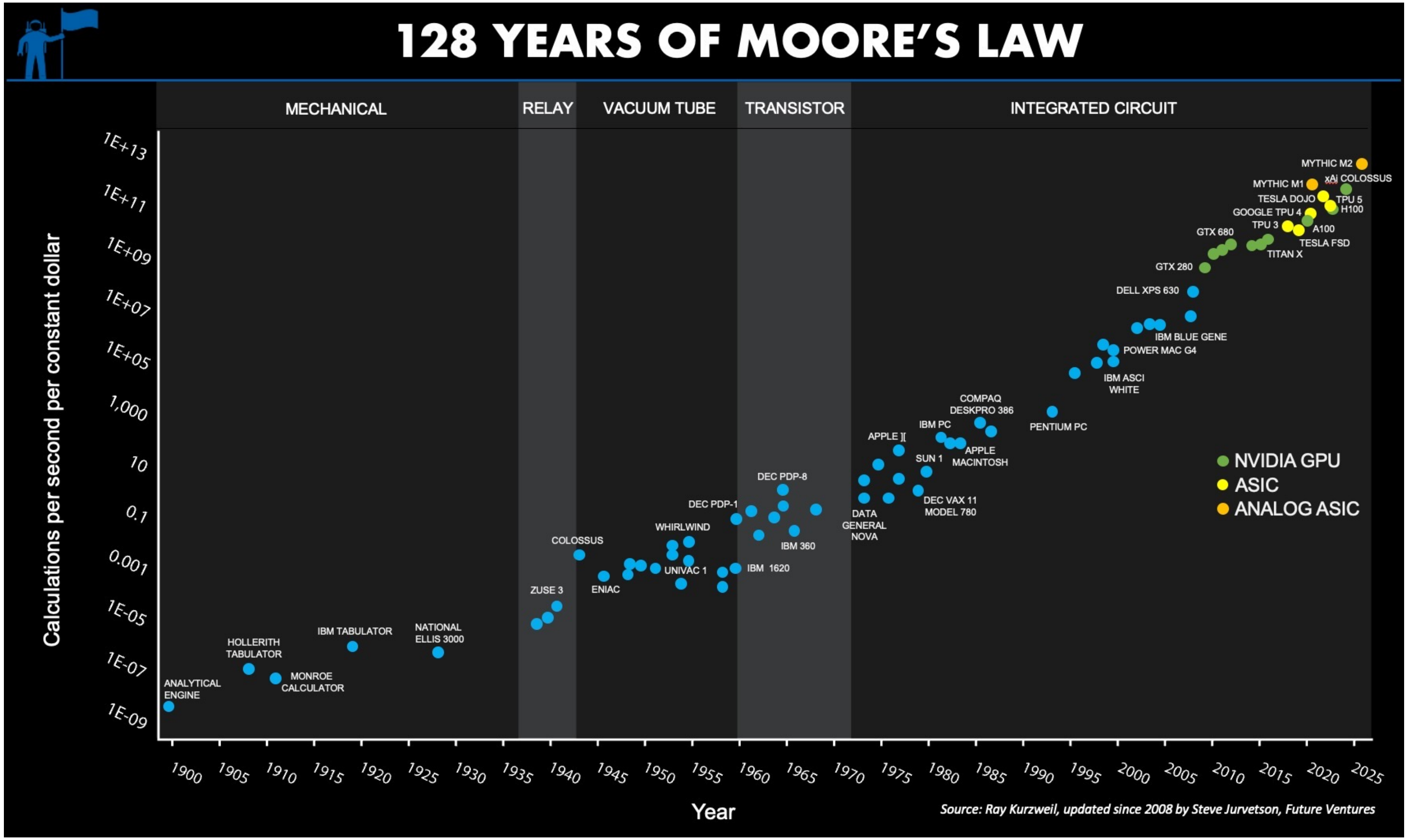
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Redes de Nueva Generación
Área de Ingeniería Telemática

Arquitectura de conmutadores Switching ASICs

Evolución del hardware

- 40 años de ley de Moore (x2 transistores cada 24 meses)
- Tick-tock
- Servidores ciclos de 2 años, networking reutilización 8-10 años (!!)
- Hoy en día producción en 3 nm y comienzos en 2nm

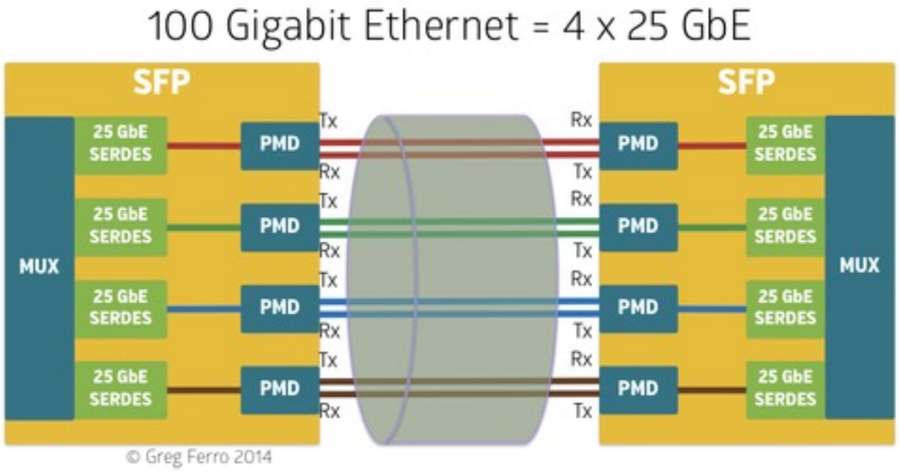


Source: Ray Kurzweil, updated since 2008 by Steve Jurvetson, Future Ventures

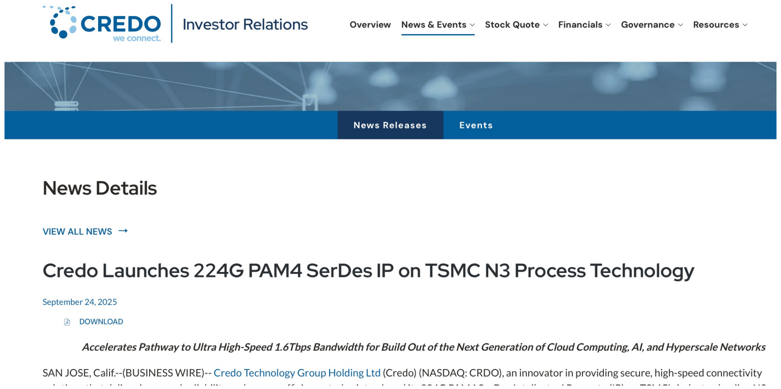
2025 Apple M3 Ultra: 184x10^9 transistores

Evolución del hardware

- Durante años ha estado desacoplada la evolución de hardware de computación del hardware de red
- Esto ha hecho que se moviera a software tareas de red (redes virtuales, VXLAN, etc) pues mejores CPUs eran más baratas que mejores switches
- Fabricantes de equipos de red están adoptando los ritmos de producción de electrónica
- Empujados por pocos grandes clientes
- Por ejemplo: donde teníamos SerDes a 10 Gb/s los tendremos posteriormente a 25 Gb/s, al mismo coste
- Esto permite interfaces 100GE (4x25) donde antes teníamos 40GE (4x10), al mismo precio
- Hoy en día SerDes a 112 Gb/s es común. Existen ya a 224Gb/s



http://www.networkcomputing.com/data-centers/25-gbe-big-deal-will-arrive/1714647938



