New Packet Fragmentation for
S-Link to Gigabit Ethernet Adapter

Joaquim E. Neves[1,2], Niko Neufeld[1], Richard Jacobsson[1] and Beat Jost[1]

[1]***EP-Division, CERN, CH-1211 Geneva 23, Switzerland***
[2]***Escola de Engenharia da Universidade do Minho, P-4800-058 Guimarães, Portugal***

ABSTRACT

High Speed Interconnect (HSI) systems are essential in order to transport the large amount of data generated by the detectors, to storage devices in high energy physics experiments.
The integration of a new packet fragmentation mechanism on the S-Link to Gigabit Ethernet Adapter will permit not only the physical interconnection between involved interfaces, but also the conversion of data formats and transmission protocols.
This paper presents and discusses simulation results of the S-Link to Gigabit Ethernet Adapter, with and without the new packet fragmentation mechanism, based on the data format adopted in the LHCb experiment.

Keywords: *Gigabit Ethernet, LHCb, Packet Fragmentation, S-Link.*

1. INTRODUCTION

In the Large Hadron Collider experiments, such as the LHCb [1,2] which will take place at CERN, the transmission between detectors, connected to the Front-end Motherboards (FEMB), and storage devices, connected to the Read-out Motherboards (ROMB), is supported by Data Acquisition and Control Systems.
Usually the throughput capacity of these transmission systems is asymmetric, since a large amount of information has to be transferred from FEMBs to ROMBs, while only a small volume of data is needed in the opposite direction for the control and management of system.
The LHCb experiment integrates the Gigabit Ethernet (GbE) [3] as the link technology, to connect the output of the FEMB electronics boards to the input of the ROMB Sub-Farm Controllers (SFCs), and the S-Link [4], as a standard interface between the FEMB and the Data Acquisition (DAQ) system.
Since the Readout Network already supports the GbE interface, an S-Link card for the FEMB, based on GbE, has been designed and will be implemented within a short time. This module, designated S-Link to Gigabit Ethernet Adapter (SGbEA), converts physical interfaces, data formats and transmission protocols between S-Link and Gigabit Ethernet interfaces.
Next, in Section 2, it is discussed the technical specification of the S-Link to Gigabit Ethernet Adapter, while Section 3 presents simulation results of its throughput capacity for two different modes of operation and packet lengths. Section 4 summarizes the main conclusions of this paper.

2. TECHNICAL SPECIFICATION

As the name specifies, the Gigabit Ethernet is able to support the transmission of variable length data packets over a standardized copper or optical medium with debits lower than 1 Gbit/s on the physical interface.
On the other hand, the S-LINK is defined as a simple FIFO-like user interface (at both ends of the transmission link), which remains independent of the technology used to implement the physical link and provides the transfer of event data and control words, error detection, optional flow control and test facilities [5].
The S-LINK specification describes the interface between the FEMB and the Link Source Card (LSC) and the interface between the Link Destination Card (LDC) and the ROMB, in either simplex or duplex version.
In the simplex version, since there is no communication path from the LDC to the LSC, the transmission is unidirectional, while, in the duplex version, the return channel between the LDC and the LSC allows the transmission of flow control commands from the ROMB

to the FEMB. In both versions a single, high-density, 64-pin connector is used to connect the FEMB to the LSC and the LDC to the ROMB, allowing, per clock cycle, the transport of a 32 bit word at the frequency of 40 MHz, which corresponds to a throughput of 1.28 Gbit/s.

S-Link to Gigabit Ethernet Adapter

There is more than one solution for the implementation of the Front-End Multiplexers (FEMs) and the Readout Units (RUs). The S-Link to Gigabit Ethernet Adapter is a possible solution to interconnect these systems, since it implements the Ethernet functionalities of the Physical and Media Access Control (MAC) Layers, and the S-link specifications. The Architecture of the LHCb Data Acquisition, based on SGbEA, is depicted in Figure 1.

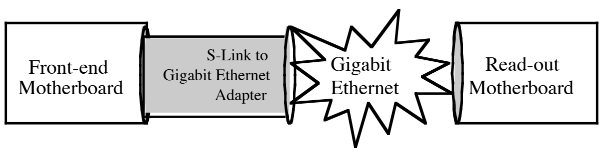

Figure 1 – Architecture of the LHCb Data Acquisition System, based on S-Link to Gigabit Ethernet Adapter.

