# SoC workshop CERN - Common Minutes - June 2019

https://indico.cern.ch/event/799275

**Notes:**

- These minutes were collected by P. Zejdl, M. Twomey, M. Fontes Medeiros, and M. Dobson.
- They provide a record of the questions and answers following the presentations. For the presentations themselves, please, refer to the indico timetable.

## Wednesday - 12th of June:

Morning session:

- Conveners: Revital, Ralf
- R. Spiwoks, Intro & motivation:
  - No questions
- K. Aernoudt, Xilinx R&D, Gregory Donzel, Xilinx FAE (Avnet Silica):
  - Q: Where is the Versal Prime product positioned?
    A: The smallest Versal Prime device has been announced with 352k Logic Cells (up to 985k Logic Cells). This may position this family as a mid-range platform (similar to Kintex UltraScale/UltraScale+).
  - Q: What is the time scale for obtaining first samples of Versal devices?
    A: General Engineering Samples scheduled for end of Q3CY2020. Production parts for mid 2021. For Early Access, please contact your Avnet Silica FAE.
  - Q: Versal doesn't have GPU, what is the reason?
    A: The first devices announced are without GPU. It is not a successor to Zynq Ultrascale+. There might be future series with GPU.
  - Q: For Versal, do you take the programmable logic from UltraScale+?
    A: This is slightly different, see https://www.xilinx.com/support/documentation/data_sheets/ds950-versal-overview.pdf .
  - Q: Can you say something about the Versal vs UltraScale+ product lines?
    A: Versal is going to be a parallel/separate product line to UltraScale+. We expect Versal to replace most of the Kintex/Virtex family applications. We still recommend Zynq 7000, if a low end SoC is needed. We are still shipping Spartan-3. There is no short-term risk. We are shipping now Virtex UltraScale+ with HBM (High Bandwidth Memory).
  - Q: Will Versal have HBM as well?
    A: It cannot be disclosed yet. Only information about Prime and AI Core Series is publicly available.
  - Q: Can you say something about programmable logic, size, ... ?
    A: Everything about Zynq UltraScale+ is a copy paste from UltraScale+, we have various sizes. The scalability of Zynq is actually in the size of the programmable

logic. Information is available and detailed at https://www.xilinx.com/support.html#documentation .

- K. Chaplin, Intel R&D:
  - No questions

Afternoon session:

- Conveners: Ralf, Revital
- F. Meijers, Evolution of CMS online system and role of SoC
  - Comment about the low usage of the board fabric interface.
  - It was commented that FireFly fibre connectors were chosen to avoid different solutions for different sub-detectors with a large range in data throughput per back-end board (avoid different designs).
  - Q: Why were the Fat Pipes on the backplane not used for the DAQ aggregation? A: Not sufficient for the sub-detectors with large data throughput per back-end boards (e.g. inner Tracker, HGCAL), but would be OK for muons for example. Today back-plane links limited to 100Gb/s.
  - Q: Are RTM's used? A: Depends on the sub-detector, it is being considered for Tracker to split the data between trigger and DAQ and for the CSC muons for a memory board to be used to store large look-up tables.
- R. Kopeliansky, ATLAS SoC - Review
  - Q: For the HTT there is a large number of ATCA blades - are they all the same? A: There are two types: the AMTP (Associative Memory Tracking Processors) and the SSTP (Second Stage Tracking Processors), where one HTT unit has several AMTPs and single SSTP. There are a few other systems in TDAQ that have more than one type of blade, but most of them are the same.
  - Q: Can you tell more about the real-time triggering in the ARM CPU planned? A: For now this is being investigated as an option ...
- R. Spiwoks, ATLAS MUCTPI and System on Chip
  - Q: Also work done on Altera? A: No, only on Xilinx
  - Comment: "hardware compiler" on slide 7, similar goals are addressed by the Cheby project (BE-CO), see https://gitlab.cern.ch/cohtdrivers/cheby/wikis/home. It aims at defining a file format to describe the HW/SW interface (the memory map), and a set of tools to generate HDL, drivers, etc.
- H. Boukabache, Fault Resilient Design for 28nm ZYNQ SoC based Radiation Protection Monitoring System Fulfilling Safety Integrity Level 2
  - Boot mechanism is interesting from a reliability point of view.
  - Q: Did you check the radiation hardness of the processor part of the SoC? A: We do not rely on the processor for the functioning of the monitoring. But we will check the hardness of the processor in the future.
- P. Van Trappen, SOC technology for embedded control and interlocking within fast pulsed systems

