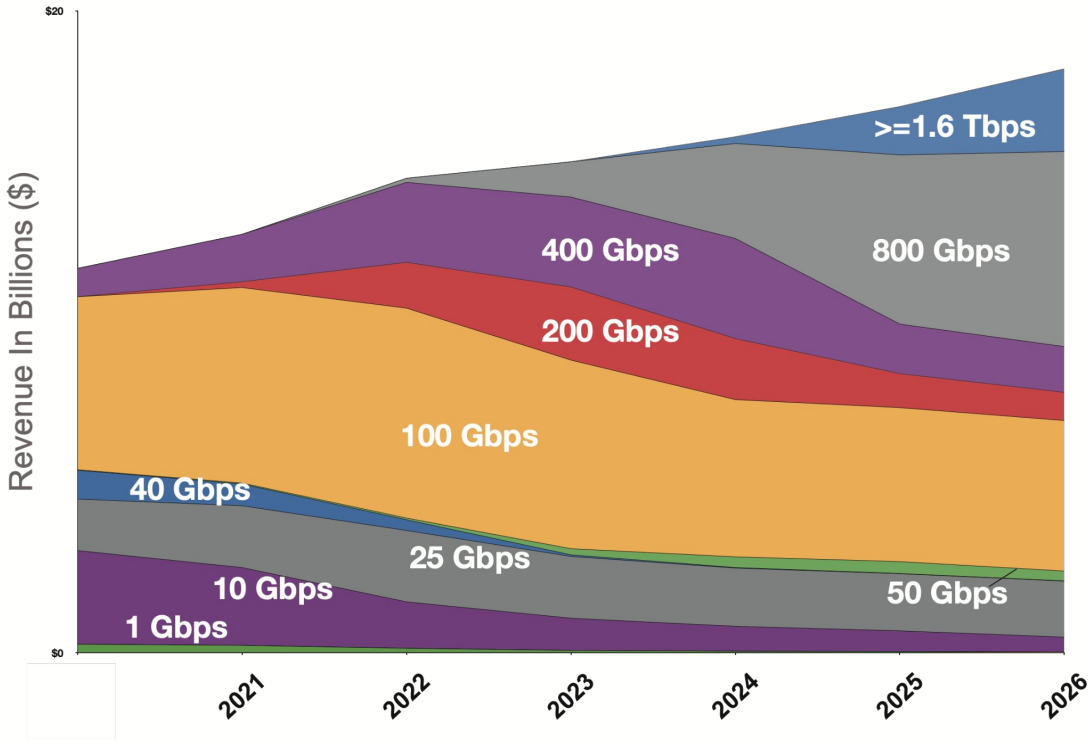
Merchant Silicon for Service Providers

Ethernet Switch Revenue Forecast

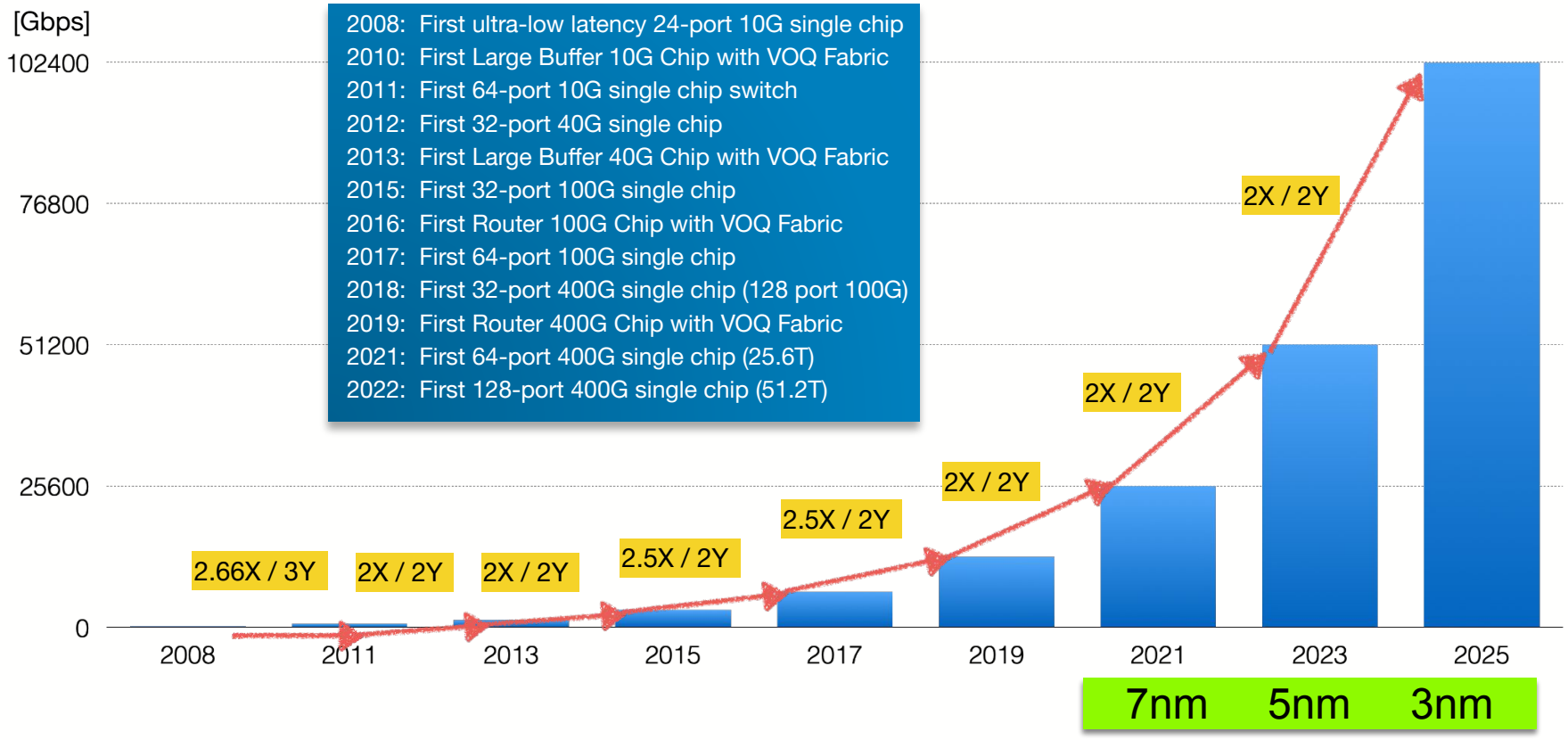


Source: Dell'Oro March 2022 - Long Term Ethernet Switch Forecast

800G and 1600G expected to be 45% of Market in 2026

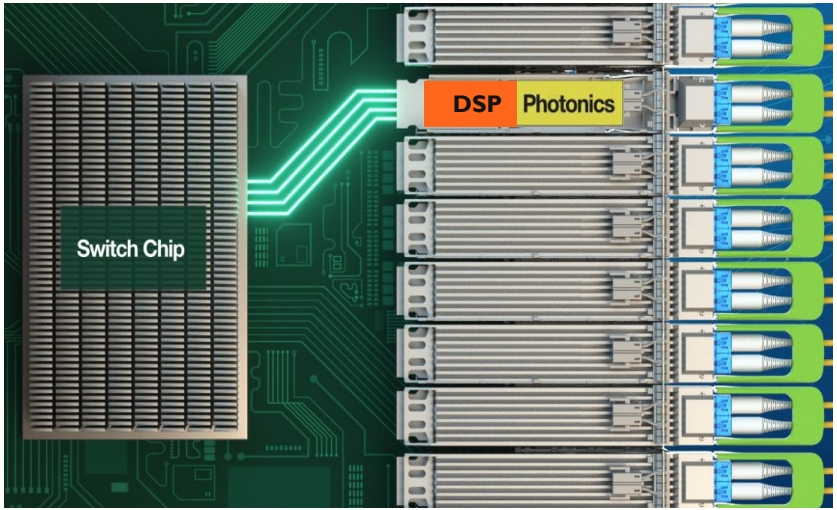
400G and below expected to be 55% of Market in 2026

