THE IMPLICATIONS FOR NETWORK SWITCH DESIGN IN A NETWORKED FTI DATA ACQUISITION SYSTEM

Nikki Cranley, Ph.D

ACRA CONTROL, Dublin, Ireland

***Abstract**:*

Switches are a critical component in any networked FTI data acquisition system in order to allow the forwarding of data from the DAU to the target destination devices such as the network recorder, PCM gateways, or ground station. Commercial off the shelf switches cannot meet the harsh operating conditions of FTI. This paper describes a hardware implementation of a crossbar switching architecture that meets the reliability and performance requirements of FTI equipment. Moreover, by combining the crossbar architecture with filtering techniques, the switch can be configured to achieve sophisticated forwarding operations. By way of illustration, a Gigabit network tap application is used to demonstrate the fundamental concepts of switching, forwarding, crossbar architecture, and filtering.

Keywords: Gigabit Network, Switching, Filtering, Network Tap

1. INTRODUCTION

Ethernet technology offers numerous benefits for networked Flight Test Instrumentation (FTI) systems such as increased data rates, flexibility, scalability and most importantly interoperability owing to the inherent interface, protocol and technological standardization. In a networked FTI system, the switch is a key component that allows data to be routed between networked Data Acquisition Units (DAU's), networked recorders, data processing and analysis stations.

Although Ethernet technology offers the ability to use commercial-off-the-shelf (COTS) networking devices, FTI networked systems have specific critical requirements that are central to the initial design of the FTI switch, in particular:

- **Environmental specification**: Unlike COTS switches, FTI switches must be able to work reliably in harsh environments.
- **Physical specification**: FTI switches must be rugged to operate in constrained environments and be designed with form factors where space and weight is a premium. With this in mind, the switch may combine functions mitigating the need for an additional external device, for example the switch may have an embedded IEEE 1588 Grandmaster. Equally the switch may be integrated into the networked DAU eliminating the need for a standalone switch thus reducing the wiring required by moving the networked DAU cluster closer to the sensors.
- **IEEE 1588 PTP Time Support**: This is essential in a distributed networked FTI to ensure coherency and simultaneous sampling in the networked DAUs. Implementing IEEE 1588 support in hardware close to the physical line level improves the performance and accuracy of time synchronization.
- **Reliability and Performance**: The primary function of the switch is to forward Ethernet packets to the appropriate destination. Unnecessary dynamic and sophisticated switching functions that compromise the switching performance should be kept at a minimum. This

is even more important as Gigabit and 10Gigabit networks become more prevalent. For example, FTI networks have static topologies therefore it may not be necessary to support dynamic routing and topology learning functions.

- **Live at power up**: The switch must be able to immediately and instantly forward packets on power-up in order to minimize latency and packet loss. The switch should not need to wait to boot up a micro-operating system or wait to populate the learned MAC routing table.

These design goals are central to the development of a reliable, flexible, and high performance FTI network switch. A crossbar switch design architecture not only meets these challenges in terms of meeting the environmental specification, reliability, and performance criteria but also provides complete flexibility to meet the needs and requirements of all applications through the use of statically configured forwarding tables and filtering mechanisms. These hardware design techniques allow for live at power-up, reliability, and high performance. This paper concludes demonstrating how these hardware design techniques can be applied to achieve a number of network topologies and applications including network taps.

The remainder of this paper is structured as follows. Section two provides an introduction to switching concepts, starting by outlining the subtle differences between routing and forwarding and how they apply to switches. Section three describes the crossbar switching fabric design that is central to this forwarding process that enables switch to forward the Ethernet frames to their destination.

2. SWITCHING OVERVIEW

In an FTI network of distributed peer DAU's, the switch is a key component that allows data to be transmitted to and from different nodes in the network. Switches comprise a number of ports to which DAU's may be connected or inter-connect switches. In general, connections in the switch utilize point-to-point full-duplex Ethernet links. The most important task of the switch is to reliably and quickly forward and route packets to their destination.

Quite often the terms routing and forwarding are used interchangeably. There is in fact a subtle relationship between the two.

