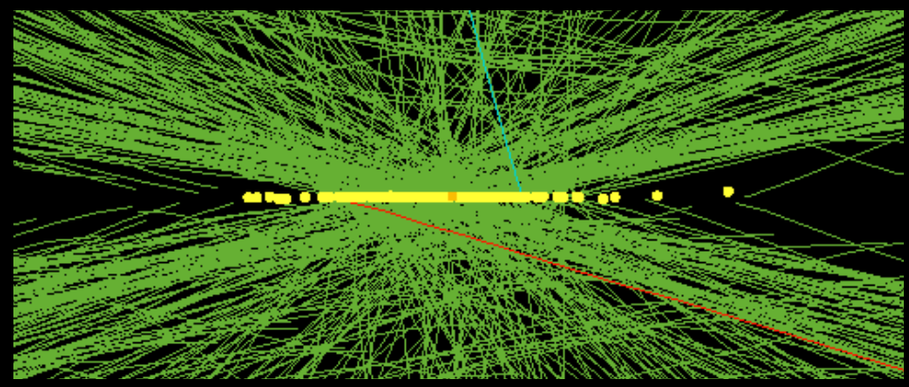
Energy Frontier

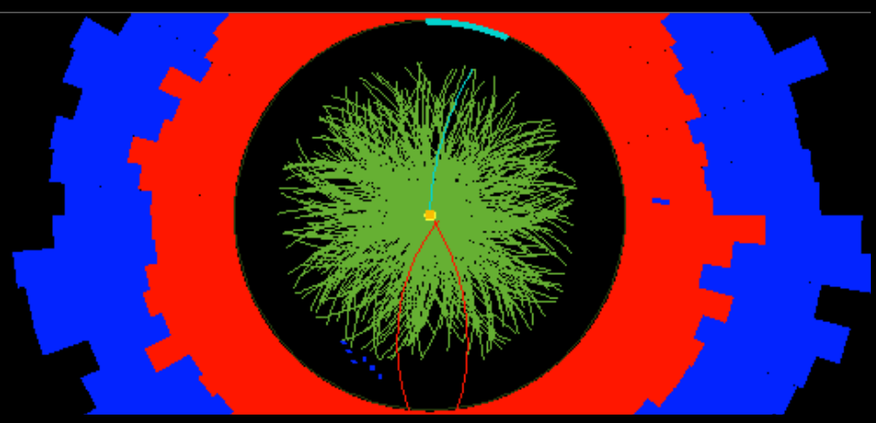
- Requires highest possible luminosity one can afford
- Initial data reduction relies on ‘simple’, ‘isolated’ signatures
  - High-P_T leptons, jets
- But then it gets more complicated quickly…

