Understanding Avaya Stackable Chassis Architecture

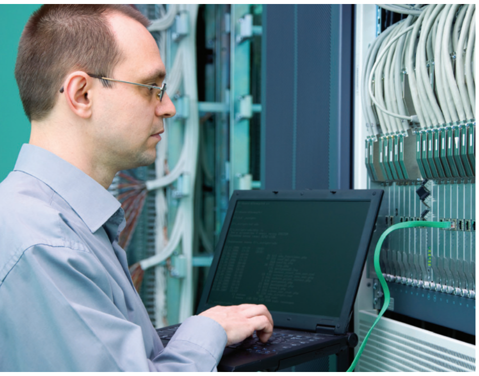
Table of Contents

Gone are the days of refreshing networking equipment every three-to-five years. Economic recession, having placed significant burdens on business and business budgets, has extended the network refresh cycle to seven-to-ten years in some instances. Yet, despite economic turmoil, a proliferation of new converged technologies – smart phones, soft clients, tablet computers, wireless LAN, network and compute virtualization, and cloud computing – continues to add pressure to the networking environment.

These new converged communication technologies are forcing the data network to mature rapidly from a simple utility for transferring files to an integral business application that can handle a full range of collaborative communications needs. Required to be more than a transport mechanism for ever-increasing loads of data, today's networks must provide high-quality, reliable transport for real-time, latency-sensitive applications, such as voice and video, while simultaneously supporting the IP phone system, video conferencing, video surveillance, storage, and an abundance of wired and wireless offerings.

Network administrators face the challenge of making equipment they purchased six years ago (for a primarily data-centric purpose) function in today's converged technological world. Further, they are asked to ensure that future procurements will be in place longer than previous purchases and that they will adapt readily to the quickening pace of convergence, unforeseen technological advancements, and disruptive technologies.

Having recognized this trend years ago, Avaya began architecting networking products that would remain viable and flourish in the face of rapid technological advancement. For example, in 2003 the Avaya Ethernet Routing Switch (ERS) 5500 series was introduced and, to this day, it remains as much a state-of-the-art component of network infrastructure as competing products introduced as



recently as one year ago. The success and durability of the ERS 5500 series are rooted in its platform architecture that, once in place, needs only a software upgrade to keep pace with the evolution of the networking industry. This software upgrade approach is more cost-effective and less resource intensive than competitive forklift approaches requiring reinstallation and reconfiguration of the network infrastructure.

Central to the architectural endurance of Avaya products is the virtual backplane architecture known as Flexible Advanced Stacking Technology (FAST). Introducing the concept of truly scalable stacking in 1998 (with its Bay Networks products), Avaya released the third generation of FAST in the ERS 5500 series product line and, with the introduction of the ERS 5600, delivered the fourth generation which remained backwards compatible with the previous generation.

What is Stacking?

Over the years, the term stacking has been used to describe (1) a method for interconnecting switches through uplink ports that allows linked switches to be managed as a single entity via a single IP address, and (2) fixed form factor boxes that mount one above the other within a rack (commonly referred to as "pileable") but offer no physical stacking mechanism for interconnection.

In data networking vernacular, the proper definition of stacking is the ability of a number of discreet switches to be physically interconnected in a resilient mechanism (that doesn't rely on Spanning Tree) that forms a single switching entity that is managed via a single IP address. Each member within the stack has awareness of the other units and shares and synchronizes forwarding and configuration data. If one unit fails, the stack continues to operate.

In Avaya Networking, stacking refers to any number of switches (currently up to eight) physically interconnected in a resilient fashion through dedicated stacking ports. These stacking ports are typically connected in a physical ring topology, with each unit connected to the unit above and below and the top-most unit connected to the bottom-most unit, forming a resilient ring. Individual switches, now physically interconnected, become a single, logical switching entity (stack) that is managed via a single IP address.

