Arista 7280R3 Switch Architecture ('A day in the life of a packet')

Figure 1: Arista 7280R3 Universal Leaf Systems

The Arista Networks 7280R3 Series are purpose-built high-performance systems available in both fixed and modular form-factors. Full featured switching and routing, combined with deep buffers and a virtual output queued architecture make the 7280R3 Series ideal Universal Leaf or Spine platforms. The 7280R3 series is the latest in the evolution of fixed systems that commenced in 2010 with the first generation systems delivering 1G and 10G as a deep buffer fixed leaf based on the Broadcom Petra silicon. Through successive incremental bandwidth and feature enhancements, the 7280R3 series supports a wide variety of systems options and interfaces, including 25G, 100G, and 400G, based on the Broadcom Jericho2 silicon family.

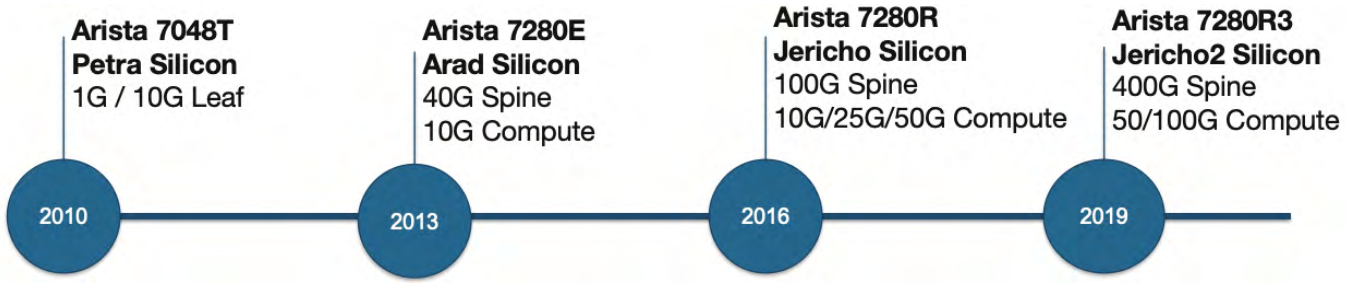


Figure 2: Arista 7280R Series Platform Generations

The 7280R3 series addresses the demands of modern networking, AI/ML workloads and rich multimedia content delivery, requiring a lossless forwarding solution. They are ideal for Universal Cloud Networks, where deep buffers and wire speed L2 and L3 forwarding are combined with advanced features for network virtualization, open and programmable monitoring and network analysis, resiliency, and architectural flexibility.

Additionally, the deep packet buffers, rich routing stack and support for highly scalable IPv4 and IPv6 tables allow for a wide range of open networking solutions, including Cloud WAN aggregation, Service Provider NFV, Internet Peering, Peering Exchanges, Overlay Networks, Content Delivery, and evolution of the network edge.

Since its introduction, the Arista 7280 series has delivered continuous improvements in performance and density with each generation, with significant increases in feature breadth and functionality and investment protection of existing deployments. With the R3-Series, the 7280 family has evolved from a best-in-class Cloud-scale switching to an Internet-scale, service provider routing platform. This white paper provides an overview of the architecture of the Arista 7280R3 Series.

Arista 7280R3: Overview

The Arista 7280R3 Series of Universal Leaf and Spine platforms represents the evolution of the 7280R family of fixed form-factor switches and are available in a choice of 1RU and 2RU fixed systems and a 4RU compact modular system. The 7280R3 has a number of differentiators compared to other fixed configuration systems. The major advantages are derived from the proven VOQ and deep buffer architecture combined with the rich EOS feature set and a programmable pipeline.

- Standards-based high-density 100G and 400G switch ideal for future proof designs and high bandwidth needs, such as in Internet Exchanges, High-Performance Storage or Content Delivery Networks (CDNs)
- Extensive L2 and L3 feature set for open multi-vendor networks with no proprietary lock-in
- Comprehensive, wire-speed, cloud data center virtualisation and carrier routing with EVPN-VXLAN, MPLS and Segment Routing
- Scalable forwarding table resources allow deployment flexibility in both large L2 and L3 environments with any-workload suitability
- Programmable packet processor for advanced features and flexible profiles enabling different use cases including peering and DCI
- Ultra-deep buffers in a compact form-factor make for an ideal spine switch where lossless performance and in-cast problems are expected, such as in big data analytics, search and IP storage.
- Line rate encryption with MACsec, IPsec or Arista TunnelSec™ encryption for simple, reliable and scalable strong layer 2 and layer 3 security for WAN, data center interconnect and for securing links between tiers in leaf and spine data center designs.
- Flexible support for 100G and 400G with a wide selection of optics and cables including 400G OSFP and QSFP-DD form-factors
- Ideal for directly connected 25GbE, 50GbE, 100Gbe and 200GbE attached storage systems, requiring high performance and predictable latency.
- Accelerated sFlow and IPFIX for network forensics
- Streaming network state for advanced analytics with Arista CloudVision®
- Hardware-assisted IEEE 1588 PTP and Synchronous Ethernet enables accurate timing solutions across Ethernet-based networks, without costly investment in separate timing networks.
- Unique monitoring and provisioning features – LANZ, DANZ, AEM,, ZTP, VM Tracer, and eAPI
- NEBS compliance and DC power supplies designed for service provider environments

The 7280R3 platform is designed for lossless behavior, in environments where large scale routing, VXLAN, DANZ, and enhanced LANZ are needed. It is ideal for networks where 100G and 400G uplinks are planned, for investment protection, and any place where intensive workloads or system scalability are a concern. In addition, all members of the 7280R3 Series support full Internet-scale peering.

At a system level the Arista 7280R3 scales from 800 Gbps to 21.6 Tbps switching capacity for up to 96 x 100G in 2 RU and 54 x 400G in 2RU providing industry-leading performance and density without compromising on features, functionality, or investment protection for small, medium and large scale use cases.

System Specifications

The tables below provide system specifications for each of the 7280R3 platforms.

Table 1: Arista 7280R3 and 7280R3K 400G Systems Port Combinations and Forwarding Metrics

7280R3 Series	7280DR3A-54 Series	7289R3 Modular Series	7280DR3A-36 Series	7280PR3-24 Series	7280DR3-24 Series
Max 400GbE Ports*	54	36	36	24	24
Max 100GbE Ports*	216	144	144	96	96
Max 50GbE Ports*	348	144	232	192	192
Max 40GbE Ports*	54	144	36	24	24
Max 25GbE Ports*	348	144	232	192	192
Max 10GbE Ports*	348	144	232	192	192
Max Total Interfaces§	348	216	232	192	192
L2/3 Throughput	21.6 Tbps	14.4 Tbps	14.4 Tbps	9.6 Tbps	9.6 Tbps
L2/3 PPS	8.1 Bpps	5.4 Bpps	5.4 Bpps	4 Bpps	4 Bpps
Latency	From 3.8 us				
Total System Buffer	24 GB	16 GB	16 GB	16 GB	16 GB
Rack Units	2	4	2	1	1
Airflow	Front to Rear				

Table 2: Arista 7280R3 and 7280R3K 400G/100G Systems Port Combinations and Forwarding Metrics

7280R3 Series	7280CR3A-24D12 Series	7280CR3A-48D6 Series	7280CR3-32P4 Series	7280CR3-32D4 Series	7280CR3A-32S	7280CR3-36S Series
Max 400GbE Ports*	12	6	4	4	2	2
Max 100GbE Ports*	72	72	48	48	32	40
Max 50GbE Ports*	112	114	96	96	48	80
Max 40GbE Ports*	36	54	36	36	32	36
Max 25GbE Ports*	116	114	96	96	48	120
Max 10GbE Ports*	116	116	96	96	48	120
Max Total Interfaces§	116	116	96	96	56	120
L2/3 Throughput	7.2 Tbps	7.2 Tbps	4.8 Tbps	4.8 Tbps	3.2 Tbps	2.4 Tbps
L2/3 PPS	2.7 Bpps	2.7 Bpps	2 Bpps	2 Bpps	1.35 Bpps	1 Bpps
Latency	From 3.8 us					
Total System Buffer	8 GB	8 GB	8 GB	8 GB	4 GB	4 GB
Rack Units	1	2	1	1	1	1
Airflow	Front to Rear and Rear to Front					

** Maximum port numbers are uni-dimensional, may require the use of break-outs and are subject to transceiver/cable capabilities.*

§ Where supported by EOS, each system supports a maximum number of interfaces. Certain configurations may impose restrictions on which physical ports can be used.

Table 3: Arista 7280R3 and 7280R3K 25G/100G Systems Port Combinations and Forwarding Metrics				
7280R3 Series	**7280CR3-96 Series**	**7280CR3A-72 Series**	**7280SR3A-48YC8 Series**	**7280SR3-48YC8 Series**
Max 100GbE Ports*	96	72	8	8
Max 50GbE Ports*	192	102	54	16
Max 40GbE Ports*	96	72	8	8
Max 25GbE Ports*	192	102	54	80
Max 10GbE Ports*	192	102	54	80
Max 1GbE Ports*	–	–	–	48
Max total interfaces§	192	116	56	80
L2/3 Throughput	9.6 Tbps	7.2 Tbps	3.6 Tbps	2.0 Tbps
L2/3 PPS	4 Bpps	2.7 Bpps	1.35 Bpps	1 Bpps
Latency	From 3.8 us			
Total System Buffer	16 GB	8 GB	4 GB	4 GB
Rack Units	2	2	1	1
Airflow	F/R	Front to Rear and Rear to Front		

** Maximum port numbers are uni-dimensional, may require the use of break-outs and are subject to transceiver/cable capabilities.*
§ Where supported by EOS, each system supports a maximum number of interfaces. Certain configurations may impose restrictions on which physical ports can be used.

Table 4: Arista 7280R3 and 7280R3K 25G/100G Systems Port Combinations and Forwarding Metrics		
7280R3 Series	**7280SR3-40YC6 Series**	**7280TR3-40C6 Series**
Max 100GbE Ports*	6	6
Max 50GbE Ports*	12	12
Max 40GbE Ports*	6	6
Max 25GbE Ports*	52	4
Max 10GbE Ports*	52	40
Max 1GbE Ports*	36	36
Max total interfaces§	52	52
L2/3 Throughput	800 Gbps	800 Gbps
L2/3 PPS	600 Mpps	600 Mpps
Latency	From 3.8 us	
Total System Buffer	2 GB	2 GB
Rack Units	1	1
Airflow	Front to Rear and Rear to Front	

** Maximum port numbers are uni-dimensional, may require the use of break-outs and are subject to transceiver/cable capabilities.*
§ Where supported by EOS, each system supports a maximum number of interfaces. Certain configurations may impose restrictions on which physical ports can be used.

Arista 7280R - Router Table Scale, Features and Functionality

In addition to increasing system capacity and performance, forwarding table sizes have continued to grow. Arista's innovative FlexRoute™ Engine enables more than 2.5 million IPv4 and IPv6 route prefixes in hardware, significantly beyond what merchant silicon enables natively. Arista EOS NetDB™ evolution of SysDB allows for increased scale and performance with industry-leading routing convergence, creating the first fixed switch system to truly be called a router. K-Series models with larger on-chip resources extend FlexRoute'scapability to over 5M routes, with the ability to contain multiple full-route table copies and ensure many years of investment protection.

The table below shows the key scale metrics of the 7280R3 and 7280R3K Series.

Table 5: Arista 7280R3 Key L2, L3 Scale Metrics[1]

	7280R3/R3A Series		7280R3K/R3AK Series			
	L3 Profile (default)	Balanced Profile	L3-XL Profile (default)	L3-XXL Profile	L3-XXXL Profile	Balanced-XL Profile
ARP Entries	88k	80k	112k	112k	80k	96k
MAC Addresses	224k	224k	256k	192k	384k	256k
IPv4 Unicast Routes	1450k	800k	2250k	2850k	3950k	1850k
Additional IPv4 Unicast Routes with FlexRoute	+ 1,792k	+ 1,792k	+ 2,048k	+ 1,536k	+ 3,072k	+ 2,048k
IPv6 Unicast Routes	433-483k	250-267k	683-750k	833-950k	1100-1317k	567-617k
Multicast Routes	128k	128k	128k	128k	128k	128k
TCAM ACL Entries (Per chip)	24k	24k	24k	24k	24k	24k
Traffic Policy ACL IPv4 Prefixes	30k	30k	430k	296k	30k	430k
Traffic Policy ACL IPv6 Prefixes	10k	10k	150k	100k	10k	150k
ECMP	512-Way	512-Way	512-Way	512-Way	512-Way	512-Way

[1]Unidimensional scaling maxima with currently available MDB profiles. Available resources depend on user configuration.

The table below shows the key scale metrics of the 7280SR3-40YC6 and 7280TR3-40C6 series:

Table 6: Arista 7280SR3-40YC6 Key L2, L3 Scale Metrics[1]

	7280SR3-40YC6 Series and 7280TR3-40C6 Series		
	L3-XL Profile (default)	L3-XXL Profile	Balanced-XL Profile
ARP Entries	72k	72k	64k
MAC Addresses	128k	96k	128k
IPv4 Unicast Routes	1100k	1300k	900k
Additional IPv4 Unicast Routes with FlexRoute	+ 1024k	+ 768k	+ 1024k
IPv6 Unicast Routes	350-383k	416-483k	283-300k
Multicast Routes	128k	128k	128k
TCAM ACL Entries (Per chip)	12k	12k	12k
Traffic Policy ACL IPv4 Prefixes	280k	187k	280k
Traffic Policy ACL IPv6 Prefixes	93k	62k	93k
ECMP	512-Way	512-Way	512-Way

[1]Unidimensional scaling maxima with currently available MDB profiles. Available resources depend on user configuration.

Arista 7280R3 - Cloud Scale And Features

Each iteration of the packet processor silicon in the Arista 7280 Series – from the first-generation in 2010 (Petra / 7048) to the latest (Jericho2 and Jericho2C+ / 7280R3 and 7280R3A) has been riding Moore's Law. This observation, that on average, there will be two times more transistors available every two years has held true for decades. The network packet processing silicon that debuted in the Arista 7048/7280 Series followed this growth curve to more than double performance and density at each generation, from 80 Gbps per packet processor (8 x 10G interfaces) to 7.2 Tbps (18 x 400G interfaces) in Jericho2C+.

In addition to delivering increased port density and performance, forwarding table sizes have continued to increase. The 7280R3 family provides operators with Internet routing scale through Arista's innovative FlexRoute™ Engine, extending forwarding table capacity beyond what merchant silicon enables natively. With the 7280R3, this innovation continues. The 7280R3 series introduces the Modular Database (MDB) to enable the flexible allocation of forwarding resources to accommodate a wide range of network deployment roles.

The MDB provides a common database of forwarding and lookup resources to the ingress and egress stages in the 7280R3 platform. These resources are allocated using forwarding profiles that ensure the optimal allocation to different tables for a wide range of networking use-cases. Unlike legacy routing silicon architectures, increasing the size of forwarding tables does not compromise the performance or latency of packet processing, which remains wire-speed. Layer 3 optimized profiles expand the routing and next-hop tables to address large scale networks where route table capacity is required, while the balanced profile is suited for leaf and spine data center applications.

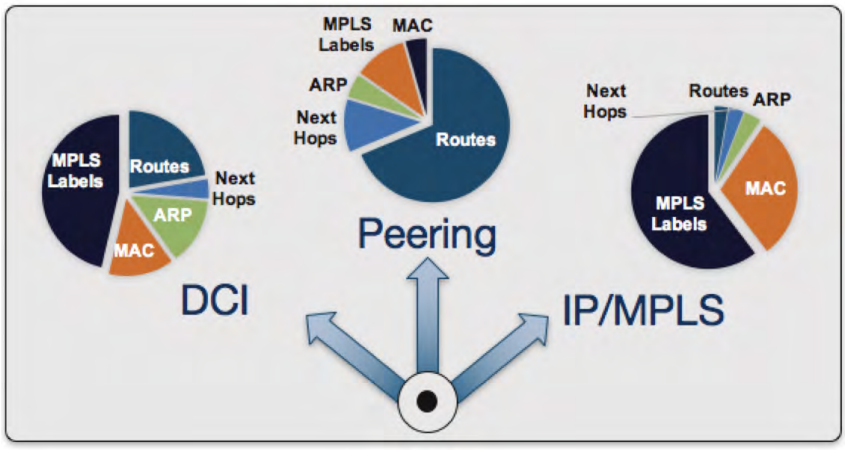


Figure 3: MDB enables a flexible range of deployment profiles.

The fungible nature of the resources within the MDB ensures that operators have the flexibility they need to standardize on a common platform across a wide range of roles with the confidence that the specific resource requirements can be allocated according to the needs of any given role. There is no need to have a separate platform for core network roles and edge roles in today's service provider networks. This enables cloud and service providers to streamline their deployments, simplify sparing and consolidate testing.

Airflow

Many models in the 7280R3 Series offer a choice of airflow direction (front-to-rear or rear-to-front). Field replaceable power and fan modules are color-coded to show the airflow direction through the switch, making it easy to identify how the switch should be installed in a rack where Hot/Cold aisle separation is required.

The red handles shown below denote hot aisle placement with Airflow exit from the switch. Blue handles denote cold aisle placement with Airflow intake to the switch.

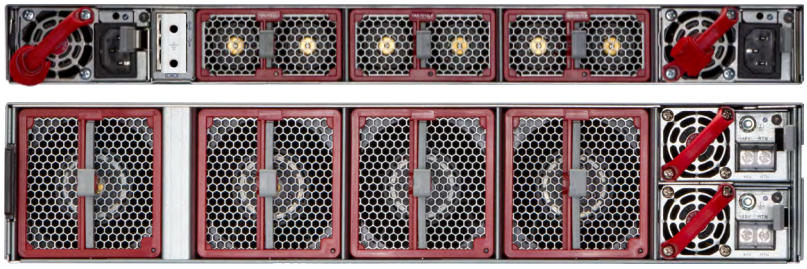


Figure 4: 7280R3 Series 1RU & 2RU Switch Chassis Rear View

Arista 7280R3 Universal Leaf System Architecture

All 7280R3 Series switches share a common system design built around a high-performance x86 CPU for the control plane. The CPU is connected to system memory, internal flash, SSD, boot flash, power supplies, fans, management I/O and peripherals, as shown below.