Routing is the mechanism of looking up a routing table to determine the best path for a packet from a given sender to reach its destination through intermediate routers. There are two forms of routing, static and dynamic. Static routing is suitable for small networks whereby the number of routes is limited and can be manually configured. Dynamic routing is more suited to larger networks with complex topologies that may change over time. Adaptive routing "learns" the network topology using routing protocols, which are then used to automatically periodically generate and populate routing tables. Routing tables contain information derived from routing algorithms. Routing protocols communicate routing information about the connected devices between neighbouring routers. The routing table maps an IP address prefix to the next hop address prefix. This is, in essence, a Layer-3 topological view of the network and is optimized for detecting and adapting to changes in the topology.

Forwarding is the mechanism of passing or forwarding a packet from one port or interface in the switch to the appropriate out-going egress interface by looking up the forwarding table. Although it is possible to supplement the forwarding table with extra information that is typically found in the routing table, such as next hop information, forwarding statistics, QoS

metrics etc., it is more common for the routing and forwarding tables to be kept separate. In this way the routing tables can be used to generate compact and efficient forwarding tables, which are optimized for hardware storage and lookup, for example using Ternary Content Addressable Memory (TCAM).

2.1. STORE-AND-FORWARD OVERVIEW

As the name implies a key function of a store-and-forward switch is to forward incoming packets to the appropriate destination interface. The mechanism by which packets are forwarded to the appropriate destination interface is called store-and-forward whereby the packets received by the switch are stored in an input queue until they reach the head of the queue. Once at the head of the queue, the switch core examines the packets' destination and through a lookup mechanism determines how to forward the packet to its intended destination forwarding the data through the switch fabric.

Before being forwarded through the switch fabric, the switch core may perform various Layer 2 (MAC-layer) and Layer 3 (IP-layer) validation checks such as:

MAC Layer 2 validation

- Ethernet Frame Validation: Every Ethernet frame that is to be forwarded is validated to ensure it is well-formed i.e. it is within the allowed frame size limits and that "known" fields in the Ethernet frame are correct. Layer 3 (IP layer) validation may also occur to ensure the correct IP version field, protocol identifiers etc are correct.
- Ethernet Frame Check Sequence (FCS) error checking: The Ethernet MAC FCS is compared against the CRC calculated by the store-and-forward switch. If the Ethernet frames own FCS differs from the calculated CRC the frame is considered to contain physical or data-link errors and is dropped. In this way, the corrupt Ethernet frame is prevented from propagating through the rest of the network.

IP Layer 3 validation

- Packet Lifetime Control: Layer 3 switches must also decrement the time-to-live (TTL) field in the IP packet header to prevent packets infinitely circulating the network in routing loops. When the TTL value reaches zero, the packet is discarded.
- Checksum Recalculation: If the Layer 3 switch modifies the TTL, the corresponding IP header checksum and Ethernet FCS need to be recalculated and updated.
- Fragmentation: Should the Maximum Transmission Unit (MTU) of the outgoing Ethernet link be smaller than the size of the packet; the packet will need to be fragmented before being forwarded.

If the Ethernet frame is determined to be valid, the switch begins the forwarding process whereby the switch core examines the packets' destination and looks up the forwarding table to determine which out-going egress port (unicast) or ports (multicast/broadcast) are to be used. Since the Destination MAC address is the first 6Bytes of the Ethernet frame, the forwarding process is generally faster than Layer-3 routing table lookup which requires dissection of the various MAC and IP layer protocol fields.

If there is no entry for a given Destination MAC address, the switch does not know where to forward the packet. In this case, the Ethernet frame is forwarded out to all ports on the switch, or flooded. As Ethernet frames are passed through the switch, the switch core updates the MAC forwarding table noting the Source MAC address and the interface on which it arrived. By doing this, the switch core is able to maintain the forwarding table – however, since MAC tables have a finite memory size, entries age out to ensure that the table is up to date.