News Details

VIEW ALL NEWS →

Credo Launches 224G PAM4 SerDes IP on TSMC N3 Process Technology

September 24, 2025

DOWNLOAD

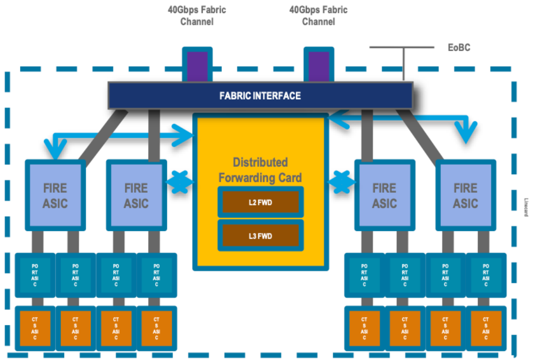
Accelerates Pathway to Ultra High-Speed 1.6Tbps Bandwidth for Build Out of the Next Generation of Cloud Computing, AI, and Hyperscale Networks

SAN JOSE, Calif.--(BUSINESS WIRE)-- Credo Technology Group Holding Ltd (Credo) (NASDAQ: CRDO), an innovator in providing secure, high-speed connectivity

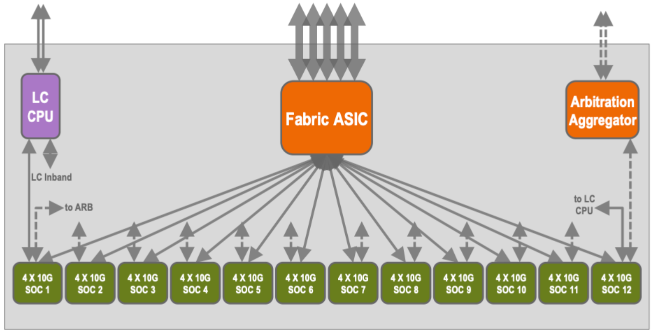
Evolución del hardware

- Lo que era un switch modular puede ser ahora un SoC
- A día de hoy SoC (Switch on Chip) a varios Tbps
- En los últimos 20 años
  - Acceso a DRAM: x90
  - Número de transistores: x8.000
  - Capacidad en switch: x30.000

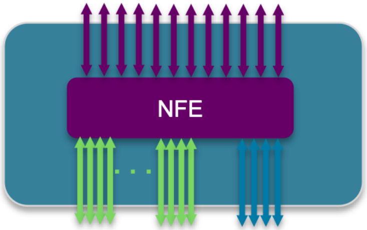
    (nº puertos x bandwidth)



32 x 10G Ports



48 x 10G Ports



64 x 10G Ports

Design Shifts Resulting from Increasing Gate Density and Bandwidth

Merchant silicon

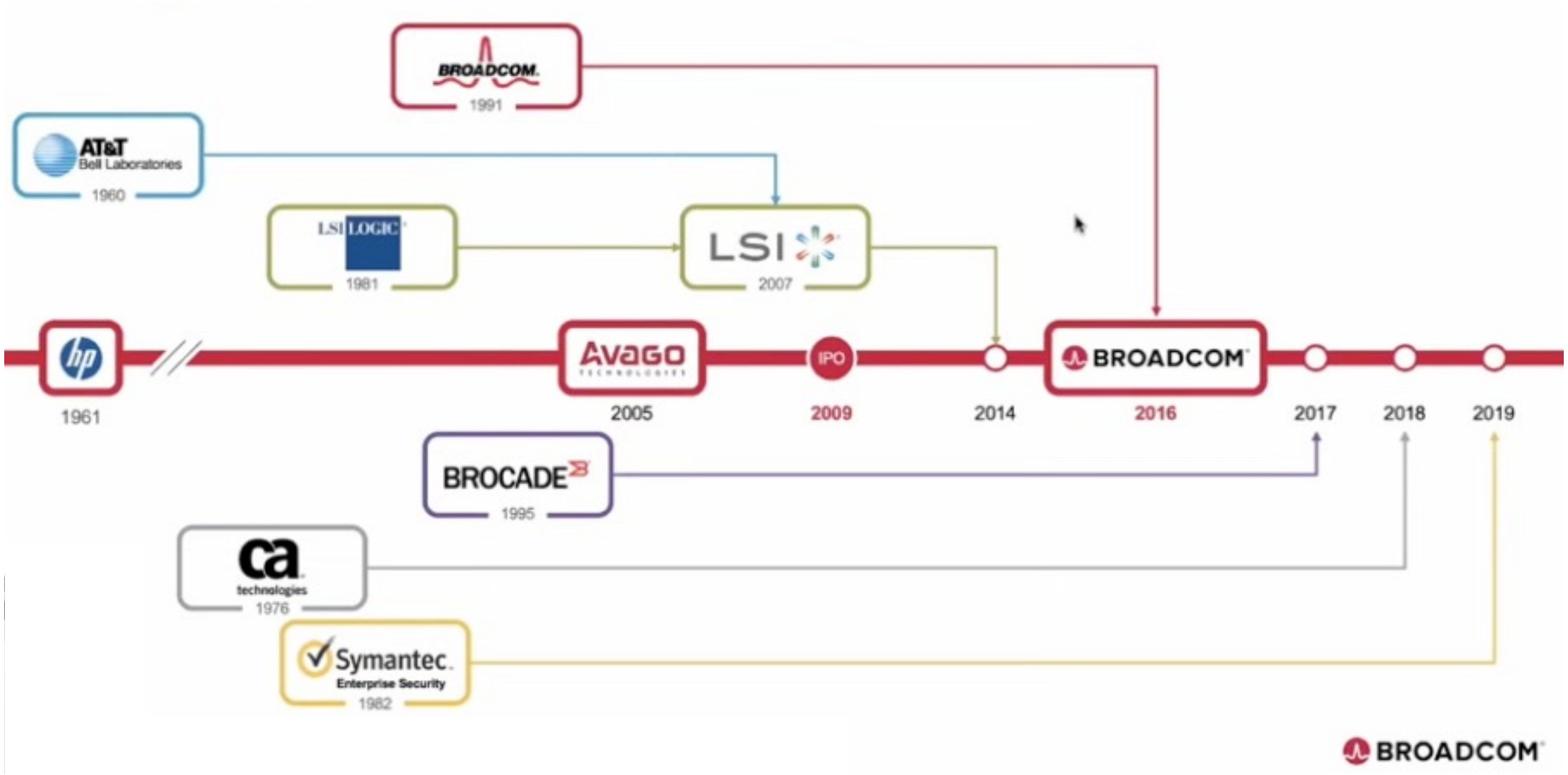
- ASICs creados por una compañía pero switches ensamblados por otra
- Broadcom
- Marvell
- Barefoot Networks (ahora Intel)
- Mellanox (ahora Nvidia)