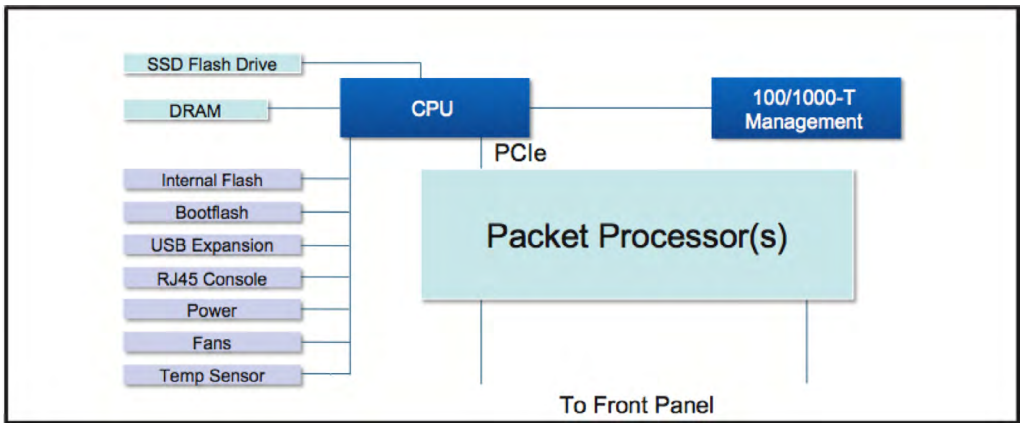


Figure 5: Arista 7280R3 High-Level System Architecture

Arista 7280R3 switches are designed for continuous operations with system-wide monitoring of both hardware and software components, simple serviceability and provisioning to prevent single points of failure.
Key high availability features include:

- 1+1 hot-swappable power supplies and hot-swap fans provide dynamic temperature control combined with N+1 redundancy
- Color-coded PSU's and fans that deliver platinum level power efficiency
- Live software patching
- Self-healing software with Stateful Fault Repair (SFR)
- Smart System Upgrade (SSU)

The x86 CPU is connected over PCIe to the 7280R3 Packet Processor(s) that handle all data plane forwarding and connect directly to the front panel ports.

All stages associated with packet forwarding are performed in an integrated system on chip (SoC) packet processor. The Jericho2 family of packet processors provide both the ingress and egress packet forwarding pipeline stages for packets that arrive or are destined to ports serviced by that packet processor. Each packet processor can perform local switching for traffic between ports on the same packet processor.

The Jericho2 packet processor family consists of multiple chip variants, each designed to meet different bandwidth and interface configurations enabling a common pipeline architecture to be implemented in systems ranging from single chip 800 Gbps platforms up to large scale multi-chip systems like the 7800R3-Series modular family.

Jericho2C+ is the highest specification member of the Jericho2 family. Based on a 7 nm fabrication process, Jericho2C+ implements a pipeline and memory design that is fully compatible and consistent with Jericho2, while providing 7.2 Tbps of front panel bandwidth and offering fully integrated wire speed AES-256-GCM bulk encryption enabling MACsec, IPsec and VXLANsec. In addition 9.6 Tbps of inter-chip fabric capacity is available to connect to other J2C+ chips or to a fabric in large scale multi-chip systems.

Each Jericho2C+ packet processor supports network interface speeds ranging from 10G to 400G delivered over a total of 144 50G PAM4 SerDes lanes that can be run from 10G to 50G individually or combined in groups to allow flexible 10G, 25G, 40G, 50G 100G, 200G, and 400G interfaces.

Each packet processor supports up to a maximum of 118 logical or physical interfaces per chip, which defines the maximum possible port density for a given product form factor. Device specific front panel layout, availability of suitable transceivers and EOS support will govern the actual amount of possible breakouts.

Jericho2 supports network interface speeds ranging from 10G to 400G for up to 4.8 Tbps of total network capacity. In addition 5.6 Tbps of inter-chip fabric capacity is available for multi-chip systems. The Jericho2 packet processor has a total of 96 50G PAM4 SerDes interfaces that can be run from 10G to 50G individually or combined in groups to allow flexible 10G, 25G, 40G, 50G 100G, 200G and 400G interfaces.

Each packet processor supports up to a maximum of 96 logical or physical interfaces per chip, which defines the maximum possible port density for a given product form-factor.

Jericho2M is a half-size variant of Jericho2C+, providing 3.6 Tbps of bandwidth to the front panel with 4.8 Tbps of fabric connectivity. It offers features consistent with Jericho2C+, including optional line-rate encryption, while providing 72 50G PAM4 SerDes lanes and up to 58 logical ports (subject to device specific implementation).

Jericho2C is a member of the Jericho2 silicon family providing 2.4 Tbps of front panel bandwidth and 2.4 Tbps of inter-chip fabric capacity per chip. It is designed for lower capacity systems with a focus on 1G to 100G network connectivity with 100G to 400G uplinks. Each Jericho2C chip supports a maximum of 32 50G PAM4 and 96 25G NRZ SerDes lanes.

The final member of the Jericho2 family, Qumran2a, is designed for single-chip, low capacity systems. Qumran2a supports 800 Gbps of switching capacity with up to 1.6 Tbps of front panel interfaces from a pool of 50G PAM4 and 25G NRZ SerDes lanes.

As with Jericho2, 50G lanes support speeds from 10G to 50G individually or may be combined in groups to support interfaces up to 400G. The 25G NRZ lanes support 1G, 10G or 25G individually or 40G/100G when combined. The broad range of SerDes speeds allows for multiple product form factors.

In many 7280R3 models gearboxes are employed to increase the front panel interface density and maximize the capabilities by converting the 50G PAM4 SerDes lanes to more lanes at lower speeds and different encoding.

Gearboxes enable systems designers to maximize the range of interfaces and interface speeds without needing to add additional packet processors, reducing overall system power consumption and heat generation. As the number of physical interfaces and supported breakout options is flexible EOS provides tools to enable both configuration and analysis of the available port combinations for each platform.

Reviewing the 7280CR3-32P4 as an example, an architecture block diagram is shown in the figure below. In this model, 4 x 400G ports are connected directly to the Jericho2 chip, while the QSFP ports are placed after 8 gearboxes that each convert sets of 8 x 50G PAM4 lanes to 16 x 25G NRZ to support 4-lane 100G and 40G interfaces. This configuration provides a total of 32 QSFP ports and 4 OSFP ports from a single J2 chip. Without the use of gearboxes, the same set of 64 x 50G lanes, each running at 25G, would only support half the number of QSFP ports.

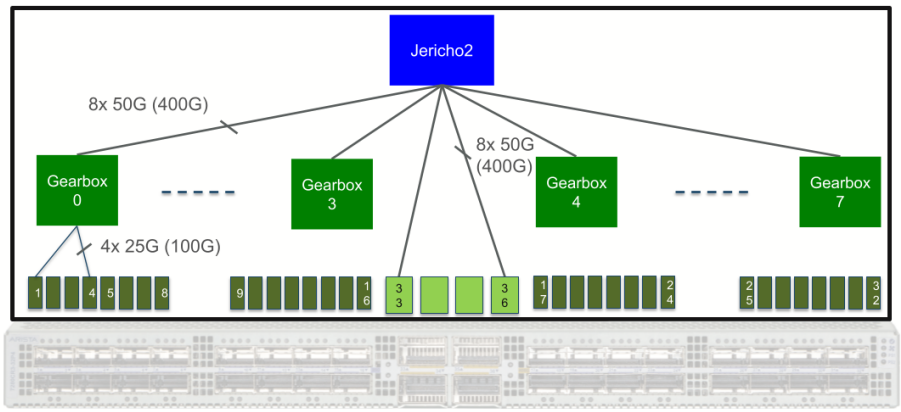


Figure 6: Arista 7280CR3-32P4 Switch Architecture

In the case of the DCS-7280CR3-32P4 each of the 32 100G ports can operate as either 1 x 100G or 1 x 40G, depending on the QSFP optics and cables used. In this configuration, only 1 logical interface is used per QSFP port (as it is a single interface).

When breaking out QSFP ports into 4 individual lanes of 10/25G, each QSFP port requires 4 logical interfaces. In order to meet the 96 logical interface capacity of the J2 chip, odd numbered ports can be operated as either 4x25G or 4x10G, with the adjacent even port disabled.

As the OSFP ports are directly connected to the J2 chip, each port is already equipped with 8 logical interfaces and can support full breakout without restrictions.

Each 400G port is capable of supporting copper, AOC as well as the range of optics available in the OSFP form-factor.

Port Identification

40G and 100G QSFP type transceivers are the same physical size and can be inserted into any QSFP based ports. It's important that switches indicate the port capabilities on the front panel so that an engineer installing transceivers inserts them into the correct port types.

The figures below show QSFP100 and QSFP-DD ports. The 100G QSFP ports are highlighted with a purple line to identify they are QSFP100 capable; these ports will support either QSFP+ (40G) or QSFP100 (100G) transceivers. The 400G capable QSFP-DD ports in the center are highlighted in orange and support 40G, 100G and 400G transceivers. The OSFP 400G ports are marked similarly. On the 7280CR3-36S ports that are 200G capable are marked in a blue color that identifies them as 200G capable. They may also be used for 100G and other speeds.

Figure 7: QSFP 100 & QSFP-DD Ports

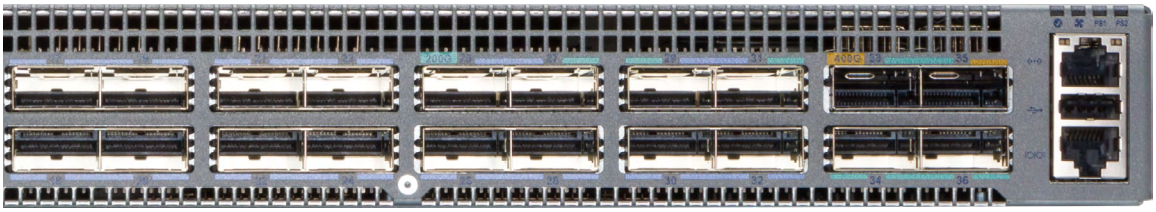


Figure 8: QSFP100, QSFP200 & QSFP-DD Ports

Selected devices include dedicated ports for the 400G-ZR OSFP Line System Amplifier transceiver.

These ports provide a convenient location to deploy OSFP-AMP-ZR amplifiers without consuming data-plane ports. Each port provides sufficient power to operate the OSFP-AMP-ZR amplifier as well as control plane connectivity for monitoring.

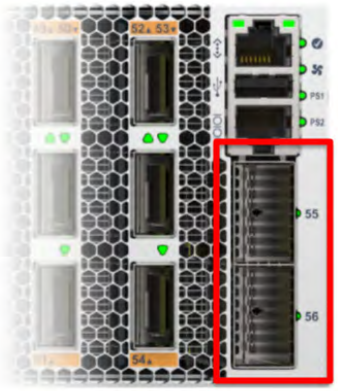


Figure 9: OSFP Line System Ports

Arista 7280R3 Universal Leaf Platform Layout

Arista 7280R3 series switches utilize high performance packet processors, with the number of packet processors varying based on the number and type of ports on the system. The packet forwarding architecture of each of these systems is essentially the same; a group of front-panel ports (different transceiver/port/speed options) are connected to each packet processor. The following diagrams show the layout of each 7280R3 system.

7280DR3A-54

The 7280DR3A-54 is a 2RU system utilizes three Jericho2C+ chips connected in a triangle with each chip providing 18 x 400G ports to the front panel for a total of 54 x 400G or 21.6 Tbps. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

In the default configuration, each port is designed to offer up to 4-way breakout when using either a 100/200G QSFP or a 400G QSFP-DD. 6 logical interfaces are available on each physical port, corresponding to lanes 1,2,3,4,5 and 7. Of these 6 interfaces, up to 4 logical interfaces can be active on each port, which provides the following connectivity choices:

- When used with a 400G QSFP-DD:
  - 1 x 400G-8
  - 2 x 200G-4
  - 4 x 100G-2
  - 2 x 100G-4
  - 4 x 50G-2
- When used with a 200G QSFP:
  - 1 x 200G-4
  - 2 x 100G-2
  - 4 x 50G-1
  - All 40/100G QSFP speeds subject to transceiver support
- When used with a 100G QSFP:
  - 1 x 100G-4
  - 2 x 50G-2
  - 4 x 25G
  - All 40 QSFP speeds subject to transceiver support
- When used with a 40G QSFP:
  - 1 x 40G
  - 4 x 10G

EOS provides alternative configuration profiles which enable different breakout capabilities up to a maximum of 348 logical ports.

Each port is capable of supporting copper or AOC as well as the range of optics available in QSFP or QSFP-DD form-factors. The 7280DR3A-54 also provides two dedicated OSFP-Line System ports. CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

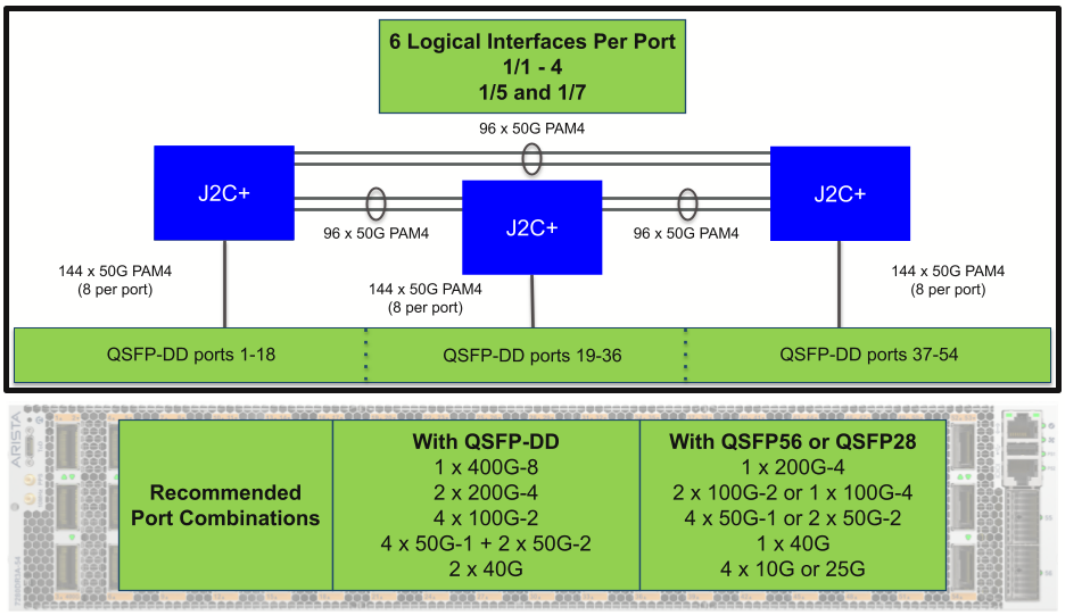


Figure 10a: 7280DR3A-54 architecture

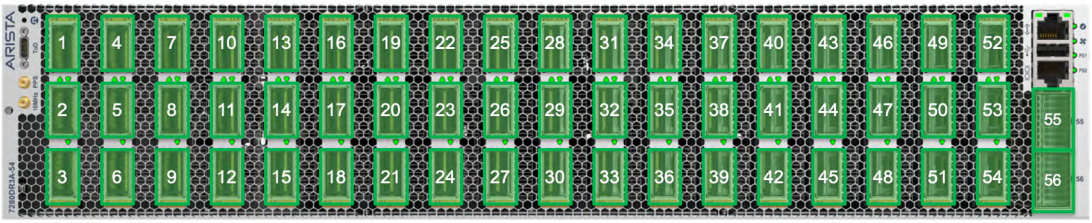
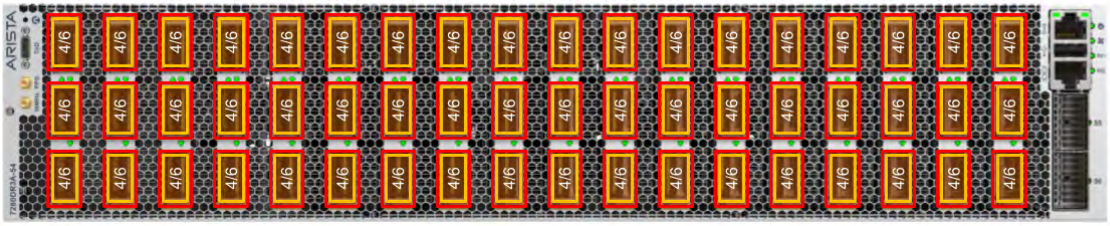


Figure 10b: 7280DR3A-54 port numbering

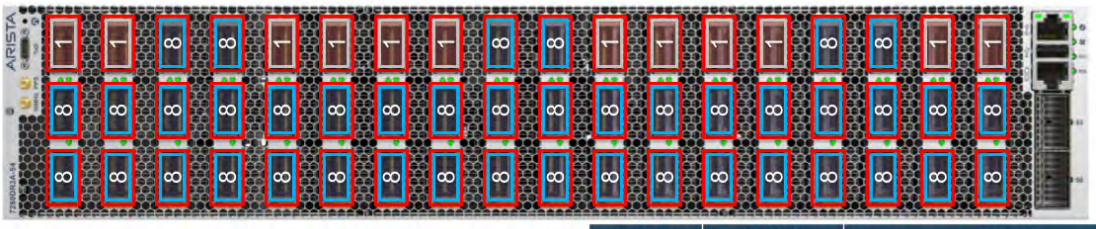


4/6 6 Logical Ports (Max 4 Active) – e.g. Eth x/1,3,5,7 or Eth x/1,2,3,4

Speed Group (Max 2 SerDes Speeds)

Port Type	Max Density	Comment
10/25G	216*	Using all ports
50G-2	216	Using all ports
50G-1	216*	Using all ports
100G-4	108	Using breakout QDD
100G-2	216	Using all ports
200G-4	108	Using all ports
400G-8	54	Using all ports

Figure 10c: 7280DR3A-54 default breakout capabilities (indicates not all signaling lanes are used)*



1 1 Logical Port – Eth x/1

8 8 Logical Ports – Eth x/1,2,3,4,5,6,7,8

Speed Group (Max 2 SerDes Speeds)

Port Type	Practical / Max Density	Comment
10/25G	336/348*	Using 8-way ports only / all ports
50G-2	168/180*	Using 8-way ports only / all ports
50G-1	336/348*	Using 8-way ports only / all ports
Any 50G	348	336 x 50G-1 using 8-way ports & 12 x (1 x 50G-1 or -2) using 1-way ports
100G-4	96*	Using QSFP28 in 1 way ports and breakout QDD on 8 way
100G-2	168/180*	Using 8-way ports only / all ports
Any 100G	180*	42 x (4 x 100G-2) on 8-way ports & 12 x (1x 100G-2 or -4) on 1-way ports
200G-4	96*	Using QSFP56 in 1 way ports and breakout QDD on 8 way
400G-8	54	Using all ports

Figure 10d: 7280DR3A-54 maximum breakout profile (indicates not all signaling lanes are used)*

7280DR3A-36

The 7280DR3A-36 2RU system utilizes two Jericho2C+ chips connected back to back, with each chip providing 18 x 400G ports to the front panel for a total of 36 x 400G or 14.4 Tbps. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

In the default configuration, each port is designed to offer up to 4-way breakout when using either a 100/200G QSFP or a 400G QSFP-DD. 6 logical interfaces are available on each physical port, corresponding to lanes 1,2,3,4,5 and 7. Of these 6 interfaces, up to 4 logical interfaces can be active on each port, which provides the following connectivity choices:

- When used with a 400G QSFP-DD:
  - 1 x 400G-8
  - 2 x 200G-4
  - 4 x 100G-2
  - 2 x 100G-4
  - 4 x 50G-2
- When used with a 200G QSFP:
  - 1 x 200G-4
  - 2 x 100G-2
  - 4 x 50G-1
  - All 40/100G QSFP speeds subject to transceiver support
- When used with a 100G QSFP:
  - 1 x 100G-4
  - 2 x 50G-2
  - 4 x 25G
  - All 40 QSFP speeds subject to transceiver support
- When used with a 40G QSFP:
  - 1 x 40G
  - 4 x 10G

EOS provides alternative configuration profiles which enable different breakout capabilities up to a maximum of 232 logical ports.

