Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

Virtual SpaceWire Networks

A Mechanism for Real-Time Applications

Acknowledgements

Thanks to the people who have made comments on early drafts

- Many requests for clearer explanations
- Many suggestions of how to make it clearer
- Some doubts, plenty of questions

Also

- Interest in a common treatment of priority between processing and communication
- Excitement at the opportunities created by VNs
- Support, up to and including full endorsement

Contents

RT requirements

- Timely and/or fast delivery of (some) data
  - Timely – to make things happen at regular intervals
    - May be part of a control loop where stability is likely to be impaired by late delivery
    - Loop time includes processing - therefore need fast delivery
      - Loop times: 100ms developing toward 10ms with some applications (e.g. Robotics) wanting 1ms
  - Fast for important, but possibly infrequent or irregular events such as failures (reports and/or retries) or unexpected events (interrupts)
    - hard (impossible?) to schedule

Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

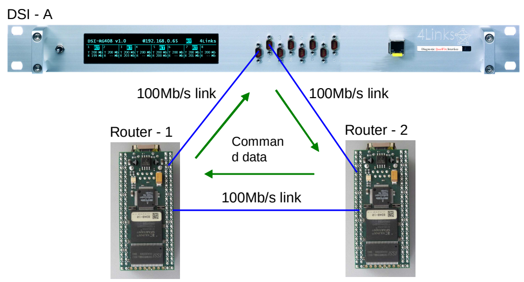
Traffic Profiles

Cause	*Size*	*Notes*
Events (interrupts)	Very small (<20B) ?	Irregular Need fast delivery
Command/Control	Small (~200B)	Likely to have delivery deadline
CCSDS	Up to 64KB (some missions ~4KB)	
Instruments	Large, up to MB's	

Traffic Profiles (2)

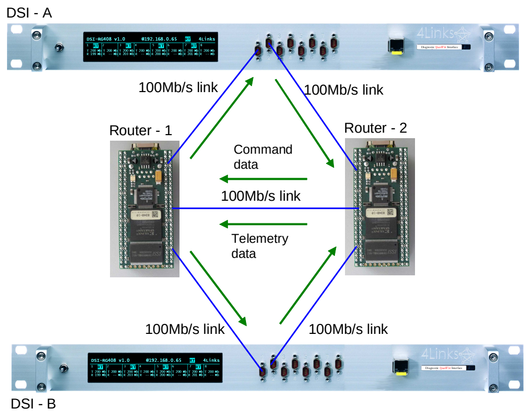
- Events and command/control are relatively low bandwidth but must be delivered quickly / before a deadline
- Data transfer requires most of the bandwidth - but can cope with some delay (not too much)

SpW - The Good



Measured latency: average 1.14 μs, jitter <0.2 μs (100Mb/s link, 248 byte command/control packets)

SpW - The Bad



Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

SpW - The Bad (2)

4Links®

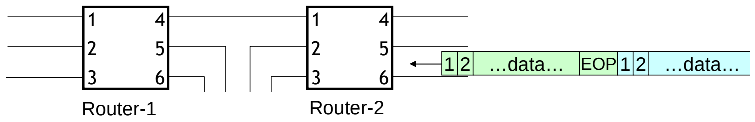
- But if a link is already busy when a packet arrives, we must wait for it to finish
- With the shared link continuously busy sending 4114-byte telemetry (CCSDS) packets

Measured latency for command/control packets is now:

average 173 μs, jitter 381μs

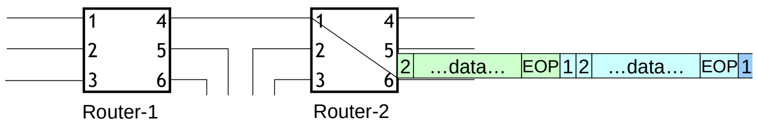
SpW - The problem

4Links®



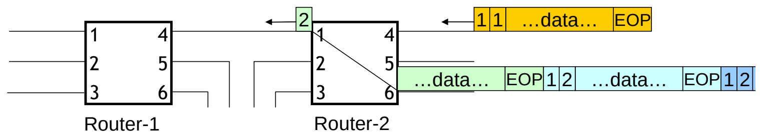
SpW - The problem (2)

