A comparative discussion of bus/crate standards and their use at CERN

Markus Joos

CERN

Please note

This presentations includes the personal opinions of several experienced colleagues:

- Alex Kluge for ALICE
- Philippe Farthouat for ATLAS
- Magnus Hansen for CMS
- Niko Neufeld and Ken Wyllie for LHCb
- Javier Serrano and Marc Vanden Eynden for the Accelerator Controls group
- Vincent Bobillier and Francois Vasey for the xTCA evaluation project

Contradictory pieces of information reflect that different view points or “cultures” exist in different areas of CERN. This leads to different solutions for problems that may appear similar.

The slides contain more material than I can present in 30 minutes. I hope you will find the bonus slides useful for off-line reading.

<u>The CERN schedule</u>
All accelerators and experiments are stopped every ~4 years for maintenance (LS = long shutdown)
LS1: 2013 - 2014
LS2: 2019 - 2020
LS3: 2024 - 2026

Standards covered in this talk

- NIM (http://www.osti.gov/scitech/servlets/purl/7120327)

- VITA (www.vita.com)
  - VMEbus
  - VXS

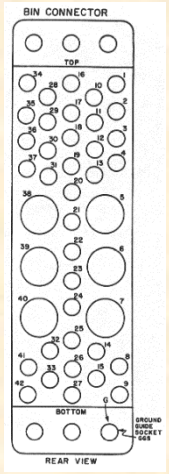
- PICMG(www.picmg.org)
  - CompactPCI (and its derivative PXI)
  - PCIe cards
  - PICMG 1.3 and 1.4
  - ATCA
  - MTCA
  - AMC

Standards NOT covered in this talk

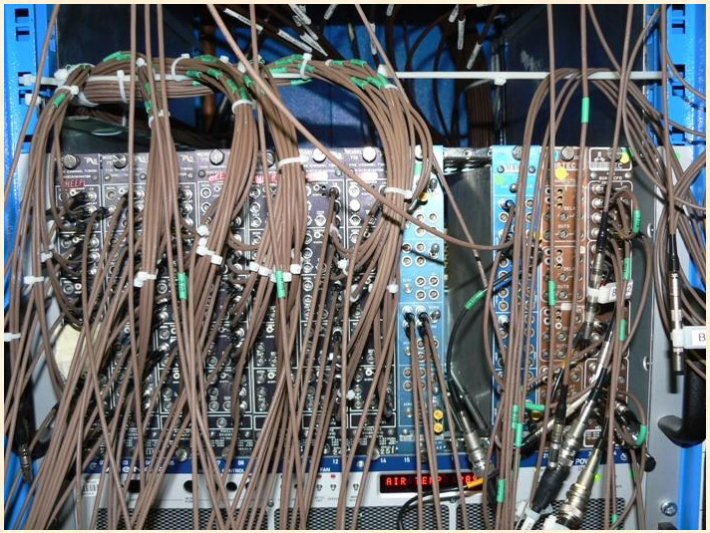
- Fastbus & Camac: Rarely used (at CERN), niche markets
- VXI: Similar to VME
- VPX: I (and my colleagues) have no experience with it

NIM

Year	1964
Module dimensions	34 x 221 x 246 mm
Connector	42 pins (29 reserved since 1964)



Connector



Special features

NIM modules (usually)

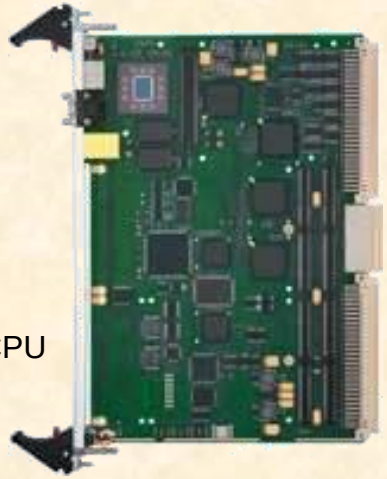
- Need no software
- Are not connected to a computer
- No data communication via the backplane
- Are used to implement trigger logic

Why was / is it successful?

- First of its kind
- Targeted at one type of application
- Simple to use and robust
- Data transfer performance is not an issue
  - No pressure to go obsolete

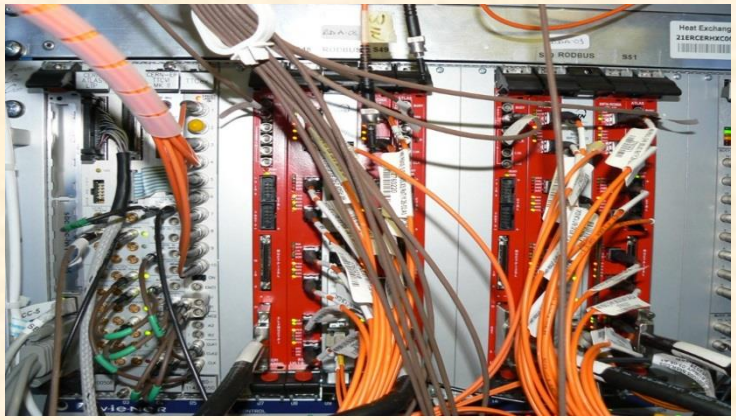
VMEbus

Year	1982
Module dimensions	3U x 20 x 160 mm 6U x 20 x 160 mm 9U x 20 x 340 mm
Connector	2 x 160 pins (plus P0 option)



Embedded CPU

ATLAS RPC read-out

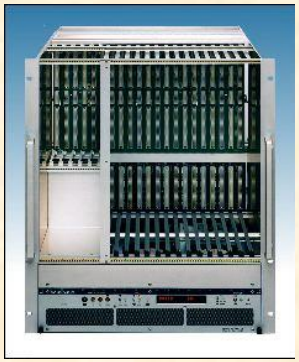


Special features

- Parallel, asynchronous bus
- Data transfer protocol (relatively) easy to implement in FPGA
- Cooling capacity limited
- Still a very large market with many companies and products

Why was / is it successful?

