Hermes - A robust, low latency, optical link protocol for synchronous data transfer at commercial asynchronous line rates

Tuesday 21 September 2021 17:20 (3 minutes)

The Phase-2 CMS Level-1 Trigger and associated upstream systems consist of more than 20,000 25Gb/s optical links, transferring almost a Pb/s synchronously between different back-end processing nodes. The stable operation of these links is essential to avoid the injection of an erroneous signals into the trigger path, potentially leading to a flood of false triggers.

The Hermes protocol, implemented on Xilinx UltraScale+ FPGAs, provides this stability while operating at asynchronous, industry standard, line rates. The protocol design as well as the performance from extensive tests are presented here

Summary (500 words)

The Phase-2 CMS Level-1 Trigger and associated upstream systems must synchronously transfer data between many different processing nodes with 25Gb/s optical links. The data transmitted on these links must remain aligned to the LHC beam to avoid accidentally swamping CMS with a large number of false triggers due to misinterpretation of link data.

Synchronization between different processing nodes is achieved by splitting the link bandwidth into a synchronous part with a fixed number of data words per LHC orbit and an asynchronous part that is used to pad the link with filler words so that the line rate is within the optical transceiver Clock & Data Recovery (CDR) operating range. This approach requires that filler words can be reliably distinguished from data words even in the presence of a bit flip. The Forward Error Correction (FEC) methods used in industry to protect against bit errors (e.g. RS-FEC(528,514) for 25G Ethernet) are not suitable due to their latency. Instead a new approach is needed, which has led to the development of the Hermes link protocol. It combines a toughened protocol layer with rapid re-alignment to minimize downtime due to a failing link. Hermes has shown to be immune to single bit flips at BERs exceeding 1 in 10^{9}.

Link metadata, CRC checksums and even the alignment markers, used to synchronize to the LHC orbit, are transmitted through the Filler bandwidth so that the entire synchronous bandwidth can be dedicated solely to carrying the data necessary for detecting interesting physics signatures. Separation of Data from Control words is done utilizing the 64b67b encoding scheme, which defines a 3-bit Header alongside every 64-bit word. Apart from word characterization, the Header allows for a simple FEC to ensure that Data and Control words can be safely distinguished. In addition, Hamming (7.4) codes are used to ensure that the type of Control words, particularly those used as Fillers, can be reliably identified in cases of single bit errors.

So far, the Hermes Protocol Firmware have undergone detailed testing by transmitting data between several Phase-2 ATCA processors, all of which have been successful. No bit errors have even been observed. The adoption of the 64b67b encoding has raised concerns about the lack of DC balance, albeit small, introduced by the 3-bit Header, since, apart from the scrambled 64-bit words, the Header will always have disparity of +1 or -1. To counter this the Header toggles between two polarity modes.

The protocol has been implemented on the GTH and GTY transceivers used by the Xilinx UltraScale+ FPGA family and has been tested extensively. The protocol design as well as results on its performance are presented here.

Primary authors: ADAMIDIS, Kosmas (University of Ioannina (GR)); ILES, Gregory Michiel (Imperial College (GB)); FOUNTAS, Costas (University of Ioannina (GR)); HOWARD, Alexander (Imperial College (GB)); BESTINTZANOS, Ioannis (University of Ioannina (GR)); Prof. MANTHOS, Nikos (University of Ioannina (GR)); Dr PAPADOPOULOS, Ioannis (University of Ioannina (GR)); VELLIDIS, Konstantinos (National and Kapodistrian University of Athens (GR)); MALLIOS, Stavros (CERN); WILLIAMS, Tom (Science and Technology Facilities Council STFC (GB))

Presenter: ADAMIDIS, Kosmas (University of Ioannina (GR))

Session Classification: Posters Optoelectronics and Links

Track Classification: Optoelectronics and Links