4Links®



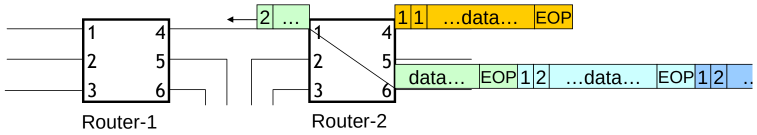
SpW - The problem (3)

4Links®



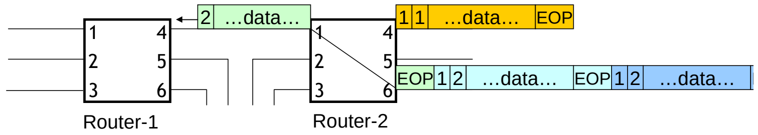
SpW - The problem (4)

4Links®



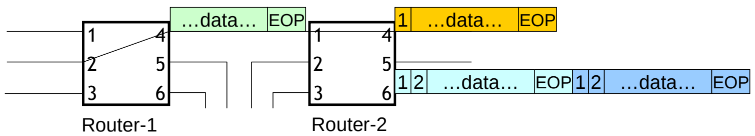
SpW - The problem (5)

4Links®



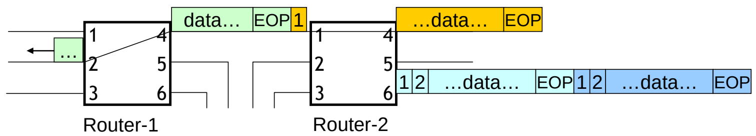
SpW - The problem (6)

4Links®



SpW - The problem (7)

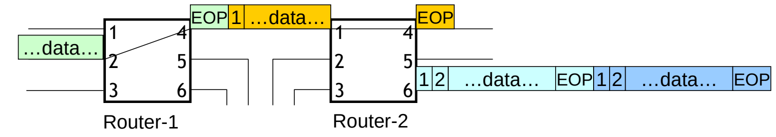
4Links®



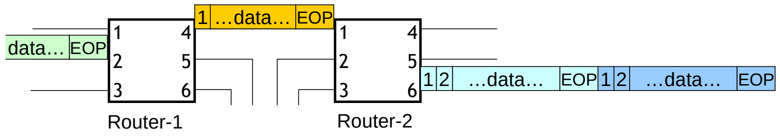
Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

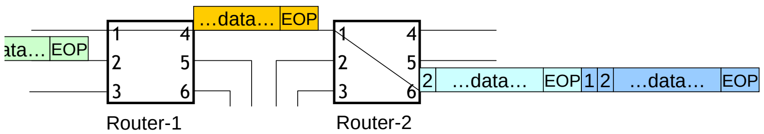
SpW - The problem (8)



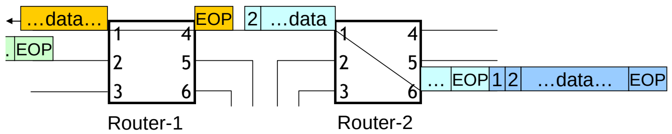
SpW - The problem (9)



SpW - The problem (10)



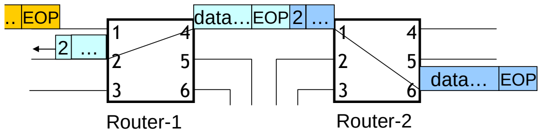
SpW - The problem (11)



Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

SpW - The problem (12)



The Theory

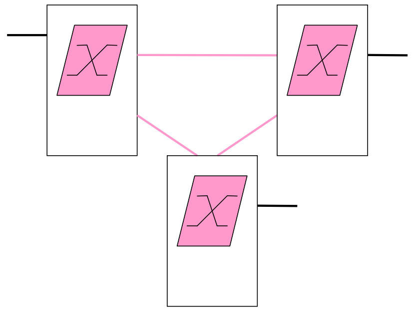
- The expected average latency (per link) is half the duration of an in-transit packet
- The expected jitter (per link) is the duration of an in-transit packet
- The in-transit packet may be telemetry and may be large (MB's)
  - Latency/jitter may be 10's to 100's of ms
  - No good for command/control packets on a 100ms (or 10ms or 1ms) update cycle

Solutions (?)