Each port is capable of supporting copper or AOC as well as the range of optics available in QSFP or QSFP-DD form-factors. The 7280DR3A-36 also provides two dedicated OSFP-Line System ports. CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

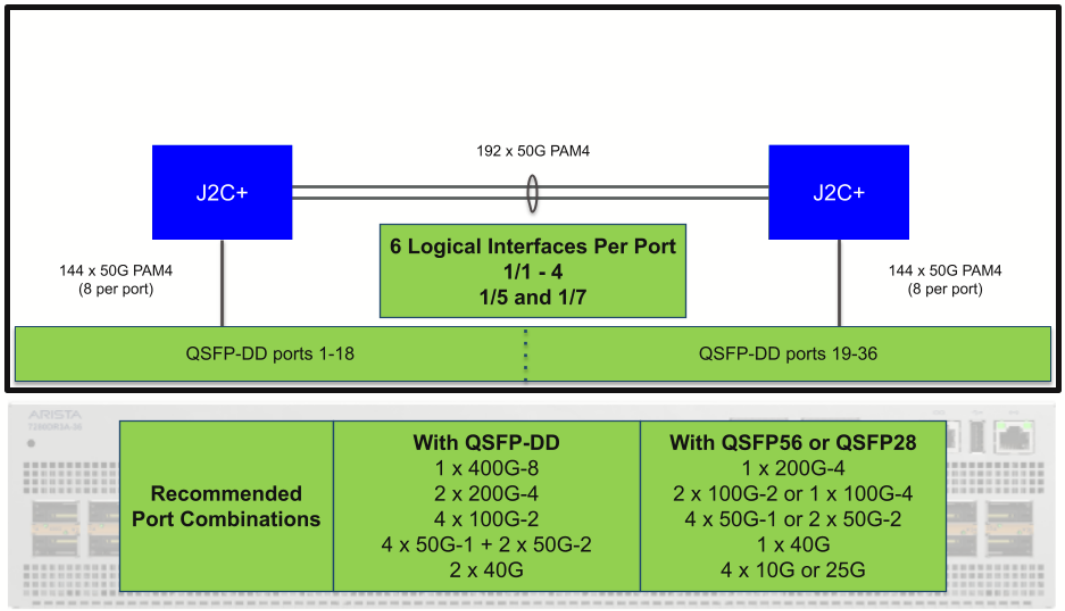


Figure 11a: 7280DR3A-36

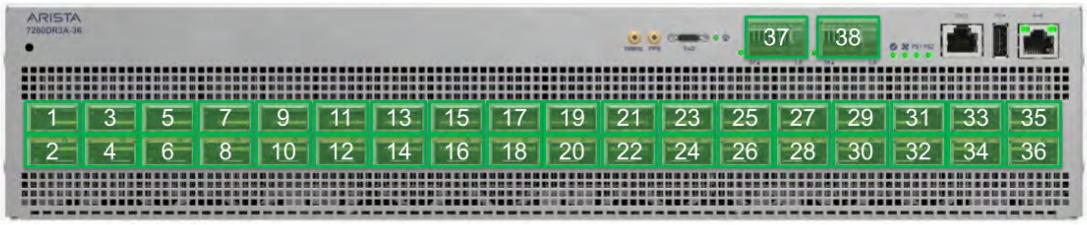


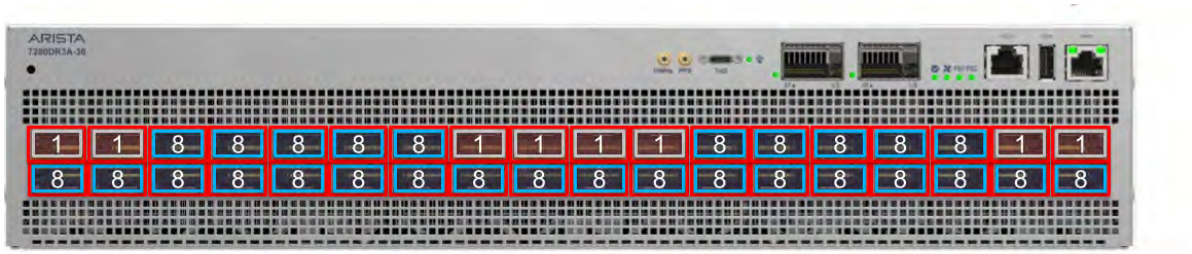
Figure 11b: 7280DR3A-36 port numbering

4/6 6 Logical Ports (Max 4 Active) – e.g. Eth x/1,3,5,7 or Eth x/1,2,3,4

Speed Group (Max 2 SerDes Speeds)

Port Type	Max Density	Comment
10/25G	144*	Using all ports
50G-2	144	Using all ports
50G-1	144*	Using all ports
100G-4	72	Using breakout QDD
100G-2	144	Using all ports
200G-4	72	Using all ports
400G-8	36	Using all ports

Figure 11c: 7280DR3A-36 default breakout capabilities (indicates not all signaling lanes are used)*



1 Logical Port – Eth x/1

8 Logical Ports – Eth x/1,2,3,4,5,6,7,8

Speed Group (Max 2 SerDes Speeds)

Port Type	Practical / Max Density	Comment
10/25G	224/232*	Using 8-way ports only / all ports
50G-2	112/120*	Using 8-way ports only / all ports
50G-1	224/232*	Using 8-way ports only / all ports
Any 50G	232	224 x 50G-1 using 8-way ports & 8 x (1 x 50G-1 or -2) using 1-way ports
100G-4	64*	Using QSFP28 in 1 way ports and breakout QDD on 8 way
100G-2	112/120*	Using 8-way ports only / all ports
Any 100G	120*	28 x (4 x 100G-2) on 8-way ports & 8 x (1x 100G-2 or -4) on 1-way ports
200G-4	64*	Using QSFP56 in 1 way ports and breakout QDD on 8 way
400G-8	36	Using all ports

Figure 11d: 7280DR3A-36 maximum breakout profile (indicates not all signaling lanes are used)*

7280PR3-24 and 7280DR3-24

The 7280PR3-24 and 7280DR3-24 are 1U systems with two Jericho2 chips to deliver 24 ports of 400G with 9.6 Tbps of non-blocking performance. The system supports up to a maximum of 192 interfaces when used in breakout mode.

Each of the 400G ports supports up to 8 unique logical interfaces and can operate as either 1 x 400G, 2x200G, 4 x 100G, 8 x 50G or 8 x 25G interfaces subject to the capability of the cable or transceiver. Each port is capable of supporting copper or AOC as well as the range of optics available in the OSFP form-factor.

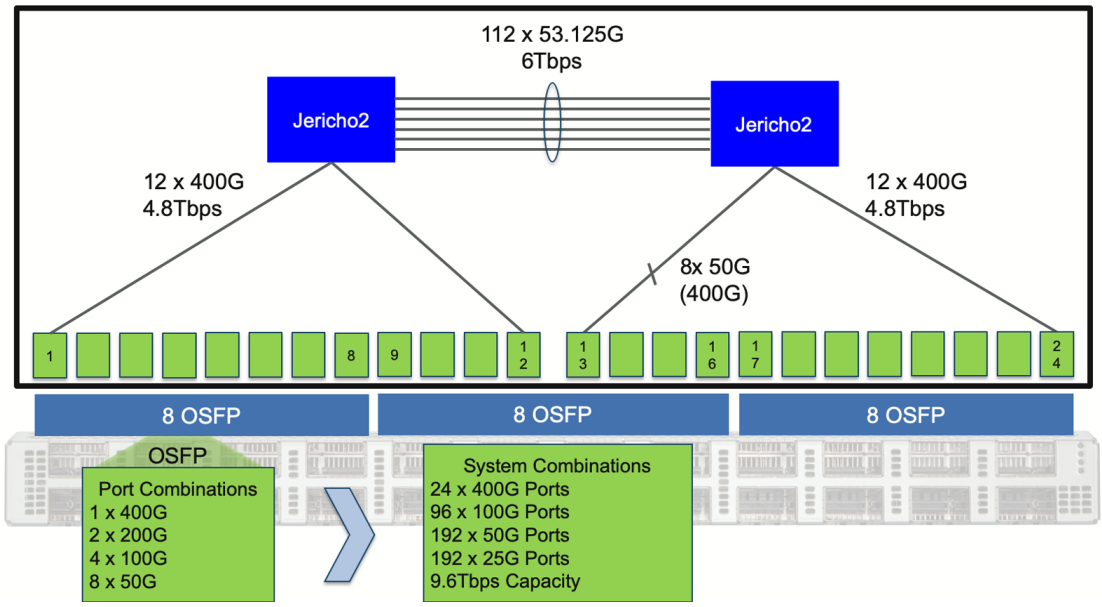


Figure 12: 7280PR3-24

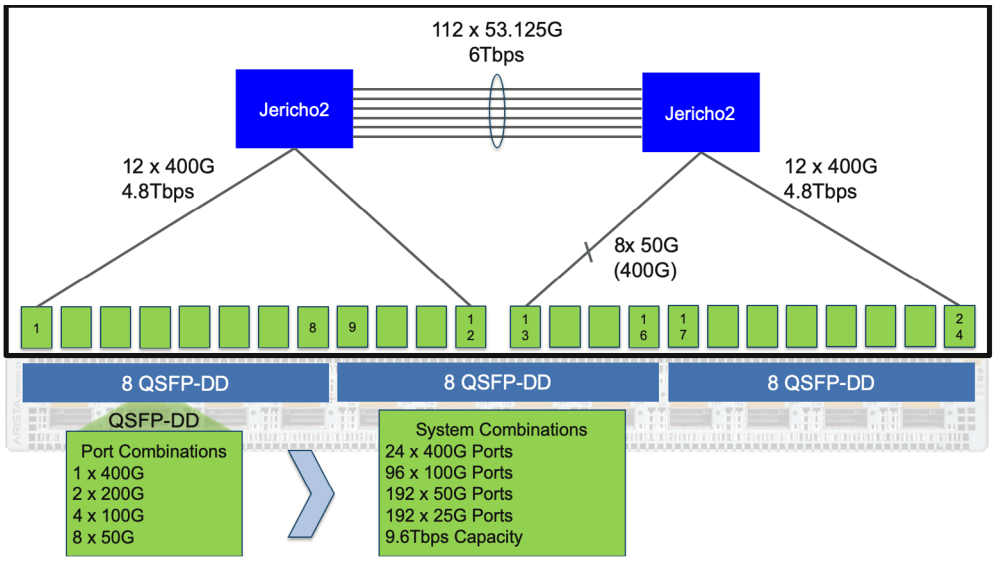


Figure 13: 7280DR3-24

7280CR3A-24D12

The 7280CR3A-24D12 is a 1U system with one Jericho2C+ chip, delivering 7.2 Tbps of non-blocking performance and up to 116 interfaces when used in breakout mode. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

With a mix of direct 50G lanes to the front panel and a total of 6 gearboxes, the system supports 24 ports of 100G QSFP and 12 ports of 400G QSFP-DD.

By default, logical ports are distributed across the front panel to provide sufficient breakout combinations for common deployments and a maximum of 116 total interfaces. Depending on the number of logical ports allocated and other speed combination restrictions, each port can support:

400G QSFP-DD Ports

- When used with a 400G QSFP-DD:
  - 1 x 400G-8
  - 2 x 200G-4
  - 4 x 100G-2
  - 2 x 100G-4
  - 8 x 50G-1*
  - 4 x 50G-2
  - 8 x 10/25G*
- When used with a 200G QSFP:
  - 1 x 200G-4
  - 2 x 100G-2
  - 4 x 50G-1
  - All 40/100G QSFP speeds subject to transceiver support
- When used with a 100G QSFP:
  - 1 x 100G-4
  - 2 x 50G-2
- When used with a 40G QSFP:
  - 1 x 40G
  - 4 x 10G

*Refer to default logical port map for restrictions

Odd/even pairs of QSFP Ports (Odd numbered ports are primary ports)

- Primary port running in 200G mode: 1 x 200G-4 (1 logical interface)
  - » Even numbered port disabled
- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 50G: 4 x 50G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

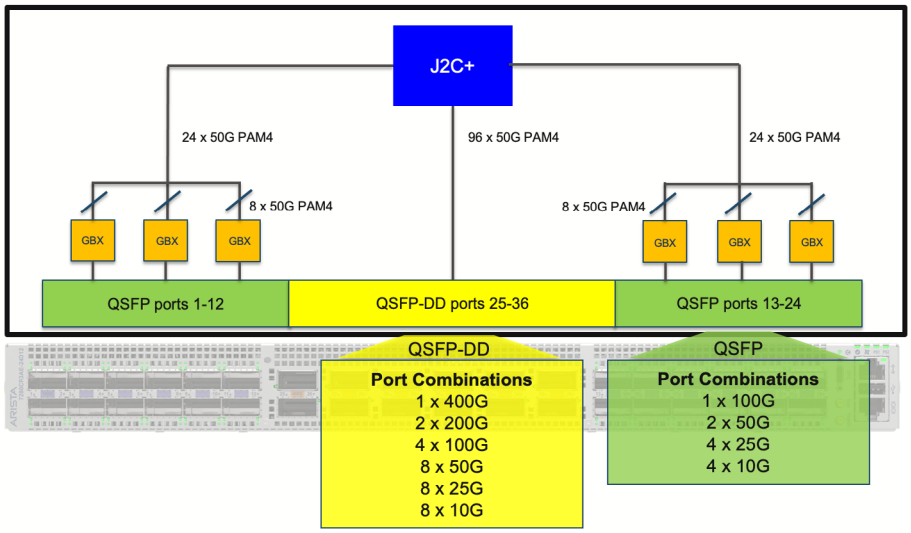


Figure 14a: 7280CR3A-24D12

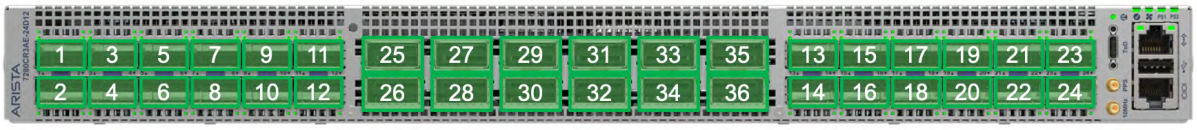


Figure 14b: 7280CR3A-24D12 port numbering

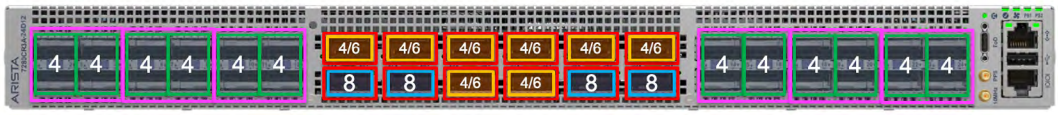
Primary Ports – Support QSA, 4-way breakouts or 200G QSFP-56

Secondary Ports (Disabled if Primary is broken out 4-ways or used for 200G QSFP-56)

Speed Group (Max 2 SerDes Speeds)

Speed Group (Max 2 SerDes Speeds, each pair [Primary (odd)/Secondary (even)] at Same Speed)

Figure 14c: 7280CR3A-24D12 speed and port groups



4 Logical Ports across each Primary/Secondary pair

4/6 6 Logical Ports (Max 4 Active) – e.g. Eth x/1,3,5,7 or Eth x/1,2,3,4

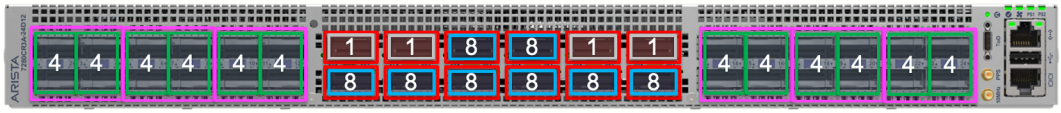
8 Logical Ports – Eth x/1,2,3,4,5,6,7,8

Speed Group (Max 2 SerDes Speeds)

Speed Group (Max 2 SerDes Speeds, each pair [Primary/Secondary] at Same Speed)

Port Type	Max Density	Comment
10/25G	112	Using all ports
50G-2	96	Using all ports
50G-1	80/112*	Using Primary QSFP28 and QDD breakout
Any 50G	112	48 x 50G-2 using QSFP28 & 8 x (4 x 50G-1 or -2) on QDD & 4 x (8 x 50G-1) on QDD
100G-4	48	Using QSFP28 and 100G-4 breakouts
100G-2	72	Using Primary QSFP with QSFP56 breakout and QDD breakout
Any 100G	72	24 x 100G-4 using QSFP28 & 12 x (4 x 100G-2) on QDD
200G-4	36	Using Primary QSFP28 and QDD breakout
400G-8	12	Only supported on Q-DD ports

Figure 14d: 7280CR3A-24D12 default breakout capabilities (indicates not all signaling lanes are used)*



4 4 Logical Ports across each Primary/Secondary pair

1 1 Logical Port – Eth x/1

8 8 Logical Ports – Eth x/1,2,3,4,5,6,7,8

Speed Group (Max 2 SerDes Speeds)

Speed Group (Max 2 SerDes Speeds, each pair [Primary/Secondary] at Same Speed)

Port Type	Max Density	Comment
10/25G	116*	Using all ports
50G-2	84*	Using all ports
50G-1	116*	Using Primary QSFP28 and QDD breakout
100G-4	44*	Using QSFP28 and 100G-4 breakouts
100G-2	60*	Using Primary QSFP with QSFP56 breakout and QDD breakout
200G-4	32*	Using Primary QSFP28 and QDD breakout
400G-8	12	Only supported on Q-DD ports

Figure 14e: 7280CR3A-24D12 maximum breakout profile (indicates not all signaling lanes are used)*

7280CR3A-48D6

The 7280CR3A-48D6 is a 2U system with one Jericho2C+ chip, delivering 7.2 Tbps of non-blocking performance and up to 116 interfaces when used in breakout mode. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

With a mix of direct 50G lanes to the front panel and a total of 12 gearboxes, the system supports 48 ports of 100G QSFP and 6 ports of 400G QSFP-DD. The 7280CR3A-48D6 also provides a dedicated OSFP-Line System port.