- Outperformed predecessor (CAMAC)
- Large number of manufacturers
- Not too complex
- Central support at CERN



9U crate

VMEbus at CERN

- CERN has a legacy of ~30 years in VMEbus
  - CERN engineers have actively participated to the development of the standard
- Accelerator controls:
  - ~900 crates in the accelerator controls systems
  - In 2015 BE group decided to buy 700 new VMEbus SBCs
- Research sector
  - ~800 crates (including spares and El. Pool inventory)
- There is still a lot of (fully supported) legacy H/W.
- E.g. Electronics Pool

Year	Type	References	Total items	Rented items
2010	Modules	48	1083	564
2015	Modules	49	1084	476
2017	Modules	44	990	405
2010	Crates (bin + fan + PS)	58	1226	692
2015	Crates (bin + fan + PS)	39	794	493
2017	Crates (bin + fan + PS)	38	727	490

Note: 1 crate = ~3 items

- Many commercial front-end functions (e.g. TDC, QDC) not available in any other format
- Still a lot of in house S/W (device drivers, libraries) and experts (despite some retirements)
  - Good support for all users
  - Long-term contracts with manufacturers

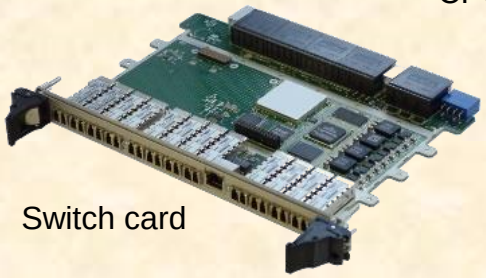
VXS

Year	2006
Module dimensions	6U and 9U (as VMEbus)
Connector	105 pins (plus VMEbus connectors)

CPU card

Special features

- Compatible with legacy VMEbus systems
- Addressed the speed limitation of VMEbus by adding high-speed (10 GHz) differential lines
- P0 backplane connector is protocol agnostic
  - Infiniband, Serial RapidIO, Ethernet and PCIe possible
- Hot-swap and IPMI possible (but not fully standardized)



Switch card

Why was / is it NOT successful?

- Had to compete with xTCA
- Did not address many shortcomings of VMEbus
  - Power, cooling, management, hot swap, module width
- Little market interest
- Backwards compatibility not necessarily an advantage

6U crate

CompactPCI (and friends)

Year	1995
Module dimensions	Same as 3U and 6U VMEbus
Connector	Various type (parallel and serial protocol)

CompactPCI crate

Special features

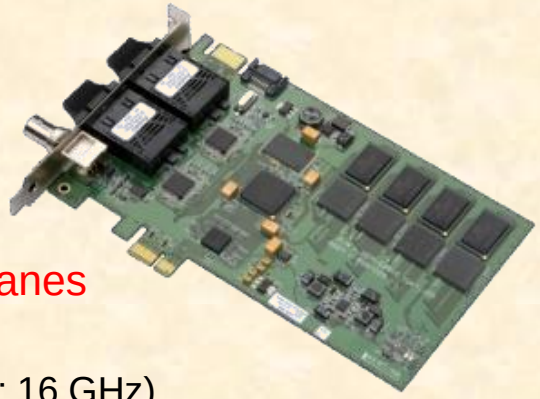
- Based on the PCI(e) protocol
- Many derivatives: CompactPCI Serial, CompactPCI PlusIO, PXI, CompactPCI Express
- S/W compatibility in PCI->PCIe migration
- Single master (scalability)

PXI system

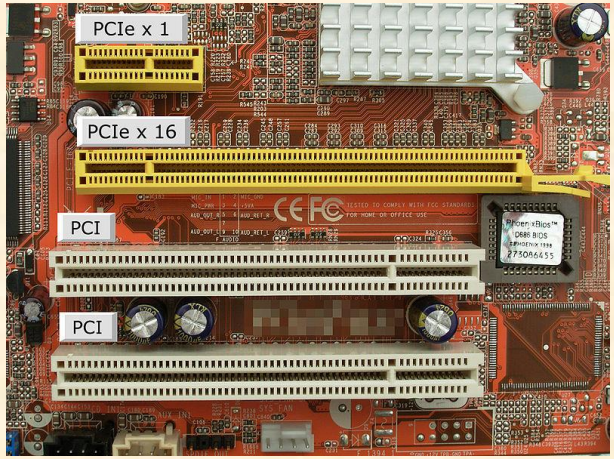
Why was / is it partially successful?

- No large performance advantage over (well established) VMEbus
- Too late to market
- Many modules for Test & Measurement (PXI)

PCIe (aka PCI Express)