- Make all packets "small" to limit the waiting time
  - What is "small"?
    - ~200B: similar size to the most demanding data
      - Large overhead for CCSDS sized packets (need to break into 20 packets)
      - Huge overhead for MB instrument packets (need to break into 5000+ packets)
    - 4KB? 64KB?
      - Adds latency to command/control
      - Still large overhead for demanding instruments
- Use two networks
  - MIL Std1553B for Command/control and SpW for Telemetry
  - SpW and SpW

Command/control Network

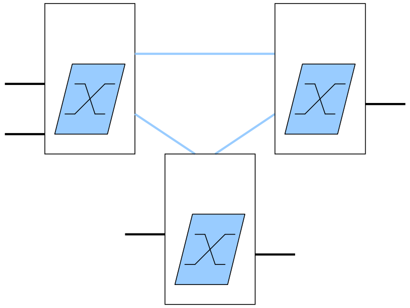
- All small packets, bounded latency requirements



Telemetry Network

4Links®

- Large / very large packets, may be considerable latency



Cost!

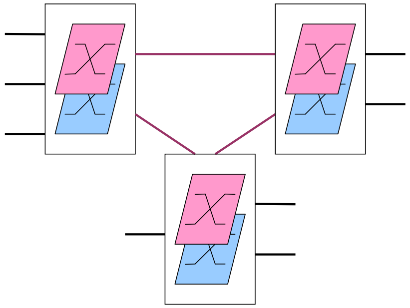
4Links®

- But this duplication costs (money, mass, power)

Virtual Networks

4Links®

- Logically two networks - but in one set of hardware



- Command/control data on one Virtual Network (red) and Telemetry data on the other Virtual Network (blue)

Removing the delays (1)

4Links®

- By allocating priority to the networks we can allow Command/control data (high-priority) to be sent even if telemetry data (low-priority) is already using the link
- The high-priority network delivers low latency - determined only by the traffic on that network
- The low-priority network has higher latency
  - (but also has access to all the remaining bandwidth)

Removing the delays (2)

- The link must be modified to prevent one network blocking traffic on the other network
- That is: we must allow data from a higher-priority network to overtake data from a lower-priority network
  - We add "Redirect" tokens to indicate that the following data belongs to a different Virtual Network
    - (use a time-code token)
  - Each Virtual Network has its own link receive buffer and flow-control counters (Add 8- to 56-bytes of buffer per Virtual Network)
    - (use another time-code token as a VN flow-control token)

Re-interpreted tokens (1)

- Redirect token

	7	6	5	4	3	2	1	0
Time code data value:	1	1	0	P	P	P	P	P

Where PPPPP indicates the Virtual Network number / priority

- Flow-control token

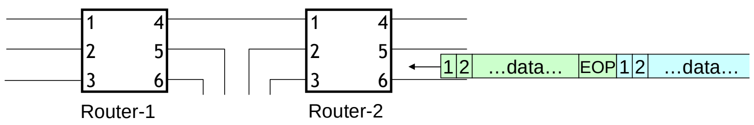
	7	6	5	4	3	2	1	0
Time code data value:	1	1	1	P	P	P	P	P

- These flow control tokens function for Virtual Networks exactly as normal flow control tokens and may be sent singly or in bursts, as many as are required to indicate the amount of free space in that Virtual Network's receive buffer

Re-interpreted tokens (2)

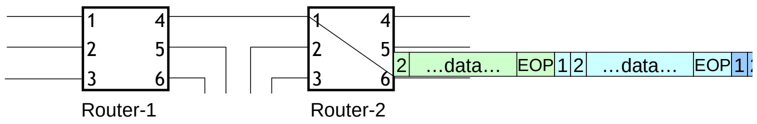
- These tokens are sent only where Virtual Networks are supported
- Detection of legacy / VN-capable devices takes place at link start (when the normal start sequence has completed and the link is in the "run" state)
- A suggested detection sequence is two non-incrementing "time-code" values
  - i.e. two Redirect tokens with equal or descending priorities
    - This is not a valid time indication and should be ignored by legacy devices

Virtual Network (1)



