Whitepaper: Advantages of a Full-Mesh Fabric Configuration vs. Dual-Star in AdvancedTCA® Systems

Matthew Dharm, JumpGen Systems LLC

8 March 2010

Overview

While the Base Channel in AdvancedTCA® systems is defined as 10/100/1000 Ethernet in a dual-star topology, the Fabric Channel has multiple data transport options and multiple topology options. Commonly, the dual-star topology is used in conjunction with dual switch cards in the chassis regardless of what data transport is being used. This deployment configuration follows a well-understood and commonly used design. However, this topology can result in several sub-optimal design limitations, especially when contrasted with the full-mesh topology. In this paper, we will examine some of the beneficial aspects of full-mesh topology as well as compatibility issues surrounding migration from dual-star to full-mesh and mixing of cards designed for both types of systems.

In this paper, we will consider a hypothetical 16 slot chassis using 10GigE as the fabric interface. While 10GigE is a common interface, it is by no means the only one in use in AdvancedTCA® systems. However, almost any fabric will see similar performance gains to what will be discussed here; 10GigE is used only as one example.

In our hypothetical example, the dual-star 16-slot system consists of 2 switch cards and 14 node cards. In the full-mesh configuration, the system consists of 16 node cards which are all identical. For the sake of simplicity, we assume that all node cards are identical.

Bandwidth Comparison

The most obvious advantage offered by a full-mesh topology over a dual-star topology is in the area of system bandwidth. However, when examining the amount of bandwidth offered by a particular configuration, we must take care to distinguish between ‘internal’ and ‘external’ system bandwidth.

Internal System Bandwidth

We use the term ‘internal’ system bandwidth in this paper to refer to the total amount of bandwidth available for nodes to communicate with each other. We exclude any bandwidth available for nodes to communicate with components outside of the AdvancedTCA® chassis.

Efficient use of this communication capacity is critical in maximizing the utilization of each node's capacity.

In a dual-star topology, each node board has 2 fabric connections (one to each switch), and the two switches are also connected together. In our hypothetical 16 slot system, that represents 580 Gbps of total internal bandwidth. This is calculated as 14 nodes x 2 links per node = 28 node links, plus one inter-switch link (total of 29 links) running full-duplex 10GigE (20 Gbps of bandwidth).

However, the full-mesh system has significantly more bandwidth available. Our hypothetical system has 2720 Gbps of bandwidth, which is approximately **4.7x more** than the dual-star approach. This is calculated as $\sum_{i=0}^{n} i$ (the total number of links in a fully connected graph) running full-duplex 10GigE.

External System Bandwidth

We use the term 'external' system bandwidth in this paper to refer to the total amount of bandwidth available for nodes to communicate with components outside of the AdvancedTCA® chassis. The total amount of external bandwidth, and its utilization level, effectively defines the total system throughput.

Two cases must be considered here. The first case is where the primary external connection for the system is on the switch cards. The other is where the primary external connection is on the node cards. Both configurations are common deployment scenarios.

For systems where the primary connection is on the node cards, the full-mesh topology offers more node cards (discussed more in the 'Increased Processing Capacity' section below) and thus a **14% increase** in external bandwidth. For systems where the primary connection is on the switch card, the full-mesh system now offers a dramatic increase in external bandwidth. Since each card is a fully-connected member of the mesh, any card can be used for an external connection. This yields a system with **8x more** external bandwidth.

Improved Failover and Redundancy

Typically, dual-star configurations are selected over single-star configurations in order to provide N+1 redundancy in the event of a failover. In single-star configurations, a failure of the switch component yields a complete system failure. Dual-star configurations eliminate this single point of failure, but replace it with a dual point of failure. While this is an improvement, it is far from optimal.

Full-mesh topologies, on the other hand, do not suffer from this problem. Since each node can communicate directly with any other node, any two functioning nodes in the system can continue system operation (albeit in a degraded mode). Put another way, the failure of any one (or two) node cards will not directly impact the ability of other node cards in the system to continue operation.