    - The presentation showed a nice table about the requirements for SysAdmins: page 11. It was mentioned in the discussion that IT are considering redoing part of the PXE/TFTP server infrastructure during 2019.
    - Concerns were raised on using a 32-bit platform instead of 64-bit for long-term support.
- T. Gorski, The CMS APx Platform using SoC Devices
    - A question from the audience led to a discussion about the reliability of Xilinx support for DDR memories since it is possible that they can evolve with time. Xilinx guaranteed the support for the low-end FPGAs but for the DDR memories.
- L. Ardila, CMS R&D for Phase-2 Tracker Back-end Electronics
    - A question was raised from the audience if the IPMC license is from CERN or a private one. They replied that they were using a private one.

## **Thursday - 13th of June:**

Morning session:

- Conveners: Revital, Marc
- P. Zejdl, K. Mor, PetaLinux - Presentation + Tutorial:
    - Q&A on first part (presentation):
    - Q: Yocto - does building the root FS work on new zynq US+
      A: Not tried yet.
    - Q: Does PL get reset if PS gets reset?
      A: The default boot loader will setup PL by default - this is a simple project which can be modified - also PL can be powered down as it is a separate power island.
    - Q: How does Xilinx handle upstream patches - do they back port into the kernel?
      A: Xilinx/Avnet (M. Hennersperger): standard release cycles - in rare cases they back port. Mainline release is recommended - but some drivers missing.
    - Q: Is the mainline kernel available to be plugged into the petalinux framework?
      A Xilinx/Avnet (M. Hennersperger): It is possible to overwrite default - can pull the kernel from alternative sources, again drivers missing.
    - Q&A on second part (demonstration):
    - Q: Workflow - to use git, they use tcl files instead of gui. Is there a standalone Xilinx command line?
      A: Petalinux and Vivado very tightly integrated. Not sure if older HDL files.
      A: Ruckus from slacktivism.
    - Q: HDF errors?
      A: Older files have failed in compile.
    - Q: Update issues - will things stop working (TCL, HDF)?
      A (Xilinx/Avnet): Petalinux okay to start - go directly to Yocto which is git aware. Add everything to the yocto recipes (include Xilinx). Nothing guaranteed - HDF files being deprecated. HDF to device tree files - Include meta-xilinx-petalinux layer.
- P. Papageorgiou, CentOS: on Xilinx Zynq US+
    - Q&A on first part (presentation):

    - Q (from IT): Would you be interested in RPM's rebuilt by IT for ARM?
      A: Yes!
    - Q: Benchmark for GCC on zynq and ARM server. Anything for PC that can be cross compiled?
      A: They will look into it.
    - Q: Why not cross-compile everything?
      A: Some packages like python which produce low-level code don't work. Qemu for cross-compile is terrible in terms of slowness of speed. Maybe we can provide an ARM server for the experiments - pool the resources, create a CentOS SIG for SoC @centos.org. ARM servers in general, are not useful in price/performance ratio, but for SoC, that may need to be re-evaluated.
    - Q&A on second part (demonstration):
    - Q: How much memory does CentOS take?
      A: 2 GB, works with smaller systems.
    - Q: How big is the root filesystem?
      A: 1.3 GB.
- M. Wittgen, CentOS cross-installation/compile and performance:
    - No questions