The SGbEA has to generate the Start Of Packet (SOP) and the End Of Packet (EOP) signals for the MAC device. It has also to stop and restart the transmission when receiving watermark flags from the MAC or S-link FIFOs. Optionally, it can process the fragmentation of the S-link Packets on Ethernet frames, as is explained below.

Operational Modes

Since within the LHCb communication links between FEMs and RUs, the data can be transported with two different formats, the SGbEA supports two operational modes: Short Packet Mode (SPM) and Long Packet Mode (LPM).

In SPM, the length of the packets generated by the FEMB is variable, but always smaller than the maximum length of the Ethernet MAC frames. On the other hand, since the header of the LHCb data format already includes the header of the MAC Frames, no protocol conversion is necessary and the S-link packet is forwarded directly from the S-link to the Gigabit Ethernet.

In LPM, the FEMB can generate data packets ranging in length from 52 Bytes to 32 Kbytes.

In this operational mode the length of the S-link packet can be higher than the maximum length of the Ethernet MAC frames. For this reason, the header of each S-link packet, which contains a field with its length, has to be memorized in the SGbEA RAM, before calculating the number of fragments into which it must be split in order to be transmitted. In this mode, the SGbEA inserts, within each fragment header, the Type/length of the frame on the Ethernet header field, together with the number of fragments and the current fragment number of the S-link packet.

Optionally, in both operational modes, the Ethernet Sources and Destination Address can be generated by the FEMB (on a packet-by-packet basis), or can be previously inserted in the SGbEA Registers, by the control system and then transferred to each packet fragment header

Physical Implementation

The first prototype of the SGbEA module has been implemented on a small daughter board over a PCI Mezzanine Card (PMC), according to the IEEE Common Mezzanine Card standard [5]. Since the next version of the SGbEA will already be connected to the FEMB, the new packet fragmentation can be fully tested, at the maximum S-link throughput, without the present constraints of the PCI Bus.

3. SIMULATION RESULTS

A model of the SGbEA Module, presented in Figure 1, has been developed for simulation purposes, together with a S-link packet generator.

Packets with different lengths and data formats were generated by this model at the S-link interface, in order to be processed by the SGbEA, described on VHSIC Hardware Description Language (VHDL), on the VisualHDL platform, supported by a Computer-aided Engineering (CAE) system.

Since preliminary results of the SGbEA simulation, with Short and Long Packet Modes, have already been reported [7], the performance of the SGbEA with and without the new packet fragmentation will subsequently be presented.

In both cases the S-link overhead, introduced by the control words used for signaling the beginning and the end of the LHCb data packets, was reduced to a minimum: one word for the start and another one for the end of packet.