3. FORWARDING CROSSBAR SWITCHING FABRIC

To realise a flexible store-and-forward switching solution, a fully inter-connected crossbar switching architecture with *N* input buses and *N* output buses may be implemented where each cross point may be either on or off [1]. The advantage of such an architecture is that it is simple and the ease of implementation using a two-state crosspoint (on or off). Moreover, having high-speed data links in the fabric lowers the switching latency compared to other switching architectures by minimizing the number of connecting points.

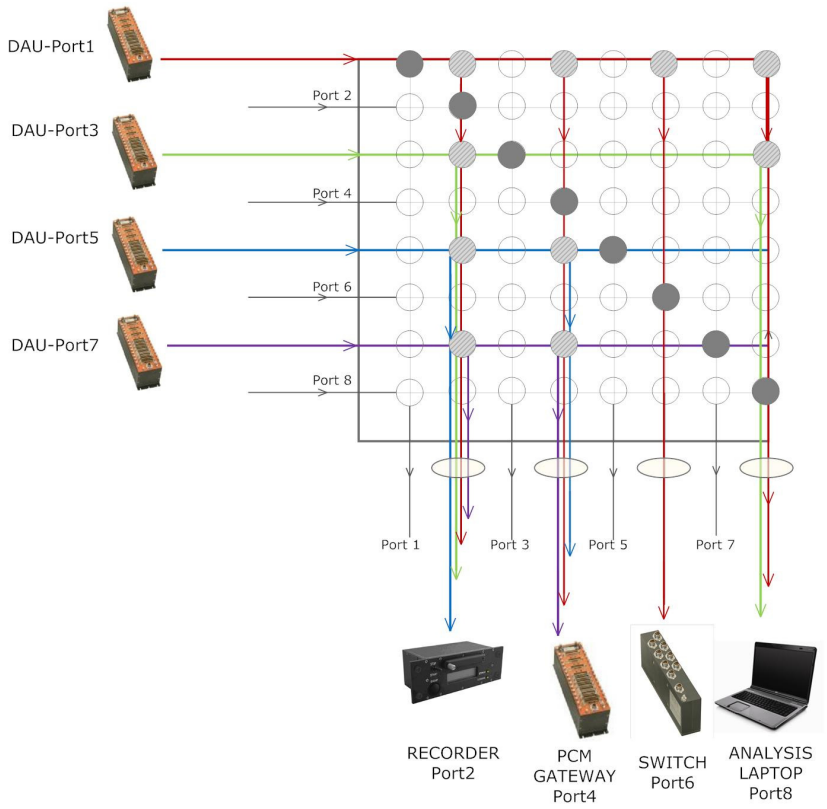
Consider an eight port switch with four DAU connected on ports 1,3,5, and 7. Ethernet frames transmitted by the DAU need to be forwarded to a number of sink devices such as a Recorder connected on port 2, a PCM gateway connected on port 4, a switch connected on port 6, and finally an analysis laptop connected on port 8. All data received by the switch must be forwarded to the recorder, while subsets of the DAU data is forwarded to the PCM gateway, Switch, and analysis PC as outlined in Table 1.

Table 1: Simple Forwarding Example

Input	**Output**
DAU-Port1	Recorder-Port2, PCM Gateway-Port4, Switch-Port6, Analysis Laptop-Port8
DAU-Port3	Recorder-Port2, Analysis Laptop-Port8
DAU-Port5	Recorder-Port2, PCM Gateway-Port4, Switch-Port6
DAU-Port7	Recorder-Port2, PCM Gateway-Port4, Switch-Port6

Figure 1 illustrates the configuration of the crossbar switching fabric required to realise this forwarding configuration between the DAU and the sink devices. The crossbar fabric comprises a fully-interconnected matrix of crosspoints between the input and output lines. Data received a given port is never forwarded back out on itself as indicated in Figure 1 by the dark grey crosspoint.

It can be seen that data transmitted by the DAU and received on port 1 is forwarded to the Recorder on port 2, PCM Gateway on port 4, Switch on port 6, and the analysis laptop on port 8. The hashed grey crosspoint indicates a connection has been made between the input and output lines. Similarly forwarding from the DAU on port 3 can be forwarded to the recorder on port 2 and Analysis Laptop on port 8 as realised through the on-crosspoints.