The stack is controlled by a Base Unit and, while any switch in the stack can be the Base Unit, initial selection is done manually via a selector switch on the rear of the unit. The Base Unit is the control point for the rest of the stack; it helps ensure that the stack forms properly. Should the Base Unit fail, the unit directly downstream from the Base Unit becomes the Temporary Base Unit and all switching operations continue normally (except for devices physically connected to the failed Base Unit). And, should a Temporary Base Unit fail, the next unit downstream can become the new Temporary Base Unit, and so forth. Once the original failed Base Unit comes back on-line (either by itself, or through replacement), it rejoins the stack and resumes switching operations but does not resume the role of Base Unit until the stack is reset.

Flexible Advanced Stacking Technology (FAST)

Now that we've described the general concept of Avaya stacking, let's focus on FAST, one of our more recent advancements. FAST is actually an architecture made up of several components including the physical hardware (stack cables and ports), the control plane, and the switching software. All components work in unison to deliver the kind of 'virtual backplane' capability (long sought after in the wiring closet, data center, and small core) that enables discreet, stackable units to take on the positive aspects of a chassis-based system without the associated expense and up-front costs.

Previous generations of stacking technology relied (and some competitive implementations still rely) upon a directional ring topology in which packets sequentially traverse the units in a stack until reaching the point of egress. For instance, if ingress packets on unit #1 of an eight unit stack are destined to egress on unit #8, the packets would traverse the stack moving from unit #1 to unit #2 to unit #3, and so on until unit #8 is reached. In this implementation, the stack bandwidth between units #2 through unit #7 is consumed by transporting the packets to unit #8. Some competing technologies even require that those same packets continue on back to the originating unit before being "stripped" from the stack (called source stripping), making even less efficient use of the stack bandwidth.

While the physical cabling of competing technologies is similar to the physical cabling of FAST-enabled devices, the similarity ends here. With FAST, packets entering the virtual backplane from a given unit in the stack are not forced to traverse through the stack in a given direction, due to FAST's bi-directional forwarding capabilities. Using the eight-unit stack example from the previous paragraph, FAST's shortest path algorithms would determine that the most efficient path for the ingress packets on unit #1 to egress on unit #8 is not through unit #2 to unit #3, and so on until unit #8 but, instead, the other direction– directly to unit #8.

This intelligent decision making has several advantages. First, it reduces the overall latency of the packet by delivering it through the shortest path. Second, it preserves the stack bandwidth between units #2 through #7, allowing other packets to traverse that portion of the stack without contending for available bandwidth. This preservation of stack bandwidth is key to allowing available stack bandwidth to be the cumulative sum of each discreet unit's stacking bandwidth, thus enabling the highest possible stack bandwidth. (See Figure 1)