MARVELL™

tsmc

Ejemplos de switching ASICs

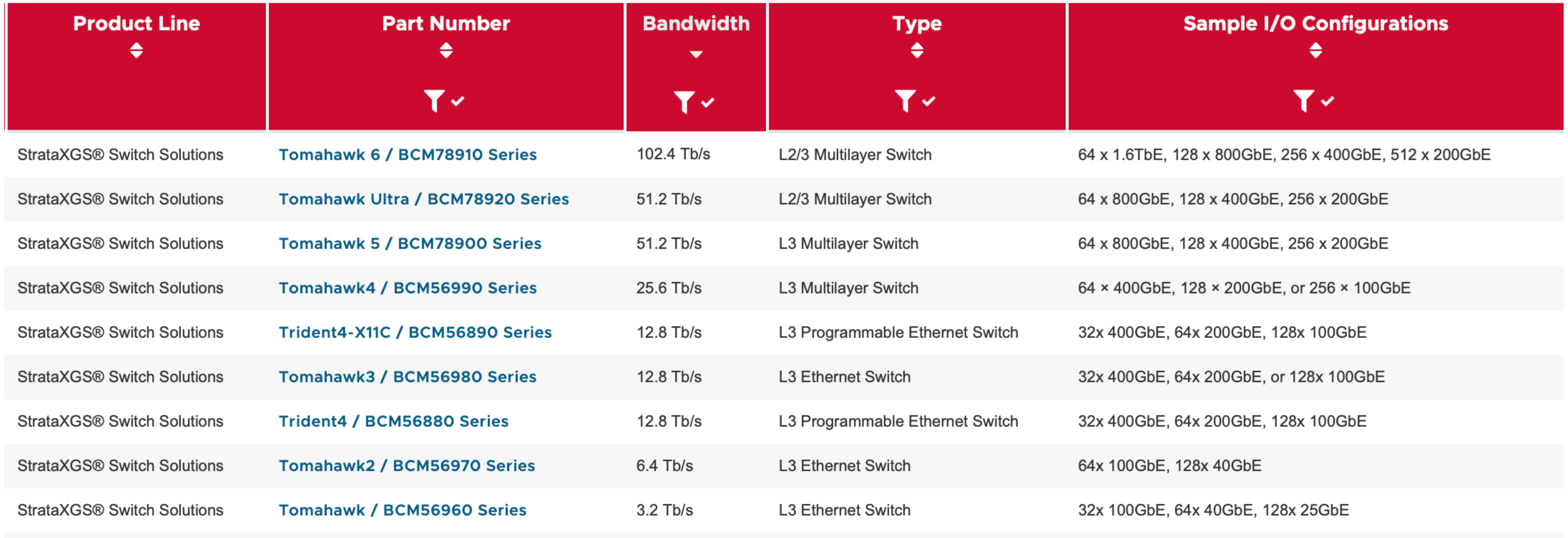
Ejemplo: Broadcom



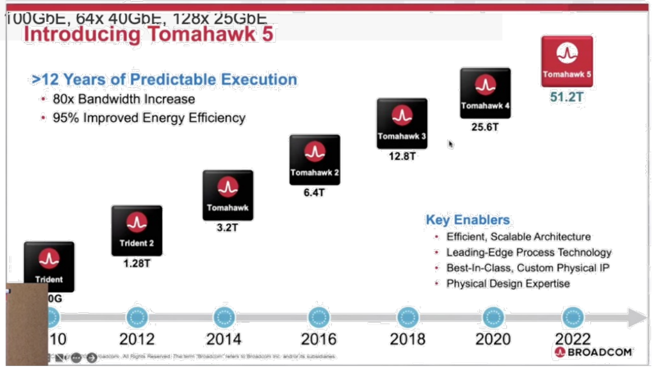
https://youtu.be/g7vcuV2_jSM?si=rkXuTUBihJLfbnoU

Broadcom

- Trident : Feature-rich (programmable for cloud Edge and Enterprise)
- Tomahawk : Hyperscale Fabrics
- Jericho : Scale-out, programmable, Deep buffered, Carrier-grade infrastructure



Product Line	Part Number	Bandwidth	Type	Sample I/O Configurations
StrataXGS® Switch Solutions	Tomahawk 6 / BCM78910 Series	102.4 Tb/s	L2/3 Multilayer Switch	64 x 1.6TbE, 128 x 800GbE, 256 x 400GbE, 512 x 200GbE
StrataXGS® Switch Solutions	Tomahawk Ultra / BCM78920 Series	51.2 Tb/s	L2/3 Multilayer Switch	64 x 800GbE, 128 x 400GbE, 256 x 200GbE
StrataXGS® Switch Solutions	Tomahawk 5 / BCM78900 Series	51.2 Tb/s	L3 Multilayer Switch	64 x 800GbE, 128 x 400GbE, 256 x 200GbE
StrataXGS® Switch Solutions	Tomahawk4 / BCM56990 Series	25.6 Tb/s	L3 Multilayer Switch	64 × 400GbE, 128 × 200GbE, or 256 × 100GbE
StrataXGS® Switch Solutions	Trident4-X11C / BCM56890 Series	12.8 Tb/s	L3 Programmable Ethernet Switch	32x 400GbE, 64x 200GbE, 128x 100GbE
StrataXGS® Switch Solutions	Tomahawk3 / BCM56980 Series	12.8 Tb/s	L3 Ethernet Switch	32x 400GbE, 64x 200GbE, or 128x 100GbE
StrataXGS® Switch Solutions	Trident4 / BCM56880 Series	12.8 Tb/s	L3 Programmable Ethernet Switch	32x 400GbE, 64x 200GbE, 128x 100GbE
StrataXGS® Switch Solutions	Tomahawk2 / BCM56970 Series	6.4 Tb/s	L3 Ethernet Switch	64x 100GbE, 128x 40GbE
StrataXGS® Switch Solutions	Tomahawk / BCM56960 Series	3.2 Tb/s	L3 Ethernet Switch	32x 100GbE, 64x 40GbE, 128x 25GbE
StrataXGS® Switch Solutions	Trident3-X7 / BCM56870 Series	3.2 Tb/s	L3 Programmable Ethernet Switch	32x 100GbE, 64x 40GbE, 128x 25GbE



Introducing Tomahawk 5

>12 Years of Predictable Execution

- 80x Bandwidth Increase
- 95% Improved Energy Efficiency

Key Enablers

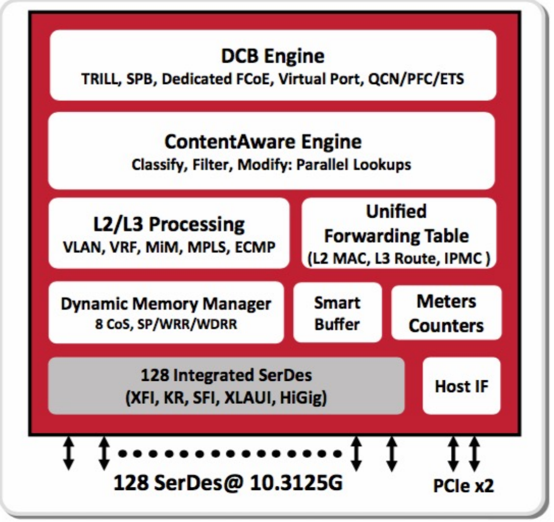
- Efficient, Scalable Architecture
- Leading-Edge Process Technology
- Best-In-Class, Custom Physical IP
- Physical Design Expertise

https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs

Broadcom Trident 2

BROADCOM.