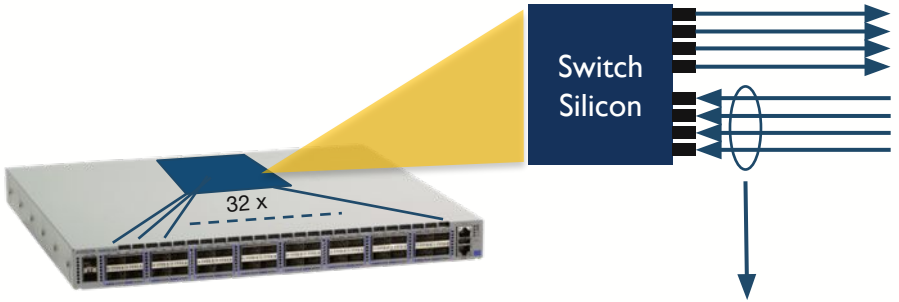
Single-chip Switch Bandwidth & Serdes Speeds



SERDES Speeds are Key to Scaling networks

- Serdes (or **Ser**ializer-**Des**erializers) refer to the technology used for high-speed chip I/O
- Serdes speeds place a fundamental limit on datacenter bandwidth
- The easiest way to go faster is (for serdes speeds) to go Faster

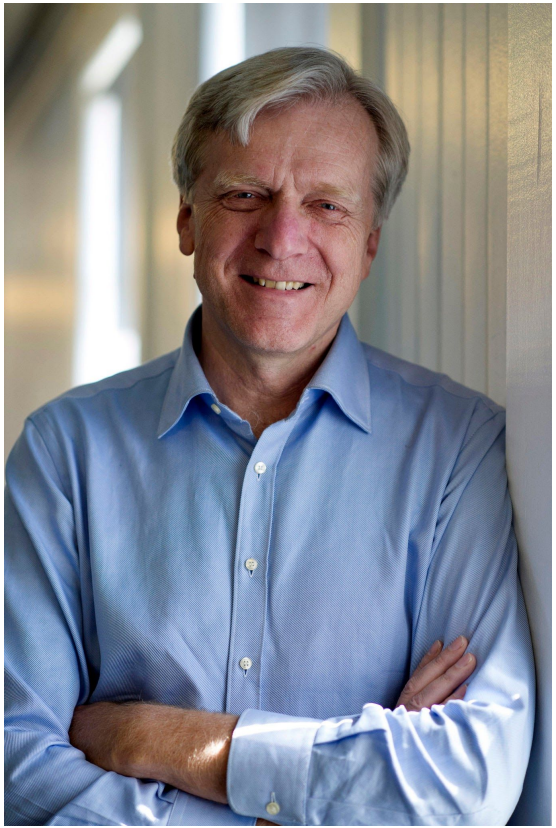




Ethernet switch chip I/O **(or serdes) speeds**:
10G -> 25G -> 50G -> 100G

ARISTA

“The easiest way to go faster is to go faster”

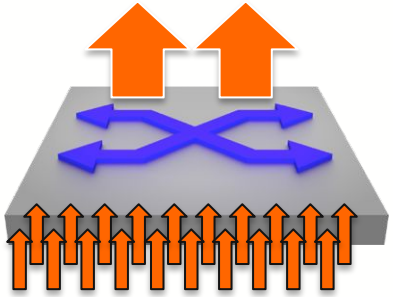


Process Technology Improvements (TSMC)

Process Node	7nm	5nm	3nm
Relative Density	1	1.5	2.25
Speed @ IsoPower	1	1.15	1.4
Power @ IsoSpeed	1	0.8	0.6
Volume Manufacturing	2019	2021	2023

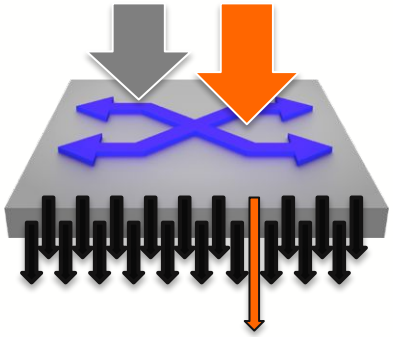
Each process generation enables more throughput, better Power Efficiency, more buffers, bigger routing tables, etc

When Buffers Matter in Provider Networks

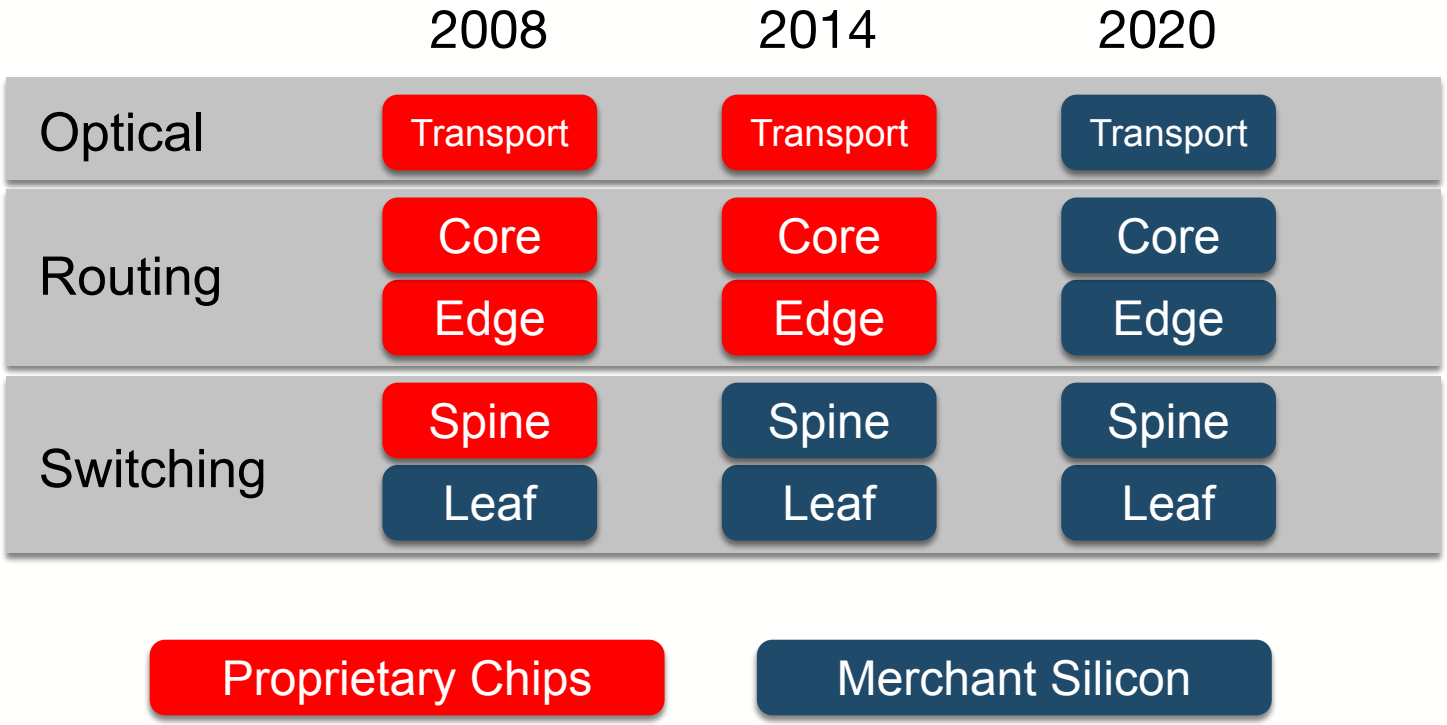


Incast (Many to Fewer)

Speed Change (Faster to Slower)