Afternoon session:

- Conveners: Ralf, Diana
- Minutes responsibles: Marc, Frans, All
- T. Vanat, Caribou - a versatile DAQ June 13, 2019
    - Q: Why do they need patches for I2C and SPI?
      A: ASIC uses the same clock as the SPI signal. Normally it would turn off, they need to keep it running so ASIC continues to run.
    - Q: How do they keep the HW Description synchronised for PS and PL
      A: Done by hand - change the bitstream for the FPGA used. The address space is kept the same for different device configurations.
    - Q: Typical throughput?
      A: No formal load test. 10's of MB/sec. DMA to be implemented. Currently use polling to collect data.
    - Q: Continuous integration?
    - A: Only for software, not for CDL/HDF.
- M. Wittgen, RCE based readout and evolution June 13, 2019
    - Q: How do they boot?
      A: Need SD card with a minimum boot. Geographical versus MAC network information via DHCP.
    - Q: Why no IPv6 support?
      A: RTEMS was using an outdated network stack from openBSD
- D. Gastler, The APOLLO Platform - Use of SoC Devices in ATLAS and CMS
    - Q: If CMS and ATLAS chose different "common" OS what happens?
      A: Each group gets delivered their common OS configured boards.

    - Q: IPBus - is it available for testing?
      A: Currently in testing - to be made available soon.
    - Q: Is IPBus DMA or memory mapped?
      A: Currently memory mapped
    - Q: Why commercial mezzanine card?
      A: Enclustra - give you all files for producing them - must buy them from them as long as they are in production - once it isn't produced, you can make your own.
- D. Miller, Zynq MPSoC in the gFEX hardware trigger in ATLAS
    - Q: Zynq in data path, i.e. calculation of missing energy?
      A: PL used for time constrained calculations. Extra calculation not in the trigger path on the PS.
    - Q: Slide 15 - 2 links - separate DCS and RunControl?
      A: Don't need 2 links, used to test for data packet loss. DCS is 100kbs. QOS via different links.
- E. Smith, Exploring the use of a Zynq MPSoC for machine learning inference in hardware triggers
    - Q: Why use FPGA versus GPU or CPU?
      A: Fixed latency, low power GPU Mali-400 - not very powerful.
    - Q: Implement tensor flow inside chip?
      A: Model is used offline to build the block.
    - Q: How many parallel algorithms can be run-in the FPGA?
      A: Not sure. Multiple processes to speed up a single algorithm. Not sure about multiple simultaneous algorithms.
    - Q: DNNdk - where is it running?
      A: In PL. Some functions can be run in the Application Processor Unit (APU) / Real-time Processor Unit (RPU).
    - Q: Is this usable for trigger purposes?
      A: Not yet, maybe in a few years.
- O. Kepka, The SoC-based Readout Drivers of the ATLAS Pixel/IBL detector
    - Q: FW registers exported to C++?
      A: Follow up offline.
    - Q: What mechanism to synchronise data between PS and PL?
      A: Continuous polling.
    - Q: What debugging is available?
      A: GDB and gcc printing routines.
- F. Carrio Argos, A SoC module for the ATLAS Tile Calorimeter PreProcessors at the HL-LHC
    - Q: What is the total amount of data downloaded by the TileCom to the detector electronics at startup?
      A: Not yet calculated. In order to gain time, the data are down-loaded and only in the end it is checked if it was successful.
- V. Andrei, ATLAS TREX - VME digital modules using Zynq MPSoC

- Q: Updating SD cards over network?
  A: Not sure if rewriting SD card via network is viable.
- Q: Comment - copy (scp) image to SD card
  A: Will look into it.