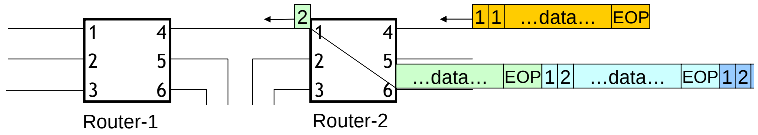
Virtual Network (2)

4Links®



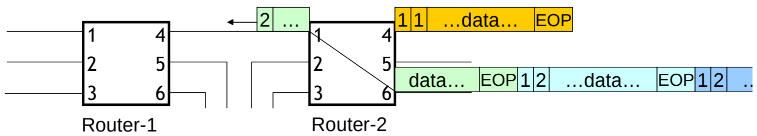
Virtual Network (3)

4Links®



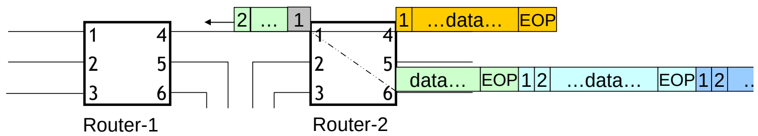
Virtual Network (4)

4Links®

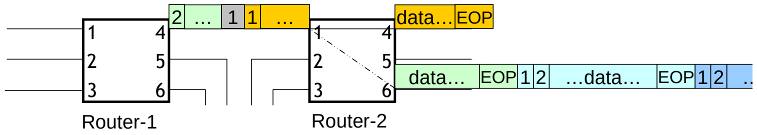


Virtual Network (5)

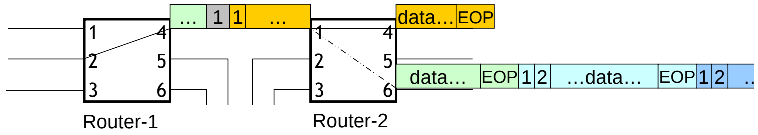
4Links®



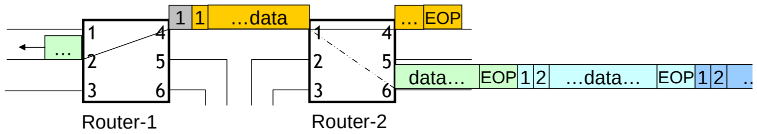
Virtual Network (6)



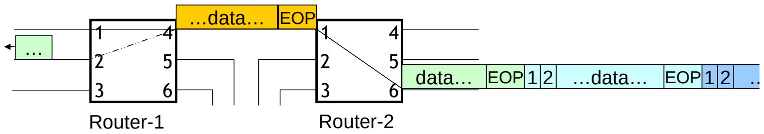
Virtual Network (7)



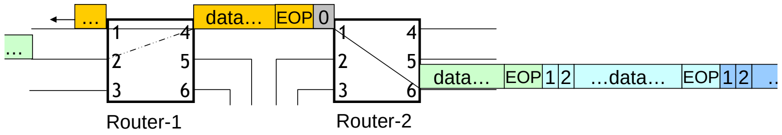
Virtual Network (8)



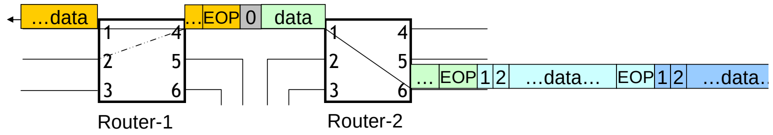
Virtual Network (9)



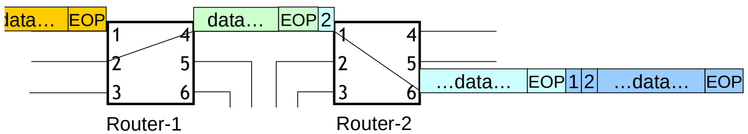
Virtual Network (10)



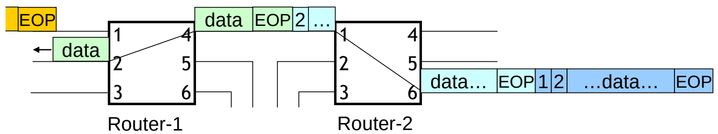
Virtual Network (11)



Virtual Network (12)



Virtual Network (13)

