Subject: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by Walter F.J. Müller on Tue, 14 Sep 2004 07:54:59 GMT
View Forum Message <> Reply to Message

During the 2nd FutureDAQ workshop (all talks in document management) some questions came up during E. Denes's talk on PCIe-AS. On the Intel PCIe website is a very good and comprehensive tutorial talk given at the PCI SIG Developers Conference 2004, see

ftp://download.intel.com/netcomms/as/devcon_as_overview.pdf

On 116 slides, many aspects of the protocol are described in quite some detail.

Now to a specific question raised during the talk:

Q: Is there a CRC protection for the address header ?

A: Yes, see on page 14 of the tutorial. The AS header is 64 bits, which include
7 bit header CRC
5 bit turn pointer
7 bit PI number
31 bit turn pool
1 bit direction (forward/backward routing)

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by David Slogsnat on Wed, 15 Sep 2004 12:59:02 GMT
View Forum Message <> Reply to Message

Walter F.J. Müller wrote on Tue, 14 September 2004 09:54During the 2nd FutureDAQ workshop (all talks in document management) some questions came up during E. Denes's talk on PCIe-AS. On the Intel PCIe website is a very good and comprehensive tutorial talk given at the PCI SIG Developers Conference 2004, see

ftp://download.intel.com/netcomms/as/devcon_as_overview.pdf

On 116 slides, many aspects of the protocol are described in quite some detail.

Now to a specific question raised during the talk:

Q: Is there a CRC protection for the address header ?

A: Yes, see on page 14 of the tutorial. The AS header is 64 bits, which include
7 bit header CRC
5 bit turn pointer
7 bit PI number
31 bit turn pool
1 bit direction (forward/backward routing)

This is true. However, things get more complicated when looking at the ASI specification:

-The Turn Pointer is not included in the header CRC.
-The final receiver of an AS Packet has to check the CRC. The intermediate switches may check it, but they don't have to.

One thing you can observe from this is that a packet may reach a wrong receiver due to an bit error in the Turn Pointer. The sender cannot be notified of this failed message transfer, since the Turn Pool in reverse direction does not lead to it.

Also, i wonder how the wrong receiver finds out that the packet was intended for another destination, since the header CRC check will not show an error!!!

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by Walter F.J. Müller on Wed, 15 Sep 2004 14:05:10 GMT
View Forum Message <> Reply to Message

David Slogsnat wrote on Wed, 15 September 2004 14:59This is true. However, things get more complicated when looking at the ASI specification:
-The Turn Pointer is not included in the header CRC.
-The final receiver of an AS Packet has to check the CRC. The intermediate switches may check it, but they don't have to.

One thing you can observe from this is that a packet may reach a wrong receiver due to an bit error in the Turn Pointer. The sender cannot be notified of this failed message transfer, since the Turn Pool in reverse direction does not lead to it.

Also, i wonder how the wrong receiver finds out that the packet was intended for another destination, since the header CRC check will not show an error!!!

How were all these header corruption issues solved in ATOLL, which uses a quite similar path addressing scheme ?

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by David Slogsnat on Wed, 15 Sep 2004 14:40:19 GMT
View Forum Message <> Reply to Message

Walter F.J. Müller wrote on Wed, 15 September 2004 16:05David Slogsnat wrote on Wed, 15 September 2004 14:59This is true. However, things get more complicated when looking at the ASI specification:
-The Turn Pointer is not included in the header CRC.
-The final receiver of an AS Packet has to check the CRC. The intermediate switches may check it, but they don't have to.

One thing you can observe from this is that a packet may reach a wrong receiver due to an bit error in the Turn Pointer. The sender cannot be notified of this failed message transfer, since the Turn Pool in reverse direction does not lead to it.

Also, i wonder how the wrong receiver finds out that the packet was intended for another destination, since the header CRC check will not show an error!!!

How were all these header corruption issues solved in ATOLL, which uses a quite similar path addressing scheme ?

ATOLL does not use something like a turn pointer. Instead, there is just a routing string, which is composed of routing bytes. The first routing byte contains the routing information for the next switch(crossbar). At every switch, the leading routing byte is removed. And so on....
Every single routing byte is parity checked at each hop(parity check is only a very small overhead). If there is an error, the packet is retransmitted on link level.

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by Walter F.J. Müller on Wed, 15 Sep 2004 16:15:10 GMT
View Forum Message <> Reply to Message

David Slogsnat wrote on Wed, 15 September 2004 16:40....
Every single routing byte is parity checked at each hop(parity check is only a very small overhead). If there is an error, the packet is retransmitted on link level.

That means, in ATOLL there is protection against single bit errors, not less, not more.