Legend:

- *Hashed grey crosspoint: A connection has been made between the input and output lines.*
- *Dark grey crosspoint: Data received on a given port is never forwarded back out on itself.*

Figure 1: Crossbar Fabric Forwarding Operation

Crossbar fabrics are inherently modular since more ports can be supported by adding more crosspoints. Crossbar switch fabrics are non-blocking in that all input and output ports can transfer packets simultaneously whilst simultaneously being able to pass packets to multiple output ports. This greatly increases the aggregate bandwidth of the switch. So although all output ports receive a copy of the incoming Ethernet frame, not all output ports are permitted to transmit the frame. This feature makes the fabric have native multicast support since if a multicast packet needs to be forwarded to multiple output ports, all the crosspoints corresponding to the input and output ports are turned on simultaneously providing a copy of the multicast packet to each output. However without proper design, the crossbar may experience Head of Line blocking. Since all ports have high-speed access to the switch fabric, this manifests when input ports attempt to create a connection with a "busy" output port. Packets waiting to be transmitted are essentially prevented or blocked packets until the output port becomes free. In this case, packets destined for a different output port that is not busy is blocked by the packets at the head of the line. One solution to overcome Head of Line blocking is the use of appropriate buffer sizes. Since blocking occurs on the output ports, the egress output port buffers should be larger than the ingress input port buffers. The larger egress port buffers allow for packets to be queued and reduce the probability of packet drops due to buffer overflow [2].

The crossbar switching fabric can be statically pre-configured setting up the appropriate crosspoints to forward packets from the input ports to the appropriate output ports. This static configuration is used to populate the static forwarding table. However, greater control and granularity may be required in terms of forwarding the packets i.e. selectively forwarding a subset of packets from a given input to a given output port. This in effect, applies a filtering function to the egress port buffer. To minimize the forwarding latency, the processing required during the forwarding process in the fabric is minimized and filtering is applied to the output port buffers. In the next section the setup of the forwarding and filtering tables is described using standardized Management Information Base (MIB) data structures.

Figure 2 illustrates the relationship between forwarding and filtering. Consider two DAU connected to the switch on ports 1 and 3, transmitting data to be forwarded to a Recorder and the PCM gateway. The DAU on port 1 is transmitting three packet streams to unique destination multicast addresses: video data stream (red), analog data stream (green), and n ARINC-429 data stream (blue). The DAU connected on port 3 is transmitting a single MIL-STD-1553 packet stream (orange) to a unique destination multicast address. The PCM Gateway needs only to be forwarded the video and the analog streams from the DAU on port 1, while the Recorder needs to be forwarded the ARINC-429 from the DAU on port 1 and the MIL-STD-1553 stream from the DAU on port 3. Once the forwarding paths between the inputs and outputs have been configured, the filter is applied whereby only those packets that meet the filter criteria are passed through.