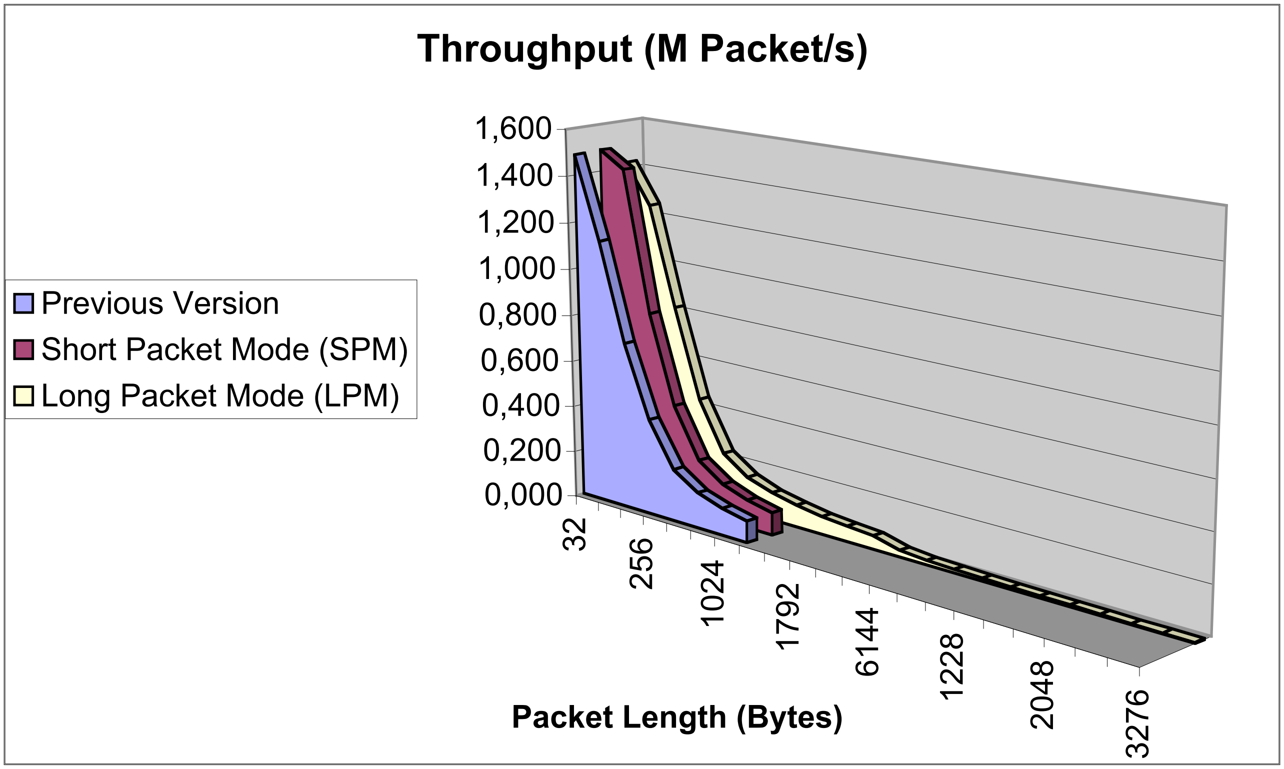

Figure 2 – Maximum throughput achieved with the previous version of the SGbEA and with the proposed Short and Long Packet Modes, for different packet length.

In another interface, at the Ethernet side, in addition to the MAC overhead, a minimum inter frame gap of 12 bytes was been guaranteed.

Maximum Throughput

Figures 2 reports the results of the maximum throughput achieved by SGbEA, for different packet lengths, with SPM, LPM and without packet fragmentation.

It is interesting to note that, in the previous version, without packet fragmentation, for a 64 byte packet, the reference value for packet length on several interfaces of the LHCb data acquisition system, the maximum throughput reached by SGbEA is approximately 1,1 M Packet/s, a value very close to the trigger rate of those interfaces.

On the other hand, with the proposed packet fragmentation, the maximum exceeds 1,4 M Packet/s with SPM, and 1.2 M Packet/s with LPM, which means that not only SPM but also LPM can be used with these packet lengths in the interfaces with a trigger rate of about 1.1 MHz.

For longer packets, the decrease of the throughput is proportional to the increase of packet length, as the overhead introduced by the fragmentation process becomes negligible. It should also be noted that, both in SPM and without packet fragmentation, the length of the LHCb data packets has to be lower than the payload of the Gigabit Ethernet MAC frame.

Relation with Ethernet Capacity

Figure 3 shows the percentage of transmission capacity, at the GbE interface, which could be reached with the previous version of the SGbEA, the Short Packet Mode proposed and the Gigabit Ethernet capacity, for different packet lengths.

As the figures show, due to the overhead of the Ethernet frames, the transfer rate achieved by packets lengths with a size to the order of tens of bytes, is clearly lower than the transmission capacity, both in the S-link and in the GbE interfaces.

In contrast, for longer packets with over a hundred bytes, the transfer rate begins to be limited by the effective GbE capacity. The Figure also shows that the proposed SPM absorbs the entire GbE capacity, independently of the packet length, while the throughput achieved by the previous version remains a little distant of that value, in particular for short length packets.

