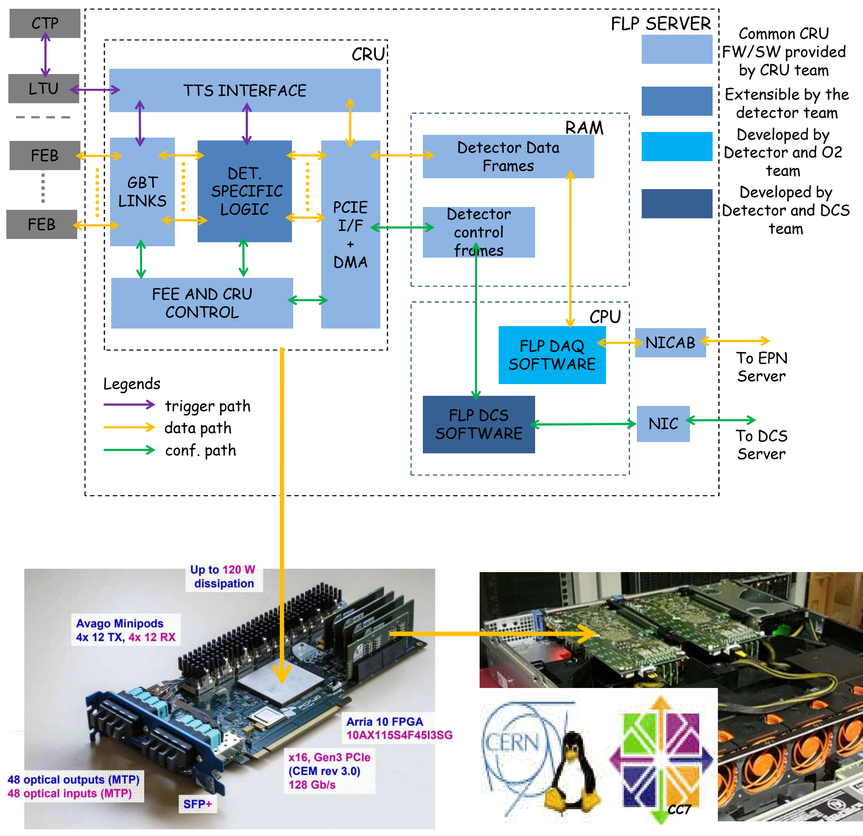
Design and implementation of custom DMA controller for the ALICE-CRU, to optimize data transfer reducing the CPU utilization

Sanjoy Mukherjee*, Jozsef Imrek, F. Costa, P.Boeschoten

CRU (Common Readout Unit) and PCIE-DMA:

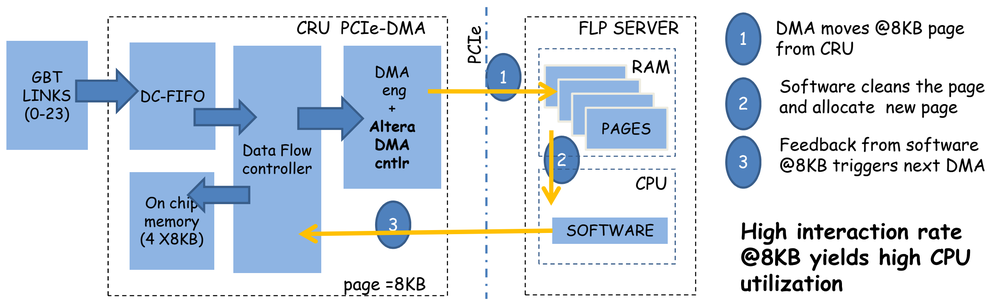
ALICE-CRU is FPGA based PCIe peripheral card which will be mounted on several FLP based servers of online-offline system (O2) to read out data of most of the ALICE sub-detectors.