Merchant Silicon Maturation



Cost of silicon design

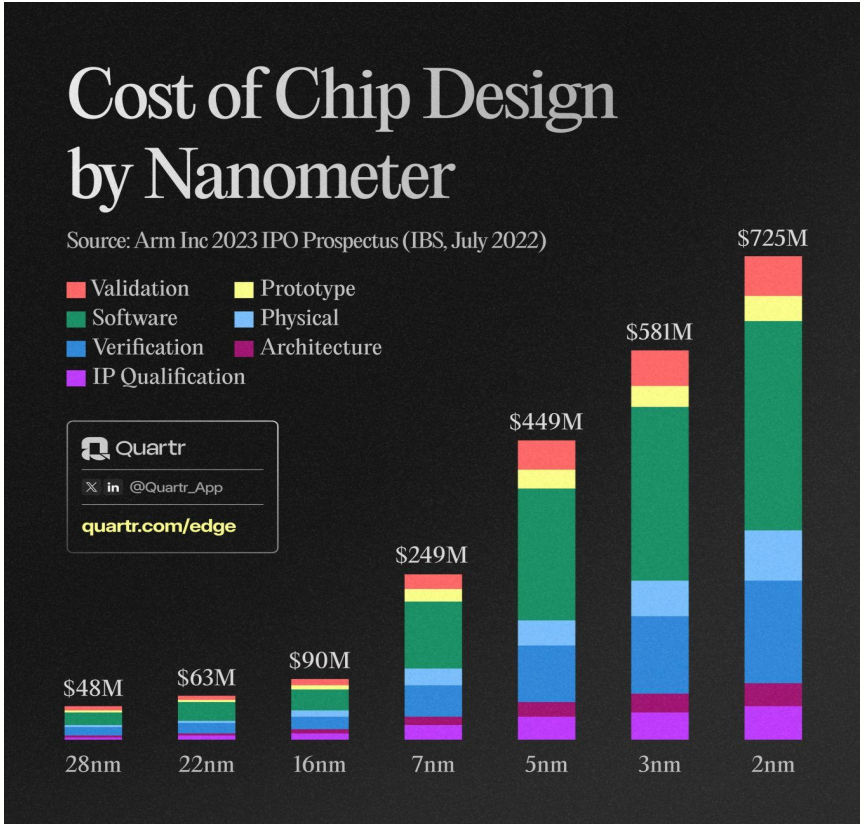
Cost of design (aka tape out) has risen significantly

- 28nm: 48M$
- 22nm: 63M$ +33%
- 16nm: 90M$ +43%
- 7nm: 249M$ +176%
- 5nm: 449M$ +80%
- 3nm: 581M$ +30%
- 2nm: 725M$ +25% (new: ~ 2nd half of 2025)

This favors merchant silicon because of higher volume

This also leads to:

- Lower cost / better economy of scale
- Faster product cycles
- Faster innovation (because of shorter cycles)
- Lower power consumption
- Better scale
- More features (not always but often)



Choices in Switching Silicon

All chip makers have access to the same technology

- same fabs and processes
- same memories, TCAMs, serdes
- same clock rate

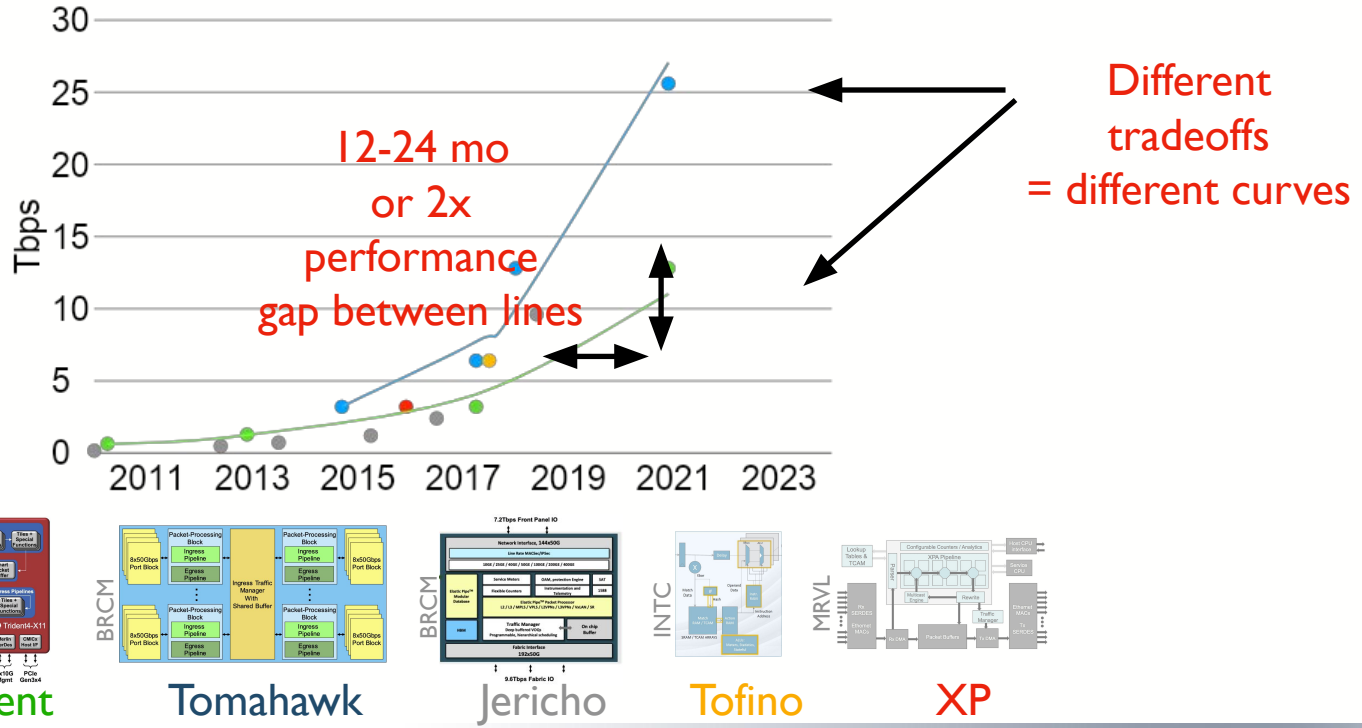
Differences arise *primarily* because of

- design tradeoffs for different use cases
- process shifts (28nm -> 16nm -> 7nm -> 5nm)
- faster innovation cycles

There is no fundamental advantage to proprietary silicon

Merchant Silicon Trajectory

10 years at Arista, across chip families



ARISTA

Domain-Specific Products for Different Networks

Trident

Enterprise application stacks

RoCEv2, EVPN, VXLAN

Rich Telemetry for deep visibility

Compute TOR for 10/25/50/100G

Flexible traffic management

128 x 100G in 4RU

Tomahawk

Cloud application stacks

Highest Switch performance

Lowest Latency

Scale Out & High Radix

High density 400G Fixed Spines

128 x 200G in 4RU

Jericho

Switch with **deep buffers**

EVPN, MPLS, SR

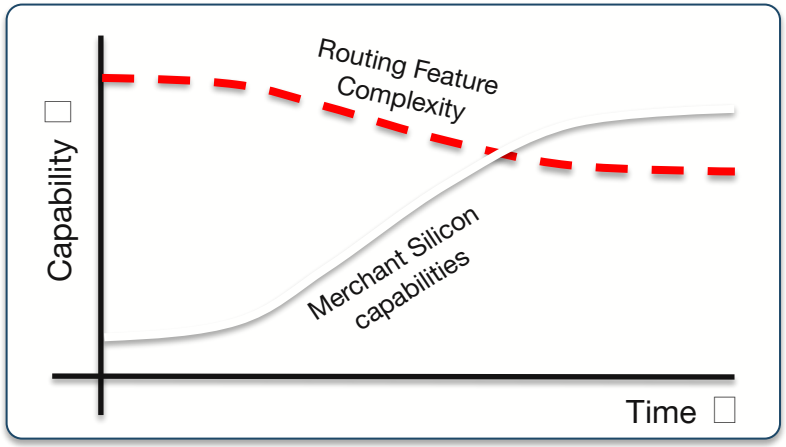
High capacity routing scale

Metro & DCI with MACSec & ZR

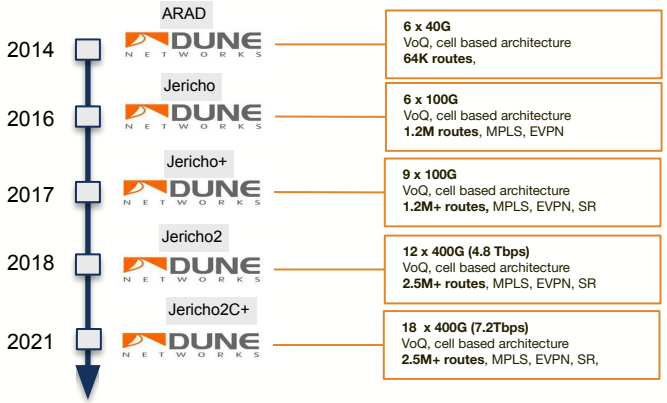
Fixed and Modular form factors

Arista: Bringing Merchant Silicon to the Routing Market

- Look at the routing market
  - The domain of the Network vendor's own in-house ASIC
  - Due to complexity of functionality and table scale requirements