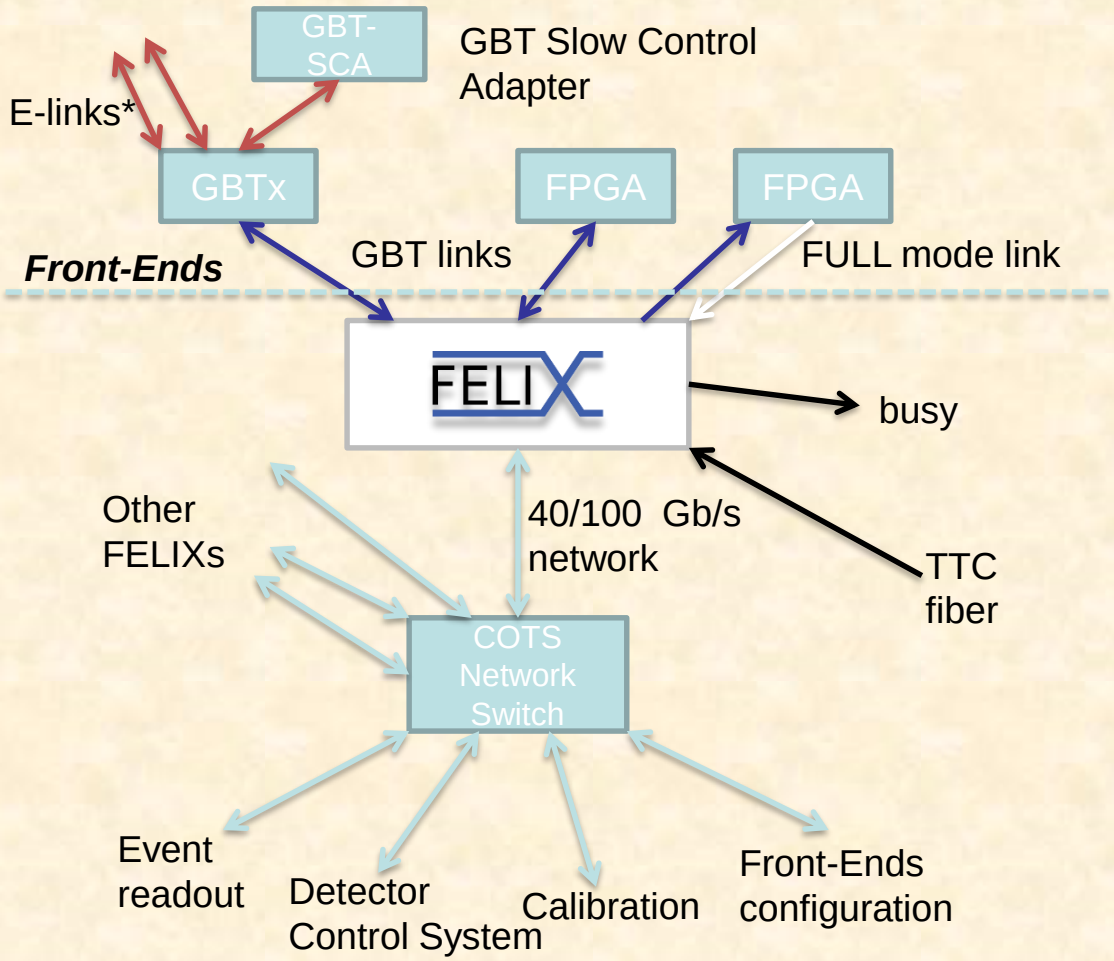


- Not a bus any more but a point-to-point link
- Data not transferred on parallel lines but on one or several serial lanes
  - Lane: One pair of LVDS lines per direction
  - Clock rate: 2.5 GHz (PCIe2.0: 5 GHz, PCIe 3.0: 8 GHz, PCIe 4.0-draft: 16 GHz)
  - 8b/10b encoding (PCIe3.0: 128/130b encoding)
  - 250 MB/s (PCIe 1.0) raw transfer rate per lane
  - Devices can support up to 32 lanes
- Protocol at the link layer has nothing to do with protocol of parallel PCI
- Fully transparent at the S/W layer



Why is it successful?
- Servers to host PCIe cards are cheap and powerful
- Little overhead for board management (no IPMI)
- Large data transfer capacity
- Reasonably large board size

PCIe – the ATLAS FLX-711



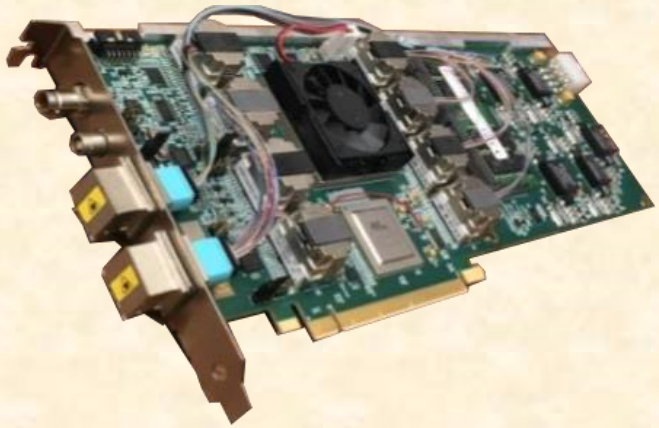
- Scalable architecture
- Routing of event data, detector control, configuration, calibration, monitoring
- Connect the ATLAS detector Front-Ends to the DAQ system, for both the to and from FE directions
- Configurable E-links in GBT Mode
- Detector independent
- TTC (Timing, Trigger and Control) distribution integrated

*** E-link:** variable-width logical link on top of the GBT protocol. Can be used to logically separate different streams on a single physical link

PCIe – FELIX components

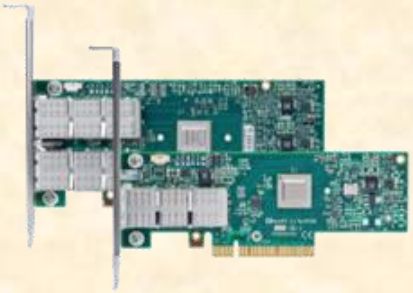
- **VC-709 from Xilinx**
  - Virtex7 X690T FPGA
  - FLX-709 or miniFelix
  - 4 optical links (SFP+)
  - Intended for FE development
  - PCIe Gen3 x8
- **TTCfx (v3) mezzanine card**
  - TTC input
  - ADN2814 for TTC clock-data recovery
  - Si5345 jitter cleaner

- **SuperMicro X10SRA-F used for development**
  - Broadwell CPU, e.g. E5-1650V4, 3.6GHz
  - PCIe Gen3 slots



- **BNL-711 from BNL**
  - Xilinx Kintex Ultrascale XCKU115
  - 48 optical links (MiniPODs)
  - FELIX Phase-1 prototype
  - TTC input ADN2814
  - SI5345 jitter cleaner
  - PCIe Gen3 x16 (2x8 with bridge)
  - Version 2.0 currently tested

- **Mellanox ConnectX-3**
  - 2x FDR/QDR Infiniband
  - 2x 10/40 GbE



PIGMG 1.3 and 1.4

Year	2005
Module dimensions	PCI(e) cards
Connector	All PCI edge connectors

Passive backplane

Special features

- Decouples PCI cards from the motherboard
- More than 7 slots per 19” chassis
- Solution for legacy PCI cards

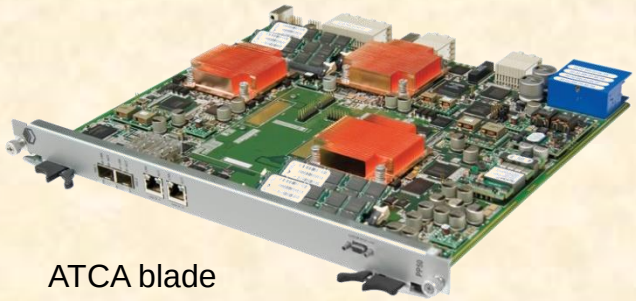
SBC

Why was / is it partially successful?