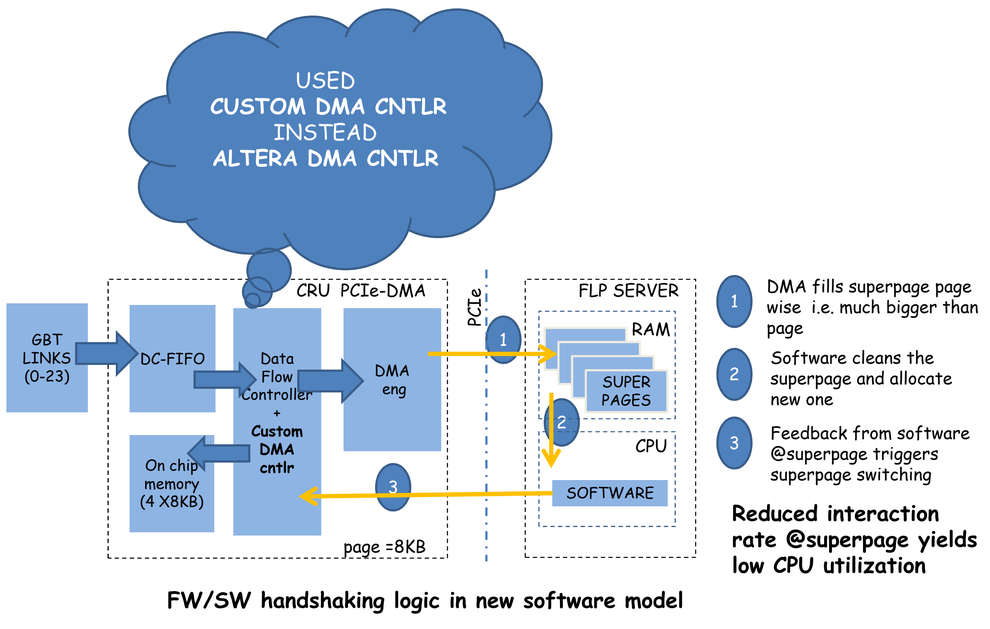
Potential challenge and solution:

During evaluation, the DMA performance was good enough in terms of **bandwidth utilization** but not in terms of **CPU utilization**.

Dedicated CPU to a single process hampers progress of others. The challenge is to **reduce the CPU utilization keeping the same performance**.



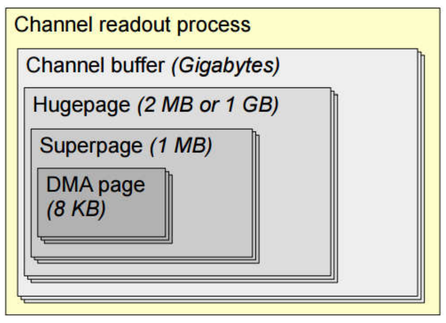
FW/SW handshaking logic in old software model



FW/SW handshaking logic in new software model

To reduce the high CPU utilization, a custom DMA controller has been designed to implement the superpage concept and optimize the handshaking logic in between software and firmware.

Memory scheme of new software model:



Note: "typical" sizes indicated

Channel buffer	A (typically large) buffer for DMA transfers from one DMA channel. Created by readout process by allocating hugepages.
Hugepage	A large CPU MMU page (either 2 MB or 1 GB on x86-64).
Superpage	A physically contiguous subdivision of a hugepage (1 MB multiple). It is passed to the driver, which will fill it with DMA pages without further intervention.
DMA page	The unit of individual DMA transfers used by the card (8 KB for CRU, configurable for C-RORC)

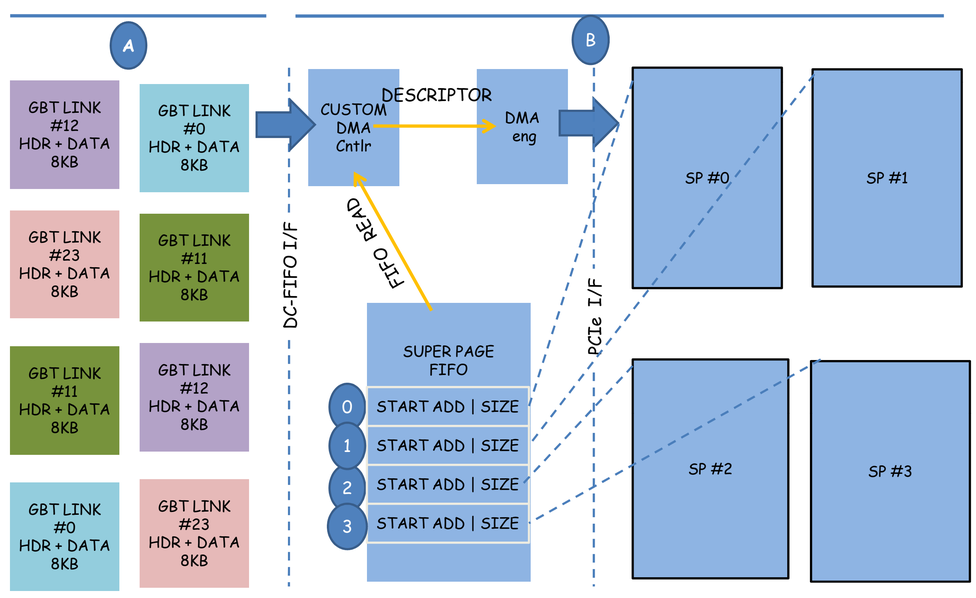
Poster: https://indico.cern.ch/event/608587/contributions/2614621/