- 1.28 Tbps con puertos 10GE/40GE
- 128 SerDes 10GE (así que un máximo de 32 puertos 40GE en base a 4x10GE)
- Cut-through y Store&Forward
- VXLAN, NVGRE, 802.1Qbg EVR, 802.1BR
- Per VM traffic shaping
- DCB PFC, QCN y ETS. FCoE
- MPLS, VPLS, ISATAP, MAC-in-MAC, TRILL, SPB, Q-in-Q



DCB Engine
TRILL, SPB, Dedicated FCoE, Virtual Port, QCN/PFC/ETS

ContentAware Engine
Classify, Filter, Modify: Parallel Lookups

L2/L3 Processing
VLAN, VRF, MiM, MPLS, ECMP

Unified Forwarding Table
(L2 MAC, L3 Route, IPMC)

Dynamic Memory Manager
8 CoS, SP/WRR/WDRR

Smart Buffer

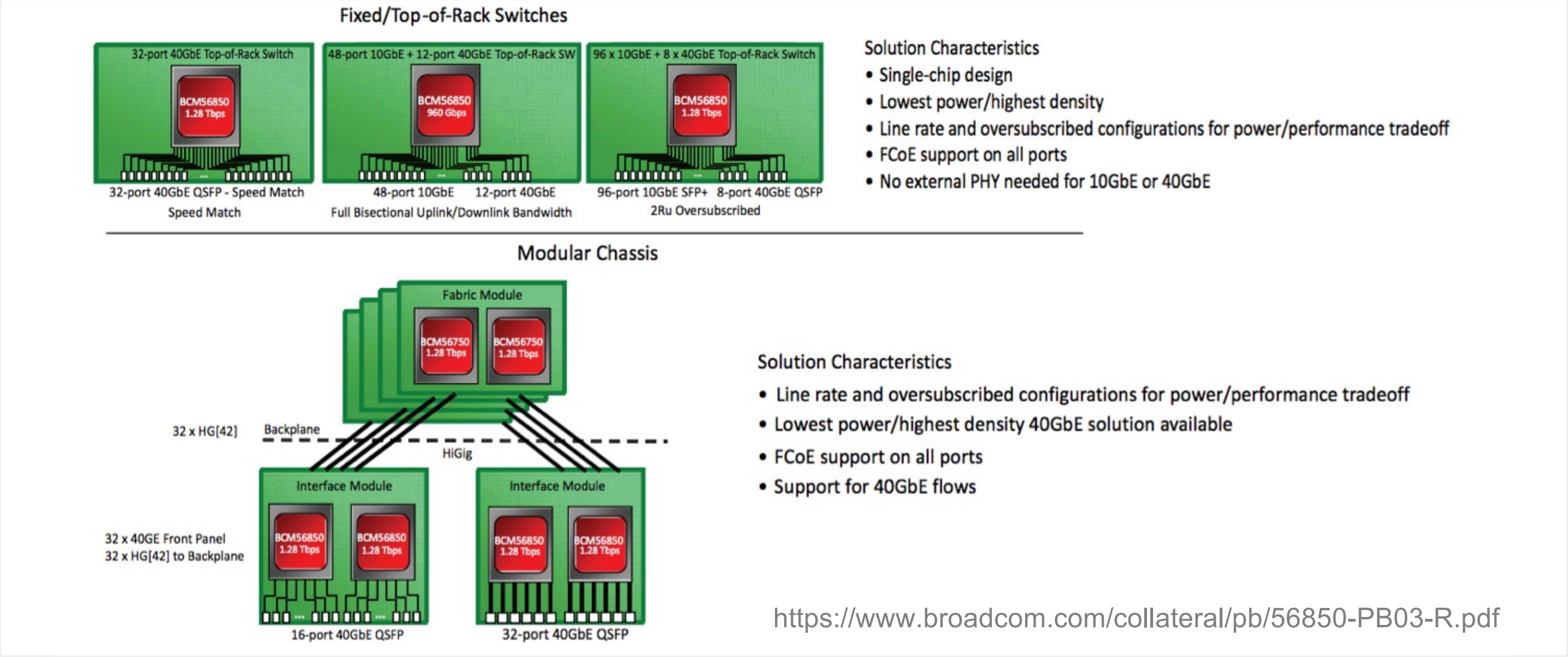
Meters Counters

128 Integrated SerDes
(XFI, KR, SFI, XLAUI, HiGig)

Host IF

128 SerDes@ 10.3125G

PCIe x2



Fixed/Top-of-Rack Switches

32-port 40GbE Top-of-Rack Switch — BCM56850 1.28 Tbps — 32-port 40GbE QSFP - Speed Match
Speed Match

48-port 10GbE + 12-port 40GbE Top-of-Rack SW — BCM56850 960 Gbps — 48-port 10GbE 12-port 40GbE
Full Bisectional Uplink/Downlink Bandwidth

96 x 10GbE + 8 x 40GbE Top-of-Rack Switch — BCM56850 1.28 Tbps — 96-port 10GbE SFP+ 8-port 40GbE QSFP
2Ru Oversubscribed

Solution Characteristics
- Single-chip design
- Lowest power/highest density
- Line rate and oversubscribed configurations for power/performance tradeoff
- FCoE support on all ports
- No external PHY needed for 10GbE or 40GbE

Modular Chassis

Fabric Module — BCM56750 1.28 Tbps — BCM56750 1.28 Tbps

32 x HG[42] Backplane HiGig

Interface Module — BCM56850 1.28 Tbps — BCM56850 1.28 Tbps — 16-port 40GbE QSFP

Interface Module — BCM56850 1.28 Tbps — BCM56850 1.28 Tbps — 32-port 40GbE QSFP

32 x 40GE Front Panel
32 x HG[42] to Backplane

Solution Characteristics
- Line rate and oversubscribed configurations for power/performance tradeoff
- Lowest power/highest density 40GbE solution available
- FCoE support on all ports
- Support for 40GbE flows