- Not one of the well known standards
- Cooling difficult and power limited
- I/O bottleneck between the PCI cards and the SBC

ATCA

Year	2001
Module dimensions	8U x 280 mm x 30.48 mm (and RTM)
Connector	Up to 200 differential pairs plus power and controls



ATCA blade

Special features

- Originally developed for the telecom market (now used in other field as well)
- Standardizes many aspects of a system: mechanics, power, cooling, remote monitoring, data transfer, clock distribution
- Focus on High Availability
- Supported by many companies
- High power density (more than 400 W per blade) and data transfer bandwidth
- Sophisticated management structure

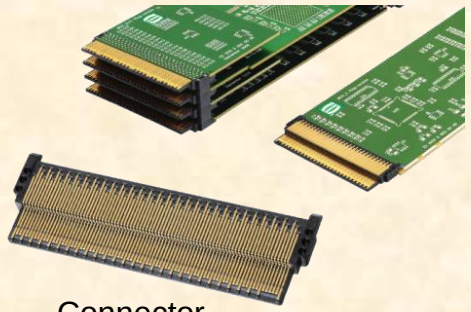
14-slot shelf

Is it successful in HEP?

- Many early adopters jumped on it
  - E.g. ATLAS
- Some got scared by (perceived) complexity or cost
- For some applications there is no (better) alternative
- Could take the place of 9U VME (mainly a form factor issue)

AMC

Year	2005
Module dimensions	74 or 149 mm wide 13, 18 or 28 mm high 180 mm deep
Connector	170 pin



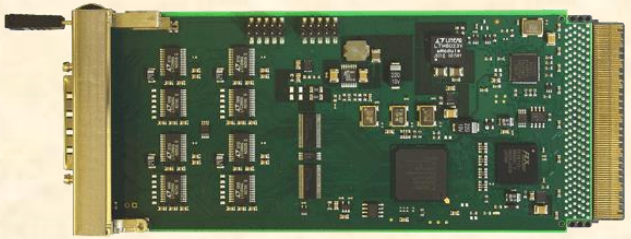
Connector

Special features

- Originally designed as hot-swap mezzanine for ATCA
- Later promoted to module for MTCA (next slide)
- Several standard (and custom) data transfer protocols

Is it successful in HEP?

- Front end functions (e.g. ADC) too slowly becoming available
- Interoperability problems have long been an issue
- Management overhead (MMC) high for a mezzanine



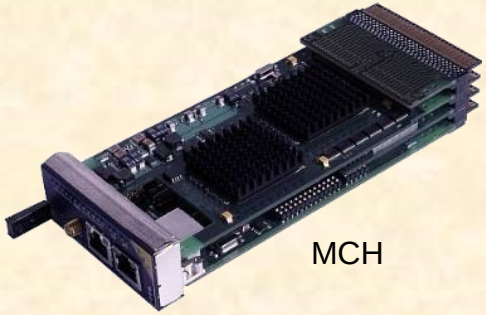
AMC card

MTCA

Year	2006
Module dimensions	All AMCs
Connector	See AMC

12 slot shelf

Special features
- Promotes AMC from “mezzanine” to “module”
- Many types of shelves
- Often “personalized” with custom backplane
- Support for RTMs added for “physics”


MCH

Is it successful in HEP?
- Used a lot in recent accelerator controls system (DESY, PAL(South Corea), ESS (Sweden)) and at CMS
- The standard is evolving
  - MTCA.4 provides RTMs
  - Other features (e.g. more power) are in the pipeline
- Complexity must not be underestimated (MMC)
  - There are open source and commercial solutions
- Could take the place of 6U VME


Double width AMC (CERN design)

Complexity is increasing

(but how can we measure that?)

By the number of pages of the standard?

Standard	Number of Pages
ATCA	660
MTCA	540
MTCA.4	100
AMC	370
IPMI 1.5	460
VME64	306
VME64X	100
VXS	60
VPX	107
NIM	75

Note: Only the base documents are listed
Sub-standards increase the volume further.
Standards for the communication protocols (PCI, Eth. Etc) are also not counted

By the number (sub)-standard documents?

Standard family	Number of documents
ATCA (with HPM)	12
MTCA (with HPM)	8
AMC	5
VME64x	10
VXS	4
VPX	19
cPCI (with Serial and Express)	21

Complexity leads to interoperability issues and long development cycles

MTCA at DESY

http://mtca.desy.de/

- Decided in 2007 to use MTCA for the XFEL controls system
- Together with SLAC (and other partners): Founded the xTCA for physics initiative
  - Result: MTCA.4
    - Introduction of RTMs to MTCA
    - Support for analog electronics
  - Result: ATCA
    - RTM Zone 3A (PICMG 3.8)
  - On the agenda for the future
    - Higher bandwidth (e.g. 40 & 100 Gbit Ethernet)
    - More power for the AMCs
    - And more…..

- 200 MTCA.4 systems used in XFEL

- The DESY business model:
  - They share “things” under NDA with industry
  - They collaborate closely with ~ 15 companies
  - New technologies are offered to other users via a license program

Further reading: http://picmg.opensystemsmedia.com/articles/microtca-culture-cots-signal-processing-desy/

The CERN accelerator controls systems – is MTCA the future?

In order to have synergies with other controls systems (e.g. DESY) MTCA will play an important role

- What we like
  - Serial communication
  - Vendor neutral (not like PXI)
  - Redundancy (but only relevant for ~5% of the systems)
  - IPMI management: remote diagnostic, SEL, automatic alarms
  - Backplane communication bandwidth (but only relevant for some applications)
  - Double width AMC form factor. RTMs will be (heavily) used
  - AMC communication protocol: PCIe (not even looking at other protocols)

- Less important
  - Cooling. VME is OK. PCIMG1.3 not OK for some PCI cards
  - Integration density: Space is available
  - Hot plug

- What we don't like
  - MTCA is an overkill in many respects and very complex
  - Too much choice within the standard
  - Market not developing fast enough (how long will the technology survive?)
  - Mechanical robustness: Edge connector, handles and some other components
  - Interoperability issues still not fully handled between vendors
  - Cost: A big issue (VME seen as cheaper; maybe a factor of 2). PXI-Express also cheaper

- Main concern
  - More open source implementations needed: MMC, MCH S/W, open source H/W (see www.ohwr.org)

How much xTCA for the upgrade of the LHC Experiments?