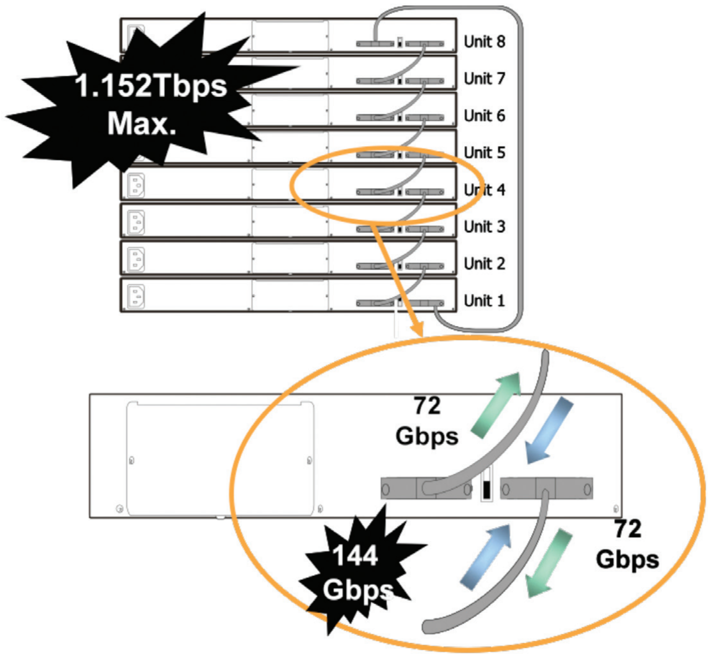


Figure 1: 8-high stack of ERS 5632FD with 1.152Tbps stacking bandwidth

It is important to note that the stacking bandwidth of an individual ERS 5600 switch is 144Gbps, with each unit in the stack able to simultaneously and independently transmit up to 72Gbps onto the stack while simultaneously receiving 72Gbps from the stack. This breaks down to 72Gbps bandwidth for the Cascade Up link (36Gbps full duplex) and 72Gbps bandwidth for the Cascade Down link (36Gbps full duplex). Since the FAST architecture is not bounded by the uni-directional nature of typical stack packet flows, stack configurations with eight ERS 5600 units can achieve up to 1.152Tbps of available stacking bandwidth (144Gbps x 8 units = 1.152Tbps).

Additionally, due to the inherent resiliency of the FAST architecture, stack cable or unit failure scenarios have a distinct advantage in preserving as much bandwidth as possible. In traditional stacking architectures, packets flow in a given direction through the stack and, if there is a failure in a cable or unit in the stack, traditional stacking technologies cause the traffic to take a counter-rotating path, causing the data to wrap back upon itself, essentially halving available stack bandwidth. The FAST architecture handles this by redirecting at either side of the failure, halving the stacking bandwidth at the two units' stack ports adjacent to the failure, but maintaining the full stack bandwidth at each of the remaining units.

For example, if there were a failure on unit #5 in the eight-unit-high ERS 5600 stack in figure 1, the traffic would wrap at the Cascade Up port on unit #4 and the Cascade Down port unit #6. This effectively eliminates one of the 72Gbps bandwidth connections on both unit #4 (Cascade Up) and unit #6 (Cascade Down), as well as the available stacking bandwidth on the failed unit #5 (Figure 2). However, both unit #4 and unit #6 retain their other 72Gbps bandwidth connections (Cascade Down and Cascade Up respectively) and, most importantly, the remaining 5 units in the stack (units #1, #2, #3, #7 and #8) still have their full 144Gbps bandwidth capacity.

This means that, unlike competing technologies that result in the halving of available stacking bandwidth, ERS 5600 stack bandwidth merely goes down by the stack bandwidth of the failed unit plus 0.5 the stack bandwidth of the two adjacent units in the stack, yielding 864Gbps bandwidth [144Gbps x 5 units (units unaffected by the failure) + 72Gbps x 2 units (units adjacent to the failed unit which have their stacking bandwidth halved) = 864Gbps]. At a minimum, the FAST architecture preserves an additional 50% of available stacking bandwidth over and above competing technologies that halve the bandwidth.

For a clear understanding of the benefits of the ERS 5600 stack, consider a traditional stacking technology claiming 64Gbps stack bandwidth. In traditional stacking technology, this is accomplished by having two bi-directional stack rings of 32Gbps. If there were a failure in unit #5 of an 8 unit stack, traffic would wrap at unit #4 and unit #6 as in the example described above. However, overall stacking bandwidth would halve, since traffic is forced to consume the other remaining ring direction, yielding 32Gbps stack bandwidth. This illustrates the benefits of FAST in preserving as much stack bandwidth as possible (864Gbps in the 5600 case vs. 32Gbps in a traditional stacking technology, 27x more available stacking bandwidth)!