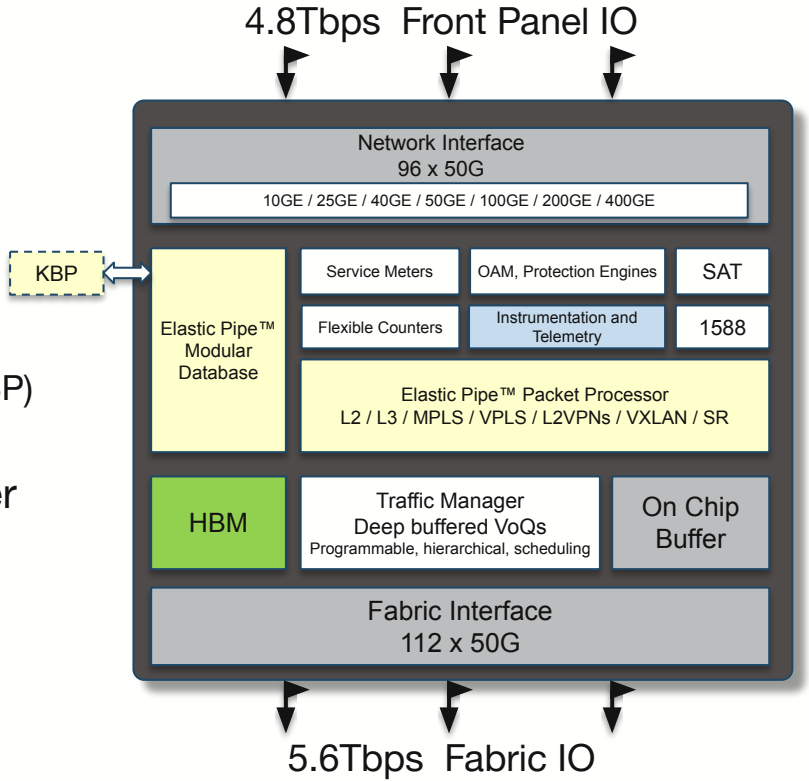


- Lines are blurring with latest Merchant silicon
  - Jericho chipset design for routing deployments
  - Market leading performance and 100G/400G density
  - Internet scale, multiple encap, Deep label stack, VoQ



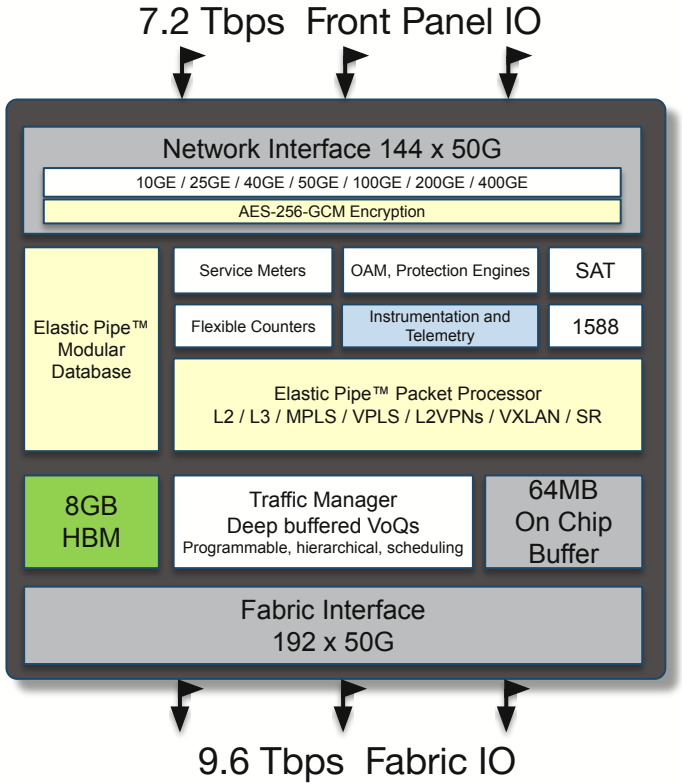
10Tbps - Jericho2

- 10Tbps of High Performance with rich features
  - Total of 208 PAM-4 50G Serdes
  - 4.8Tbps Network I/O and 2Bpps packet processing
  - Flexible Network Interfaces - 10G to 400G
- Flexible Lookup Tables and Programmable Pipeline
  - Fungible on chip tables allow multiple use case profiles
  - Off-chip expandability with External table expansion (KBP)
  - Flexible Pipeline allows reconfiguration of forwarding
- Hierarchical Traffic Management with Deep Buffer
  - 8GB High Bandwidth Memory (HBM)
  - 32MB On Chip Buffer
- Network Instrumentation and Telemetry
  - Hardware Accelerator
  - Monitor of large numbers of sessions



16.8 Tbps - Jericho2C+

- 16.8 Tbps of High Performance with rich features
  - Total of 336 PAM-4 50G SerDes
  - 7.2Tbps Network I/O and 2.7Bpps packet processing
  - Flexible Network Interfaces - 10G to 400G
  - Integrated TunnelSec Encryption (MACsec, IPsec, VXLANsec)
- Flexible Lookup Tables and Programmable Pipeline
  - Fungible on chip tables allow multiple use case profiles
  - Off-chip expandability with External table expansion (KBP)
  - Flexible Pipeline allows reconfiguration of forwarding
- Hierarchical Traffic Management with Deep Buffer
  - 8GB High Bandwidth Memory (HBM)
  - 64MB On Chip Buffer
- Network Instrumentation and Telemetry
  - Hardware Accelerator
  - Monitor of large numbers of sessions



Consistent System Resources: J2C+/J2/J2C/Q2C

	KAPS tbd		BIG KAPS			
Profile	**L3 (default)**	**Balanced**	**L3-XL (default)**	**L3-XXL**	**L3-XXXL**	**Balanced-XL**
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
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

Maximum values dependent on shared resources / user configuration

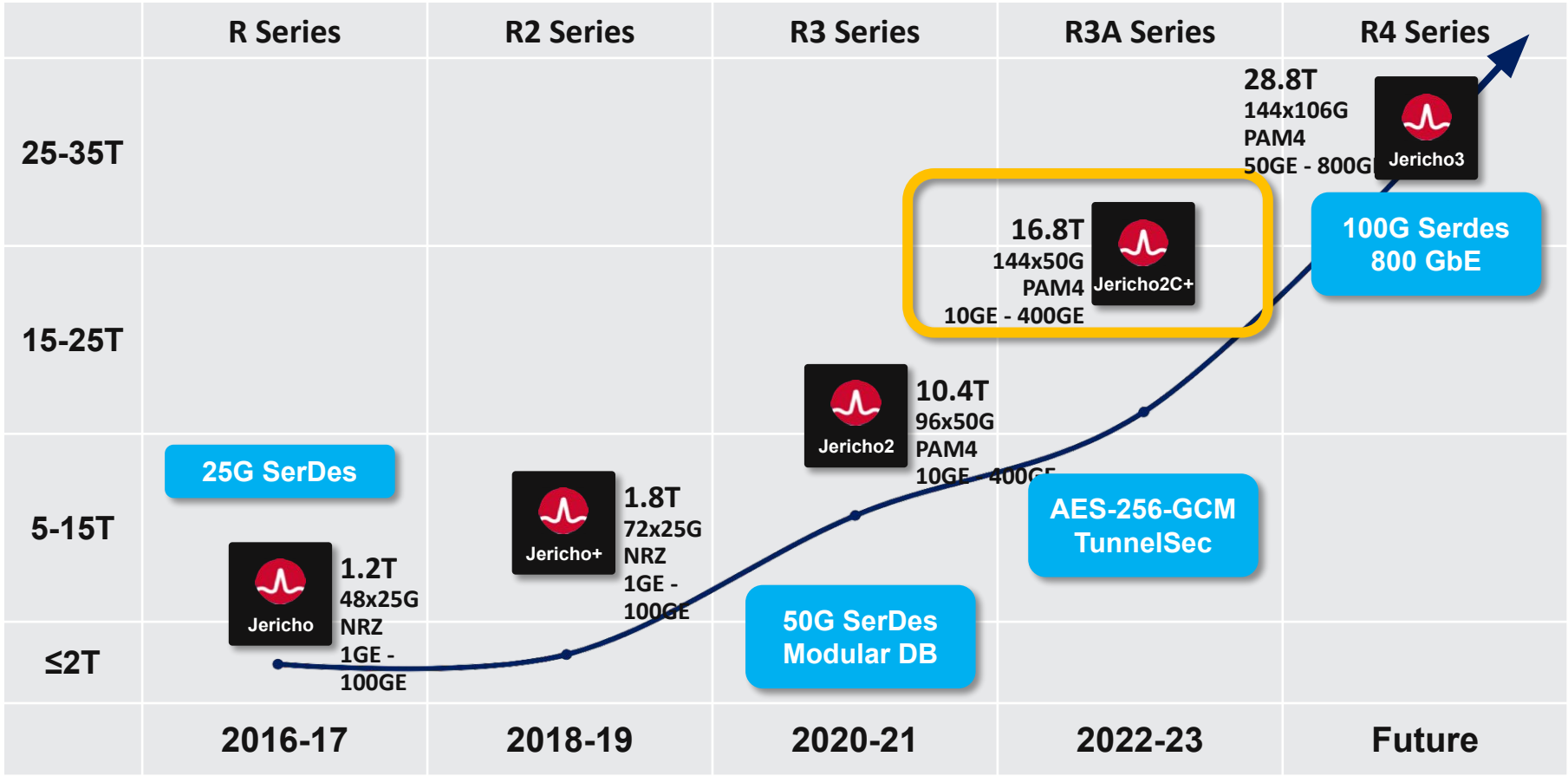
Jericho2 hardware resources are fungible. Values shown are unidimensional maxima for default profiles