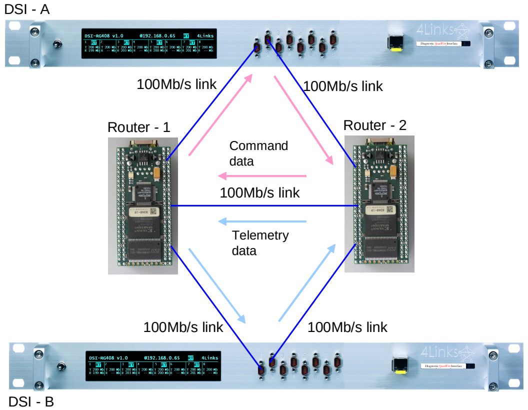


Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

Virtual Networks

4Links



VN performance

4Links

- Command/control packets on the high-priority, low-latency, network (red)
- Telemetry packets on the low-priority, higher latency, network

Measured latency for Command/control is now:
average 1.4 µs, jitter 0.3 µs

- *While the Telemetry data (4114-byte packets) is flowing at virtually the full bandwidth of the shared link (80Mb/s payload on 100Mb/s link)*

(without the telemetry traffic we measured latency:
average 1.14 µs, jitter 0.2 µs)

Re-interpreted tokens (3)

4Links

- The Virtual Network / priority field may be divided into priority and network-at-that-priority fields

	7	6	5	4	3	2	1	0
Time code data value:	1	1	0	P	P	N	N	N

Where PP indicates the priority and
NNN indicates the network-at-that-priority

- This allow us to have networks at the same priority, as well as at different priorities
  - The number of priorities and networks-within-priority is still to be determined (input needed from users)

Equal Priority Networks

4Links

- Networks at the same priority can be interleaved on an equal basis
- Multi-speed networks are supported
  - Without the need for large buffers in the routers (or anywhere else)
- This is useful for sharing bandwidth
- For example, we can share a 100Mb/s link between several 10Mb/s data streams
  - We have an overhead switching between streams
    - a 100Mb/s link can support ~8 (9?) * 10Mb/s streams

Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

Bandwidth Sharing

- An example …
- Merging two 10Mb/s (link speed, 8Mb/s data) streams carrying 1000-byte packets onto a shared 100Mb/s link we find …
  - Existing SpaceWire single network (interleaving packets)
    - Data throughput: 4.2 + 4.2Mb/s
      - Sequencing packets doubles the time required for each to be sent - halving the data rate (there is some gain, ~5%, due to buffering, 56 bytes of 1000 - also ~5%, allowing some data to be sent in faster bursts)
  - Using a Virtual Network for each stream (interleaving bytes)
    - Data throughput: 8 + 8Mb/s

"Babbling Idiot" nodes

- A "babbling idiot" node transmits data in an uncontrolled manner - possibly continuously
- The amount of data it can inject into a network is hard limited by the SpW link transmit bit rate
- With Virtual Networks, low data-rate nodes can be given a low-bit-rate connection (because we do not block a link when a packet arrives slowly)
  - We could even limit the rate at which the network issues flow-control tokens to nodes to further limit their transmit rate

Virtual Networks do …

- VN's do deliver high-priority data with (very) low latency
- VN's do allow higher-priority data to overtake lower-priority data
- VN's do have a simple conceptual model
- VN's do have a simple timing model
- VN's do allow multi-speed systems
- VN's do control "babbling idiot" nodes

Virtual Networks do not …

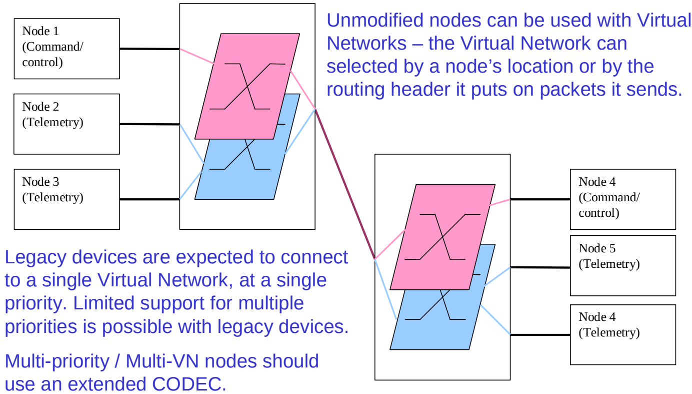
- VN's do not need to break any packets into smaller packets (or re-assemble them)
- VN's do not not need to distribute accurate time
  - To divide the bandwidth into units
    - That might not be fully used
- VN's do not need to "kill" packets
- VN's do not require (single-priority) end nodes to be changed at all

Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

Using "Legacy" Devices

4Links



Unmodified nodes can be used with Virtual Networks – the Virtual Network can selected by a node's location or by the routing header it puts on packets it sends.