- All experiments have looked at xTCA for various upgrade projects and took different roads….
  - ALICE: No xTCA (but PCIe cards in servers and still VMEbus)
  - ATLAS: ATCA
  - CMS: MTCA (and later also ATCA)
  - LHCb: No xTCA (but PCIe cards in servers)

xTCA features in they eyes of the LHC experiments:

xTCA feature	ALICE	ATLAS	CMS	LHCb
Redundancy of I/O modules	Not important	Not important	Not important	Not important
Board space	MTCA sufficient	ATCA needed	MTCA and ATCA needed	PCIe sufficient
Cooling	Server PC sufficient	Important. Up to 400 W per blade	Important. Design for 200W per blade. Max. limit: 400W	Server PC sufficient
Integration density	Minor advantage	Not important	Minor advantage	Not important
Hot Plug	Not important	Not important	Used but not crucial	Not important
Costing	Chosen solution cheaper	Not an issue	Good deal	Chosen solution cheaper
xTCA strong points	None	Cooling, card size, PSU, IPMI (powerful but complex)	Good (but complex) system standards. PSU redundancy	Cooling and PSU quality. PCs may be less reliable

Policy of CERN (research sector) with respect to xTCA

Requirements
- technically and commercially mature technologies
- maintainability for 30 years

Specification and qualification
Shelfs and PSUs for LHC experiments (HL-LHC upgrade)

Evaluations
Next slide…

Development of S/W and H/W
CERN has developed expert knowledge in xTCA control hardware and firmware. It intends to make this expertise available to the community in the form of proven, interoperable, documented and supported MMC and IPMC reference designs which users can integrate into their designs.

Communication
In order to facilitate the exchange of information CERN is hosting the xTCA interest group:
https://twiki.cern.ch/twiki/bin/view/XTCA/WebHome
- List your projects
- Platform for discussion
- Regular meetings

Support
- Offer expert centralized support to LHC experiments in the field of crates and power supplies for modular electronics.
    - This support will cover VMEbus, MTCA and ATCA form factors and will include centralized procurement, inventory and maintenance of equipment in LHC experiments.
- The service:
    - is currently running for VMEbus
    - has started in 2016 for MTCA
    - will start in 2018 for ATCA

CERN research sector xTCA evaluation project

Mandates / missions:
- Evaluate commercial products (reports are available)
- Specify and qualify common H/W platforms
- Develop components (H/W and S/W) of common interest
- Build up a support structure for xTCA shelves and power supplies
- Provide consultation to internal and external teams from the experiments

Examples of results:
- Powering options for ATCA
  - We recommend single N+1 redundant high-efficient AC-DC rectifier system in each rack (as opposed to one per shelf or group of racks)
  - Follow trends in other domains. E.g. 400 VDC
- ATCA cooling options
  - Legacy (VMEbus) rack use vertical cooling. xTCA system have a front to back airflow
  - We have made a very detailed comparison of the two options including simulations and measurements
  - Conclusion: Not a clear winner but vertical cooling has some advantages (e.g. keep existing rack infrastructure)
- MTCA interoperability problems
  - Devices from different vendors were tested for interoperability and compliance with the MTCA/AMC standards
  - Many problems have been found and solved with the help of the manufacturers (lack of product maturity)
    - Due to the complexity of the standards (xTCA and IPMI) it is easy to make mistakes
    - Participation in Interoperability Workshops at DESY helped to iron out problems
    - (Much) time has to be invested into systems test in order to avoid “bad surprises” in operation
    - We have also invested (much) money into test S/W from Polaris Networks. This was very helpful.
- Controllers
  - EP/ESE has developed a MMC and a IPMC mezzanine
    - Documentation and support is available

More information:
https://espace.cern.ch/ph-dep-ESE-BE-ATCAEvaluationProject/SitePages/Home.aspx
https://espace.cern.ch/ph-dep-ESE-BE-uTCAEvaluationProject/default.aspx

Behind the horizon….

Sep 2014: PICMG announces GEN4

http://www.picmg.org/gen4-new-high-performance-platform/

Main target: Communication Service Providers (CSP)

Quote: “GEN4 systems will provide the capacity and density required by the next decade of Internet application and traffic growth, serving the modular platform marketplace from 2015 through at least 2025.”

Advertised features:

- System throughput (to hundreds of terabits/s)
- Module bandwidth (to tens of terabits/s)
- Storage capacity in exabytes
- Module cooling capacity (over 2000 Watts, with fluid cooling options)
- Not H/W compatible with ATCA

Questions:

- (When) will it happen?
- Will it fragment the market?
- Is it relevant for our community?

Putting it all together - 1

So, what is the right standard for your project?

This obviously depends on your requirements:

- Power & cooling
- Bandwidth & latency
- Availability of commercial products
- Existing infrastructure (S/W and H/W) and expertise in your collaboration
- Start and duration of the project
- Scalability requirements
- Reliability and availability requirements