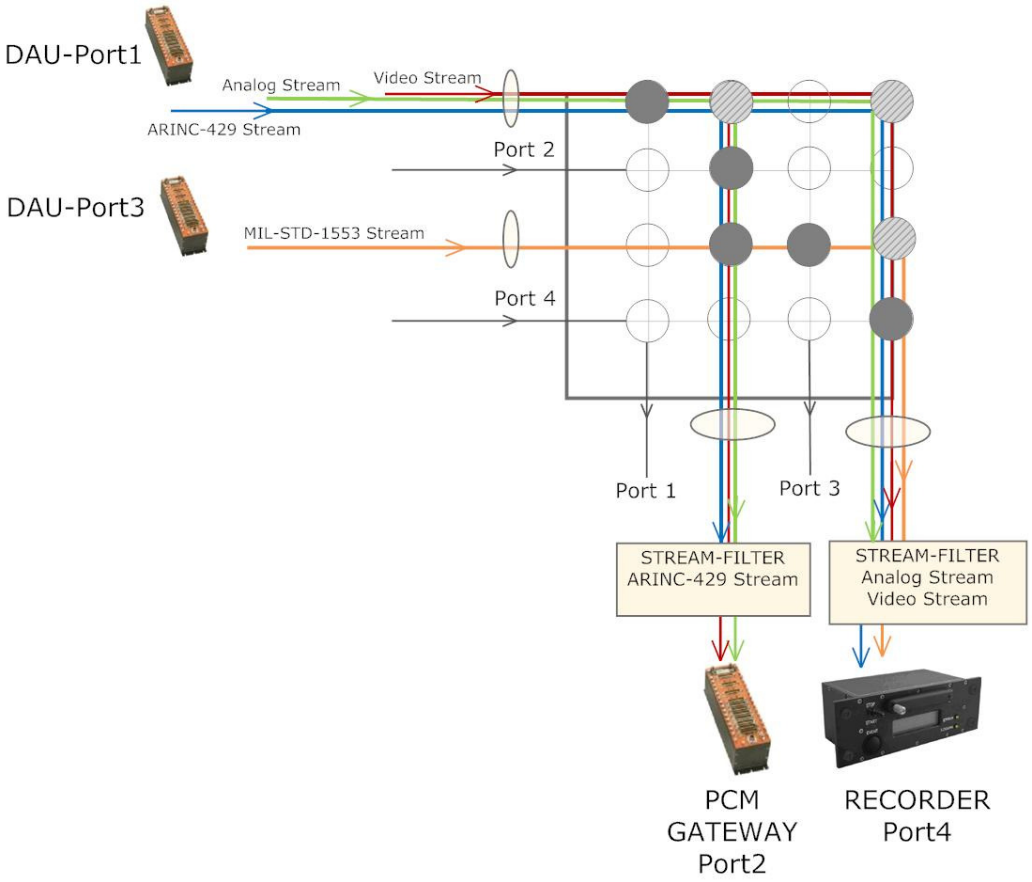


Figure 2: Forwarding and Filtering Operation

4. CONFIGURING FORWARDING AND FILTERING USING SNMP

The forwarding and filtering tables can be configured in the switch using the Simple Network Management Protocol (SNMP) through the dot1dStatic subtree in the Bridge Management Information Base (MIB) [3, 4] as defined in RFC4188 ((Request For Comments) or the Extended Bridge MIB for bridges with traffic classes, multicast filtering and Virtual LAN extensions [5] as defined in RFC4363.

The main difference between these two RFC standards is that the *dot1dStaticTable* in RFC4188 was replaced in the Extended Bridge MIB (q-bridge) with two tables: the *dot1qStaticUnicastTable* and *dot1qStaticMulticastTable* allowing for an indexing scheme that is compatible with the 802.1Q specification [6] where the unicast table is indexed by the Forwarding DataBase (FDB) ID and the multicast table is indexed by VLAN ID.

For simplicity, the *dot1dStatic* subtree in the Bridge MIB shall be discussed, which can be used to configure destination-address forwarding and filtering. The *dot1dStaticTable* subtree is a table that contains the filtering information. This allows the set of ports to which frames received from specific ports and containing specific destination addresses (unicast, multicast, and broadcast) are allowed to be forwarded and filtered. Each entry in the dot1dStaticTable comprises a 4-tuple { Destination MAC Address, Receive Port Interface, Bitvector of the Allowed Outgoing Ports, Entry Persistence }:

- **Destination MAC Address**: The destination MAC address in a frame to which this entry's filtering information applies.
- **Receive Port Interface**: The port number on which the frame must be received in order for this entry's filtering information to apply.
- **Bitvector of the Allowed Outgoing Ports:** The set of ports to which this frame is allowed to be forwarded. The bitvector is used to represent the on/off state or forwarding for each port in the switch.
- **Entry Persistence**: States the persistence of this forwarding/filtering entry indicating if it is permanent until removed, deleted on reset, deleted on timeout.

There is a similar construct, also in the Bridge MIB, called the *dot1dTpFdbTable* which is a table of unicast entries for which the bridge has forwarding and/or filtering information and is used by the switch to determine how to propagate a received frame.