Jericho2C+ - The Engine for 400Gbps

Jericho2c+ Lowest Cost, Power & RU per Gbps	Up to 50% Improvement from previous generation
Jericho2c+ 400Gbps Strong Encryption	MACsec, IPsec and VXLANsec at 10-400Gbps
Jericho2c+ Dense 400G ZR/ZR+ for WAN/DCI	Broad ZR/ZR+ Support with integrated Line System Ports
Jericho2c+ Rich DC and WAN Feature Set Large Scale Resources	Consistent Jericho2 Feature-set with dedicated 8GB HBM Deep Buffers
Jericho2c+ Flexible Product Choice	All Models Available in 3 Scale Configurations

Complete Portfolio - Uncompromised Features and Scale

Jericho based Portfolio

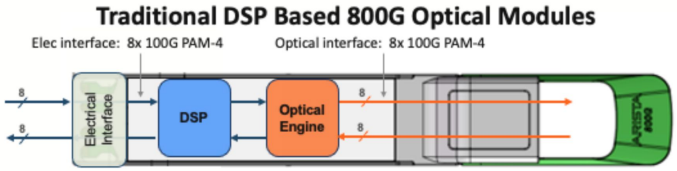


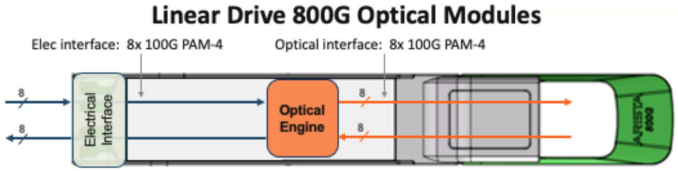
Rate Adapting 1G optics

- Support 1G-LX and 1G-SX on platforms that have a minimum port speed of 10G
  - e.g. J2 based platforms have a minimum port speed of 10G
- Connect to other devices that use CL37 (optical) autoneg when the used platform does NOT support CL37 autoneg
  - some platforms support 1/10/25G but don’t support CL37 autoneg

What are Linear Drive Optics Modules?

1. Linear Drive means no DSP or CDR in transceiver
   - Just a linear driver to provide required modulator voltage
2. Requires a high-performance switch SERDES
   - And very careful signal integrity design
3. Achieves power savings similar to direct drive CPO
   - While retaining the many advantages of pluggable optics modules
   - Opportunity to cut optics module power by 50% and system power by up to 25%



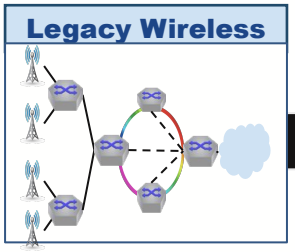


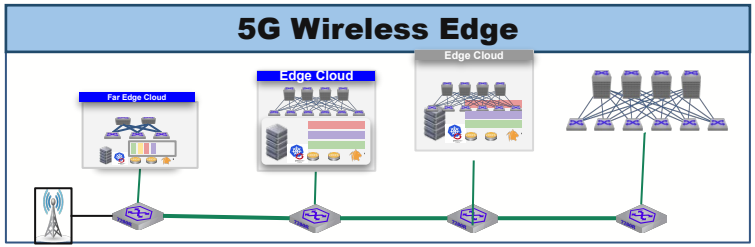
How is AI transforming the service provider market?

- AI networks need VoQ based deep buffer fabrics
- Perfect fit for Jericho chips
- next version (J3) already in
- made for high amount of 800G interfaces
- also beneficial for SPs as BW demands still >>
- also SPs benefit from enhanced telemetry functions