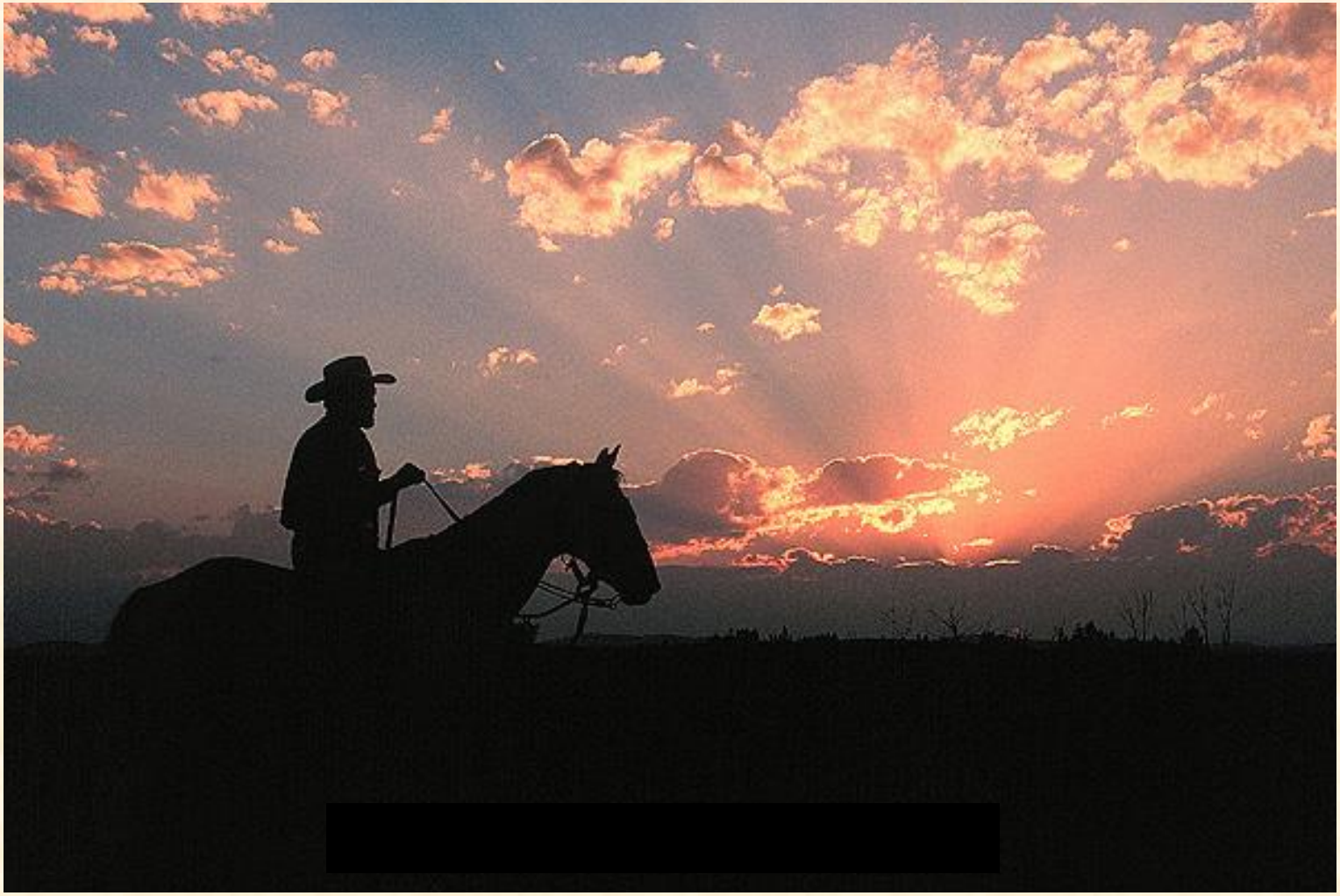
- CERN has a large legacy of VMEbus and, to a lesser extent, of other crate and systems standards
  - Much of this equipment will be around for at least another ~10 years
  - Engineers will be needed to keep these systems alive and to plan the migration to newer technologies
  - "Old" technologies will not be excluded for "new" projects if they fit
- xTCA will play an important role in the future (accelerator and experiments)

Putting it all together - 2

How to give xTCA a future in HEP?

- ATCA and MTCA share many concepts
  - Exploit this for your controls and DAQ S/W
- In terms of age and performance there are no competing standards
  - ….also not in terms of complexity…..
- To master this complexity our community should:
  - Invest in product evaluations and interoperability testing
  - Develop and share designs
    - Some (well document and supported) open S/W and H/W solutions are already available
      - Do not reinvent the wheel
      - Add your own designs to the open domain
  - Share information (e.g. via the xTCA interest group)
  - Avoid using “exotic” technologies
    - E.g.: Prefer PCIe for AMC communication over SRIO
  - Consider contributing (more) to the standardization process

The End



Bonus slides

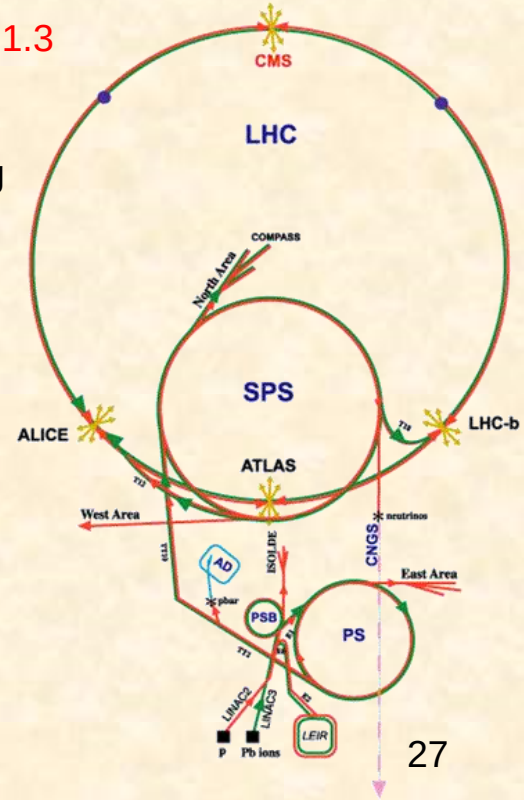
The CERN accelerator controls systems – current situation

VME

- Still designing new VMEbus cards (but less than in the past)
- Extensive use of the “carrier-mezzanine” concept
- 450 new crates installed during LS1 in all accelerators (except LHC)
- 900 crates in operation
- Not yet decided when to abandon VMEbus
- CERN Controls Coordination Committee: Give recommendation for future platforms for accelerator controls at CERN
    - Mandate: Look for future strategy (not necessarily replace VMEbus)
    - Candidate new technologies: MTCA, PC + PCIe cards, PXI, PICMG 1.3 / 1.4
    - Timescale for deliverable: 18-20 month from now
    - Timescale for implementation: ~2019 (prototypes), before and during ~2024 (bulk)
    - Full phase out of VMEbus: not before ~2030
    - Scope: All accelerators. Not a mass replacement (too expensive)
        - Decisions will be taken project by project

Other standards

- PICMG 1.3: 700 systems
- 6U Compact PCI: 50 systems
    - Will be migrated to VMEbus
- PXI: 100-200 systems



The LHC Experiments – 2

Current situation:
- VME (~60 crates), Compact PCI (~15 crates) and MTCA (~60 shelfs) in use

Future:
- VME (at least) until LS3
- ATCA: 100 systems in LS3

Why MTCA in the first place?
- AMC concept interesting (card and mezzanine)
- Spontaneous decision

Current situation:
- 9U cards (VMEbus format) with commercial credit card computer

Future:
- LS2: New On-detector electronics (does the job of the 9U cards)
- LS2: ~500 custom PCI40 cards link detector to event builder
- Main motivation: PC gives a lot of flexibility with respect to N/W technology (Ethernet, Infiniband or Intel Omnipath) and cheap memory.