CMS large pile-up event



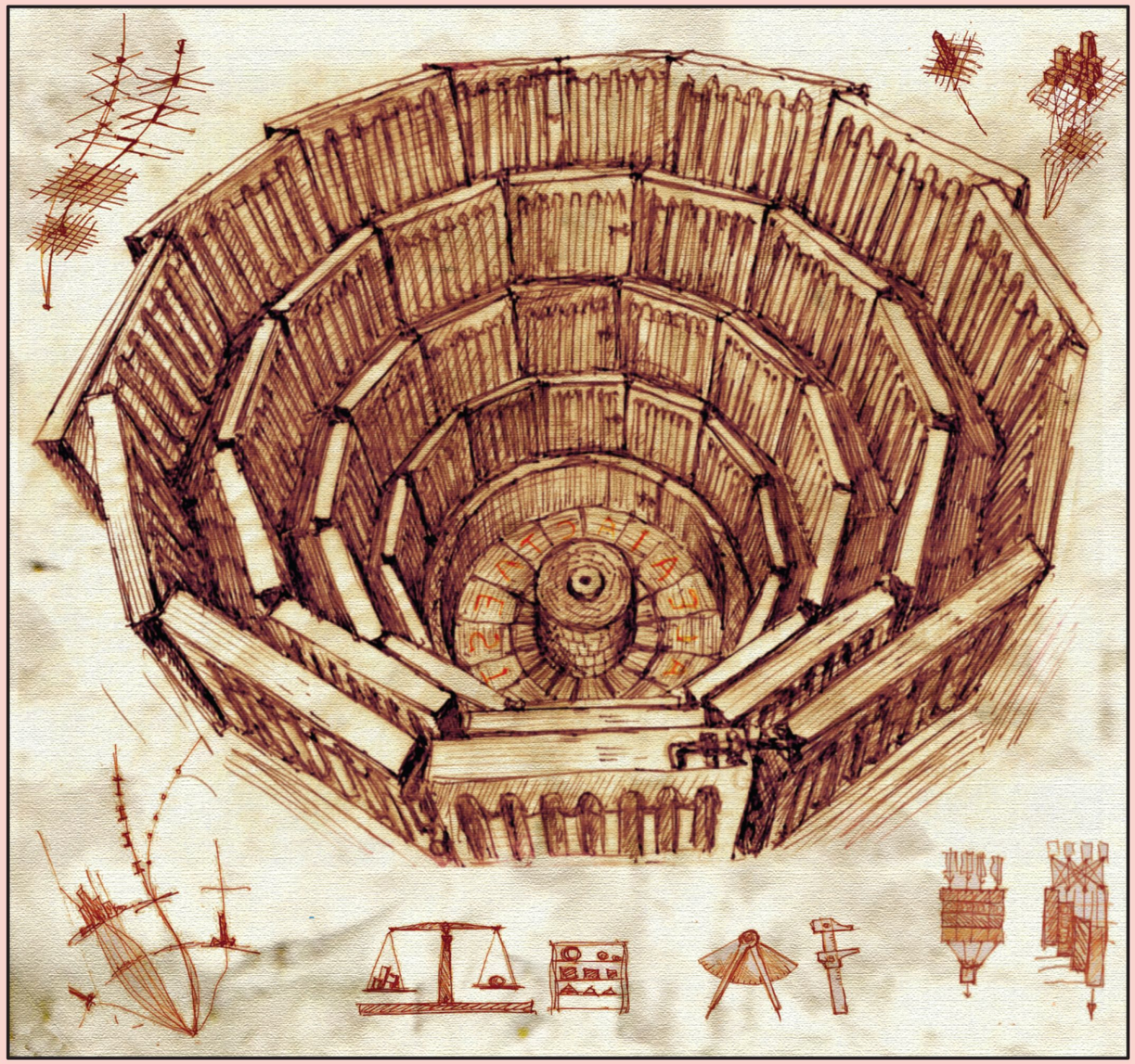






Only tracking can discriminate amongst piled-up interactions …
… so triggers need tracking to remain selective

CMS



CMS TECHNICAL DESIGN REPORT FOR THE LEVEL-1 TRIGGER UPGRADE

- "The specification for CMS electronics was to operate at a L1 trigger rate *up to 100 kHz*, and … need major upgrades to exceed this rate. With the increase in energy, luminosity and pile-up, a substantial increase in trigger thresholds will be required to fit …, especially for pile-up sensitive *multi-object triggers*."

- "The CMS Upgrade Phase 2 Trigger R&D centers on two key components. The first one is the addition of a L1 tracking trigger … The second R&D focus point is to study the option of a significant increase of L1 rate (up to 1MHz) , L1 latency and HLT output rate."

Bandwidth Requirements…

- Even Moore's law won't allow 'affordable' trigger-less CMS / Atlas in 2022 — think O(0.5M) GBT links — and 'backwards (in)compatibility'

	Event-size [kB]	Rate [kHz]	Bandwidth [Gb/s]	Year [CE]
ALICE	20000	50	8000	2019
ATLAS	4000	200	6400	2022
CMS	4000	1000	32000	2022
LHCb	100	40000	32000	2019

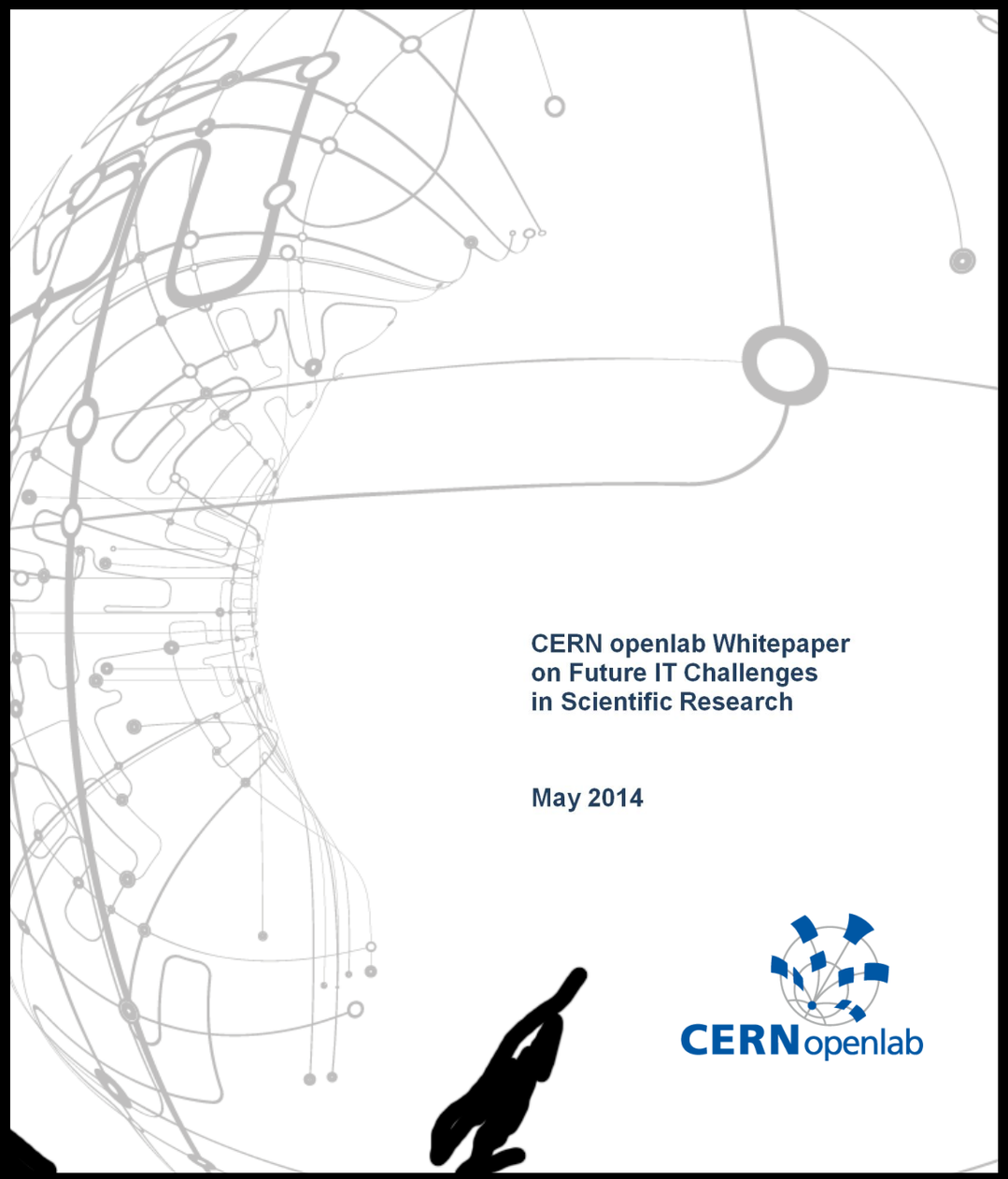
ECFA TDOC 2013 technology trends - N. Neufeld