## **Friday - 14th of June:**

Morning session:

- Convener: Revital
- J. Wyngaard, R. Knowlton, Balena as a deployment stack - Presentation + Demo
  - Q: Each application would run under hypervisor?
    A: Yes, each application would have their own docker container.
  - Q: What is the performance hit?
    A: Not sure, most likely not for real time applications.
  - Q: How does the docker work with things like memory mapping?
    A: Balena handles hooks using their own “docker” engine which allows direct exposure to the hardware.
  - Q: git deployment is automatic, how much control?
    A: Full control of the entire pipeline.
  - Q: If you have many endpoints how do you know the deployment was successful?
    A: Query the devices and get a hash of the commit to make sure they are all the same.
  - Q: Can you see what has been run over time?
    A: Not sure.
  - Q: How do Balena and openstack compare?
    A: openstack is a hypervisor and an image, Balena has no hypervisor, just image engine.
- S. Schlenker, SoCs and Control System Integration
  - Q: Does it really reduce software development effort?
    A: Rebuilding using quasar should be fairly simple. Have developer groups to work on specific boards.
  - Q: Not clear separation between two domains?
    A: DCS just needs to know that a change has been made, SCADA can generate new interfaces. Server provides common interface.
- P. Nikiel, OPC-UA on Zynq: Walk-through, experience and a demo - Tutorial part I
  - Q: How is this linked to programmable logic?
    A: To be shown.
  - Q: How does this become a WinCC OA data point?
    A: To be shown (in the part 4 - using fwQuasar).
  - Q: Where does the switch object get instantiated?
    A: In the config.xml file.

    - Q: How fast can you retrieve data? 10Hz?
      A: OPC-UA was measured to have about 15% overhead on pure media bandwidth (example: on 100MB/s network you pass 80MB/s of data).
    - Q: Framework difference FESA vs quasar?
      A: They serve a similar purpose but are different. quasar is much leaner and OPC-UA is an industry standard. FESA is bound to a timing model while quasar is not related to timing at all. quasar is simpler to use on embedded targets and has built-in support to interface to WinCC OA.
    - Q: How to synchronise without timing?
      A: Don't actually want to synchronize to LHC timing.
    - Q: OPC UA has subscription model - has this been implemented?
      A: Yes, UAexpert uses subscription. Pure Publish/Subscribe available. Currently no benefit to DCS Scada system.
- P. Nikiel, OPC-UA on Zynq: Walk-through, experience and a demo - Tutorial part II
    - Q: AXI/Wishbone where does it show up? in /sys?
      A: These are memory mapped, or custom kernel driver. UIO has a kernel and user space section. can use interrupts.
    - Q: What is used to create the SD card image.
      A: Petalinux.
    - Q: Mechanism behind data collection - polling?
      A: Yes, all methods supported: pub/sub, request-reply (for which it can go via cache read or device read) etc.
    - Q: Why do you have to rebuild each time? Why can't you do a partial recompile?
      A: I'm using image rebuild every time because it is faster for me at this workshop - with Yocto you actually can create RPMs etc.
    - Q: Could this be used in other projects?
      A: Yes, quasar is used in multiple projects (about 20 in active development or production), including embedded, FPGA, server-grade etc...

<u>Afternoon session:</u>

- Conveners: Diana, Marc
- G. Donzel (Avnet Silica), M. Hennersperger (Avnet Silica), Xilinx Cybersecurity offering in Industrial Internet of Things (IIoT)
    - Q: Anti-tamper support? Controlled technology, can't be purchased - SECMON?
      A: Not sure - to be checked. Same for PUF.
- K. Chaplin, Intel/Altera - Secure Device Manager for Stratix10 & Agilex,
  Not able to attend, white paper attached to indico page, please contact Kris.Chaplin@intel.com.
- S. Lueders, CERN IT/Security
    - Q: Why should I bother?
      A: Lateral movement. What does a firewall do? - let traffic through, otherwise cut the cable. Once you are in, you can have lateral movement.