SP Access Networks Evolution

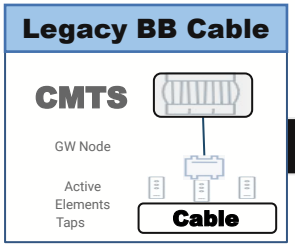
Closed vendor-centric

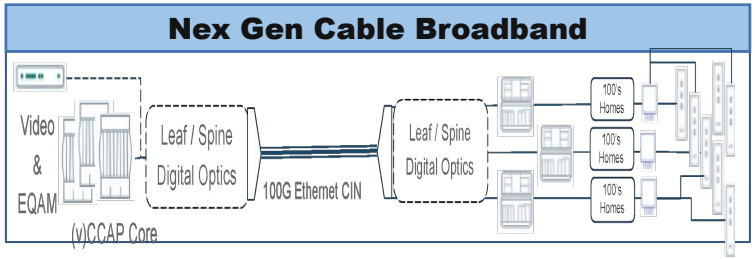




Open software-centric

Capped scale

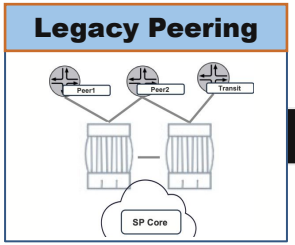


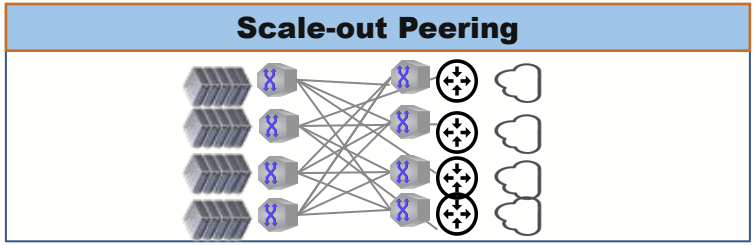


Elastic scale

Complex protocols

Open & simplified protocols



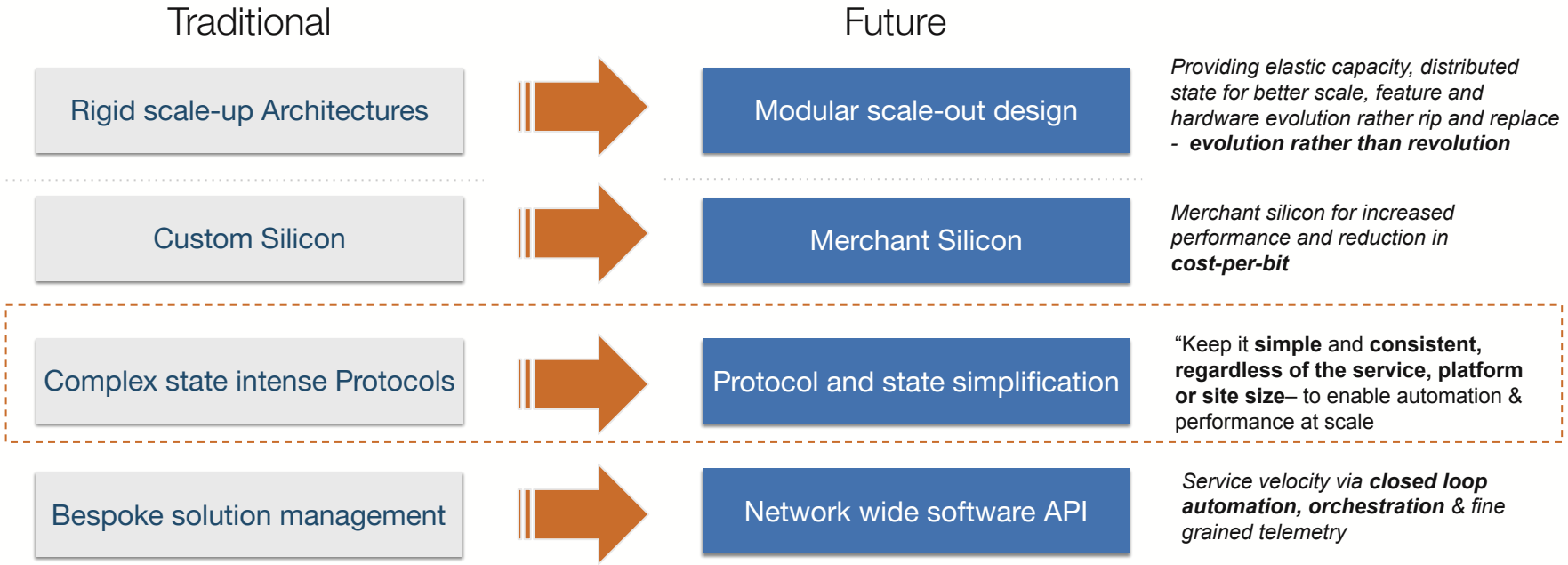


Difficult to deploy & manage

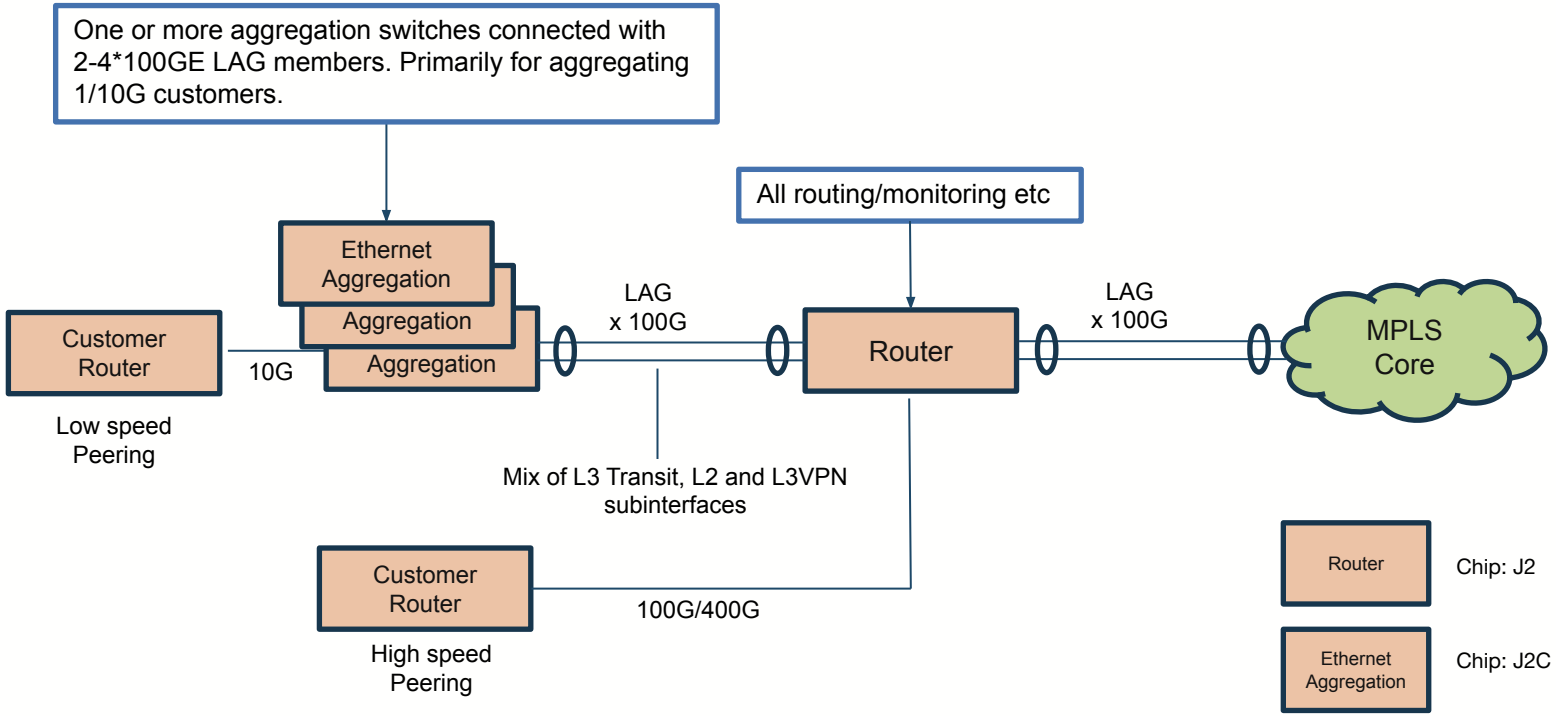
Telemetry driven management

Simplify – The network Edge and core

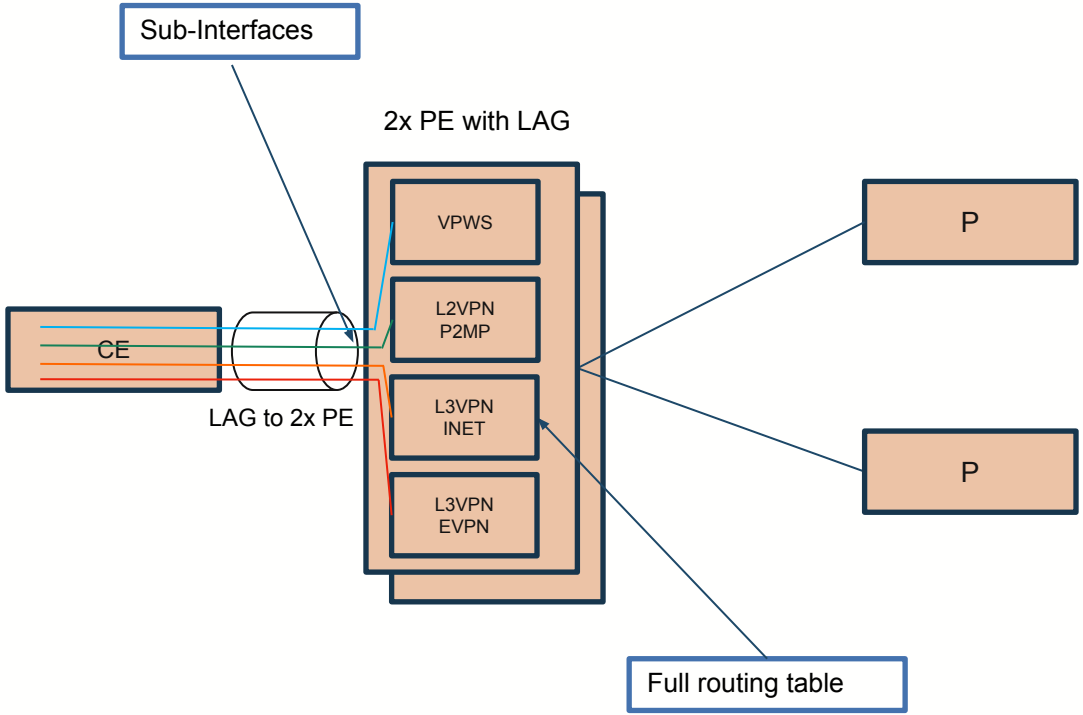
- To accelerate the adoption of high performance Merchant Silicon
  - Lessons learnt from the Cloud on how to scale without linearly growing CapEx/OpEx costs