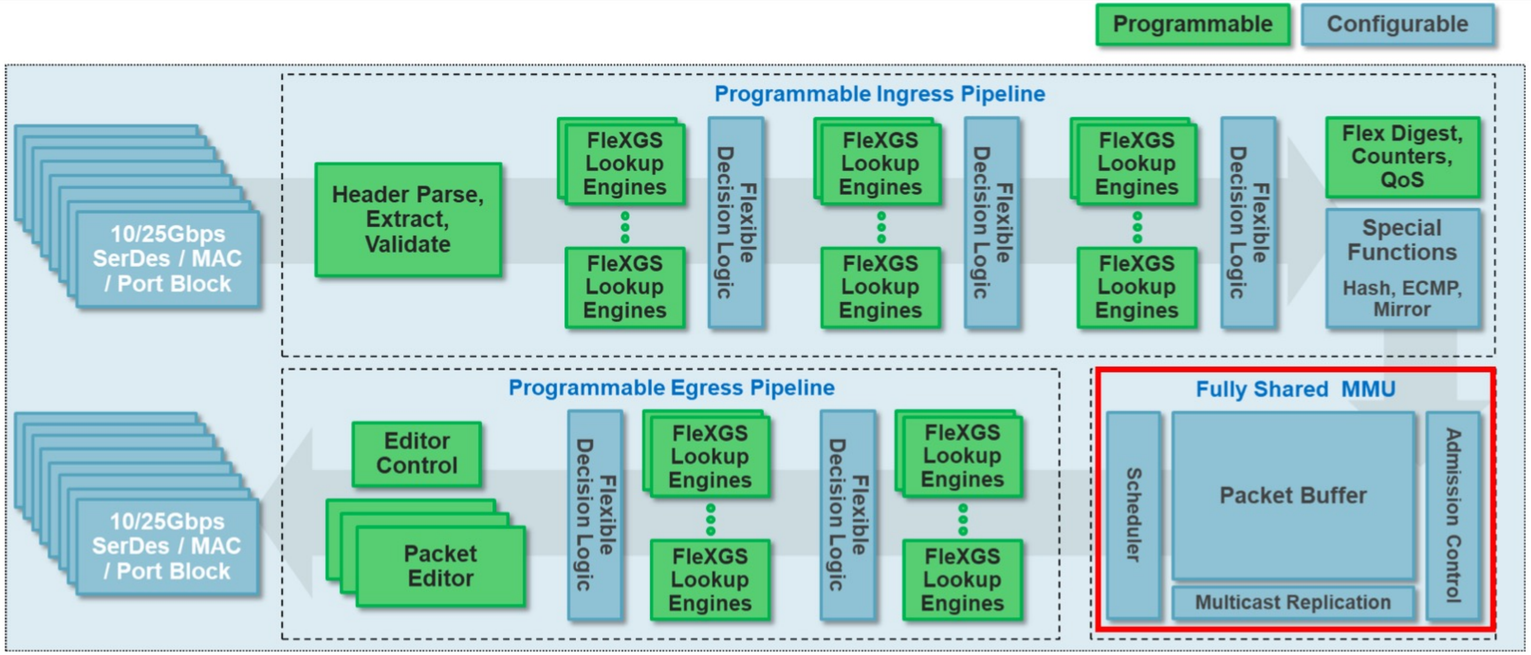
https://www.broadcom.com/collateral/pb/56850-PB03-R.pdf

Broadcom Trident 3

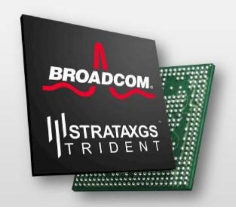
- Conmutación a 3.2 Tbps para paquetes a partir de 250 bytes
- Para paquetes de 64 bytes da un throughput de 2 Tbps
- 32 x 100GE, cada uno divisible en 4x10GE, 4x25GE, 2x50GE o 1x40GE

- 32 MB fully shared packet buffer
- SerDes 25Gbps
- Support for new overlays and tunneling such as GENEVE, NSH, VXLAN, GPE, MPLS, MPLS over GRE/UDP, GUE, ILA and PPPoE

Figure 2. Broadcom Trident 3 Switch Silicon Internal Architecture Diagram



Source: Enterprise Strategy Group



https://docs.broadcom.com/doc/12395356

https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs/bcm56870-series

Trident 2 y Trident 3

- Memoria (SRAM, TCAM) particionable para diferentes usos del switch (muchas MACs, muchas rutas IPv4, etc)

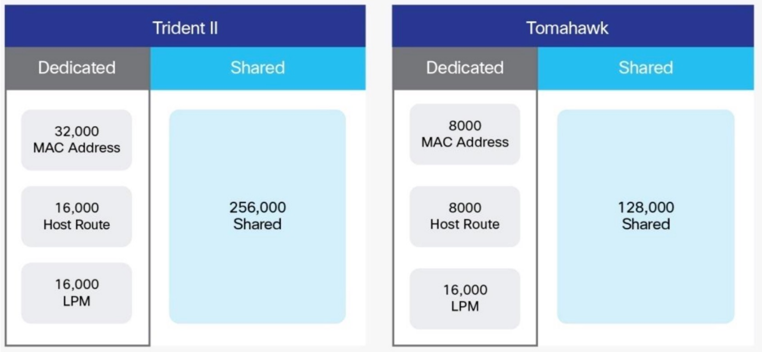
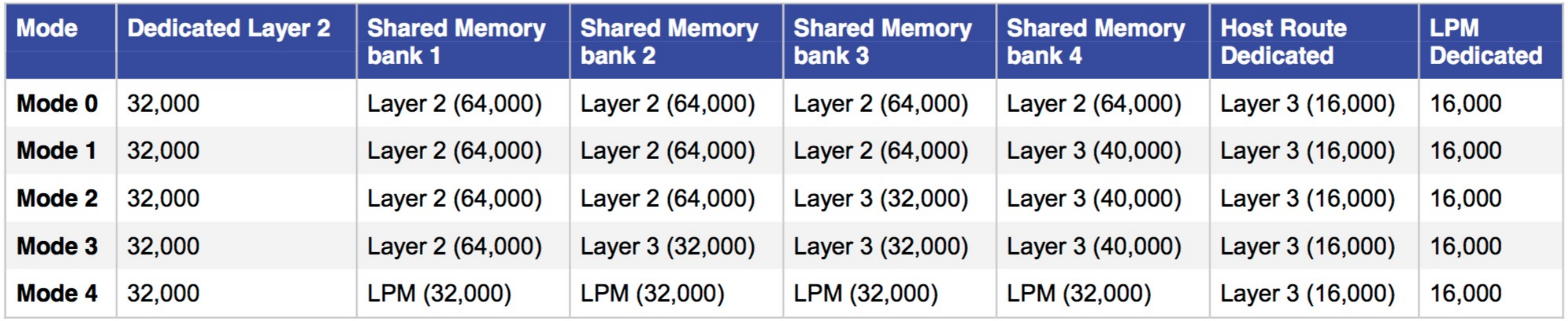
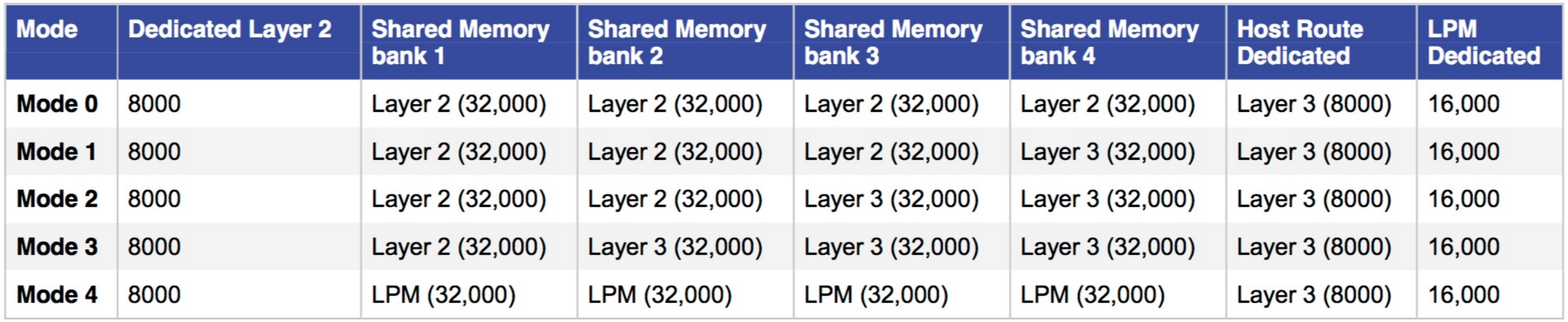