By default, logical ports are distributed across the front panel to provide sufficient breakout combinations for common deployments and a maximum of up to 116 total interfaces. Depending on the number of logical ports allocated and other speed combination restrictions, each port can support:

400G QSFP-DD Ports

- When used with a 400G QSFP-DD:
  - 1 x 400G-8
  - 2 x 200G-4
  - 4 x 100G-2
  - 2 x 100G-4
  - 4 x 50G-2
- When used with a 200G QSFP:
  - 1 x 200G-4
  - 2 x 100G-2
  - 4 x 50G-1
  - All 40/100G QSFP speeds subject to transceiver support
- When used with a 100G QSFP:
  - 1 x 100G-4
  - 2 x 50G-2
  - 4 x 25G
  - All 40 QSFP speeds subject to transceiver support
- When used with a 40G QSFP:
  - 1 x 40G
  - 4 x 10G

Odd/even pairs of 100G QSFP Ports (Odd numbered ports are primary ports)

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - Even numbered port disabled

CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

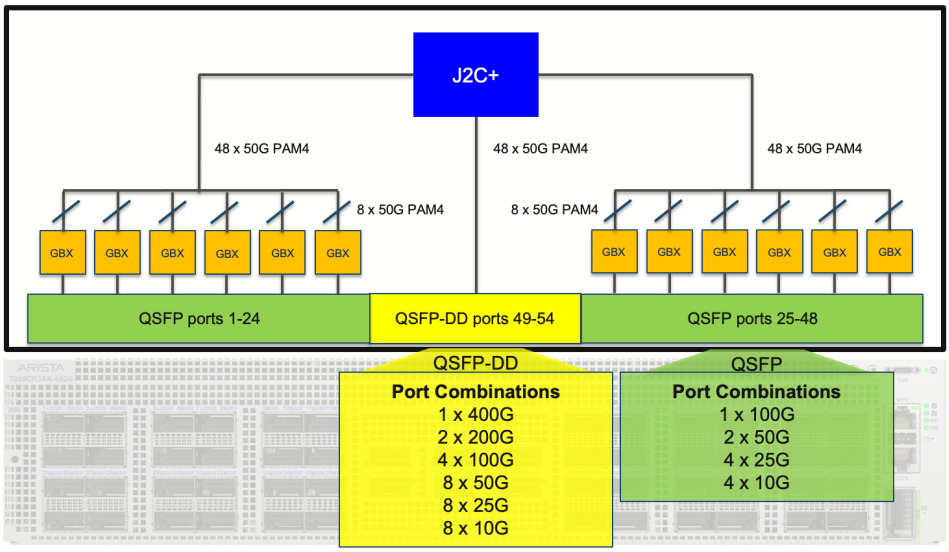


Figure 15a: 7280CR3A-48D6

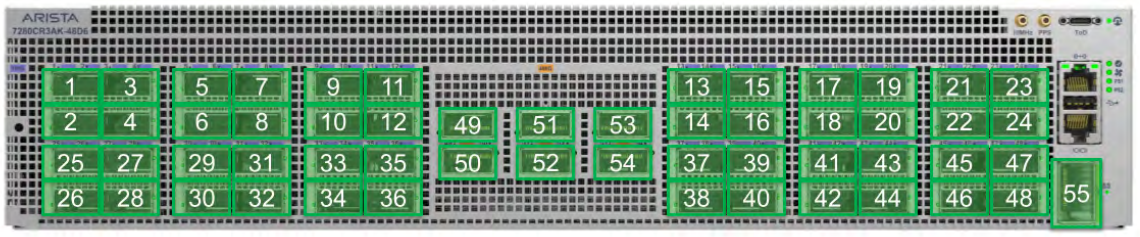


Figure 15b: 7280CR3A-48D6 port numbering

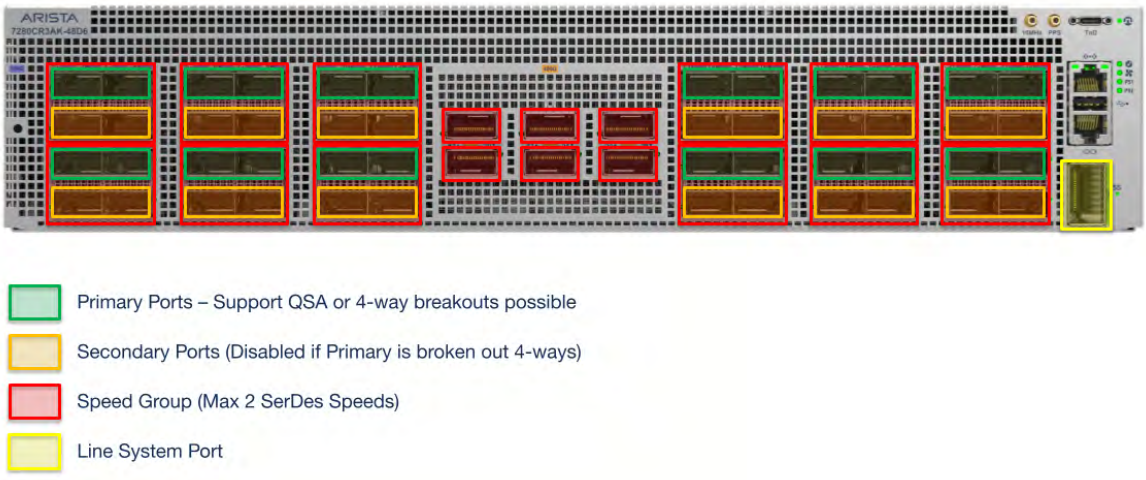
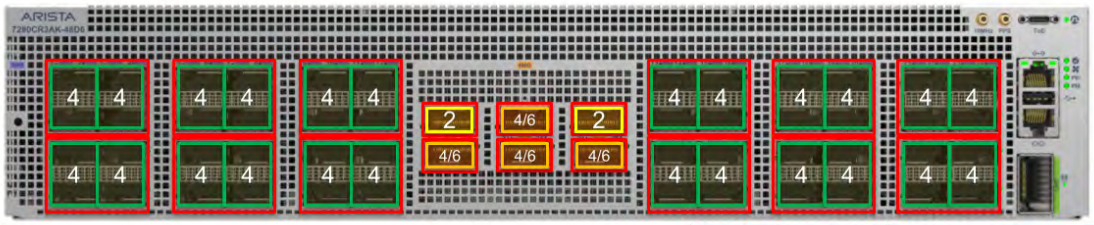


Figure 15c: 7280CR3A-48D6 speed and port groups



4	4 Logical Ports across each Primary/Secondary pair
4/6	6 Logical Ports (Max 4 Active) – e.g. Eth x/1,3,5,7 or Eth x/1,2,3,4
2	Two Logical Ports – Eth x/1, x/5**
	Speed Group (Max 2 SerDes Speeds)

Port Type	Practical / Max Density	Comment
10/25G	114	Using all 4-way ports + QSA in 2 way ports
50G-2	114*	Using all 4-way ports + first lane of QSFP28 in 2-way ports
50G-1	18*	Only supported on Q-DD ports. Single 50G-1 on 2-way ports using 1/4 lanes of QSFP56
Any 50G	114	96 x 50G-2 using QSFP28 & 4 x (4 x 50G-1 or -2) on 4-way QDD & 2 x (1 x 50G-1 or -2) on 2-way QDD
100G-4	60	Using QSFP28 and breakout QDD to 2 x 100G-4
100G-2	16/18*	4 per 4-way QDD & 1 per 2-way QDD
Any 100G	68	48 x 100G-4 using QSFP28 & 2 x (2 x 100G-4) on 2-way QDD & 4 x (4 x 100G-2) on 4-way QDD
200G-4	12	Only supported on Q-DD ports
400G-8	6	Only supported on Q-DD ports

Figure 15d: 7280CR3A-48D6 default breakout capabilities (indicates not all signaling lanes are used)*

7280CR3-32P4

The 7280CR3-32P4 system utilizes a single Jericho2 chip which supports up to 96 individual interfaces when used in breakout mode.

A total of 8 gearboxes allows the system to support 32 ports of 100G and a diverse range of optics. Four logical interfaces are assigned to each odd/even pair of QSFP ports, which allows for combinations including:

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

MACsec enabled variants of the platform implement in-line encryption in the gearboxes, providing line-rate encryption to the 100G QSFP ports.

Each of the four 400G ports supports up to 8 unique logical interfaces and can operate as either 1 x 400G, 2x200G, 4 x 100G, 8 x 50G or 8 x 25G interfaces subject to the capability of the cable or transceiver. Each port is capable of supporting copper or AOC as well as the range of optics available in the OSFP form-factor.

CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

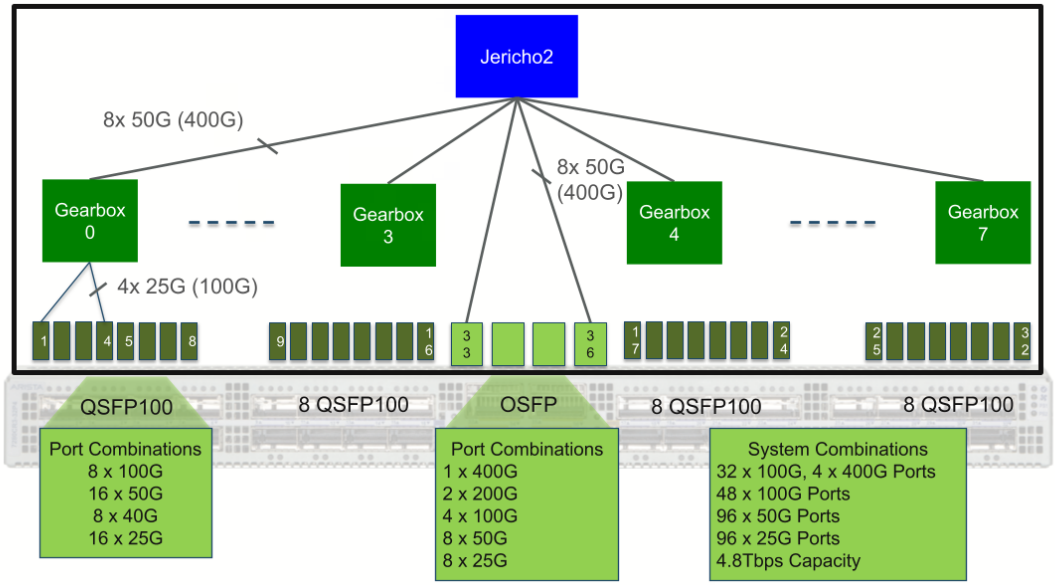


Figure 16: 7280CR3-32P4

7280CR3-32D4

The 7280CR3-32D4 system is the QSFP-DD version of the 7280CR3-32P4. The packet processor to port assignment is identical. Further, each 400G port is capable of supporting copper or AOC as well as the range of optics available in the QSFP-DD form-factor.

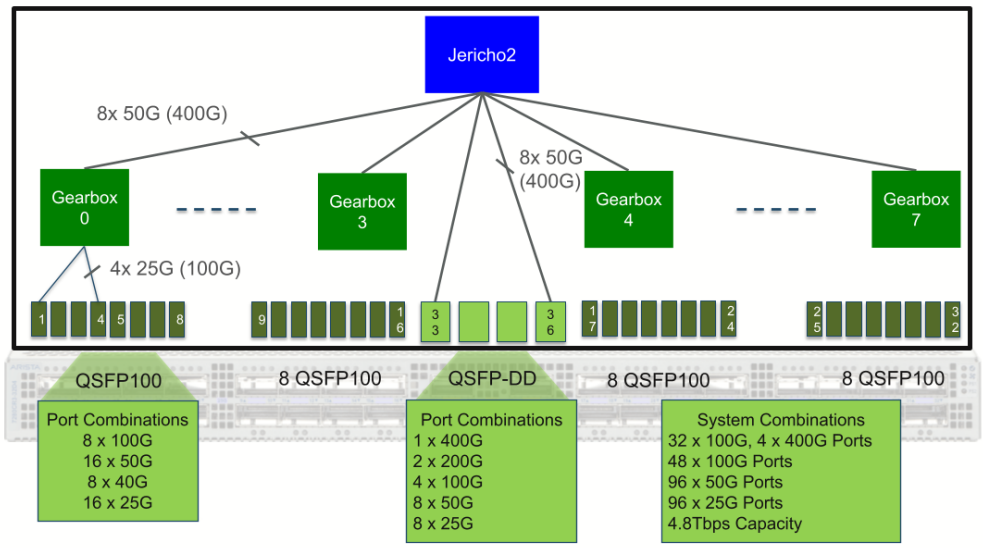


Figure 17: 7280CR3-32D4

7280CR3-96

The 7280CR3-96 is a 2U system with two Jericho2 chips to deliver 9.6 Tbps of non-blocking performance and up to 192 interfaces when used in breakout mode.

A total of 24 gearboxes allow the system to support 96 ports of 100G and a diverse range of optics. Four logical interfaces are assigned to each pair of QSFP ports, which allows for combinations including:

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

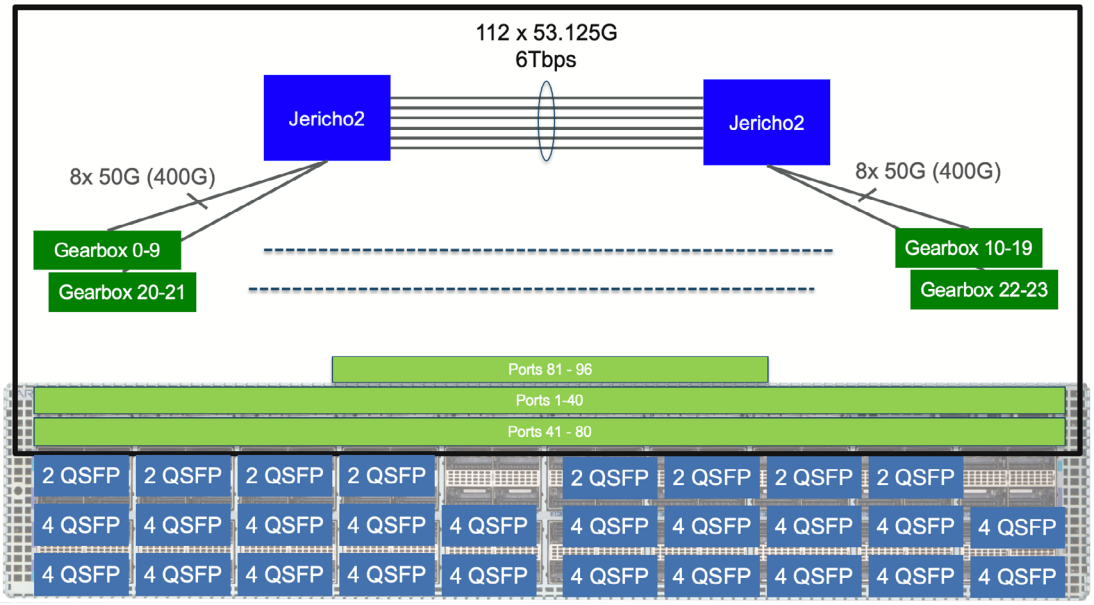


Figure 18: 7280CR3-96

7280CR3A-72

The 7280CR3A-72 is a 2U system with one Jericho2C+ chip, delivering 7.2 Tbps of non-blocking performance and up to 116 interfaces when used in breakout mode. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

A total of 18 gearboxes allow the system to support 72 ports of 100G QSFP and a diverse range of optics. Four logical interfaces are assigned to each odd/even pair of QSFP ports, which allows for breakout combinations up to 116 total interfaces including:

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)

- Running the primary port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Secondary port disabled
- Running the primary as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Secondary port disabled

CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

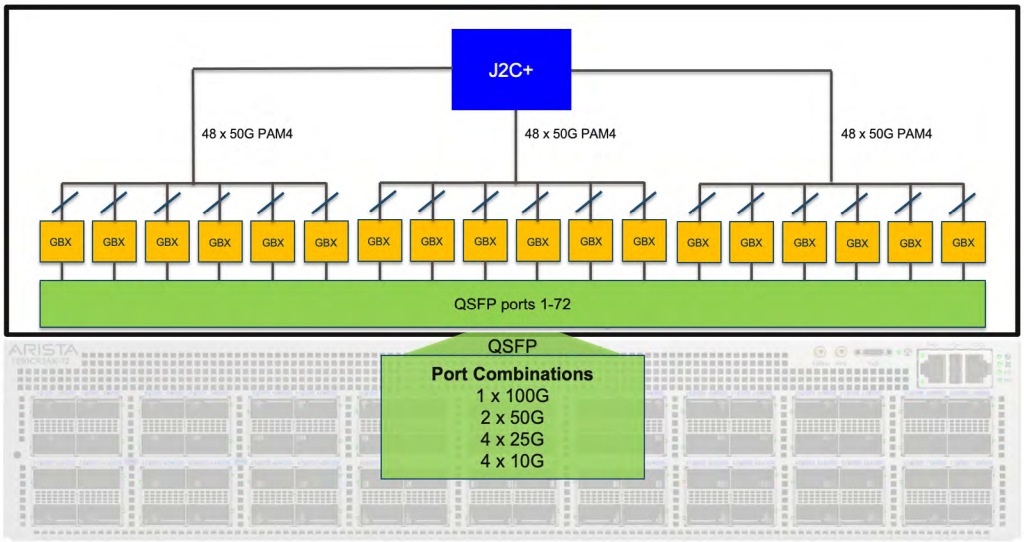


Figure 19a: 7280CR3A-72 architecture

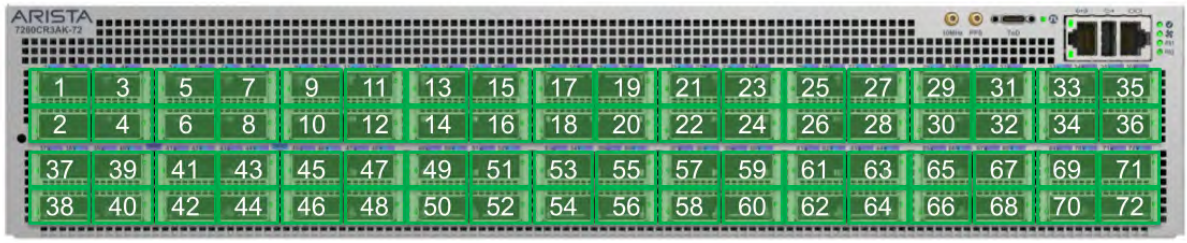


Figure 19b: 7280CR3A-72 port numbering

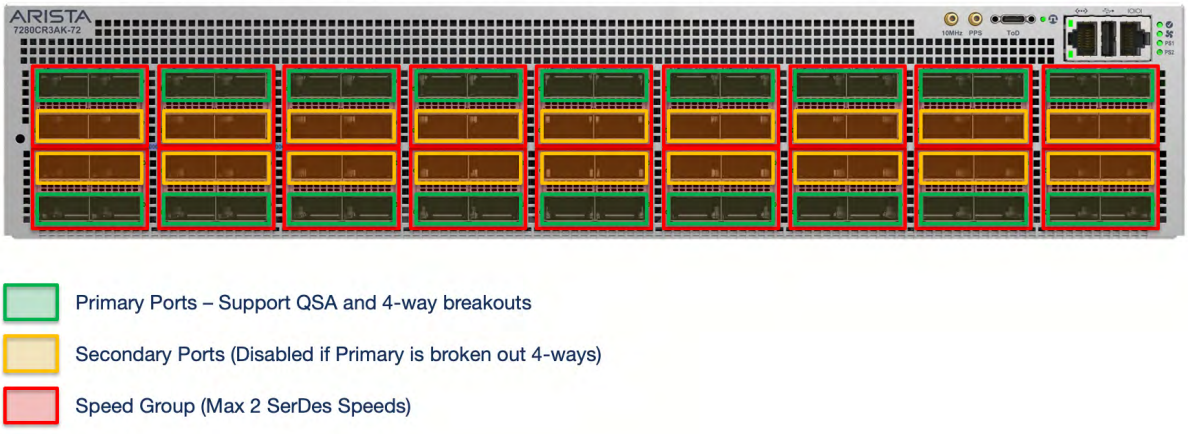
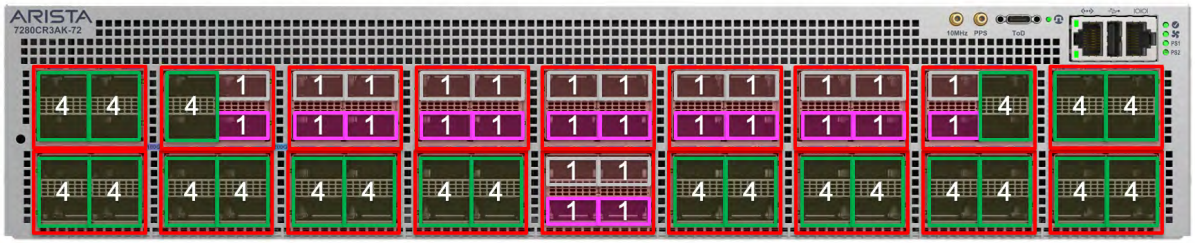


Figure 19c: 7280CR3A-72 speed and port groups



4 Logical Ports across each Primary/Secondary pair

1 Logical Port – Eth x/1

1 Logical Port – Eth x/1 – Supports 1 x 100G-4 or 40G-4 only

Speed Group (Max 2 SerDes Speeds)

Port Type	Practical / Max Density	Comment
10/25G	88/102*	Using all ports except secondary 1-way QSFP ports
50G-2	88/102*	Using all ports except secondary 1-way QSFP ports
100G-4	72	Using all ports
100G-2	44/58*	Using all primary ports
200G-4	0	Not Supported
400G-8	0	Not Supported

Figure 19d: 7280CR3A-72 default breakout capabilities (indicates not all signaling lanes are used)*

7280CR3A-32S

The 7280CR3A-32S is a 1U system with one Jericho2M chip, delivering 3.6 Tbps of non-blocking performance and up to 56 interfaces when used in breakout mode. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

The system uses 9 gearboxes to provide 32 ports of mixed 100G, 200G and 400G QSFP with up to 56 logical interfaces available through the use of breakouts.