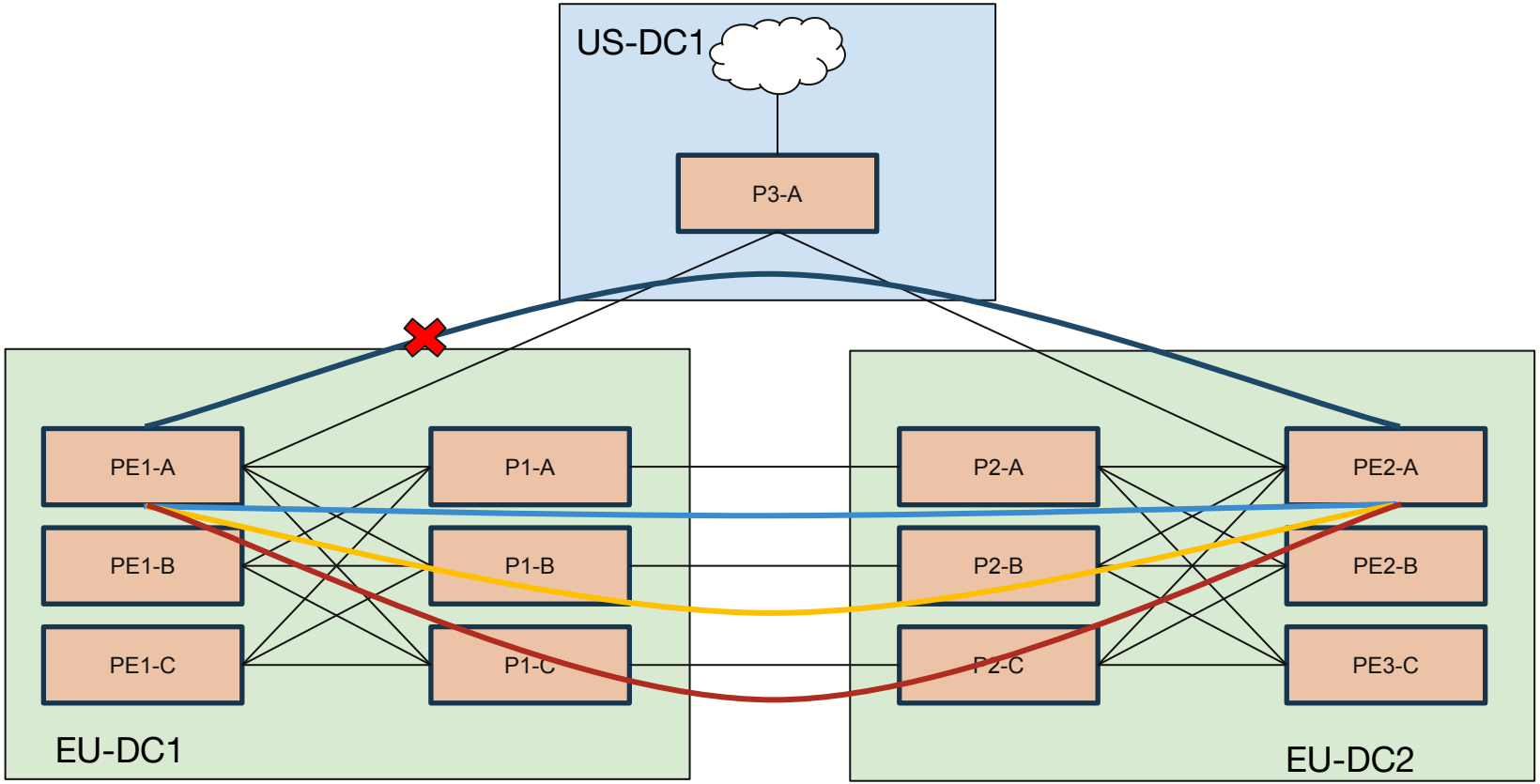
Example Deployment at the Service Edge



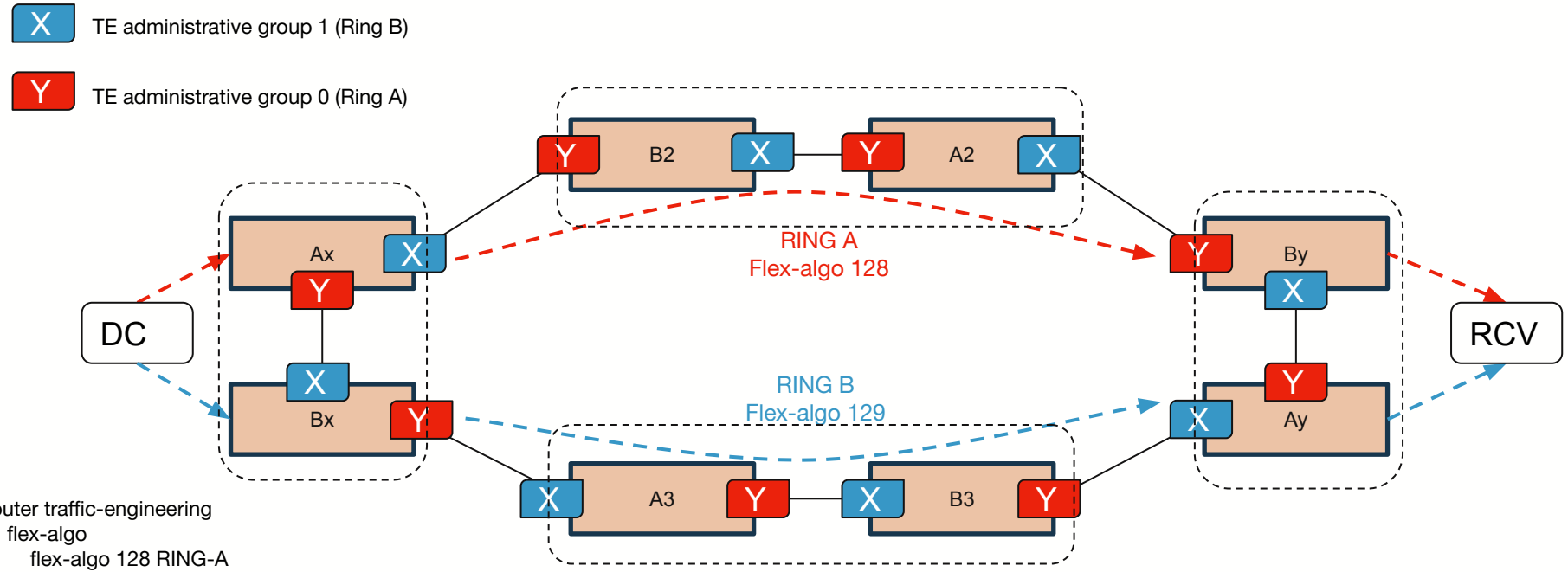
Example Deployment at the Service Edge



Example Deployment at the core



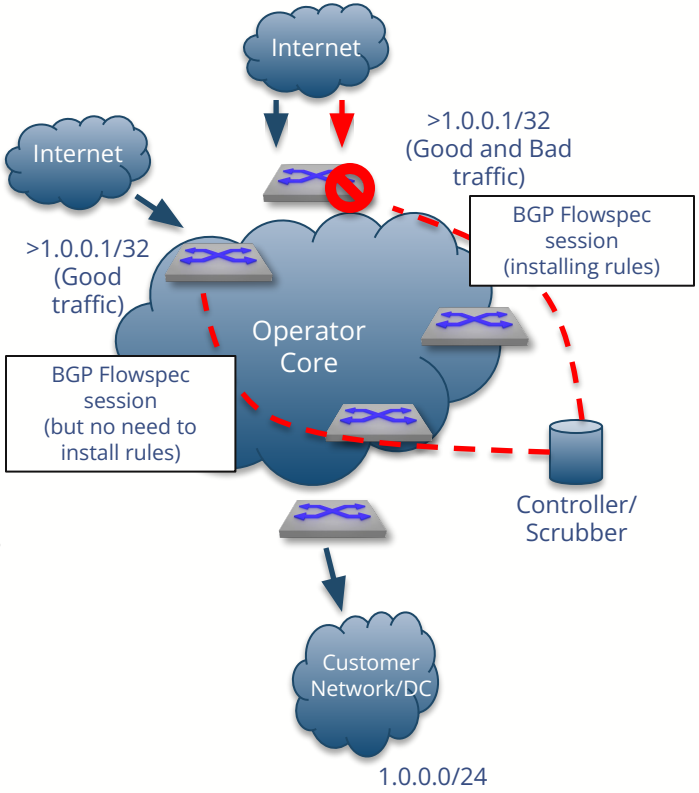
Example Deployment with Traffic-Engineering




```
router traffic-engineering
   flex-algo
      flex-algo 128 RING-A
         administrative-group include all 0
         color 101
      !
      flex-algo 129 RING-B
         administrative-group include all 1
         color 201
```