Increased Processing Capacity

One of the biggest drawbacks of dual-star in AdvancedTCA® systems is the assignment of two slots in the chassis for exclusive use a fabric switch slots. In our hypothetical 16 slots system, this means that only 14 slots can be occupied by node cards. In a full-mesh system, however, all 16 slots can be node cards, yielding a **14% increase** in the number of node cards per chassis.

The difference is even more dramatic in smaller systems. For example, a dual-star 5-slot system has 3 nodes and 2 switches. A full-mesh system of the same size has 5 nodes, which is a **66% increase** in the number of node cards per system.

In addition to the significant increase in the number of nodes per chassis, there is also a significant cost savings which comes from not having to purchase dedicated switch cards.

Compatibility with existing Dual-Star Deployments

A significant concern of system architects who are considering dual-star vs. full-mesh topologies is compatibility with existing system deployments. Organizations which have made a significant investment in one configuration or the other may be reluctant to change. These concerns are, however, easily mitigated in AdvancedTCA® systems.

The first aspect of AdvancedTCA® which reduced the risks and costs of a migration to full-mesh systems is the fact that dual-star backplane routing is a proper subset of the full-mesh backplane routing. This is covered more completely in the *PICMG® 3.0 Revision 3.0 AdvancedTCA® Base Specification* in section 6.6.3 *Fabric Interface Backplane configurations and Channel routing*. In summary, however, this means that full-mesh and dual-star cards can be used together in both full-mesh and dual-star system configurations.

Full-Mesh Cards in Dual-Star Systems

First, let's look at the case of a full-mesh card being introduced into a dual-star system. If the full-mesh card is inserted into Logical Slot 1 or 2 (which are defined to be the switch slots for dual-star fabric and also for the base interface), the full-mesh card is connected to every other card in the system. If will need to take on some switching role to provide connectivity to other cards in the system, but this is part of the design function of a full-mesh card. If the full-mesh card is inserted into a node slot, then the card will be connected to the 2 system switch slots on its first two fabric interfaces; the other fabric interfaces will simply be unconnected. In this configuration, the full-mesh card acts just like any other dual-star card.

Dual-Star Cards in Full-Mesh Systems

Next, let's consider a dual-star card inserted into a full-mesh system. The dual-star card will only be able to communicate directly with the full-mesh cards in Logical Slots 1 and 2 (which are where switch cards would go in a dual-slot system). However, the full-mesh cards in those slots are directly connected to every other card in the system, and thus able to route traffic directly to other cards. Of course, this means that a dual-star card cannot be used in Logical Slot 1 or 2 in either a dual-star or full-mesh system.

The Base Channel

Finally, we need to consider the Base Channel of an AdvancedTCA® system. As mentioned earlier in the document, the Base Channel is defined to always be a dual-star configuration. The switches for this dual-star configuration are located in Logical Slots 1 and 2. Therefore, in order for a full-mesh card to function properly in any slot, that card also needs to implement a Base Channel switch. For system slots where the Base Channel switch is not needed, those switch lanes can be potentially configured for other uses or simply left unconnected.

Summary

The following table gives a quick comparison of dual-star and full-mesh systems.

	Internal Bandwidth		**External Bandwidth**		**Processing Performance**	
System Size	**5 Slot**	**16 Slot**	**5 Slot**	**16 Slot**	**5 Slot**	**16 Slot**
Gain from Full-Mesh	2.1x more	4.7x more	2.5x more	8x more	66% more	14% more

Conclusion

Full-mesh and dual-star are by no means the only choices for system configuration in AdvancedTCA® systems. They are, however, the most common choices. And, while there may be a penalty for full-mesh architectures in system configuration complexity and possible software complexity (depending on the application), the dramatic increase in bandwidth, processing power, and reliability make a clear case for the superiority of full-mesh over dual-star systems.