The first 28 ports of QSFP are organized into Primary/Secondary pairs (odd ports are primary). The default port map provides the following capabilities per pair:

- The primary (odd numbered) port running as 1 x 200G-4 with a QSFP56 transceiver (secondary port disabled)
- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)

In addition, the first 6 pairs of ports (ports 1-12) can provide breakout to 10G or 25G subject to as follows:

- The primary (odd numbered) port configured as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- The odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

The final 4 ports (29-32) allow for both 200G and 400G configurations as follows:

- The primary (odd numbered) port running as 1 x 400G-8 with a QSFP-DD transceiver (secondary port disabled)
- Both ports running in 200G mode: 2 x 200G-4 (2 logical interfaces)
- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)

In addition, each pair of ports provides 8 logical interfaces, allowing both ports to be broken out 4-ways or the primary port to be broken out 8-ways subject to transceiver capabilities.

CLI tools provide further platform specific details on the combinations of interface speeds available across the system that are configurable. If available in EOS, alternate L1 Profiles may be applied to access different interface and logical port mappings.

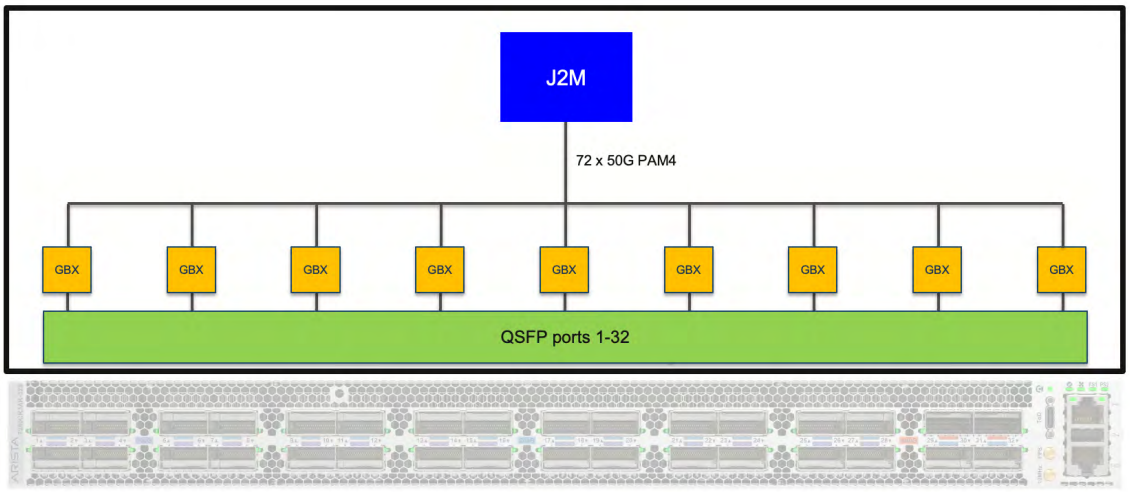


Figure 20a: 7280CR3A-32S architecture

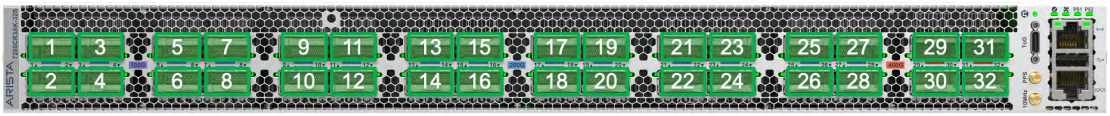
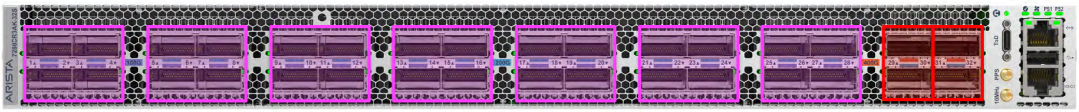


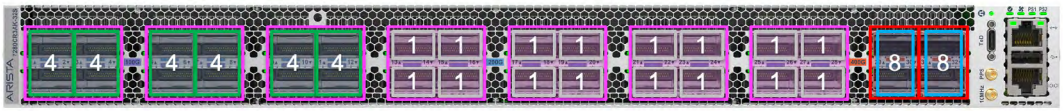
Figure 20b: 7280CR3A-32S port numbering



Speed Group (Max 2 SerDes Speeds)

Speed Group (Max 2 SerDes Speeds, each pair [Primary (odd) /Secondary (even)] at Same Speed)

Figure 20c: 7280CR3A-32S speed and port groups



1 Logical Port – Eth x/1

4 Logical Ports across each Primary/Secondary pair

8 Logical Ports – Eth x/1,2,3,4,5,6,7,8

Speed Group (Max 2 SerDes Speeds)

Speed Group (Max 2 SerDes Speeds, each pair [Primary/Secondary] at Same Speed)

Port Type	Max Density	Comment
10/25G	48*	Using odd numbered (primary) ports with maximum breakout
50G-2	48*	Using all ports
50G-1	48*	Using all ports
100G-4	32	Using QSFP28
100G-2	28	Using Primary QSFP with QSFP56 breakout and QDD breakout
Any 100G	36	28 x 100G-4 using QSFP28 & 2 x (4 x 100G-2) on QDD
200G-4	18	Using Primary QSFP56 and QSFP56 or QDD breakout
400G-8	2	Only supported on Q-DD ports

Figure 20d: 7280CR3A-32S default breakout capabilities (indicates not all signaling lanes are used)*

7280CR3-36S

The 7280CR3-36S is a 1U system with a single Jericho2C chip delivering 2.4 Tbps of non-blocking performance and up to 120 logical interfaces when used in breakout mode.

By deploying all available 25G and 50G SerDes lanes to the front panel, in combination with gearboxes, the system supports flexible interface configurations that support speeds from 1G to 400G.

As a 2.4 Tbps system the 7280CR3-36S provides interface configuration options that allow provisioning of up to 4 Tbps of front panel connectivity, for oversubscribed applications. The following section expands on some of the port configuration options.

Each of the first 24 QSFP ports support the following configurations:

QSFP Ports 1-24			
Interface Type	**Number per Port**	**Description**	**Restriction**
100G-4	1	QSFP100	None
50G-2	2	Dual 50G interfaces, each using 2 lanes of a single QSFP100	None
40G	1	QSFP+	None
25G	4	4-way breakout of QSFP100	None
10G	4	4-way breakout of QSFP+	None
1G	1	1G SFP - requires QSFP-SFP adapter	None

Each port in the group 25 to 32 supports the following configurations:

QSFP Ports 25-32			
Interface Type	**Number per Port**	**Description**	**Restriction**
200G-4	1	QSFP200 (Odd numbered port only)	Even Port Disabled
100G-2	2	Dual 100G-2 interfaces, each using 2 lanes of a single QSFP200 (Odd numbered port only)	Even Port Disabled
100G-4	1	QSFP100	None
50G-1	4	4-way breakout of QSFP200 (Odd numbered port only)	Even Port Disabled
50G-2	2	Dual 50G interfaces, each using 2 lanes of a single QSFP100	None
40G	1	QSFP+	None
25G	4	4-way breakout of QSFP100 (Odd numbered port only)	Even Port Disabled
10G	4	4-way breakout of QSFP+ (Odd numbered port only)	Even Port Disabled

Each port in the group 33 to 36 supports the following configurations:

QSFP Ports 33-36			
Interface Type	**Number per Port**	**Description**	**Restriction**
400G-8	1	QSFP-DD (Odd numbered port only)	Even Port Disabled
200G-4	2	2-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
200G-4	1	QSFP200	None
100G-2	4	4-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
100G-2	2	Dual 100G-2 interfaces, each using 2 lanes of a single QSFP200	None
100G-4	2	2-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
100G-4	1	QSFP100	None
50G-1	4	4-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
50G-2	4	4-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
50G-2	2	Dual 50G interfaces, each using 2 lanes of a single QSFP100	None
40G	2	2-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
40G	1	QSFP+	None
25G	4	4-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled
10G	4	4-way breakout of QSFP-DD (Odd numbered port only)	Even Port Disabled

The complete set of breakout options are subject to a system maximum of 120 logical interfaces and the capability of the cable or transceiver installed. Each port is capable of supporting copper cables, Active Optical Cables (AOC) and the full range of optics available in the 100/200/400G QSFP form factors. Where applicable, QSFP-SFP adapters may be used to support individual lower speed transceivers including 1, 10 and 25G parts.

CLI tools provide further platform specific details on the combinations of interface speeds available across the system that are configurable.

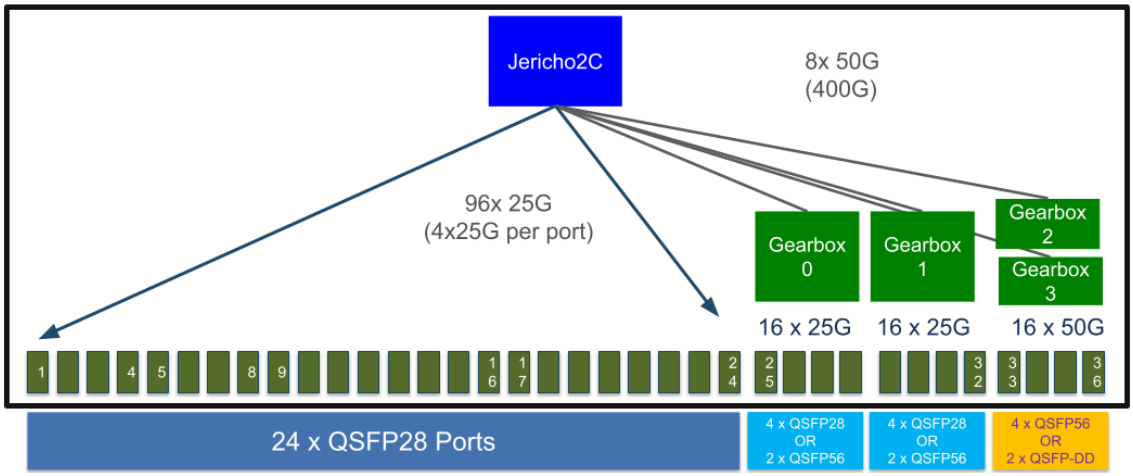


Figure 21: 7280CR3-36S

7280SR3A-48YC8

The 7280SR3A-48YC8 is a 1U system with one Jericho2M chip, delivering 3.6 Tbps of non-blocking performance and up to 56 logical interfaces. The system is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models which share the same physical interface properties.

The system provides 48 ports of 10G, 25G, 50G SFP and four ports of QSFP56 directly connected to the J2M chip. A single gearbox drives the remaining four QSFP ports enabling a choice of either two more QSFP56 or four QSFP28. This flexibility provides for up to 1.2 Tbps of uplink capacity over 6 x 200G QSFP56 or up to 8 independent uplink ports (four QSFP56 and four QSFP28). Each physical port is mapped to one of the 56 logical interfaces.

Port speed support is as follows:

- Each SFP port supports SFP+, SFP28 and SFP56 transceivers. 1G speeds may be supported using Rate Adapting transceivers.
- QSFP56 ports 49-52 support QSFP+, QSFP28 and QSFP56 transceivers.
- QSFP ports 53 -56 support:
  - » 40G, 100G or 200G in ports 53 and 55 (54 and 56 disabled)
  - » 40G or 100G in ports 53-56

CLI tools provide further platform specific details on the combinations of interface speeds available across the system that are configurable. If available in EOS, alternate L1 Profiles may be applied to access different interface and logical port mappings.

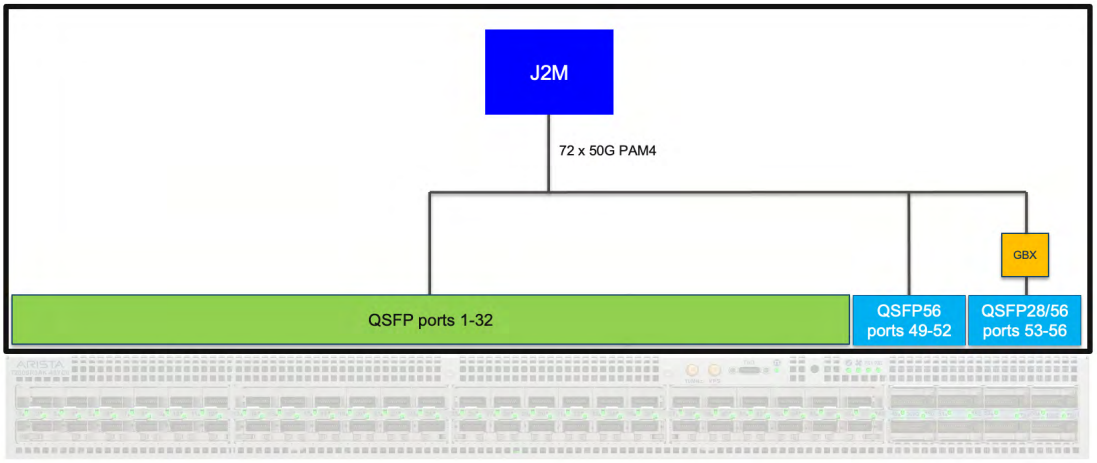


Figure 22a: 7280SR3A-48YC8 architecture

Figure 22b: 7280SR3A-48YC8 port numbering

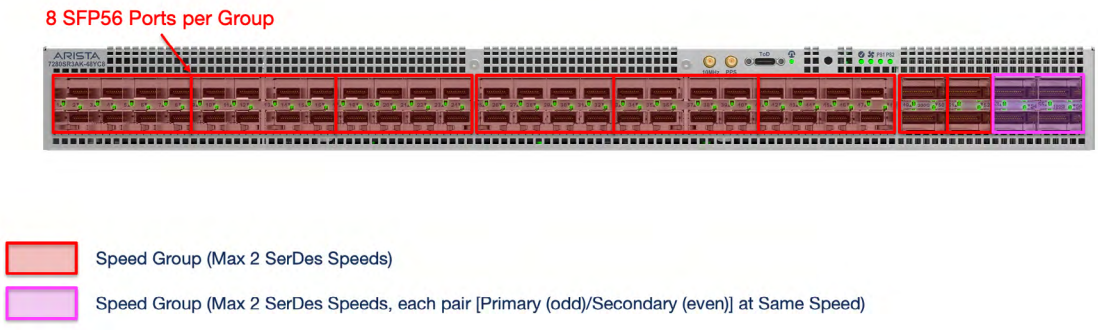
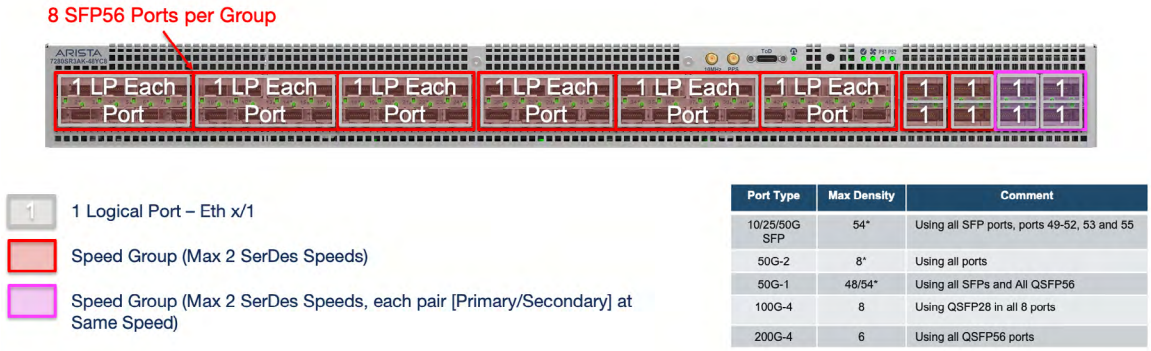


Figure 22c: 7280SR3A-48YC8 speed and port groups



Port Type	Max Density	Comment
10/25/50G SFP	54*	Using all SFP ports, ports 49-52, 53 and 55
50G-2	8*	Using all ports
50G-1	48/54*	Using all SFPs and All QSFP56
100G-4	8	Using QSFP28 in all 8 ports
200G-4	6	Using all QSFP56 ports

Figure 22d: 7280SR3A-48YC8 default breakout capabilities (indicates not all signaling lanes are used)*

7280SR3-48YC8

The 7280SR3-48YC8 and 7280SR3K-48YC8 are 1U systems with a single Jericho2C chip providing 2 Tbps of capacity across a mixture of 48 SFP and 8 QSFP ports.