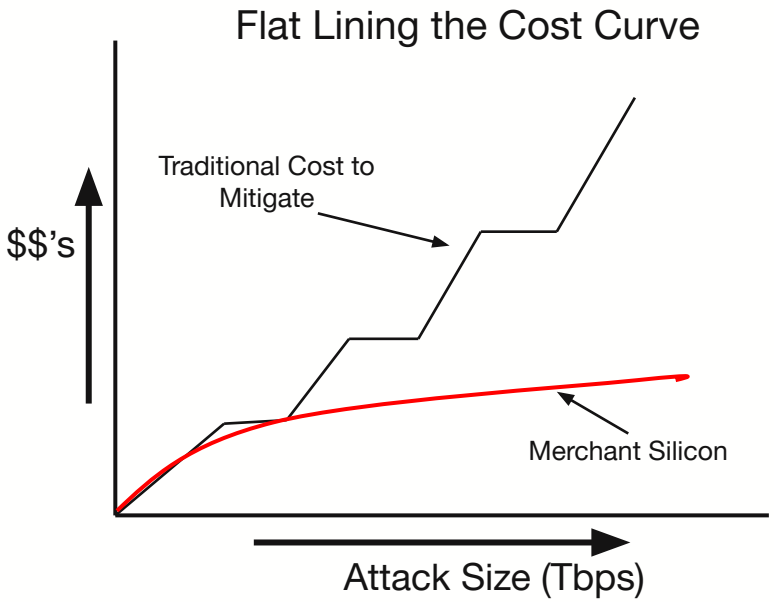
DDoS mitigation using BGP FlowSpec

- Capacity
  - The Nodes capacity regards TCAM and other hardware dependency easy controllable since Flowspec sessions only where it make sense, and also controlled to each peer(s)
- Flowspec installed only where needed
  - Flowspec specific to Peer/Node/Customer interfaces where traffic enters => Flowspec installed only where it's needed
- Mix of actions
  - Drop, rate-limit, relay can be used based on demands
  - Example stages of them (start with relay, follow by rate-limit and perhaps in the end just drop in case shape not enough)



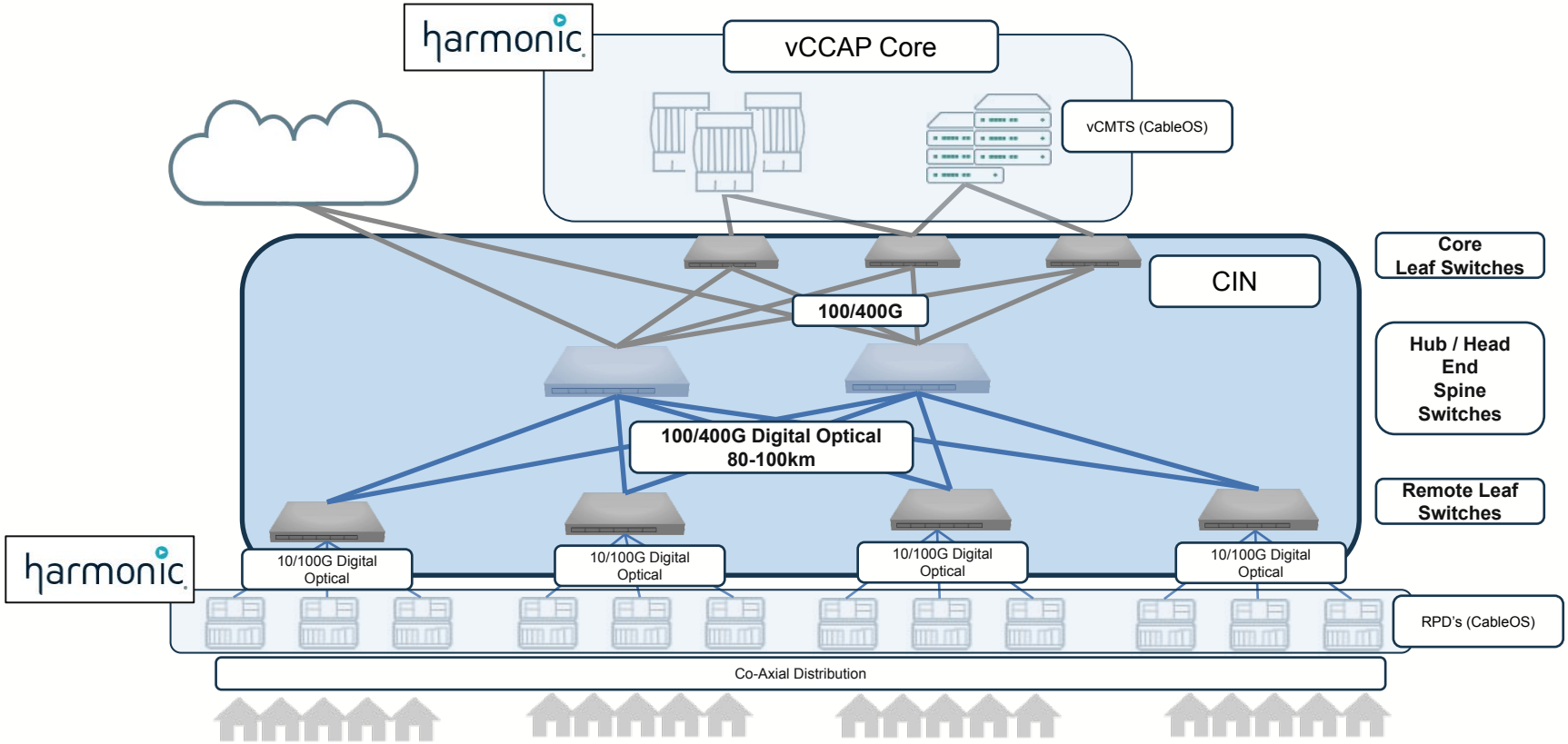
Merchant Silicon Security Advantage for Service Providers

- Cost to Mitigate no longer directly proportional to the size of DDOS Attacks
- High Scale ACL Support
- Elastic Resources to expand or contract based on Volumetric Attacks
- Fine Grain Telemetry



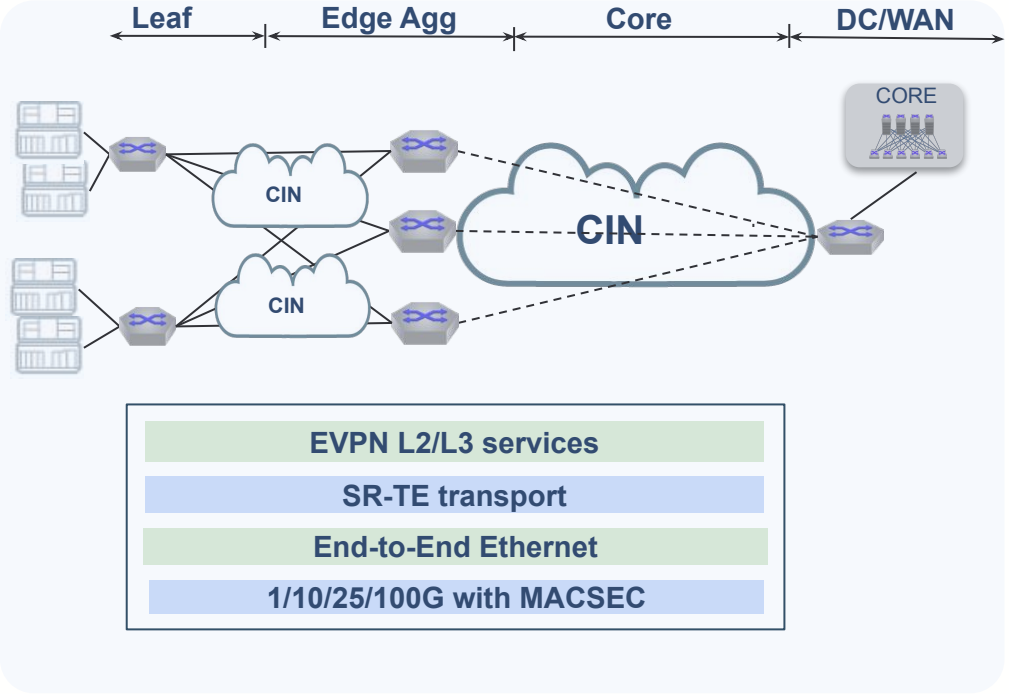
Merchant Silicon bringing Cost Effective DDOS Mitigation Solution

Common R-PHY CIN Topology



ARISTA

Towards Efficient and Simplified CIN Transport



- Simplified transport and services
  - SR/SR-TE transport
  - EVPN for L2/L3 services
  - Infrastructure slicing with Flex Algo
- CIN-Transport Optimized Routers
  - Compressed environmental footprint
  - Timing/Sync-E support
  - Ultra Low Latency components
  - 1G → 400G with variety of Optics
  - (r)ECMP tunable hashing
- Operational simplicity
  - Hitless software upgrade
  - Full programmability with fine-grained telemetry

Thank You

arista.com