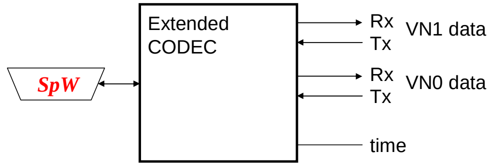
Legacy devices are expected to connect to a single Virtual Network, at a single priority. Limited support for multiple priorities is possible with legacy devices.

Multi-priority / Multi-VN nodes should use an extended CODEC.

Multi-Priority Nodes

4Links

- In order to support multiple priorities at a single node it would be better to use an extended CODEC that takes the Virtual Networks to the user as separate data streams (maybe as simple as an extended memory-mapped interface)



VN Timing Model

4Links

- Let the priorities be 0 … P, where P is the highest priority (P may be >1)
  - Priority P has reserved total network bandwidth, data will not be delayed by any traffic except that at level P
    - expectation (but not requirement): this traffic will be in small packets
  - Priority X (X<P) has access to all network bandwidth except that being used (note: *used*, not *reserved*) at priority >X, data will not be delayed by any traffic except that at levels >=X
    - Lower priority data can flow on a link being used by a higher priority packet if there is a pause in the higher-priority traffic

QoS provision

4Links

- Best effort delivery
  - Data may be lost
  - Lower-priority data may be delayed by higher-priority data
- Reserved
  - The whole of the network bandwidth is available to ("reserved" for) the highest-priority Virtual Network
- End-to-end flow control

Virtual SpaceWire Network Presentation

Barry M Cook, 4Links, SpaceWire WG Meeting 13, 2009-09-14/15

QoS non-provision (1)

- Assured and Guaranteed services are not offered
  - They might not often be used due to the very low bit error rate of SpaceWire
    - CERN measured IEEE1355 bit error rates $<10^{-17}$ (copper interconnect)
    - The bit error rate used to support the design of RDDP was, we are told, an estimate and was not measured
    - National Semiconductor have published tests on LVDS to $<10^{-12}$; Maxim to $<10^{-13}$ - in both cases the tests were stopped due to time limits, not observed errors
    - DS encoding is very tolerant of signal distortion and would be expected to show lower error rates than obtained from traditional BER testing

QoS non-provision (2)

- A bit error rate $<1.3 \times 10^{-14}$ has been measured at 2.5Gb/s in optical fibre components tested for SpaceFibre
- Each application will have a different view of the necessary provision
  - Simple acknowledgement / resend (low performance)
  - Sliding window (much higher performance)
  - RMAP
  - None - some real-time applications, such as control loops, are better served by losing data than re-sending it
- We consider it preferable to let applications select and implement their own choice of algorithm (as is done, for example, with extreme success, by the Internet)

For Consideration

- "End-to-end flow control" is a way to avoid blocking a link when a receiving node buffer fills
  - (Need not be a problem with Virtual Networks)
- It might be better to "pull" data into a known-size buffer than to "push" it into an unknown buffer
- Receivers should request data (no more than will fit into a buffer) rather than transmitters send it unrequested
  - Sliding window protocol (as used in TCP)
  - RMAP read

SpaceFibre

- Virtual Networks, with their mixed-speed network support, provide a way to unify SpaceWire and SpaceFibre
  - without the complication and cost of frames and segmentation
  - with a wide range of mixed speeds possible
  - and with a consistent protocol across the whole range

Miscellaneous items

4Links®

- Time distribution - unchanged
- Plug-and-play - fully supported
- Network status - as at present
  - available from routers via their control port
- Link errors - as at present
  - but a link error now affects more than one network
- Priority inversion - not a problem within a VN
  - must be considered when building multi-priority nodes