AS uses the PCIe physical link layer, which is serial with 8b/10b coding. The consequence is, that a single bit error on the medium will give in many cases an invalid code word, and will thus be detected, but in other cases, gives a different code word leading to multiple data bit errors. For examplefor the later case is that a single bit error in bit e can turn D0.1- into D9.1-, which when decoded gives a two bit error (two 0 are turned into 1). Since two bits are flipped, this isn't detected with a parity bit.

From this I'd conclude, that the ATOLL protection method actually doesn't protect against all single bit errors on the medium when a serial link with 8b/10b is used, thus a parity bit doesn't seem to be the prefect solution either.

Subject: What is the BER on copper ? How likely are problems ?
Posted by Walter F.J. Müller on Wed, 15 Sep 2004 17:40:15 GMT
View Forum Message <> Reply to Message

Error protection is certainly a good feature. The question is, how often do errors happen, how likely is a malfunction if there is no error protection ?

In the "Hardware" talk given by G. Kopcsay on the 2003 BlueGene/L workshop in Reno you'll find on page 17 some numbers for the serial links used in BlueGene/L. Bottom line is, that the BER is around 10-18.

If this is translated into an AS system with at 10 Tb/s thoughput, one gets roughly one bit error per day. About half will be caught by the coding layer, and with reasonable packet sizes less than 20% will be in the AS header, so there is a potential problem at the once every 2 week level.

If somebody has BER numbers for 2.5 Gbps links over 20-30" FR4 plus two connectors, which is a typical backplane scenario, please post a pointer to this information.

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by David Slogsnat on Fri, 17 Sep 2004 13:52:15 GMT
View Forum Message <> Reply to Message

Walter F.J. Müller wrote on Wed, 15 September 2004 18:15

That means, in ATOLL there is protection against single bit errors, not less, not more.

AS uses the PCIe physical link layer, which is serial with 8b/10b coding. The consequence is, that a single bit error on the medium will give in many cases an invalid code word, and will thus be detected, but in other cases, gives a different code word leading to multiple data bit errors. For examplefor the later case is that a single bit error in bit e can turn D0.1- into D9.1-, which when decoded gives a two bit error (two 0 are turned into 1). Since two bits are flipped, this isn't detected with a parity bit.

From this I'd conclude, that the ATOLL protection method actually doesn't protect against all single bit errors on the medium when a serial link with 8b/10b is used, thus a parity bit doesn't seem to be the prefect solution either.

The ATOLL protection method works very very fine for parallel links. And that is the link ATOLL has been designed to use (and it is the link that is used currently).
You are absolutely right that this method is not as good when used with 8b/10b coding over serial links. But since we don't even have such a system yet, I can't tell you from a practical standpoint wether it is sufficient or not....

Which leads us to another problem: what do you mean by a perfect solution (and what do I mean by sufficient) ?
There is always a trade-off between costs of error protection (in terms of additional delay, additional silicon, higher data redundancy...) and fault-tolerance. The level you want to have in the NIC hardware depends strongly on your needs, i.e. your application.

So as we are discussing this topic, an important question that should be answered beforehand is: what are actually the needs and requirements for our application?

Subject: Re: PCIe-AS - Tutorial -- Is there a CRC for the address header ?
Posted by Walter F.J. Müller on Sun, 19 Sep 2004 18:57:25 GMT
View Forum Message <> Reply to Message

David Slogsnat wrote on Wed, 15 September 2004 14:59
This is true. However, things get more complicated when looking at the ASI specification:
-The Turn Pointer is not included in the header CRC.
-The final receiver of an AS Packet has to check the CRC. The intermediate switches may

check it, but they don't have to.

All true. However, so far we ignored that ASI is a transaction layer build on top of the PCIe physical and data link layer. This adds
a transaction layer sequence number
a link layer CRC (LCRC)
and a link layer ACK/NAK
The LCRC is checked and regenerated at the link level.

So a switch will eventually realize that a packet is corrupted. The question is now what happens if such a corrupted packet is being already cut-through forwarded. I guess, that the only possible action in this case is to
make sure that the outgoing packet gets a bad LCRC too, to that it will eventually be dropped
send a data link level NAK, so that the sender will eventually retransmit.
If this is true, the exclusion of some fields in the header CRC isn't a real issue. However, the above is a guess (or hope), I don't know what the standart says for this case.

Subject: Re: Error recovery - the requirements
Posted by Walter F.J. Müller on Sun, 19 Sep 2004 19:05:04 GMT
View Forum Message <> Reply to Message

[quote title=David Slogsnat wrote on Fri, 17 September 2004 15:52]Walter F.J. Müller wrote on Wed, 15 September 2004 18:15So as we are discussing this topic, an important question that should be answered beforehand is: what are actually the needs and requirements for our application?

Agreed. For BNet, assuming a pure data push approach, it be probably o.k. to detect but not recover transmission errors. For PNet, assuming that at this level also request-response type protocols are used, we better recover errors. That's my guess, and I wonder what other viewpoints are.