Current situation:
- VMEbus used to house custom electronics
- Few ATCA shelves installed in LS1 (CSC, Topo, FTK)

Future:
- VMEbus: until LS3. Then FELIX will replace the remaining VMEbus systems
- Several sub systems will add ATCA systems between the detector and FELIX for a total of 100-110 shelves

Current situation:
- Custom electronics in VMEbus (50 systems), in PCIe form factor (~ 400 cards), and in non-standard formats

Future:
- LS2: Upgrade for runs 3 & 4. Read-out and store (pre-processed data) at 50 kHz (no trigger)
  - Need: new F/E electronics (on detector), GBT links, and ~ 500 PCIe40 cards (same as in LHCb) as common read-out and data concentrator units (CRUs)

One manufacturer's perspective

Note: This slide has been contributed in 2015 by a manufacturer of modular electronics and represents their view. Other companies may have different viewpoints. In order to get an accurate picture of the likely future trends one must talk to many companies.

ATCA

The market was predicted to be worth billions of USD: this piece from 2008 predicted $1.7B plus another $700m for AdvancedMC modules and is fairly typical; from 2003, this study thought the market would be $3.7B by 2007! However, the macroeconomic picture changed very quickly: some of the biggest TEMs like Nortel and Motorola disappeared, others like Lucent, Alcatel, Nokia and Siemens merged and more nimble (read lower cost) entrants like Huawei and UTStarCom grew quickly. We are not aware of the current size of the AdvancedTCA market but one of the largest vendors (Radisys) is a public company and disclosed $91m ATCA revenue in 2014.

AMC

In the early days of MicroTCA, Motorola Computer Group made a prediction that MicroTCA would become a modern day equivalent of VMEbus by appealing to many different market sectors. This hasn't happened, in reality MicroTCA works very well for those applications that need multiple CPU, DSP and FPGA processors in a small physical form with good management. Typical markets include telecoms test equipment, adjunct telecoms boxes to add features to existing deployments and high speed physics experiments. Most other embedded applications, especially the high volume ones like medical imaging, digital signage and gaming machines typically only use a single processor mounted to a custom I/O panel and these use other lower cost products like motherboards and COM Express modules.

This is not all bad news. One company disclosed in their interim half year results for 2015 that their sales into telecommunication applications have grown by 91% (compared to the same period last year) due mainly to AdvancedMC sales. There are a number of solid customers for AMC modules but they are probably best described as niche applications. This is driving some companies to show a strong roadmap for AMC and to invest in additional products to expand their portfolio.

VMEbus

Still a trendsetter in the embedded market. Many predicted a steady demise but we've seen incredible resistance. This is because the majority of our VMEbus customers (apart from the physics community) are defence related and they cannot easily swap chassis or obsolete equipment. We have released new VMEbus processor boards in 2015. Moving forward we are about to announce another new VME board and we have a roadmap with the expectation of doing another round of VMEbus boards based on a future silicon revision. We believe that we're going to be shipping VME boards until at least 2030 and are seeing strong demand across multiple programs largely driven by the military market

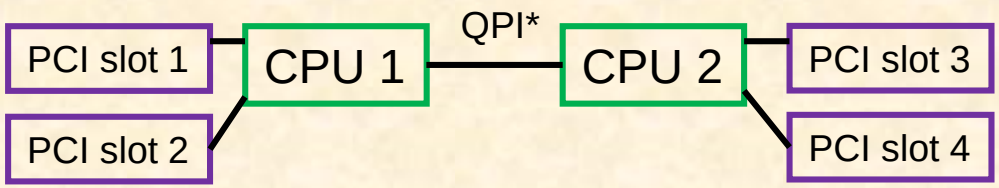
PCIe performance

- Data is transferred in frames:

Start 1 byte	Sequence 2 bytes	Header 12 or 16 bytes	Payload 0 - 4096 bytes	ECRC 4 bytes	LCRC 4 bytes	End 1 byte

- Note:
  - H/W may limit max payload size (typically 128, 256 or 512 bytes)
  - Every data packet has to be acknowledged (additional overhead)
  - Read transactions may cause additional delays
- The actual performance may be as low as ~15% of the theoretical maximum
- Achieving more than ~80% link efficiency is difficult

- The topology of the system (PC motherboard) matters as well
- You may have to use process / tread affinity in order to tie your I/O code to the CPU that connects directly to your I/O cards



*** Intel QuickPath Interconnect**

xTCA degrees of freedom (not necessarily a complete list)

- ATCA
  - Communication protocol(s) on the fabric channels
  - Routing of the fabric channels on the backplane (network topology)
  - Connection between front board and RTM
  - Degree of redundancy
- AMC
  - Card height (13, 18 & 28 mm)
  - Card width (74 & 149 mm)
  - Communication protocols (currently 4 options)
  - JTAG support
- uTCA
  - AMC height & width
  - Degree of redundancy (MCH, PSU, cooling)
  - Routing of the fabric channels on the backplane (custom backplanes)
  - JTAG support
  - Connectivity of MCH to backplane (1 to 4 tongues) and type of communication protocol on the fat pipes
  - Rear transition modules (MTCA.4)

xTCA issues

- The operation of an xTCA system requires a complex, standard compliant S/W infrastructure
  - Efforts to provide open source management S/W for xTCA: OpenSAF, SAForum
- As many features of the standard(s) are optional, products from different vendors may not be compatible
  - Efforts to insure interoperability of xTCA products: CP-TA, SCOPE alliance
  - Interoperability workshops
- Sub-standards for use in “physics”
  - ATCA 3.8: Standardizes RTMs and clock signals
  - MTCA.4: Adds RTMs (and other features) to MTCA. AMCs communicate via PCIe
- The market does not yet provide lots of front end modules for physics DAQ
  - See: http://mtca.desy.de/
- There is little information available about the system performance (end to end H/W performance and S/W overhead) of the data transfer links