Table 1. Broadcom Trident 2 Forwarding Tables



Mode	Dedicated Layer 2	Shared Memory bank 1	Shared Memory bank 2	Shared Memory bank 3	Shared Memory bank 4	Host Route Dedicated	LPM Dedicated
Mode 0	32,000	Layer 2 (64,000)	Layer 2 (64,000)	Layer 2 (64,000)	Layer 2 (64,000)	Layer 3 (16,000)	16,000
Mode 1	32,000	Layer 2 (64,000)	Layer 2 (64,000)	Layer 2 (64,000)	Layer 3 (40,000)	Layer 3 (16,000)	16,000
Mode 2	32,000	Layer 2 (64,000)	Layer 2 (64,000)	Layer 3 (32,000)	Layer 3 (40,000)	Layer 3 (16,000)	16,000
Mode 3	32,000	Layer 2 (64,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (40,000)	Layer 3 (16,000)	16,000
Mode 4	32,000	LPM (32,000)	LPM (32,000)	LPM (32,000)	LPM (32,000)	Layer 3 (16,000)	16,000

Table 2. Broadcom Tomahawk Forwarding Tables

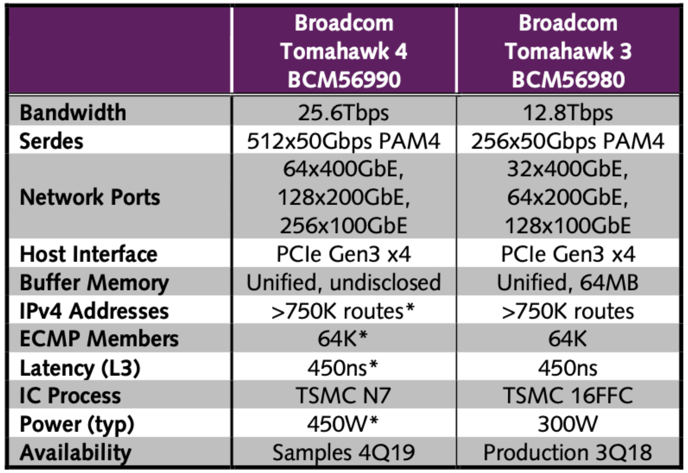


Mode	Dedicated Layer 2	Shared Memory bank 1	Shared Memory bank 2	Shared Memory bank 3	Shared Memory bank 4	Host Route Dedicated	LPM Dedicated
Mode 0	8000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (8000)	16,000
Mode 1	8000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (8000)	16,000
Mode 2	8000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (8000)	16,000
Mode 3	8000	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (8000)	16,000
Mode 4	8000	LPM (32,000)	LPM (32,000)	LPM (32,000)	LPM (32,000)	Layer 3 (8000)	16,000

http://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/white-paper-c11-736863.pdf

Tomahawk 4

- 2.5 años de desarrollo
- 25.6 Tb/s
- 7nm, 31.000 millones de transistores
- 8.000 pins
- 512 x 50G PAM-4 SerDes
- 4 cores ARM 1GHz (para telemetría)



	Broadcom Tomahawk 4 BCM56990	Broadcom Tomahawk 3 BCM56980
Bandwidth	25.6Tbps	12.8Tbps
Serdes	512x50Gbps PAM4	256x50Gbps PAM4
Network Ports	64x400GbE, 128x200GbE, 256x100GbE	32x400GbE, 64x200GbE, 128x100GbE
Host Interface	PCIe Gen3 x4	PCIe Gen3 x4
Buffer Memory	Unified, undisclosed	Unified, 64MB
IPv4 Addresses	>750K routes*	>750K routes
ECMP Members	64K*	64K
Latency (L3)	450ns*	450ns
IC Process	TSMC N7	TSMC 16FFC
Power (typ)	450W*	300W
Availability	Samples 4Q19	Production 3Q18

Table 1. Comparison of Tomahawk switch generations. Externally, Tomahawk 4 is nearly identical to its predecessor but has twice the port count. (Source: Broadcom, except *The Linley Group estimate)

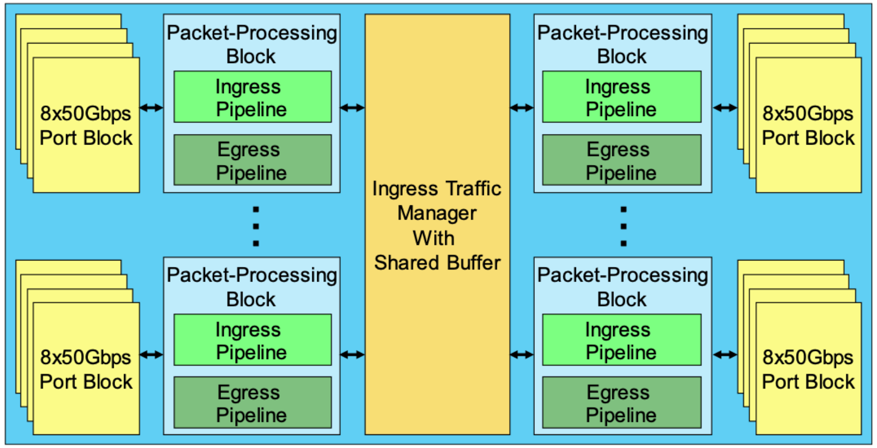
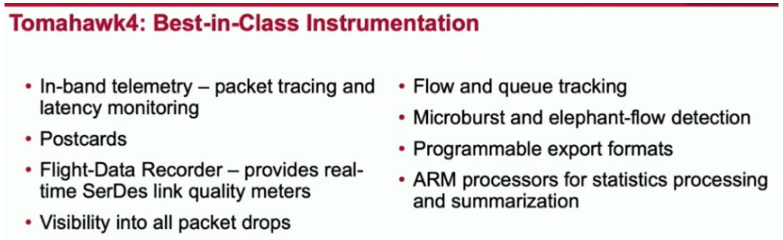


Figure 1. Tomahawk 4 switch chip. The top-level architecture carries over from Tomahawk 3, but the new chip instantiates 64 port blocks, each with eight serdes.