Firmware implementation:

Two versions of custom DMA controller represents two memory models:

A. ALL LINKS ARE IN SAME SET OF SUPERPAGES.
B. EACH LINK HAS DEDICATED SET OF SUPERPAGES.

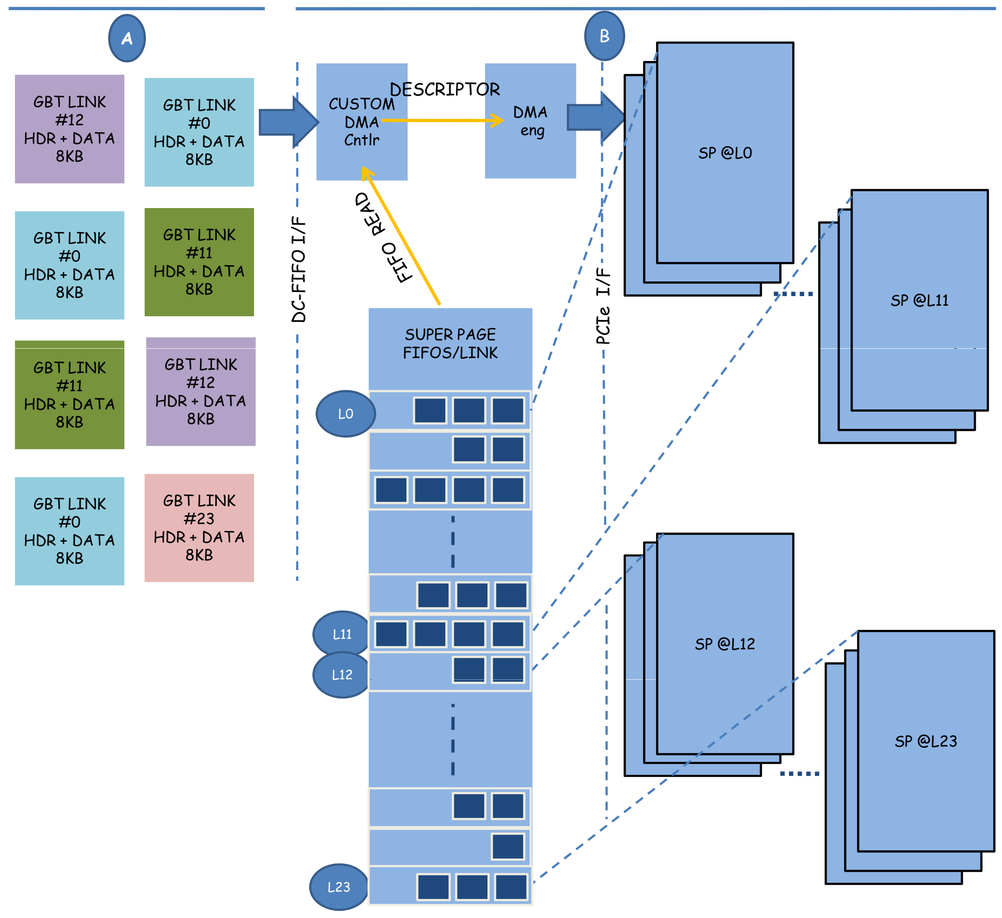
ALL LINKS ARE IN SAME SET OF SUPERPAGES:



A — 8 KB pages from different GBT links has been pushed towards DMA Engine.