    - Q: Local DHCP server?
      A: This will provide IP address to those.
    - Q: For SoC, why isolation? Gateways?
      A: Network is a better protection - different layers. Every layer should have protection: device protection, gateway, WAN.
- T. Oulevey, State of Linux distributions at CERN and developments in the community by CERN IT Linux Team
    - Q: Can all needs for different groups be concentrated?
      A: Concentrate the workflow, e.g. everybody rebuild from common source (gitlab). Common workflows. Provide common packages from vendors in a single place. Common toolchains.
    - Q: What can be done to make these systems more likely to be connected to a technical controlled network?
      A: Might be possible - what needs to be done for certification - vendor participation.
    - Q: RedHat removed support for older raid cards - can't boot.
      A: Community is providing patched kernels - CentOS-plus kernel. Ask to have features added back - with justification they will provide new kernel. Not currently supported by IT.
- D. Scannicchio, M. Dobson, SysAdmin - Networking, Security & System-administration - Overview of ATLAS and CMS
    - Q: Can ClientID be implemented?
      A: Most DHCP servers support it. It would be possible; database (LanDB) would need to be updated.
    - Q: Shelf addresses?
      A: We would need to set this up - not currently supported in ATLAS.
    - Q: Where do we get MAC addresses?
      A: IT can provide MACs for boards produced by CERN or used at CERN. Very large pool, should request them early.
    - Common systems like OPC will need experts to steer development and provide expertise (I.e build systems). What happens in the future when the developers are gone, and these systems still need to be updated
    - Q: Why use common bus rather than front ports?
      A: Shared bus limits the number of physical connections needed to the outside world: not monitored by IT, controlled by experiment
    - Q: What about CAEN crates?
      A: Limited to only being seen on local experimental switch. Downside still seen by other devices on the switch.
    - Q: Does using ACL/Inter-domain sets seem okay for DCS?
      A: That is the way it is heading. Feedback via redmine/jira ticketing - a way will be found to share the info.
    - Q: Difference between ACL and inter-domain set?
      A: Access control lists are device lists and who they can talk to. Inter-domain set

allows communication between domain lists. Each list can contain multiple nodes.
    - Q: What would need to be done to connect a SoC directly to the ATCN?
      A: Do you really want that? Running an OS is supported by IT and the SysAdmins. CentOS - need to see the kernel issue (which one). CentOS7 kernel for x86 and aarch64 are not the same version. CentOS8 may be rebased to be the same.
    - Q: How do we maintain? Systems get frozen and may need to be hidden. Can we mitigate?
      A: Use common systems tool sets/OS/Firmware toolchains. Git build systems.
- Discussion on system issues
    - One member from the audience pointed out the difficulties of setting up the SoC environment. It was mentioned that we have different ways to configure the SoC; the tutorials/demos are easy to see but in practice nothing is ever that simple to configure.
    - A question was raised about the custom-made switch (by CMS) for the ATCA crates: Is the custom made switch IT compliant? For specific needs it is allowed but not managed by IT. Can it be mass produced? The board design would be available and could be produced. And how fast is the production? That depends on many facts, least of which is the fact that the board is not yet tested!
    - From the audience a strong point was mentioned to the operations side. It was mentioned that the software and hardware developers workflow are completely different. On hardware, once the board is completed and deployed there is not a strong sensibility about maintaining it in operations. On the other hand, from a software perspective it is constantly evolving in operations. This also raised the delicate point about who is going to maintain the SoC software infrastructure in the long term since the manpower is scarce. Additionally, the expertise of the firmware development gets easily lost as time passes and it needs to be foreseen a mechanism to contemplate the future developments and prevent this from happening.
    - Question from the audience on what would it take to connect the SoC directly to the technical network? Various approaches were mentioned, through the use of gateways PC and also creating groups at the routing level with the use of ACLs. However, it was mentioned that if you need to communicate with more than 20 different devices additional configuration is needed since the network ACL size is not infinite.
    - From a security perspective it was clear that if the SoC is a security risk for any infrastructure it will get cut from the network.