Steps to Define a Common Application Transport Header

Richard Hughes-Jones
DANTE Delivery of Advanced Network Technology to Europe
8th International eVLBI Workshop, Madrid , 22-26 June 2009

Requirements

- Must be able to determine the length of a message.
- Must be able to discover missing, out of order, or duplicated data (messages).
- Should be independent of the physics headers (VDIF)
- Should be no interaction with the data (e.g. bit stuffing).
- Should not exclude different lower level network transport protocols or mechanisms.

Network Transport Protocols

- Raw Ethernet

Frame header	Application data	FCS

- UDP/IP

Frame header	IP header	UDP header	Application data	FCS

- TCP/IP

Frame header	IP header	TCP header	Application data	FCS

- Private Hardware Framing

Frame header	Application data	FCS

- Our Context

Frame header	IP header	UDP header	ATH	VDIF	Application data	FCS

Raw Ethernet

Frame header	Application data	FCS

- Raw Ethernet frames provide best effort delivery over the Link Layer.
- It is unreliable:
  - Packet may be lost
  - Duplicated
  - Out of order
- Raw Ethernet does give the application the length of the message delivered (which is the frame).

UDP/IP

Frame header	IP header	UDP header	Application data	FCS

- UDP is a connection less service over IP
- Provides best effort delivery, but it is unreliable:
  - Packet may be lost
  - Duplicated
  - Out of order
- UDP does give the application the length of the message delivered (which is the packet)

TCP/IP

Frame header	IP header	TCP header	Application data	FCS

- Connection orientated service over IP
- Reliable end-to-end Byte Stream delivered over unreliable network
- TCP provides
  - Data buffering
  - Flow control
  - Error detection & handling
  - Limits network congestion
- TCP takes care of:
  - Lost packets
  - Duplicated packets
  - Out of order packets
- But TCP does not give the application the length of the message delivered (TCP deals with byte streams).

Private Hardware Framing

Frame header	Application data	FCS

- I assume it IS framed
  - E-Merlin 30 Gig links are (ALMA160 bits, sync, sequence no. & checksum)
  - Chip-chip comms. may not be (but PCI-e is).
  - Widar Station Board appears to have CRCs interleaved/in parallel with data.
- It is most probably a bit unreliable:
  - Packet may be lost / corrupted
  - Duplicated
  - Out of order
- Framing does give the application the length of the message delivered (ie the frame).

Requirements Matrix

Requirement Protocol	Provide length of a message	Detect lost / out of order / duplicated
Raw Ethernet	Done by layer 2 framing	Need a frame sequence number.
UDP/IP	Done by layer 2 framing and IP header	Need a packet sequence number.
TCP/IP	Need a byte count	Done by transport protocol
Private Hardware Framing	Done by framing	Need a frame sequence number.

Proposal for the ATH

- The ATH will be placed on the network in little-endian (Intel x86) byte order
  - Most of our systems have PC / disk based influence
  - Compatible with the VDIFSpecification
- UDP, Raw Ethernet, Private Framing
  - A 63 bit packet or frame sequence number bits {0,62}.
  - Start from 0 for each new data flow.
  - Set bit 63 to 1 to denote a non-data packet.
- TCP
  - A 31 bit number giving the message length in bytes bits {0,30}
  - Set bit 61 to 1 to denote a non-data message.
- Tested using ATH&VDIF over UDP/IP with eMerlinIN FPGA and PC code.

8192 bytes 8 Gig
~1 10^{10} /day
$2^{31} = 2.1\ 10^9$

Open Questions

- Should there be 1 ATH version with both length & sequence number?
  - Seems a bit of a duplication
  - But I don’t like 2 versions
  - They don’t need to interwork anyway
- Have we got the bit-length correct?
- Do we define non-data packets
  - e.g. the data ACK
  - Controls
- ???

ANY QUESTIONS ?