Details about CMS

Current situation:
- VME: ~60 crates, mostly 9U for detector readout and trigger
- Compact PCI: ~15 crates, DAQ only
- MTCA:
    - ~60 shelfs already installed
    - Custom AMCs: 8 active designs and some prototypes

Future:
- VME (at least) until LS3. then replaced by ATCA. In LS3 MTCA will also be replaced by ATCA
- Readout Architecture: detector -> xTCA -> PC -> HLT

Why MTCA in the first place:
- AMC concept interesting. Spontaneous decision
- RTMs not used: prefer one big card over two small

ATCA
- To be introduced in LS3, AMC OK for Virtex7, more powerful FPGAs may need ATCA (for power dissipation).
    - 100 systems will be installed in LS3
- The power and cooling of the computing rack is limited to ~11 kW -> one ATCA shelf per rack (two if blades require less power)
- Cooling vertical (better air circulation, systems can be protected with doors). Considering to extend the length of the ATCA cards for better cooling

xCTA feature	CMS position
Redundancy	PSU yes, cooling unclear – depends on cost, data I/O: no
Board space	AMC at the limit, ATCA OK for some designs. VME PCBs are too thin (especially for 9U for mechanical stability)
Cooling	Important. VME poor.
Integration density	Counting room: space available, High density useful for performance and cost.
Hot Plug	Used. Helps but not essential. Does not improve the amount of physics data (a failure of a single I/O card) makes it impossible to use the rest of the data.
Costing	Not more expensive, MTCA quite cheap
xTCA strong points	Many mandatory features (e.g, Ethernet controlled), good system standard. BAD: Complexity
backplane	Performance not too critical for MTCA. Not OK for certain read-out tasks, use front / rear panel I/O per card. MTCA protocol – full custom, ATCA – no plan yet. Ethernet complicated (TCP/IP in FPGA difficult and Ethernet is not deterministic), use full custom protocol unless commercial cards have to be used. ATCA cards may get equipped with a COMexpress / Zync controller

Details about LHCb

Current situation:

- 9U cards (VMEbus format) with commercial credit card computer used to process data coming from the detector before going into HLT

Future (to be installed in LS2):

- The electronics on the detector will do much of the job of the 9U cards
- LHCb will use a farm of PCs in order to do the 1st level event building
- Architecture: Detector -> GBT -> 500 units of the PCIe card and maybe 300 for ALICE (Gen3, x16) -> PC
    - Crucial: 100 Gbit/s from PCIe card to PC
- Custom H/W: PCI40 (Prototype)
- Main motivation for the design: PC gives a lot of flexibility with respect to N/W technology (Ethernet, Infiniband or Intel Omnipath) and cheap memory.
    - ATCA solution: backplane would not have been used (too little bandwidth for a N->1 system). Also many other ATCA features would not have been used.

xCTA feature	LHCb position
Redundancy	Full redundancy not possible (FE links are not redundant). PCs should be reliable enough. PCs have redundant PSU and cooling.
Board space	One FPGA and optic transceivers fit onto a PCIe card
Cooling	Use a server that is optimized for graphics GPUs. (PCI40: ~100W)
Integration density	Not an issue as space was available.
Hot Plug	One warm spare per rack, In case of a failure of a PCIe card: replace the entire PC
Costing	ATCA seen as expensive at system level
xTCA strong points	Cooling and PSU quality. PCs may be less reliable

Details about ATLAS

Current situation:
- 6U and 9U cards (VMEbus) used to process data coming from the detector before going into HLT

Future:
- VME: until LS3. Then FELIX will replace the remaining VMEbus systems
- Architecture:
  - Detector->FELIX->HLT
  - Detector->ATCA->FELIX->HLT
- Schedule
  - Few ATCA shelves installed in LS1 (CSC, Topo, FTK)
  - LS2: ATCA for NSW, L1 trigger, Lar (only trigger)
  - LS3: ATCA for L0/L1 trigger, Tile, FTK, Lar (readout)
  - Other: S/W on the basis of the FELIX

xCTA feature	ATLAS position
Redundancy	PSU yes, cooling under discussion, shelf manager yes, DAQ electronics: no
Board space	Trigger will need large card. Lar: ATCA carrier for 4 AMCs Tile: pre-processor FTK Topo CSC ROD General: all designs will have several FPGAs. Backplane not used a lot. Input and output via front panel
Cooling	Important!!! For big FPGAs (VME cannot do it). Power up to 400 W (FTK)
Integration density	Not an issue
Hot Plug	Not used (the system needs to be fully functional for ATLAS to work as it should)
Costing	Not that much on an issue; ATLAS will have 100-110 shelves
xTCA strong points	Cooling, card size, PSU, IPMI (powerful but complex)

Details about ALICE

Current situation:

- Detector->(VMEbus / in-cavern custom electronics)->DDL (optic)->PCI card -> PC
- 50 VME systems in use
- Still a lot of VMEbus heritage / culture

Future:

- VME will be kept for mechanics, cooling and power even for new develpments
- Some VME will remain beyond LS3
- Schedule
    - LS2: Upgrade for run 3&4. Read-out and store (pre-processed data) at 50 kHz (no trigger)
    - Need: Need: new F/E electronics (on detector), GBT links, and ~ 500 PCIe40 cards (same as in LHCb) as common read-out and data concentrator units (CRUs)

xCTA feature	ALICE position
Redundancy	Small loss of data is acceptable, no need to spend money.
Board space	MTCA card size sufficient. ATCA only proposed for compatibility with LHCb
Cooling	
Integration density	Space was available. Trigger & timing distribution would have been easier
Hot Plug	Would not have helped a lot
Costing	Chosen solution cheaper due to read-out requirement
xTCA strong points	Architecture: better separation of DAQ and (FPGA) electronics -> clearer organization, worse interplay of components