- For Atlas / CMS, 'challenging solutions' (e.g. L1 track triggers) remain a necessity…

- But what if you had to design, *from scratch*, a CMS / Atlas - like experiment, and could assume 2030 computing & interconnects?

CERN White Paper

- “Data acquisition is where instruments meet IT systems. “
- “Costs and complexity must be reduced by replacing custom electronics with high-performance commodity processors and efficient software.”



Trends In Triggering

- Don’t — unless you must
- Commoditization is the Future (it has been for a long time![*])
- Design hardware and software together

(*) ACCMOR collaboration, C.Daum et.al. 1981

Trends In Triggering

- Don’t — unless you must
- Commoditization is the Future (it has been for a long time![*])
- Design hardware and software together

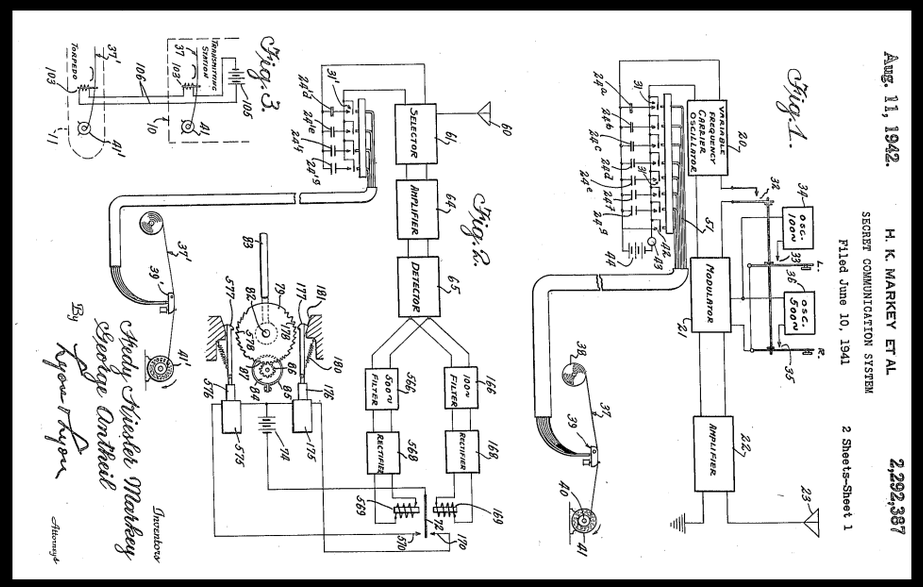
(*) ACCMOR collaboration, C.Daum et.al. 1981

Hedy Lamarr: Frequency Hopping

- http://www.google.com/patents?vid=USPAT2292387

- Military communications in 1941 — using ‘mechanical’ hopping

- Nowadays: mobile phones!



Summary interconnects and networks

- By the end of LS2 links with > 100 Gbit/s will be readily available
- Prices for networking in the local area are dropping steadily
  - The entry of silicon-photonics should reduce the price of optical links even more
- In non-radiation environments all network and link needs of LHC experiments will be satisfied by industry

	Event-size [kB]	Rate [kHz]	Bandwidth [Gb/s]	Year [CE]
ALICE	20000	50	8000	2019
ATLAS	4000	200	6400	2022
CMS	4000	1000	32000	2022
LHCb	100	40000	32000	2019

Hardware Trigger on Calorimeter Et...