Each of the 48 SFP28 port is capable of supporting 1/10/25G, while the QSFP28 ports support 100G or 40G and may also be broken out to support 4 x 10/25G each for a total of 80 ports in the system

Each port is capable of supporting copper or AOC as well as the range of optics available in the SFP and QSFP form factors.

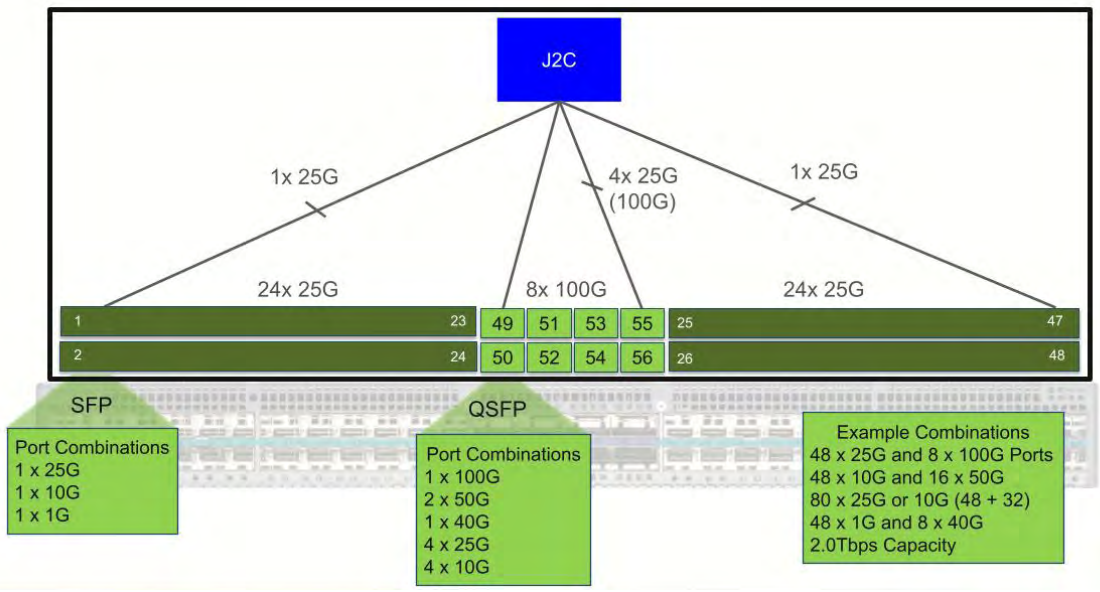


Figure 23: 7280SR3-48YC8

7280SR3M-48YC8

The 7280SR3M-48YC8 and 7280SR3MK-48YC8 are 1U systems with a single Jericho2C chip providing 2 Tbps of capacity across a mixture of 48 SFP and 8 QSFP ports. Line-rate MACsec encryption is available on all interfaces, implemented by discrete encryption engines that each support groups of QSFP28 and SFP28 ports.

The 48 SFP28 ports are organized in groups of four ports. Each group is capable of supporting a common speed of 1/10/25G, while the QSFP28 ports support 100G or 40G and may also be broken out to support 4 x 10/25G each for a total of 80 ports in the system.

Each port is capable of supporting copper or AOC as well as the range of optics available in the SFP and QSFP form factors.

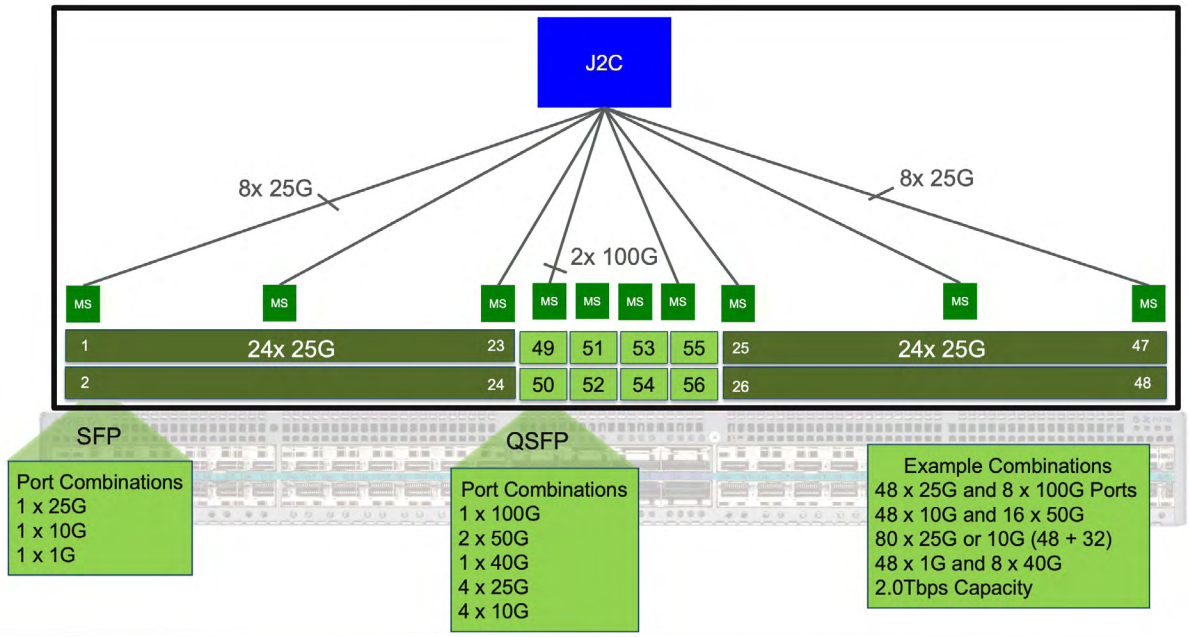


Figure 24a :7280SR3M-48YC8

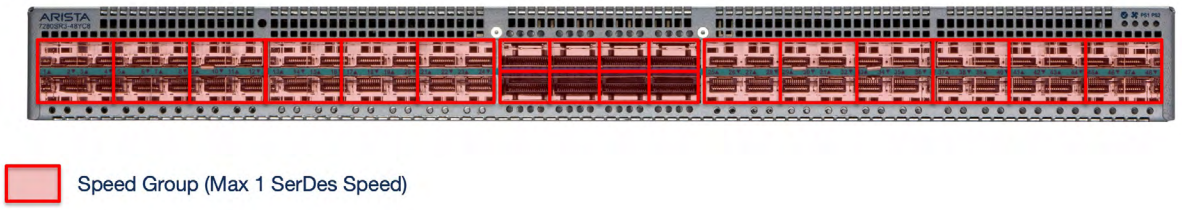


Figure 24b :7280SR3M-48YC8

7280SR3-40YC6

The 7280SR3-40YC6 and 7280SR3E-40YC6 are 1U systems with a single Qumran2a chip providing 1.6Tbps of front panel capacity across a mixture of 40 SFP and 6 QSFP ports. The 7280SR3E-40YC6 enhances the standard model with support for Synchronous Ethernet on all front panel interfaces.

The first 36 SFP28 ports are connected directly to the Qumran chip and are capable of supporting 1/10/25G. A gearbox sits behind ports 37-40, converting the chip's 50G lanes to 10G or 25G.

SFP Ports 37-40 are capable of 10G or 25G operation, while each odd/even pair of QSFP ports support the following configurations up to a system maximum of 52 logical interfaces.

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

Each port is capable of supporting copper or AOC as well as the range of optics available in the SFP and QSFP form factors. CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

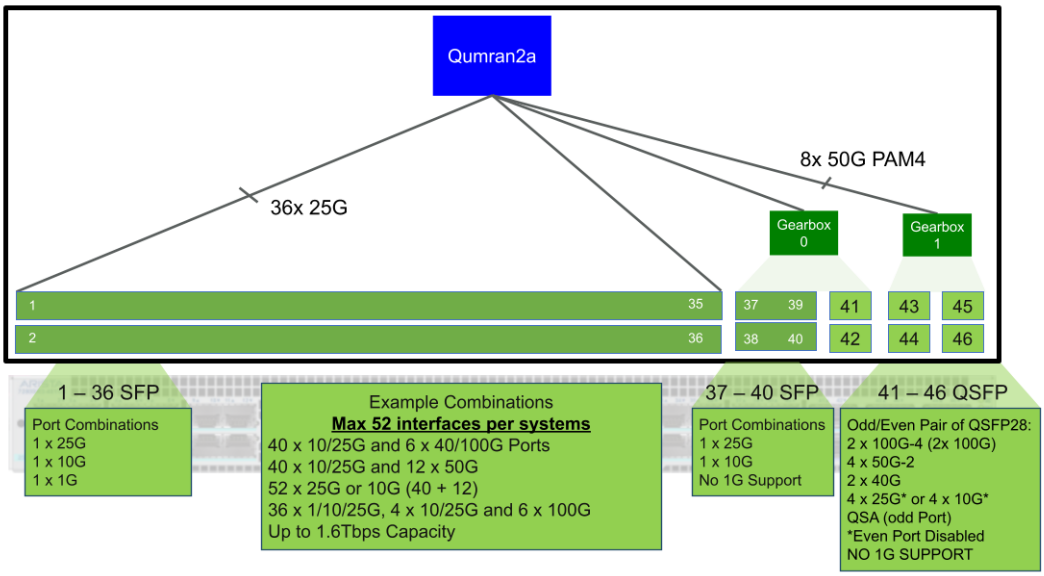


Figure 25: 7280SR3-40YC6

7280TR3-40C6

The 7280TR3-40C6 is a 1U system directly related to the 7280SR3-40YC6. A single Qumran2a chip provides 1 Tbps of front panel capacity across a mixture of 40 RJ45 and 6 QSFP ports.

The first 36 RJ45 ports are connected directly to the Qumran chip and are capable of supporting 1/10GBASE-T. A gearbox sits behind ports 37-40, converting the chip's 50G lanes to 10G (1G is not supported).The QSFP ports support the following configurations up to a system maximum of 52 logical interfaces.

- Both ports running in 100G mode: 2 x 100G-4 (2 logical interfaces)
- Both ports running in 40G mode: 2 x 40G (2 logical interfaces)
- Breaking both ports into 2x50G: 4 x 50G-2 (4 logical interfaces)
- Running the odd numbered port as a breakout to 25G: 4 x 25G (4 logical interfaces)
  - » Even numbered port disabled
- Running the odd numbered as a breakout to 10G: 4 x 10G (4 logical interfaces)
  - » Even numbered port disabled

Each transceiver port is capable of supporting copper or AOC as well as the range of optics available in QSFP form factor. CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

Arista 7280R3 Compact Modular System

The Arista 7280R3 compact modular system is a 4 RU device optimized for operations and maintenance while providing flexibility in the type and number of interfaces deployed. Nine interface module slots enable flexible configurations of up to 36 ports of 400G, 144 ports of 100G, or up to 144 ports of 25G. All other elements of the system are also field replaceable.
The 7289R3A switch card (7289R3A-SC) employs two Jericho2C+ chips delivering 3.2Tbps to each of the nine line card slots for a total of 14.4 Tbps and 5.4 Bpps. In common with other Jericho2C+ devices, the switch card is available in a choice of three levels of scale and functionality with Standard (R3A), Encryption (R3AM) and Large Scale + Encryption (R3AK) models available.

Four line card configurations are available:

- 4 Ports of 400G OSFP
  - » Common equipment with DCS-7368X4 and DCS-7358X4
- 4 Ports of 400G QSFP-DD
  - » Common equipment with DCS-7368X4 and DCS-7358X4
- 16 Ports of 100G QSFP
  - » Common equipment with DCS-7358X4
- 16 Ports of 25G SFP
  - » Common equipment with DCS-7368X4 and DCS-7358X4

There are no restrictions in which slots each line card can occupy or how many of each line card type can be installed.

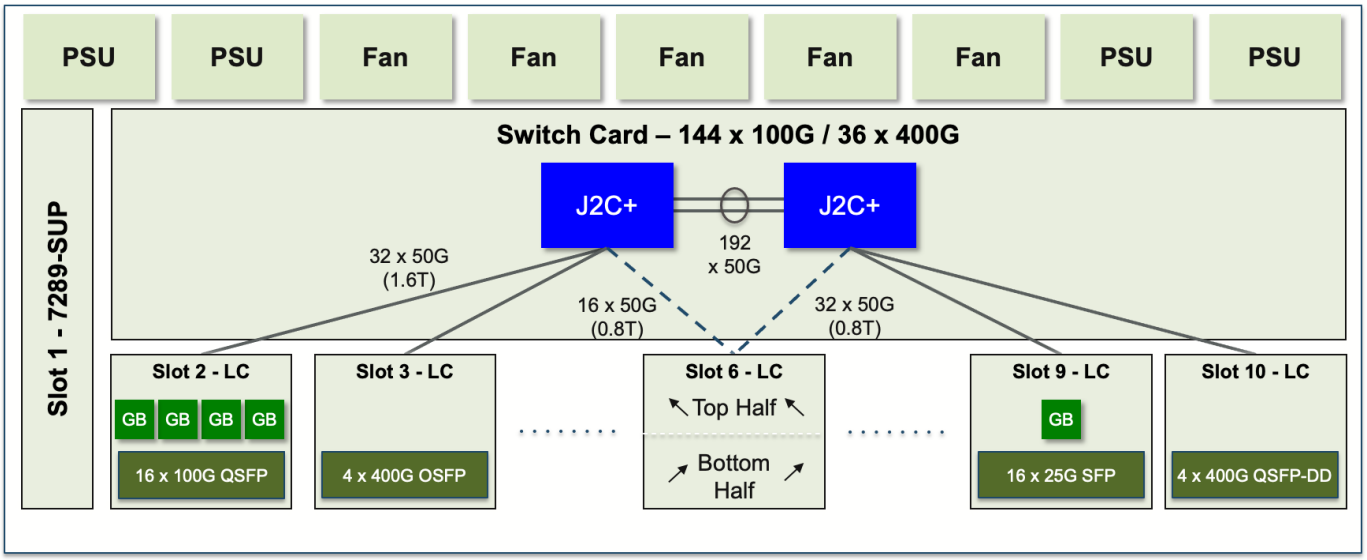


Figure 26: 7280R3 Compact Modular system

With the default configuration, the following breakout modes are supported on each line card:

4 Port 400G Line Cards (OSFP or QSFP-DD)

Each port supports the following:

- When used with a 400G OSFP (or QSFP-DD):
  - » 1 x 400G-8
  - » 2 x 200G-4
  - » 4 x 100G-2
  - » 2 x 100G-4
  - » 4 x 50G-2
- When used with a 200G QSFP:
  - » 1 x 200G-4
  - » 2 x 100G-2
  - » 4 x 50G-1
  - » All 40/100G QSFP speeds subject to transceiver support

- When used with a 100G QSFP:
  - 1 x 100G-4
  - 2 x 50G-2
  - 4 x 25G
  - All 40 QSFP speeds subject to transceiver support
- When used with a 40G QSFP:
  - 1 x 40G
  - 4 x 10G

16 Port 100G Line Cards

QSFP ports are organized in rows, with the odd numbered port on the left and the even numbered port on the right. Even numbered ports are primary ports.

The line card supports the following configurations:

- When used with 200G QSFP:
  - 4 x 200G-4 installed in ports 2, 6, 10 and 14
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)
  - Up to 8 x 100G-2 in ports 2, 6, 10 and 14
    - Ports 2, 6, 10 and 14 configured for 2 x 100G-2
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)
  - Up to 16 x 50G-1 in ports 2, 6, 10 and 14
    - Ports 2, 6, 10 and 14 configured for 4 x 50G-1
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)
  - Up to 20 x 10/25G
    - QSFP+/28 in ports 2, 6, 10 and 14 configured for 4 x 10/25G
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 4, 8, 12 and 16 available for 1 x 10/25G SFP using QSFP-SFP adapters (QSA)
    - Odd numbered ports 3, 7, 11 and 15 disabled
- When used with 100G QSFP:
  - Up to 16 x 100G-4 installed
  - Up to 16 x 50G-2 and 8 x 100G-4
    - Ports 1, 2,5, 6, 9, 10 and 13, 14 configured for 2 x 50G-2
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)
  - Up to 16 x 25G and 8 x 100G-4
    - Ports 2, 6, 10 and 14 configured for 4 x 25G
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)

- When used with 40G QSFP+:
  - » Up to 16 x 40G installed
  - » Up to 16 x 10G and 8 x 40G
    - Ports 2, 6, 10 and 14 configured for 4 x 10G
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 3, 4, 7, 8, 11, 12, 15 and 16 available for 100G QSFP (100G-4) or 40G QSFP+ (40G)
- When used with QSFP-SFP adapters (QSA):
  - » Up to 8 total QSAs in ports 2, 4, 6, 8, 10, 12 and 14
    - For each QSA installed in an even port, the adjacent odd port is disabled.
  - » Ports without QSA operate as defined above.16 Port 25G Line Cards
- For maximum breakout to 10/25G:
  - » Up to 20 x 10/25G
    - QSFP+/28 in ports 2, 6, 10 and 14 configured for 4 x 10/25G
    - Odd numbered ports 1, 5, 9 and 13 disabled
    - Ports 4, 8, 12 and 16 available for 1 x 10/25G SFP using QSFP-SFP adapters (QSA)
    - Odd numbered ports 3, 7, 11 and 15 disabled

16 Port 25G Line Cards

- All ports support 10G or 25G SFP

Each port is capable of supporting copper or AOC as well as the range of optics available in the SFP and QSFP form factors. CLI tools provide further platform specific details on the combinations of interface speeds available across the system.

Scaling the Control Plane

A central CPU complex on the 7280R3 Series switches is used exclusively for control-plane and management functions; all data-plane forwarding logic occurs at the packet processor level. Arista EOS®, the control-plane software for all Arista switches, executes on multi-core x86 CPUs with multiple gigabytes of DRAM. Each system is equipped with resources appropriate for its target use cases, with larger CPUs and more memory provisioned for platforms focused on high end routing deployments. As EOS is multi-threaded, runs on a Linux kernel and is extensible, the large RAM and fast multi-core CPUs provide for operating an efficient control plane with headroom for running 3rd party software, either within the same Linux instance as EOS or within a guest virtual machine.

Out-of-band management is available via a serial console port and/or the 10/100/1000 Ethernet management interface. The 7280R3 Series also offers USB2.0 interfaces that can be used for a variety of functions including the transferring of images or logs. **Arista**

7280R3 Universal Leaf Platform Packet Forwarding Pipeline

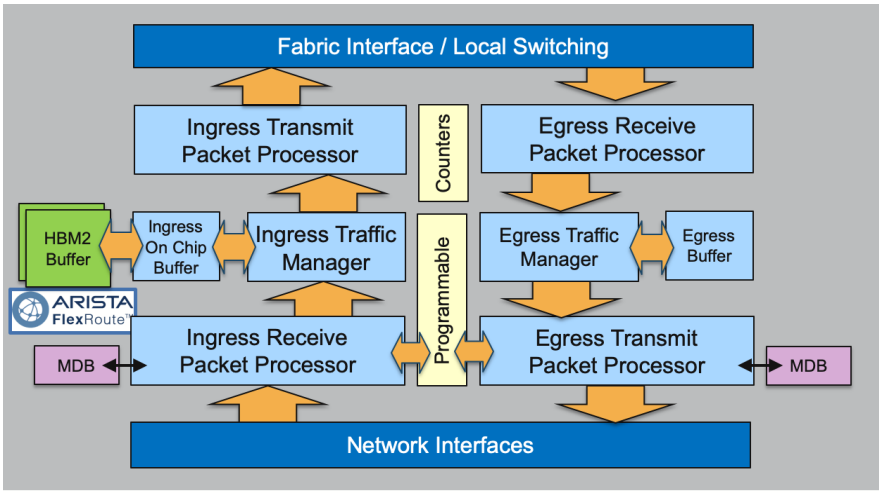


Figure 27: Packet forwarding pipeline stages inside a packet processor on an Arista 7280R3 Series

Each packet processor is a System-on-Chip (SoC) that provides all the ingress and egress forwarding pipeline stages for packets to or from the front panel input ports connected to that packet processor. Forwarding is always hardware-based and never falls back to software/CPU forwarding.