5. CASE STUDY - GIGABIT NETWORK TAP APPLICATION

A network tap is a hardware device with at least three ports which provides a way to access and monitor the data flowing across a network link in order to monitor all data traffic between two tap points A and B in the network and creating a copy of that data through the monitor port. This type of monitoring is often called “sniff and mirror” operation in the switch whereby the “mirror” port forwards an identical copy of the data that is transmitted through its paired “sniff” port. The requirements for this network tap application are that it should passively tap into a Gigabit Ethernet network on the aircraft so that the data could be recorded whilst allowing access to a subset of the data to be relayed to the ground via PCM Gateway. It must be guaranteed that the tap does not forward any data to the tapped network including all network operation protocols such as IGMP, PTP, SNMP and so on. Moreover, handshaking operations such as auto-negotiation with the tapped network should be disabled. Finally, the network tap must be reliable, high performance and be live-at-power-up.

Consider a Gigabit Ethernet link between points Ethernet tap points A and B which needs to be monitored with a copy of the data flow between these two points being forwarded to a recorder connected to port 3 and a filtered subset of the monitored data is forwarded to an PCM gateway with a 100BaseTX link. Since the Tap ports A and B are Gigabit, the Recorder port must also be Gigabit. However the PCM gateway has a 100BaseTX link, only a subset of the data can be forwarded to the Gateway. The link speed for each port in the switch is statically configured to 1000BaseT (Tap-A, Tap-B, Recorder) and 100BaseTX (PCM Gateway) respectively.

The forwarding table must ensure that data from Tap-A is forwarded to all ports i.e. Tap-B, the Recorder, and the PCM gateway. Similarly, the data from Tap-B is forwarded to all ports. However it must be guaranteed that no data from the recorder or the PCM Gateway is forwarded to either Tap-A or Tap-B. Data can, however, be forwarded between the recorder and the PCM Gateway.

The filtered subset forwarded to the PCM Gateway comprises the Blue stream from Tap-A, the Orange stream from Tap-B, and Purple stream from the recorder. All other packet streams are discarded in the filter. This configuration is summarized in Table 2 and the corresponding crossbar switching fabric forwarding configuration with filtering is illustrated in Figure 3.

Table 2: Gigabit Network Tap Setup Summary

	Port Name	**Link Speed**	**Input Stream(s)**	**Output Stream(s)**	**Filter**
1	Tap port A	1000BaseT	• Video (Red), • ARINC-429 (Blue), • Analog (Green)	• MIL-STD-1553 (Orange)	Not required
2	Tap port B	1000BaseT	• MIL-STD-1553 (Orange)	• Video (Red), • ARINC-429 (Blue), • Analog (Green)	Not required
3	Recorder	1000BaseT	• SNMP Status (Purple)	• All	No
4	PCM Gateway	100BaseTX	• PCM Stream Grey	• All	Yes Allow only ARINC-429 (Blue from Tap-A), MIL-STD-1553 (Orange from Tap-B), SNMP Status (Purple from Recorder port 3)