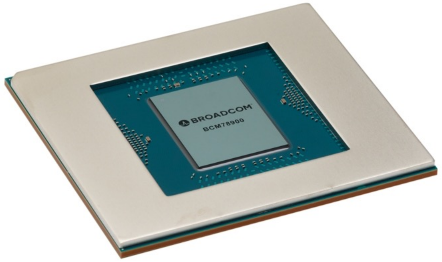


Tomahawk4: Best-in-Class Instrumentation

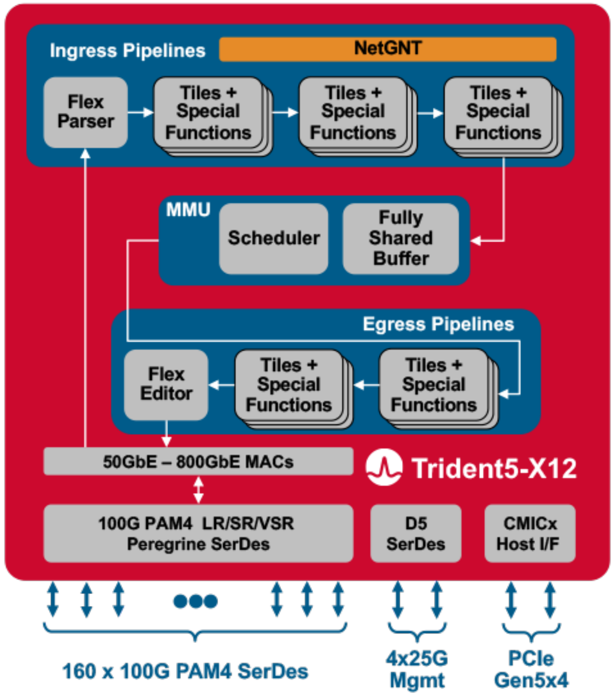
- In-band telemetry – packet tracing and latency monitoring
- Postcards
- Flight-Data Recorder – provides real-time SerDes link quality meters
- Visibility into all packet drops
- Flow and queue tracking
- Microburst and elephant-flow detection
- Programmable export formats
- ARM processors for statistics processing and summarization

https://docs.broadcom.com/doc/12398014
https://youtu.be/B-COGMbaUg4

Tomahawk 5



- Up to 51.2 Tb/s on a single chip
- 64 × 800GbE, 128 × 400GbE, or 256 × 200GbE
- 64 integrated SerDes cores, each with eight integrated 106-Gb/s PAM4
- L2 and L3 switching, routing, and tunneling
- Support for Clos and non-Clos topologies such as torus, Dragonfly, Dragonfly+, and Megafly
- 9.352 pins
- Telemetría programable (6 cores ARM)



Ingress Pipelines

NetGNT

Flex Parser

Tiles + Special Functions

Tiles + Special Functions

Tiles + Special Functions

MMU

Scheduler

Fully Shared Buffer

Egress Pipelines

Flex Editor

Tiles + Special Functions

Tiles + Special Functions

50GbE – 800GbE MACs

Trident5-X12

100G PAM4 LR/SR/VSR Peregrine SerDes

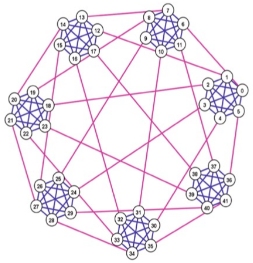
D5 SerDes

CMICx Host I/F

160 x 100G PAM4 SerDes

4x25G Mgmt

PCIe Gen5x4



(a) "Canonical" Dragonfly with $a = 6$, $g = 7$, $h = 1$.

Teh, M.Y., Wilke, J.J., Bergman, K., Rumley, S. (2017). Design Space Exploration of the Dragonfly Topology. In: Kunkel, J., Yokota, R., Taufer, M., Shalf, J. (eds) High Performance Computing. ISC High Performance 2017. Lecture Notes in Computer Science(), vol 10524. Springer, Cham. https://doi.org/10.1007/978-3-319-67630-2_5

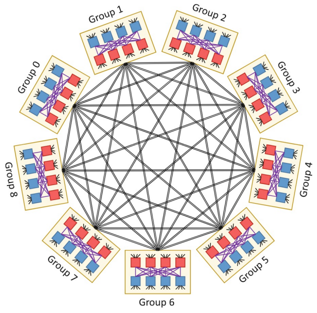
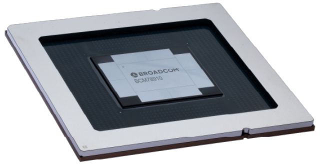


Fig. 4. Example Megafly topology.

Flajslik, M., Borch, E., Parker, M.A. (2018). Megafly: A Topology for Exascale Systems. In: Yokota, R., Weiland, M., Keyes, D., Trinitis, C. (eds) High Performance Computing. ISC High Performance 2018. Lecture Notes in Computer Science(), vol 10876. Springer, Cham. https://doi.org/10.1007/978-3-319-92040-5_15

Tomahawk 6

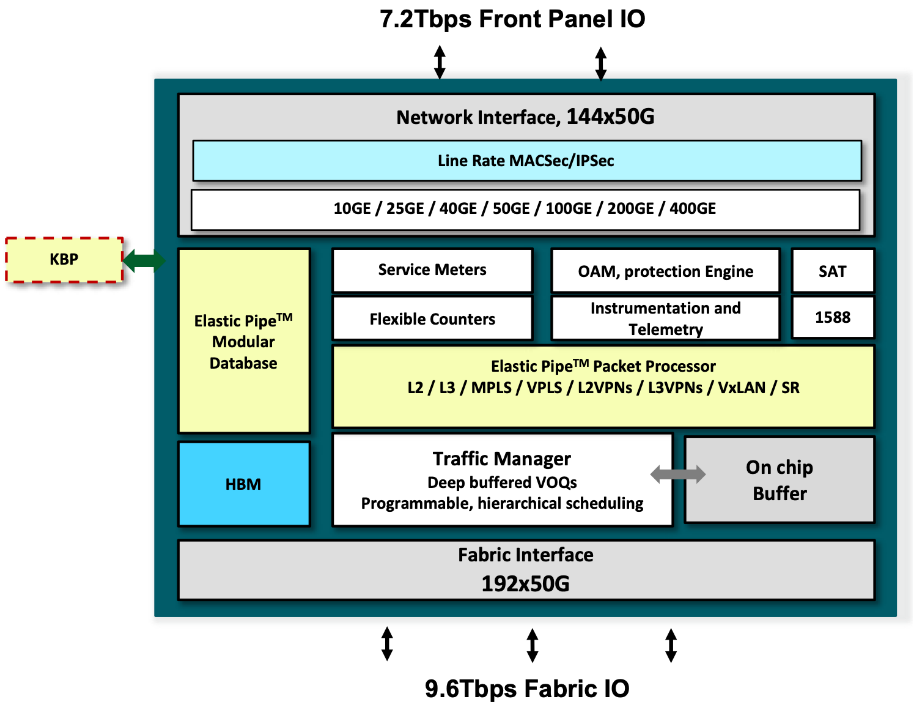


- Up to 1.6TbE, 128 × 800GbE, 256 × 400GbE, and 512 × 200GbE ports
- 128 x 106.25G PAM4 SerDes cores or 64 x 212.5G PAM4 SerDes
- up to 102.4 Tb/s on a single chip