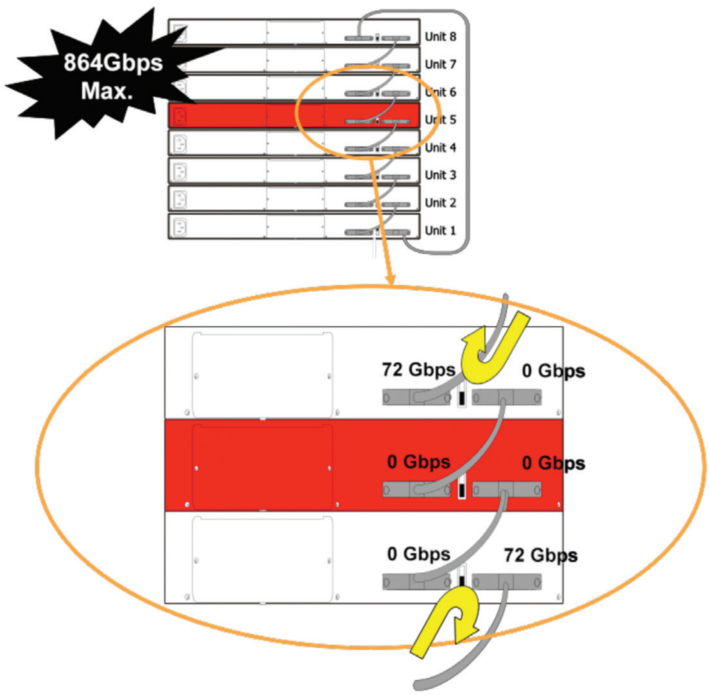


Figure 2: Advantages of FAST in failure scenarios

This flexible stacking capability delivers the 'virtual backplane' that enables these stackable switches to compete with modular chassis designs. It also allows the ERS 5000s to be deployed in a wide range of scenarios ranging from wiring closet, aggregation layer, data center top of rack with horizontal stacking, and small core deployments. This gives network administrators great flexibility in choosing a robust platform that will grow with increased network demands and the ability to simplify their training and sparing requirements.

This same FAST architecture is present in the latest generations of Avaya stackable products: The ERS 5600, ERS 5500, ERS 4000, and ERS 3500* series stackable switches. The available stacking bandwidth decreases respectively as you move down the product families, but the underlying intelligent architecture remains the same, making each switch family well positioned in its respective market segment. The ERS 5500 series scales up to 640 Gbps, ERS 4000 Series scales up to 384 Gbps, the ERS 3500 scales to 80 Gbps, and the ERS 2500 scales to 32 Gbps of virtual backplane capacity in a stack of eight units.

* Stacking will be available on the ERS 3500 in the v5.1 release

Control Plane

Stacking is a very complex technology that requires extensive knowledge of the other switches within a given stack. Requiring more than just robust hardware architecture, the key to successful stacking technology lies within the control plane. Although the specifics that make up the FAST control plane are largely confidential, some elements are discussed briefly here.

The control plane consists of a number of components, including a messaging protocol, that allows all units within the stack to synchronize and exchange configuration and forwarding information between the other switches in the stack. Control plane messages are prioritized over other packets to ensure continued exchange of information in stack congestion scenarios. However, the high stacking bandwidth inherent in our switches and the efficient usage thereof, greatly reduces the odds of congestions vis a vis other stacking technologies.

It is this messaging protocol that allows some of the more complex features, such as Auto Unit Replacement (AUR), New Unit Quick Config, and Switch Clustering, to bring increased levels of resiliency and ease of use to our stacking products. It also simplifies operations by enabling features (such as Stack Health Monitor and Stack Recovery) should there be issues on the stack.

In addition to the messaging protocol, shortest path and load balancing algorithms are implemented to provide more efficient usage of available stack bandwidth. Other portions of the control plane allow the ERS 5500 and ERS 5600 switches, which have different stacking bandwidth capabilities, to stack together and auto-adjust the stack bandwidth to the unit's upstream and downstream neighbors.

For example, consider an eight unit stack comprised of 4 ERS 5600 and 4 ERS 5500 switches. The ERS 5600s are units #1 through #4 and the ERS 5500s are units #5 through #8. The ERS 5600s are capable of 144Gbps stacking (as illustrated earlier) but the ERS 5500s are only capable of 80Gbps stacking (40Gbps upstream and 40Gbps downstream).

Traditional stacking technologies would typically revert down to the lowest common denominator (80Gbps), or worse, not allow them to stack together at all. The intelligence and flexibility of the FAST architecture allows the units in the stack to detect what their upstream and downstream neighbor's stacking capabilities are, and adjust accordingly to match.