The steps involved at each of the logical stages of the packet forwarding pipeline are outlined below.

Stage 1: Networking Interface (Ingress)

When packets/frames enter the switch, the first block they arrive at is the Network Interface stage. This is responsible for implementing the Physical Layer (PHY) interface and Ethernet Media Access Control (MAC) layer on the switch and any Forward Error Correction (FEC).

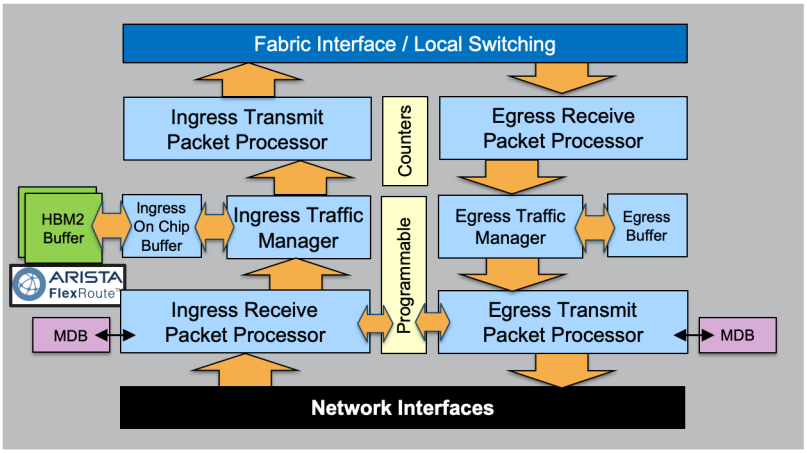


Figure 28: Packet Processor stage 1 (ingress): Network Interface

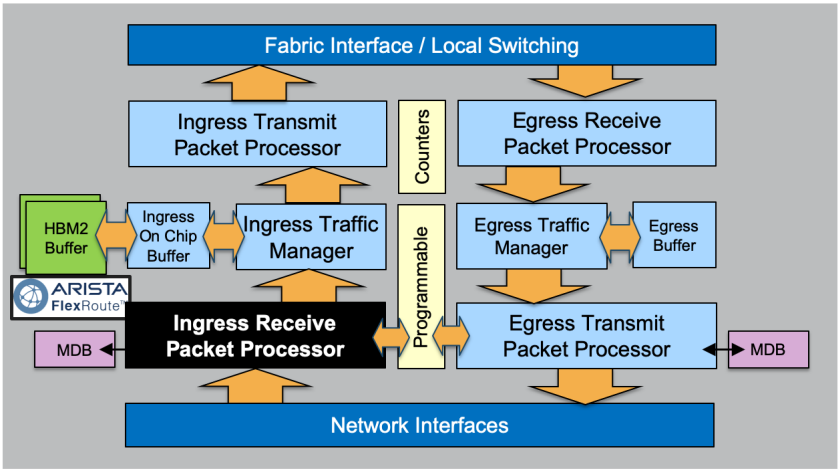
The PHY layer is responsible for the transmission and reception of bitstreams across physical connections including encoding, multiplexing, synchronization, clock recovery and serialization of the data on the wire for whatever speed/type Ethernet interface is configured.

Programmable lane mappings are used to map the physical lanes to logical ports based on the interface type and configuration. Lane mappings are used for breakout of 4x25G and 2x50G on 100G ports.

If a valid bitstream is received at the PHY then the data is sent to the MAC layer. On input, the MAC layer is responsible for turning the bitstream into frames/packets: checking for errors (FCS, Inter-frame gap, detect frame preamble) and finding the start of frame and end of frame delimiters.

Stage 2: Ingress Receive Packet Processor

The Ingress Receive Packet Processor stage is responsible for forwarding decisions. It is the stage where all forwarding lookups are performed.



- Packet Parsing
- SMAC/DMAC/ DIP lookups
- Forwarding table lookups
- Tunnel Decap
- Ingress ACL
- Resolution of forwarding action

Figure 29: Packet Processor stage 2 (ingress): Ingress Receive Packet Processor

Before any forwarding can take place, packet or frame headers must be parsed and fields for forwarding decisions extracted. Key fields include L2 Source and Destination MAC addresses [SMAC, DMAC], VLAN headers, Source and Destination IP Addresses [SIP, DIP], class of service (COS), DSCP and so on. The Arista 7280R3 packet parser supports many tunnel formats (MPLS, IPinIP, GRE, VXLAN, MPLSoGRE) as well as parsing Ethernet and IP headers under a multi-label stack. The parser is flexible and extensible such that it can support future protocols and new forwarding models. As IPv6 deployments grow and traffic shifts to being native IPv6 it is anticipated that more IPv4 traffic will be tunneled over the aforementioned tunnel formats using IPv6. Jericho2 provides the additional parsing flexibility necessary to enable this next generation of tunnel formats.

After parsing the relevant encapsulation fields, the DMAC is evaluated to see if it matches the device's MAC address for the physical or logical interface. If it's a tunneled packet and is destined to a tunnel endpoint on the device, it is decapsulated within its appropriate virtual routing instance and packet processing continues on the inner packet/frame headers. If it's a candidate for L3 processing (DMAC matches the device's relevant physical or logical MAC address) then the forwarding pipeline continues down the layer 3 (routing) pipeline, otherwise forwarding continues on the layer 2 (bridging) pipeline.

In the layer 2 (bridging) case, the packet processor performs SMAC and DMAC lookup in the MAC table for the VLAN. SMAC lookup is used to learn (and can trigger a hardware MAC-learn or MAC-move update), DMAC (if present) is used for L2 forwarding and if not present will result in the frame being flooded to all ports within the VLAN, subject to storm-control thresholds for the port.

In the layer 3 (routing) case, the packet processor performs a lookup on the Destination IP address (DIP) within the VRF and if there is a match it knows what port to send the frame to and what packet processor it needs to send the frame to. If the DIP matches a subnet local to the switch for which there is no host route entry, the switch will initiate an ARP request to learn the MAC address for where to send the packet. If there is no matching entry at all the packet is dropped. IP TTL decrement also occurs as part of this stage. Additionally, VXLAN Routing can be performed within a single pass through this stage.

For unicast traffic, the end result from a forwarding lookup match is a pointer to a Forwarding Equivalence Class (FEC) or FEC group (Link Aggregation, Equal Cost Multipathing [ECMP] or Unequal Cost Multipathing [UCMP]). In the case of a FEC group, the fields which are configured for load balancing calculations are used to derive a single matching entry. The final matching adjacency entry provides details on where to send the packet (egress packet processor, output interface and a pointer to the output encapsulation/ MAC rewrite on the egress packet processor).

For multicast traffic, the logic is similar except that the adjacency entry provides a Multicast ID, which indicates a replication requirement for both local (ingress) multicast destinations on local ports, as well as whether there are packet processors in the system that require packet replication via multicast replication in the fabric modules. By default, the Arista 7280R3 Series operates in egress multicast replication but can be configured for ingress multicast replication as well.

The forwarding pipeline always remains in the hardware data-plane. There are no features that can be enabled that cause the packet forwarding to drop out of the hardware-based forwarding path. In cases where software assistance is required (e.g. traffic destined within a L3 subnet but for which the switch has not yet seen the end device provide an ARP and doesn't have the L3-to-L2 glue entry), hardware rate limiters and Control Plane Policing are employed to protect the control-plane from potential denial of service attacks.

In parallel with forwarding table lookups, there are also Ingress ACL lookups (Port ACLs, Routed ACLs) for applying security and QoS lookups to apply Quality of Service. All lookups are ultimately resolved using strength based resolution (some actions are complementary and multiple actions are applied, some actions override others) but ultimately the outcome of this stage is a resolved forwarding action.

Counters are available within this stage providing accounting and statistics on ACLs, VLAN and sub-interfaces, as well as a range of tunnel and next-hop group types. The R3-series systems provide significant gains in overall counter scale and flexibility in allocation over previous generations, providing a 5X increase in scale in some dimensions. The criticality of flexibility in counter scaling cannot be overstated as operators migrate to next generation technologies such as Segment Routing and the use of various overlay tunnel technologies that rely upon fine-grained network utilization information to accurately place network workloads.

Data plane counters are available in real-time via streaming telemetry using NetDB to export using gRPC with OpenConfig.

Arista FlexRoute™ Engine

One of the key characteristics of the Arista 7280R3 Universal Leaf platform is the FlexRoute Engine, an Arista innovation which enables Internet-scale L3 routing tables with significant power consumption savings over legacy IP routing longest prefix match lookups. This in turn enables higher port densities and performance with power and cooling advantages when compared to legacy service provider routing platforms.

Arista's FlexRoute Engine is used for both IPv4 and IPv6 Longest Prefix Match (LPM) lookups without partitioning table resources. It is optimized around the Internet routing table, its prefix distribution and projected growth. FlexRoute enables scale to millions of prefixes, providing headroom for internet table growth for many years.

In addition to large table support, FlexRoute enables very fast route programming and reprogramming (tens of thousands of prefixes per second), and does so in a manner that is non-disruptive to other prefixes while forwarding table updates are taking place.

All Arista 7280R3 Series systems take advantage of the multi-stage programmable forwarding pipeline to provide a flexible and scalable solution for access control, secure policy based networking and telemetry in today's cloud networks. ACLs are not constrained by the size of fixed hardware tables, but can leverage the forwarding lookup capabilities of the packet processor to trigger a wide range of traffic management actions.

sFlow

The programmable packet processing pipeline on the 7280R3 platform enables a range of new telemetry capabilities for network operators. In addition to new counter capabilities, flow instrumentation capabilities are enhanced through the availability of hardware accelerated sFlow. As network operators deploy various tunnel overlay technologies in their network, sFlow provides an encapsulation independent means of getting visibility into high-volume traffic flows and enables operators to more effectively manage and steer traffic to maximize utilization. The programmable pipeline provides these capabilities inline without requiring an additional coprocessor. Sampling granularity of 1:100 on 100G and 400G interfaces can be realized on all interfaces.

Inband Network Telemetry (INT)

As a complement to sFlow, INT provides operators with a standards-based means of getting insight into per hop latency, paths, congestion and drops. This information can be correlated to allow an analysis of hotspots, path topology to influence traffic engineering decisions. INT provides operators with a data plane aware complement to standard IP/MPLS troubleshooting tools. Where ping and traceroute cannot necessarily confirm whether or not a flow traverses a specific interface in a port-channel, INT provides operators with a path and node traversal details by processing inband OAM frames and annotating these frames with metadata to provide detailed path and transit quality details. The programmable pipeline in the R3-series systems provides the ability to facilitate this packet processing inline.

Stage 3: Ingress Traffic Manager

The Ingress Traffic Manager stage is responsible for packet queuing and scheduling.

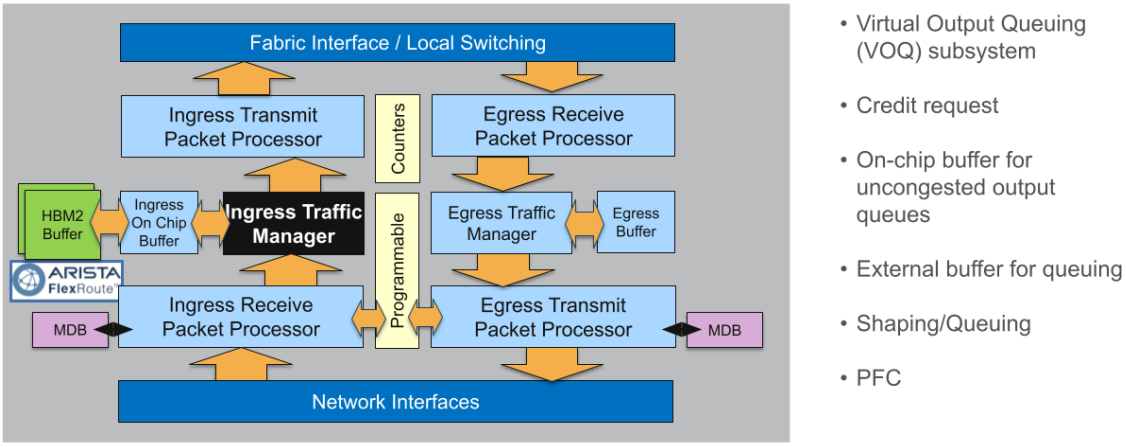


Figure 30: Packet Processor stage 3 (ingress): Ingress Traffic Manager

Arista 7280R3 Universal Leaf platforms utilize Virtual Output Queuing (VOQ) where the majority of the buffering within the switch is on the ingress port. While the physical buffer is on the input packet processor, it represents packets queued on the output side (hence, the term virtual output queuing). VOQ is a technique that allows buffers to be balanced across sources contending for a congested output port and ensures fairness and QoS policies can be implemented in a distributed forwarding system.

When a packet arrives into the Ingress Traffic Manager, a VOQ credit request is forwarded to the egress port processor requesting a transmission slot on the output port. Packets are queued on ingress until such time as a VOQ grant message is returned (from the Egress Traffic Manager on the output port) indicating that the Ingress Traffic Manager can forward the frame to the egress packet processor.

While the VOQ request/grant credit cycle is underway, the packet is queued in input buffers. A combination of on-chip memory (up to 64 MB) and external memory (8 GB) per packet processor is used to store packets while awaiting the VOQ grant. The memory is used such that traffic destined to uncongested outputs (egress VOQ is empty) will go into on-chip memory (head of the queue) otherwise external buffer memory is utilized. The external buffer memory is used because it's impractical to build sufficiently large buffers on-chip due to the very large chip die area that would be consumed.

While there is up to 24 GB buffer memory per system, the majority of the buffer is allocated in a dynamic manner wherever it is required across potentially millions of VOQs per system:

- ~30% buffer reserved for traffic per Traffic Class per Output Port.
- ~15% buffer for multi-destination traffic.
- ~55% available as a dynamic buffer pool.

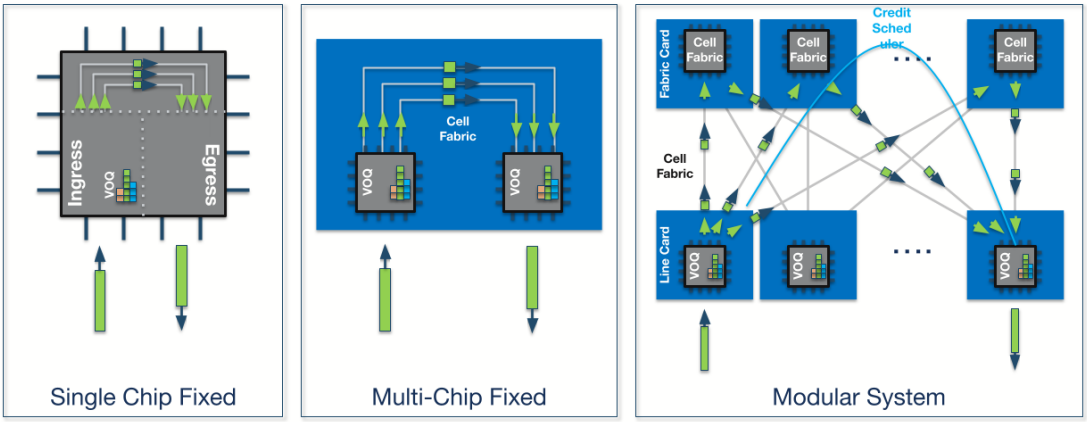


Figure 31: Physical Buffer on Ingress allocated as Virtual Output Queues

The dynamic pool enables the majority of the buffer to be used in an intelligent manner based on real-time contention and congestion on output ports. While there are potentially hundreds of gigabytes of buffer memory, individual VOQ limits are applied such that a single VOQ doesn't result in excess latency or queuing on a given output port. The default allocations (configurable) are as per Table 3:

Table 7: Default per-VOQ Output Port Limits

Output Port Characteristic	Maximum Packet Buffer Depth (MB)	Maximum Packet Buffer Depth (msec)
VOQ for a 10G output port	50 MB	40 msec
VOQ for a 25G output port	125 MB	40 msec
VOQ for a 40G output port	200 MB	40 msec
VOQ for a 50G output port	250 MB	40 msec
VOQ for a 100G output port	500 MB	40 msec
VOQ for a 400G output port	500 MB	10 msec

The VOQ subsystem enables buffers that are dynamic, intelligent and deep so that there is always packet buffer space available for new flows, even under congestion and heavy load scenarios. There is always complete fairness in the system, with QoS policy always enforced in a distributed forwarding system. This enables any application workload to be deployed – existing or future – and provides the basis for deployment in Content Delivery Networks (CDNs), service providers, internet edge, converged storage, hyper-converged systems, big data/analytics, enterprise and cloud providers. The VOQ subsystem enables maximum fairness and goodput for applications with any traffic profile, be it any-cast, in-cast, mice or elephant flows, or any flow size in between.

7280R3 Deep Packet Buffers

As with previous generations, the 7280R3 series systems utilize on-chip buffers (32MB with Jericho2, 64MB with Jericho2C+) in conjunction with flexible packet buffer memory (8GB of HBM2 per packet processor). The on-chip buffers are used for non-congested forwarding and seamlessly utilize the HBM2 packet buffers for instantaneous or sustained periods of congestion. Buffers are allocated per VOQ and require no tuning. It's further worth noting that during congestion, packets are transmitted directly from the HBM2 packet buffer to the destination packet processor.

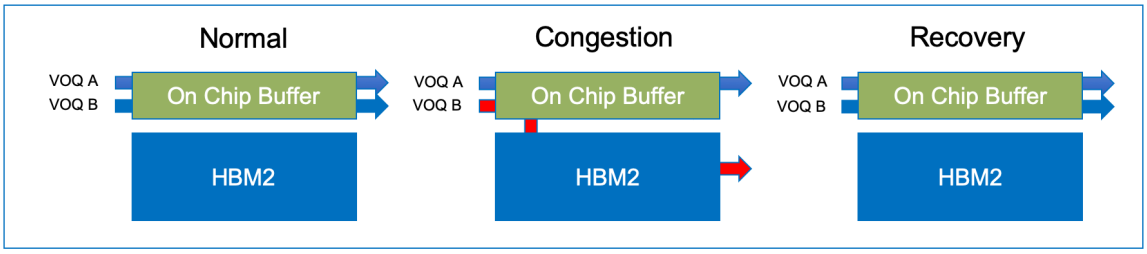


Figure 32: Packet buffer memory access

HBM2 memory is integrated directly into the Jericho2 packet processor this provides a reliable interface to the Jericho2 packet processor and eliminates the need for additional high-speed memory interconnects as does HMC or GDDR. This results in upwards of a 43% reduction in power utilization than the equivalent GDDR memory.