Changes to the SpW Standard

4Links®

- Physical level: No change
- Signal level: No change
- Character level: No change (if time-codes are used for Redirect and VN flow-control)
- Exchange level: Define use of additional characters for flow-control and for redirecting partial packets; Define protocol sequence for enabling use of virtual networks
- Packet level: No change
- Network level: No change
- Application level: No change

Contextual Remarks

4Links®

- Virtual Networks are a low-level mechanism for separating different classes of traffic in one physical network
- They are especially useful in supporting time-critical and real-time traffic
- They can remove the need for higher level protocols
  - such as the need to impose time slots onto all traffic
- Some higher level protocols are still required
  - For example, the concept of a virtual channel is still useful when a node provides services to several other nodes
  - Simple time scheduling - for a subset if the traffic

SpW-T (synch) cf. VN

4Links®

Feature	SpW-T (synch)	Virtual Network
Low latency data delivery	No 10ms at 200Mb/s 200ms at 10Mb/s	Yes 1 µs at 200Mb/s 20µs at 10Mb/s
Low jitter delivery	Yes ~10 µs (?)	Yes ~10 µs
Handles irregular / rare traffic	No Cannot be efficiently scheduled	Yes With very fast delivery

SpW-T (synch) v. VN (2)

4Links

Feature	SpW-T (synch)	Virtual Network
Good network utilisation	No Max 81% (peak) Time slots may be highly under-used	Yes Max 100% (peak) No unusable time
Large packet support	No Must break large packets	Yes Send unchanged

SpW-T (synch) v. VN (3)

4Links

Feature	SpW-T (synch)	Virtual Network
Self-timed	No Needs accurate time distribution	Yes
Low-overhead retries	No Must reserve time for retries	Yes Ad hoc retries when required

SpW-T (synch) v. VN (4)

4Links

Feature	SpW-T (synch)	Virtual Network
"Babbling idiot" node control	No	Yes
Mixed speed networks	No	Yes
Common model for SpaceWire and SpaceFibre	No	Yes

SpW-T (synch) v. VN (5)

4Links

Feature	SpW-T (synch)	Virtual Network
Able to use existing routers	No The whole network must support time slots	Yes For non-real-time traffic
Compatible with existing SpW devices	Maybe If they have CPU's able to respond fast enough	Yes For single-priority nodes

Conclusions

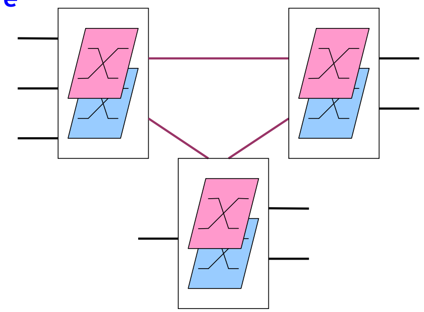
4Links®

- Virtual Networks provide a simple model that provides Real-Time characteristics
  - Existing nodes can be re-used if they are single-priority (either Real-Time or non-Real-Time)
  - An extended CODEC will allow mixed priority nodes
- The Virtual Network model unifies SpaceWire and SpaceFibre
  - We aim to have a SpaceFibre Virtual Network demonstration at the next WG
- SpaceWire hardware exists as a proof-of-concept demonstrator
  - Evaluation boards will be available soon - with commercial and flight silicon (Actel) available later

Conclusions (2)

4Links®

- Virtual Networks provide two, or more, separate networks of equal or ordered priority - in one set of hardware



Finally (via Dick Pountain)

4Links®

"There are two ways of constructing a software design: one way is to make it so simple that there are obviously no deficiencies and the other is to make it so complicated that there are no obvious deficiencies."
Professor Sir C.A.R."Tony" Hoare *(inventor of Quicksort and CSP)*

"Make it as simple as possible, but not simpler"
Albert Einstein

"Entia non sunt multiplicanda praeter necessitatem"
William of Occam *(the principle known as Occam's Razor)*

"A designer knows that he has arrived at perfection not when there is no longer anything to add, but when there is no longer anything to take away."
Antoine de Saint-Exupéry

"Our life is frittered away by detail - Simplify, simplify".
Henry David Thoreau

"I have made this letter longer than usual, because I lack the time to make it short."
Blaise Pascal