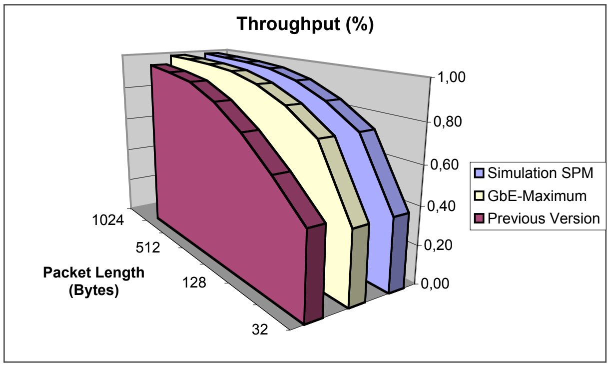

Figure 3 – Relationship between the throughput, achieved with the previous version of the SGbEA, the Short Packet Mode proposed and Gigabit Ethernet capacity, for different packet lengths.

Figure 4 shows the percentage of transmission capacity, at the GbE interface, which could be reached with the previous version of the SGbEA, the Long Packet Mode proposed and Gigabit Ethernet capacity, for different packet lengths.

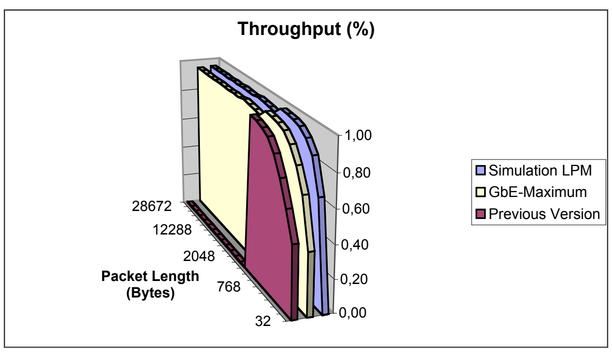

Figure 4 – Relationship between the throughput, achieved with the previous version of the SGbEA, the Long Packet Mode proposed and Gigabit Ethernet capacity, for different packet lengths.

As is depicted in Figure 4, the effective occupancy of the Ethernet payload increases with the length of the packet with short length packets (as was observed in the SPM simulations), and remains constant for longer packets. Figure 4 shows that only the proposed LPM is able to transfer packets with lengths longer than the payload of the Gigabit Ethernet MAC frame, and the GbE occupancy is almost filled.

4. CONCLUSIONS

A new packet fragmentation mechanism, for the S-Link to Gigabit Ethernet Adapter, has been developed to be used on LHC experiments at CERN.

As the simulation results show, this technique suppresses the drawbacks of the previsions of this system, which did not process any kind of packet fragmentation of S-link data.

The new version of the S-Link to Gigabit Ethernet Adapter, with this packet fragmentation mechanism can transmit Short and/or Long Packets over the LHCb Data Acquisition System, based on Gigabit Ethernet.

5. REFERENCES

[1] "LHCb Online System - Data acquisition and experiment control", Technical Design Report, CERN/LHCC 2001-40, December 2001. http://lhcb-comp.web.cern.ch/lhcb-comp/General/Publications/onlinetdr.pdf

[2] T. Nakada, "CP Violation", CERN Academic Training lecture series, Geneva, Feb 97. http://lhcb-doc.web.cern.ch/lhcb-doc/presentations/B-physics%20and%20CP%20violation/lecture/cpviol.htm

[3] "IEEE Std Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications", *IEEE - Std Association*, Piscataway, NJ 08855-1331, USA, 2002. http://shop.ieee.org

[4] O. Boyle et al., "The S-LINK Interface Specification", ECP-Division, Geneva, March 97, http://hsi.web.cern.ch/HSI/s-link/spec/spec

[5] H. C. van der Bij et al., "S-LINK, a Data Link Interface Specification for the LHC Era", *IEEE Transactions on Nuclear Science*, Volume: 44 Issue: 3 Part: 1, Anaheim, CA, USA, June 97.

[6] "IEEE P1386 Standard Mechanics for a Common Mezzanine Card Family: CMC", *Institute of Electrical and Electronics Engineers*, Piscataway, NJ 08855-1331, USA.

[7] J. E. Neves et al.,"S-Link to Gigabit Ethernet Adapter: - New Frame Segmentation for LHCb Data Acquisition System", Conference on Computing in High Energy and Nuclear Physics - CHEP03, La Jolla, CA, USA, March 2003.