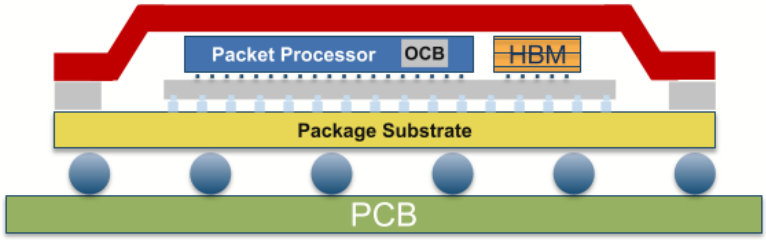


Figure 33: HBM memory packaging integration

Stage 4: Ingress Transmit Packet Processor

The Ingress Transmit Packet Processor stage is responsible for transferring frames from the input packet processor to the relevant output packet processor. Frames arrive at this stage once the output port has signaled, via a VOQ grant message, that it is the allocated slot for a given input packet processor to transmit the packet.

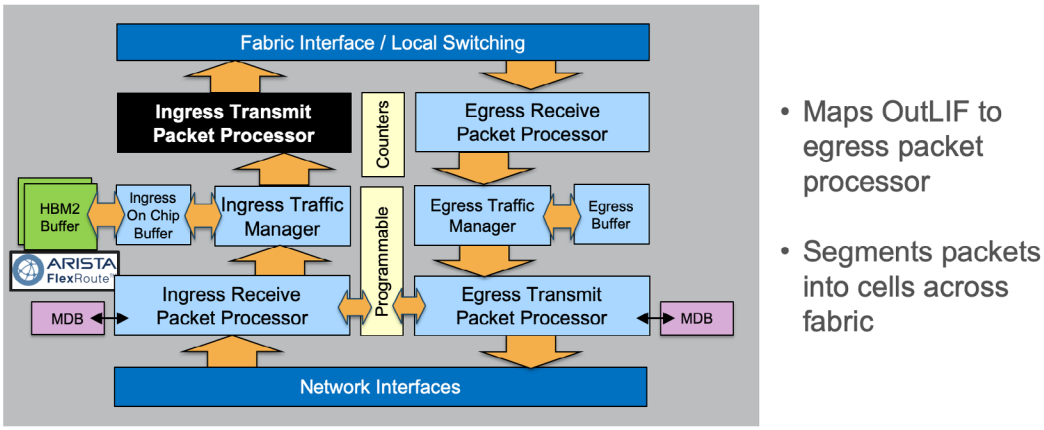


Figure 34: Packet Processor stage 4 (ingress): Ingress Transmit Packet Processor

All available paths are used in parallel to transfer the frame or packet to the output packet processor, with the original packet segmented into variable-sized cells which are forwarded across the fabric links simultaneously. This mechanism reduces serialization to at most 256 bytes at 50Gbps and ensures there are no hot spots as every flow is always evenly balanced across all paths. Since a packet is only transferred once there is a VOQ grant, there are guaranteed to be resources to process the frame on the egress packet processor.

Each cell has a header added to the front for the receiving packet processor to be able to reassemble and maintain in-order delivery. Forward Error Correction (FEC) is also enabled for traffic across the fabric modules, both to correct errors (if they occur) but also to help monitor data-plane components of the system for any problems.

Packets destined to ports on the same packet processor are switched locally and do not use fabric bandwidth resources, but otherwise aren't processed any differently in terms of the VOQ subsystem.

Stage 5: Egress Receive Packet Processor

The Egress Receive Packet Processor stage is responsible for reassembling cells back into packets/frames. This is also the stage that takes a multicast packet/frame and replicates it when there are multiple locally attached receivers on this output packet processor.

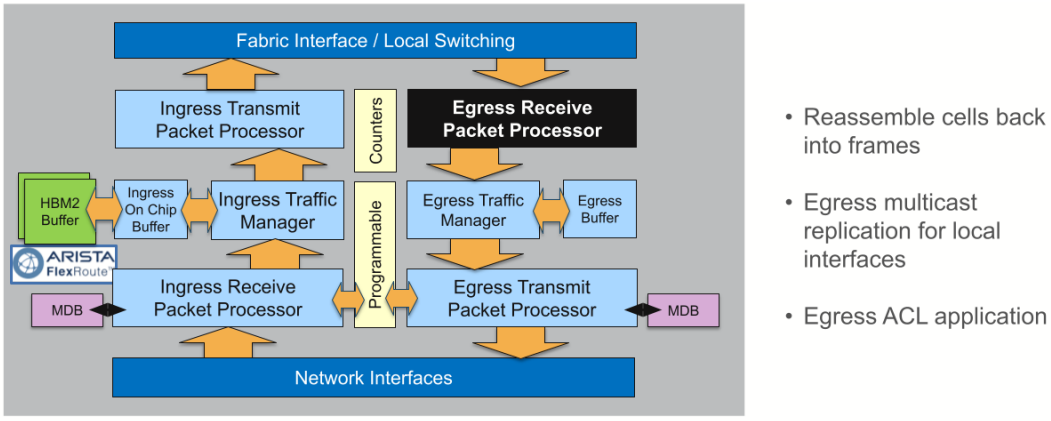


Figure 35: Packet Processor stage 5 (egress): Egress Receive Packet Processor

This stage ensures that there is no frame or packet reordering in the system. It also provides the data-plane health tracer, validating reachability messages from all other packet processors across all paths in the system.

Egress ACLs are also performed at this stage based on the packet header updates, and once the packet passes all checks, it is transmitted on the output port.

Stage 6: Egress Traffic Manager

The Egress Traffic Manager stage is responsible for the granting of VOQ credit requests from input packet processors and managing egress queues.

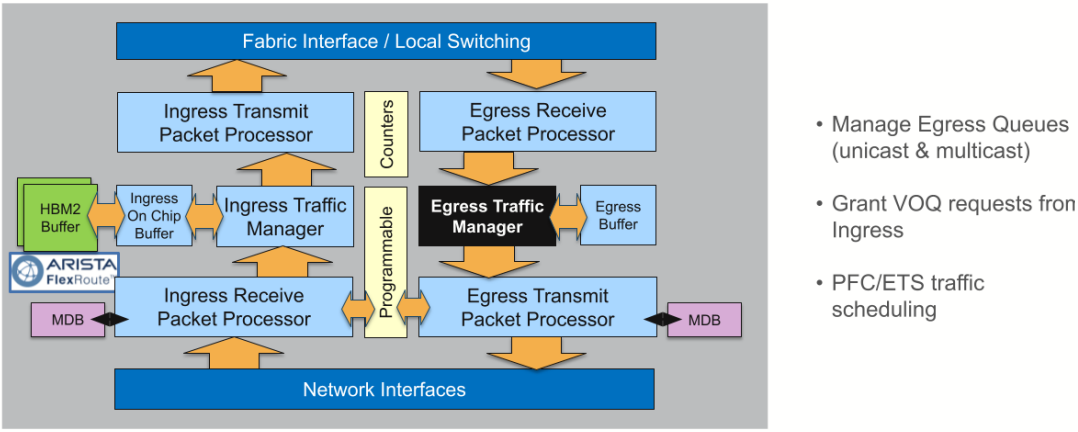


Figure 36: Packet Processor stage 6 (egress): Egress Receive Packet Processor

When an ingress packet processor requests to schedule a packet to the egress packet processor it is the Egress Traffic Manager stage that receives the request. If the output port is not congested then it will grant the request immediately. If there is congestion it will fairly balance the service requests between contending input ports, within the constraints of QoS configuration policy (e.g. output port shaping) while also conforming to PFC/ETS traffic scheduling policies on the output port. Scheduling between multiple contending inputs for the same queue can be configured to weighted fair queuing (WFQ) or round-robin.

The Egress Traffic Manager stage is also responsible for managing egress buffering within the system. There is an additional 32MB on-chip buffer used for egress queuing. This buffer is primarily reserved for multicast traffic as unicast traffic has a minimal requirement for egress buffering due to the large ingress VOQ buffer and fair adaptive dynamic thresholds are utilized as a pool of buffers for the output ports.

Stage 7: Egress Transmit Packet Processor

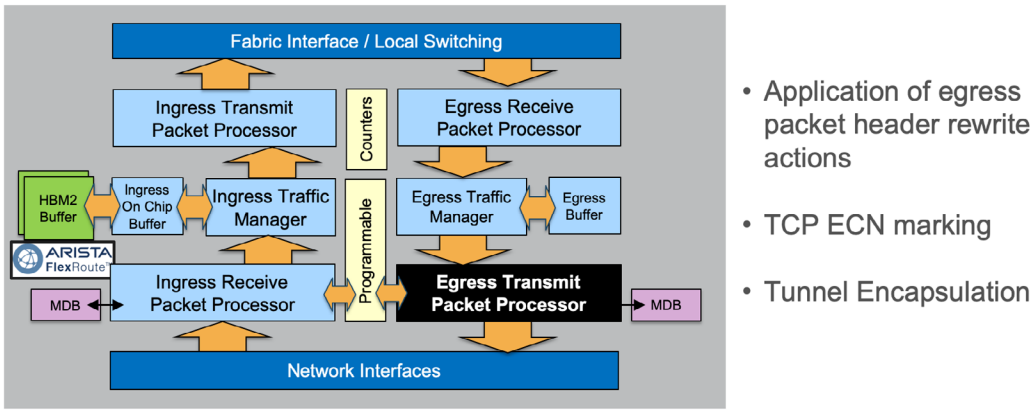


Figure 37: Packet Processor stage 7 (egress): Egress Transmit Packet Processor

In this stage, any packet header updates such as updating the next-hop DMAC, Dot1q updates and tunnel encapsulation operations are performed based on packet header rewrite instructions passed from the Input Receive Packet Processor stage. Decoupling the packet forwarding on ingress from the packet rewrite on egress provides the ability to increase the next-hop and tunnel scale of the system as these resources are programmed in a distributed manner.

This stage can also optionally set TCP Explicit Congestion Notification (ECN) bits based on whether there was contention on the output port and the time the packet spent queued within the system from input to output. Flexible Counters are available at this stage and can provide packet and byte counters on a variety of tables.

Stage 8: Network Interface (Egress)
Just as packets/frames entering the switch went through the Ethernet MAC and PHY layer with the flexibility of multi-speed interfaces, the same mechanism is used on packet/frame transmission. Packets/frames are transmitted onto the wire as a bit stream in compliance with IEEE 802.3 standards.

Arista EOS: A Platform For Scale, Stability and Extensibility

At the core of the Arista 7280R3 Universal Leaf platform is Arista EOS® (Extensible Operating System). Built from the ground up using innovative core technologies since our founding in 2004, EOS contains more than 8 million lines of code and years of advanced distributed systems software engineering. EOS is built to be open and standards-based and its modern architecture delivers better reliability and is uniquely programmable at all system levels.

EOS has been built to address two fundamental issues that exist in cloud networks: the need for non-stop availability and the need for high feature velocity coupled to high quality software. Drawing on our engineers experience in building networking products over more than 30 years, and on the state-of-the-art in open systems technology and distributed systems, Arista started from a clean sheet of paper to build an operating system suitable for the cloud era.

At its foundation, EOS uses a unique multi-process state-sharing architecture which separates system state information from packet forwarding and from protocol processing and application logic. In EOS, system state and data is stored and maintained in a highly efficient System Database (SysDB). The data stored in SysDB is accessed using an automated publish/subscribe/notify model. This architecturally distinct design principle supports self-healing resiliency in our software, eases software maintenance and enables module independence. This results in higher software quality overall, and accelerates time-to-market for the new features that customers require.

Arista EOS contrasts with the legacy approach to building network operating systems developed in the 1990s that relied upon embedding system state within each independent process, relying on extensive use of inter-process communications (IPC) mechanisms to maintain state across the system, with a manual integration of subsystems. These legacy system architectures lack an automated structured core like SysDB. In legacy network operating systems, as dynamic events occur in large networks or in the face of a system process failure and restart, recovery can be difficult if not impossible.

Additionally, as legacy network operating systems attempt to adapt to industry demands, such as streaming telemetry, individual subsystems must be manually extended to support state export into a system that was never designed to facilitate cloud-scale export mechanisms. As such, stabilizing and adapting to a wide range of telemetry and control protocols remains an ongoing challenge complicating integration and delaying migration to next-generation management interfaces for operators.

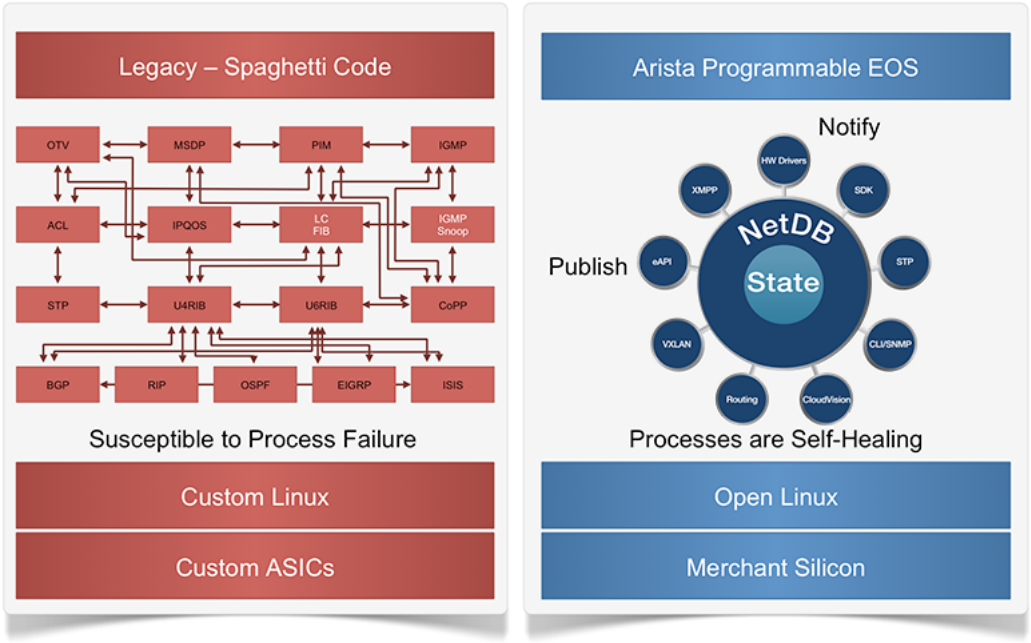


Figure 38: Legacy approaches to network operating systems (left), Arista EOS (right)

Arista took to heart the lessons of the open source world and built EOS on top of an unmodified Linux kernel maintaining full, secured access to the Linux shell and utilities. This allows EOS to utilize the security, feature development and tools of the vibrant Linux community on an on-going basis. This is in contrast to legacy approaches where the original OS kernel is modified or based on older and less well-maintained versions of Unix. This has made it possible for EOS to natively support things like Docker Containers to simplify the development and deployment of applications on Arista switches. Arista EOS represents a simple but powerful architectural approach that results in a higher quality platform on which Arista is able to continuously deliver significant new features to customers.

EOS is extensible with open APIs available at every level: management plane, control-plane and data-plane. Service-level and application-level extensibility can be achieved with access to all Linux operating system facilities including shell-level access. Arista EOS can be extended with Linux applications and a growing number of open source management tools to meet the needs of network engineering and operations.

Open APIs such as EOS API (eAPI), OpenConfig and EOS-SDK provide well-documented and widely used programmatic access to configuration, management and monitoring that can stream real-time network telemetry, providing a superior alternative to traditional polling mechanisms.

The NetDB evolution of SysDB extends the core EOS architecture in the following ways:

- NetDB NetTable enables EOS to scale to new limits. It scales the routing stack to hold millions of routes or tunnels with sub-second convergence.
- NetDB Network Central enables system state to be streamed and stored as historical data in a central repository such as CloudVision, HBase or other third party systems. This ability to take network state and efficiently and flexibly export it, is crucial for scalable network analysis, debugging, monitoring, forensics and capacity planning. This simplifies workload orchestration and provides a single interface for third party controllers.
- NetDB Replication enables state streaming to a variety of telemetry systems in a manner that automatically tolerates failures, and adapts the rate of update propagation to match the capability of the receiver to process those updates.

The evolution of SysDB to NetDB builds on the core principles that have been the foundation of the success of EOS: openness, programmability, and quality on a single build of EOS runs across all of our products.

System Health Tracer And Integrity Checks

Just as significant engineering effort has been invested in the software architecture of Arista EOS, the same level of detail has gone into system health and integrity checks within the system. There are numerous subsystems on Arista 7280R3 Universal Leaf platform switches that validate and track the system health and integrity on a continual basis:

- All memories where code executes (control-plane and data-plane) are ECC protected; single bit errors are detected and corrected automatically, double bit errors are detected.
- All data-plane forwarding tables are parity protected with shadow copies kept in ECC protected memory on the control-plane. Continual hardware table validation verifies that the hardware tables are valid and truthful.
- All data-plane packet buffers are protected using CRC32 checksums from the time a packet/frame arrives, and at the time it leaves the switch. The checksum is validated at multiple points through the forwarding pipeline to ensure no corruption has happened, or if there has been a problem, rapidly facilitate its isolation.
- Forward Error Correction (FEC) is also utilized for traffic across the fabric modules, both to correct errors (if they occur) but also to help monitor data-plane components of the system for problems.
- Data-plane forwarding elements are continually testing and checking reachability with all other forwarding elements in the system. This is to ensure that if there are issues they can be accurately and proactively resolved.

Conclusion

Designed to address the demands of the world's largest cloud and service providers the Arista 7280R3 Series switches continue to provide operators with a proven, industry leading, platform to evolve their network capabilities. By combining industry leading 400G density with Internet scale service capabilities and next generation packet processing functionality at the optimum intersection of performance and power utilization.

The 7280R3 leverages the proven architecture that has made the previous generations of the product so successful; focus on efficient system design, reliability and flexibility. This trend continues with innovations in the packet processors powering the R3-series, enabling operators to use the 7280R3 in an ever wider range of roles with a single hardware platform.

Arista's EOS network operating system continues to lead the industry in openness, extensibility and software quality. EOS has been leading the industry in telemetry innovations through the availability of NetDB and enabled operators to truly automate their network deployments through rich programmatic interfaces and support for industry standards such as OpenConfig.

Given the cloud scale hardware and software capabilities of the 7280R3, it makes the ideal platform for a range of applications. The 7280R3 is ideally suited for cloud scale data centers, Service Provider WAN backbones and Peering edges as well as large enterprise networks.

Santa Clara—Corporate Headquarters
5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500
Fax: +1-408-538-8920
Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office 1390
Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office
Global Tech Park, Tower A, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062