B — Custom DMA controller pushed multiple links data into same superpage based on the start address of superpage in FIFO.

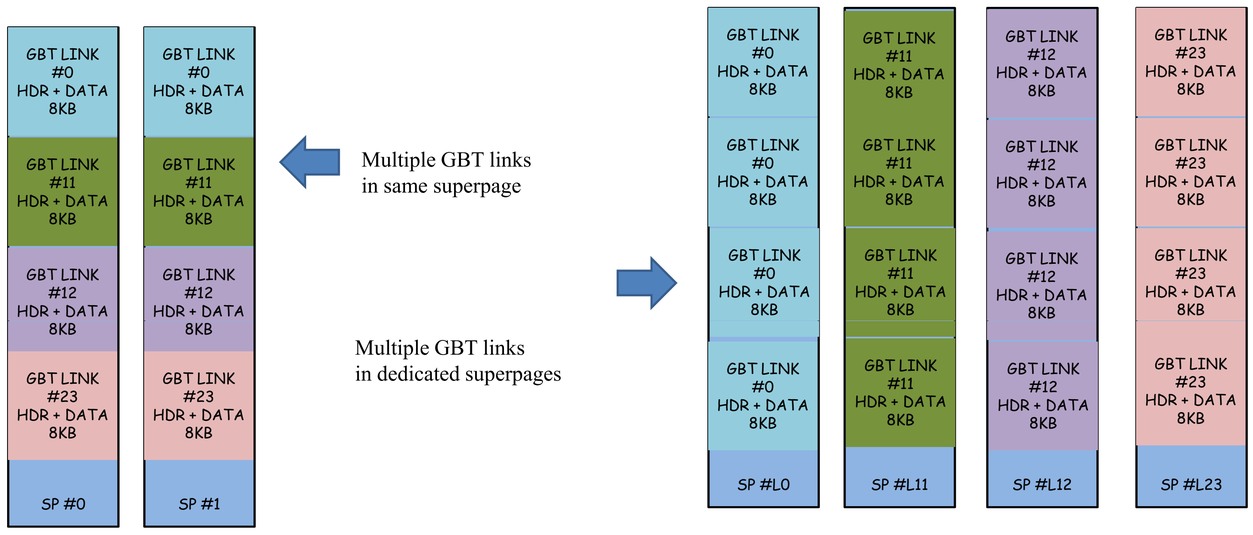
EACH LINK HAS DEDICATED SET OF SUPERPAGES:



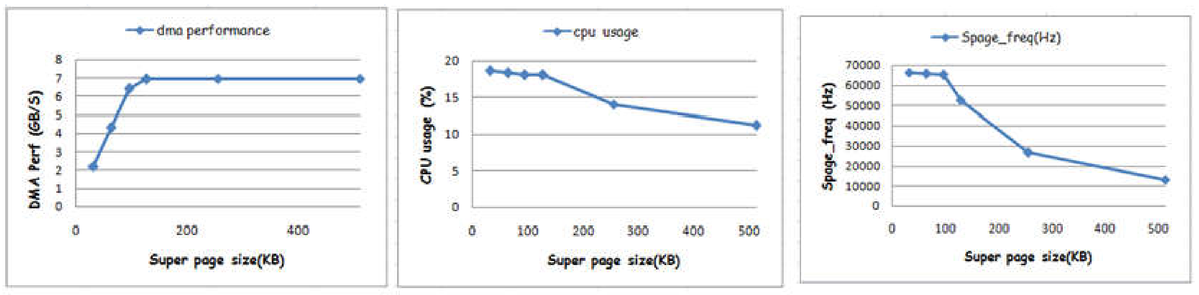
A — 8 KB pages from different GBT links has been pushed towards DMA Engine.

B — Custom DMA controller pushed each link data into dedicated superpage for that link based on the start addresses of superpages in FIFO.

Difference in memory model:



Result:



Topical Workshop on Electronics for Particle Physics (TWEPP 2017)
September 11-15, 2017
UC Santa Cruz, SCIPP 1156 ,High St Santa Cruz, CA 95064
Organized by SCIPP (Santa Cruz Institute for Particle Physics)
with support from CERN.