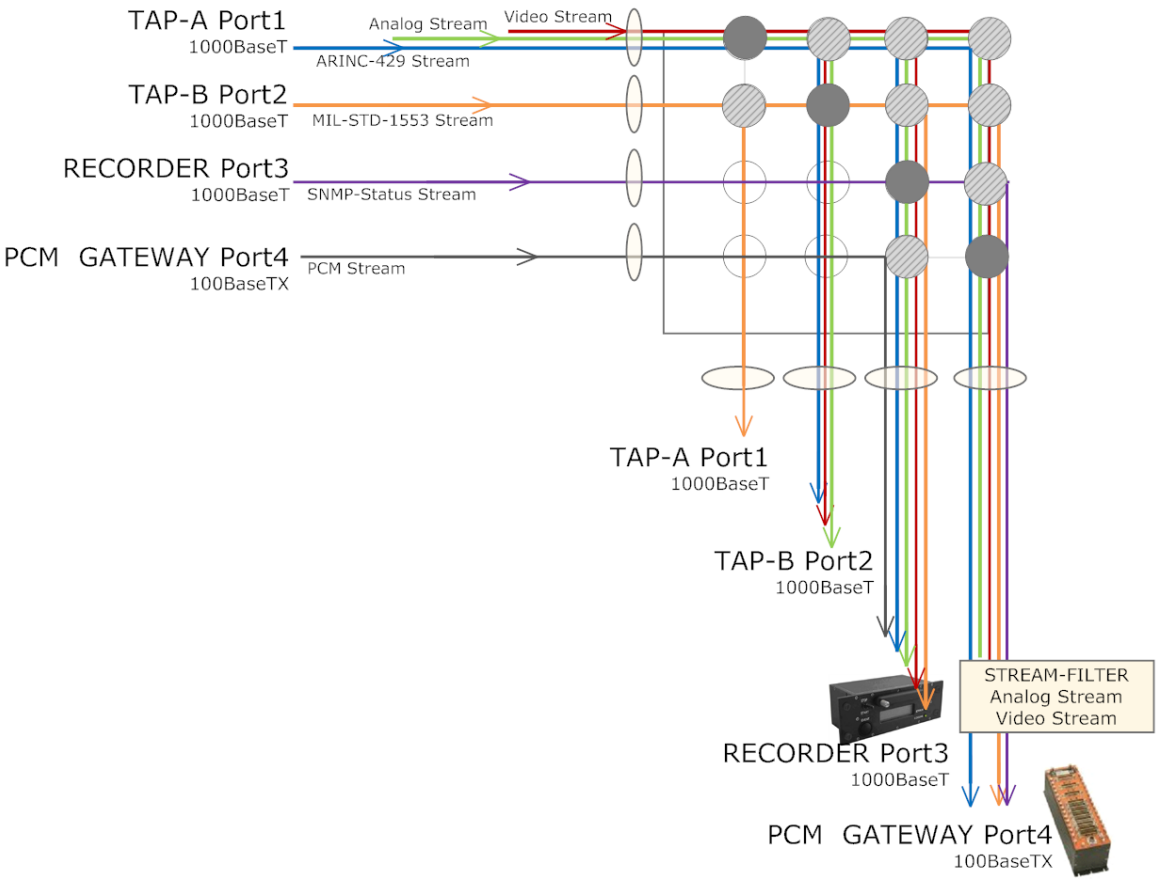


Figure 3: Four-port Gigabit Network Tap

6. CONCLUSIONS

The network switch is a key component in the networked FTI data acquisition system as it provides a mechanism to interconnect the network nodes such as data acquisition units (DAU), recorders, PCM gateways, and analysis PCs together. At its most simplistic implementation the network switch must be able to receive data from the DAU and forward the aggregated data flows to known destination sink devices, such as the recorder and PCM gateway. This approach is ensures live-at-power up operation and is reliable and deterministic however it lacks flexibility when often more complex selective forwarding is required. This paper describes a crossbar switching architecture that provides a means of supporting complex sophisticated forwarding and filtering while ensuring reliable and deterministic switching behaviour. This paper concludes providing a real-life application of an FTI Gigabit network tap device that demonstrates the configuration of the crossbar switching architecture to support the passive monitoring and recording of the aircraft Gigabit network whilst also allowing for a subset of the monitored data to be filtered for real-time transmission to the ground via a PCM gateway.

7. REFERENCES

[1] Medhi, D., Ramasamy, K., "Network routing: algorithms, protocols, and architectures", Morgan Kaufmann (April 12, 2007), ISBN-10: 9780120885886

[2] Barnes, D., Sakandar, B., "Cisco LAN Switching Fundamentals", Cisco Press (July 15, 2004), ISBN: 1-58705-089-7

[3] IETF RFC 4188, "Definitions of Managed Objects for Bridges"

[4] IEEE Std. 802.1D 2004, "Media Access Control (MAC) Bridges"

[5] IETF RFC 4363, "Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering, and Virtual LAN Extensions"

[6] IEEE Std. 802.1Q 2003, "Virtual Bridged Local Area Networks",