In this example, only the boundary units (those adjacent to a dissimilar stacking bandwidth) would adjust their stacking downward. Unit #1 would detect that unit #2 is another ERS 5600 and, therefore, allow the Cascade Down port to operate at 72Gbps, but it would also detect that unit #8 (its upstream neighbor) is an ERS 5500 and would then adjust the Cascade Up port to operate at 40Gbps. Unit #2 and #3 would have all their stack ports operate at 72Gbps, since units on either side of them are ERS 5600s. Unit #4 would detect that unit #3 is an ERS 5600 and allow the Cascade Up port to operate at 72Gbps, but it would also detect that unit #5 (its downstream neighbor) is an ERS 5500 and then would adjust the Cascade Down port to operate at 40Gbps. Units #5 through #8 would have all their Cascade ports operate at 40Gbps.

As illustrated in Figure 3, this unique intelligence allows for a mixed stack with 832Gbps stacking bandwidth vs. the alternative 640Gbps if the lowest common stacking denominator was selected for the entire stack.

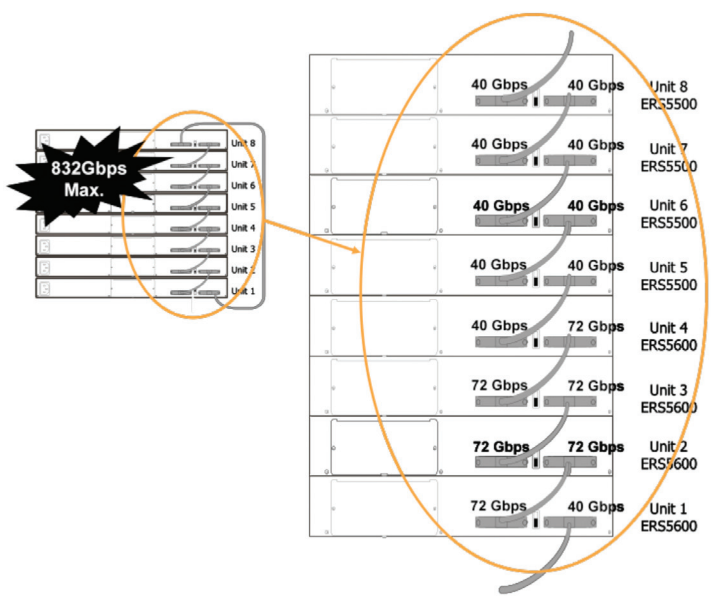


Figure 3: Maximization of stack bandwidth in a mixed ERS 5500/ERS 5600 stack

This sophisticated level of stack intelligence has several benefits. It allows the customer to migrate from one product (the ERS 5500) to a newer product (the ERS 5600) while preserving their previous investment. It also enables optimal

About Avaya

Avaya is a global provider of business collaboration and communications solutions, providing unified communications, contact centers, networking and related services to companies of all sizes around the world. For more information please visit **www.avaya.com**.

usage of the stack bandwidth by allowing the maximum bandwidth possible at each point of demarcation to the virtual backplane.

Switching Software

The operational switching software is layered on top of the control plane and physical hardware and delivers a strong suite of Layer 2 and Layer 3 features. With a code-base that draws on more than 14 years of feature development (beginning with the BayStack 350), the switching software is feature-rich, robust, and able to serve the needs of the most demanding wiring closet applications. Most packet handling decisions (Layer 2 and Layer 3 forwarding, VLAN tagging, Quality of Service, etc.) are made to the packet, in hardware, at the point of ingress. However, other points of software control require collaboration with the stack's Base Unit, requiring a reliable hardware platform and robust control plane architecture.

Trifecta

Perhaps the benefits of Avaya stacking architecture are most evident in the ability to deploy this single product line in three very different environments – the trifecta. ERS 5000 series products are equally well suited for high density enterprise wiring closets, data center top of rack, and small core applications. This flexibility, primarily due to the high bandwidth stacking architecture, makes it possible to deliver a highly scalable, efficient, always-on platform that can be used in a variety of different deployments today, and the as-yet-unknown deployments of tomorrow.

06/12 • DN7087