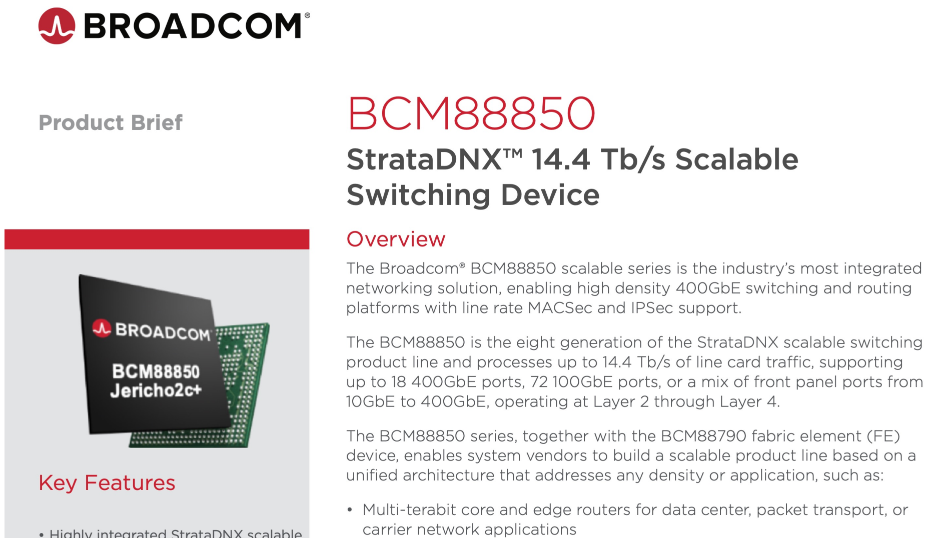
Jericho2

- Jericho2c+ : 14.4Tb/s
- Hasta 18 puertos 100GE
- High Bandwidth Memory (HBM): 8GB, en el mismo encapsulado
- Pipeline reprogramable (C++)
- Jericho2 + Ramon: permite construir single-stage system 900Tb/s

Figure 1: BCM88850 Block Diagram


7.2Tbps Front Panel IO
Network Interface, 144x50G
Line Rate MACSec/IPSec
10GE / 25GE / 40GE / 50GE / 100GE / 200GE / 400GE
KBP
Elastic Pipe™ Modular Database
Service Meters
OAM, protection Engine
SAT
Flexible Counters
Instrumentation and Telemetry
1588
Elastic Pipe™ Packet Processor
L2 / L3 / MPLS / VPLS / L2VPNs / L3VPNs / VxLAN / SR
HBM
Traffic Manager
Deep buffered VOQs
Programmable, hierarchical scheduling
On chip Buffer
Fabric Interface
192x50G
9.6Tbps Fabric IO

https://docs.broadcom.com/doc/88850-PB1-PUB



BROADCOM®

Product Brief

BCM88850

StrataDNX™ 14.4 Tb/s Scalable Switching Device

Overview

The Broadcom® BCM88850 scalable series is the industry's most integrated networking solution, enabling high density 400GbE switching and routing platforms with line rate MACSec and IPSec support.

The BCM88850 is the eight generation of the StrataDNX scalable switching product line and processes up to 14.4 Tb/s of line card traffic, supporting up to 18 400GbE ports, 72 100GbE ports, or a mix of front panel ports from 10GbE to 400GbE, operating at Layer 2 through Layer 4.

The BCM88850 series, together with the BCM88790 fabric element (FE) device, enables system vendors to build a scalable product line based on a unified architecture that addresses any density or application, such as:

- Multi-terabit core and edge routers for data center, packet transport, or carrier network applications

Key Features

- Highly integrated StrataDNX scalable

Jericho2 : Ejemplos

- Arista 7280CR3-32P4

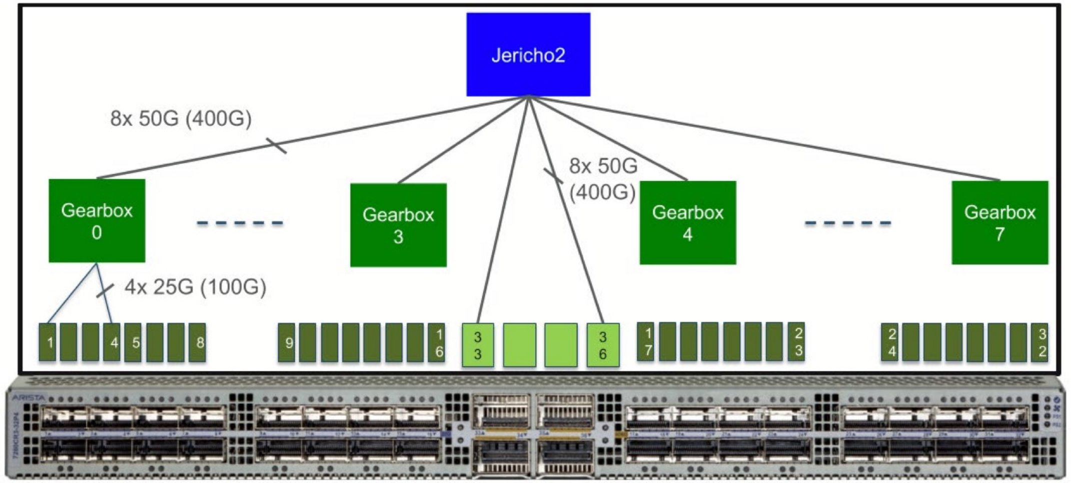
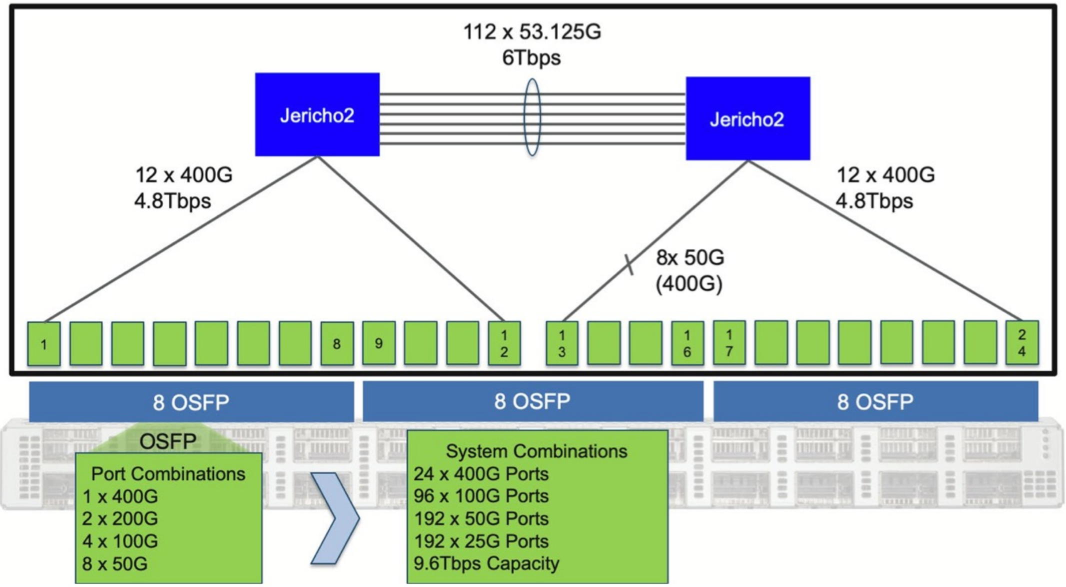


Figure 6: Arista 7280CR3-32P4 Switch Architecture

- 7280PR23-24



Port Combinations
1 x 400G
2 x 200G
4 x 100G
8 x 50G

System Combinations
24 x 400G Ports
96 x 100G Ports
192 x 50G Ports
192 x 25G Ports
9.6Tbps Capacity

Figure 11: 7280PR3-24

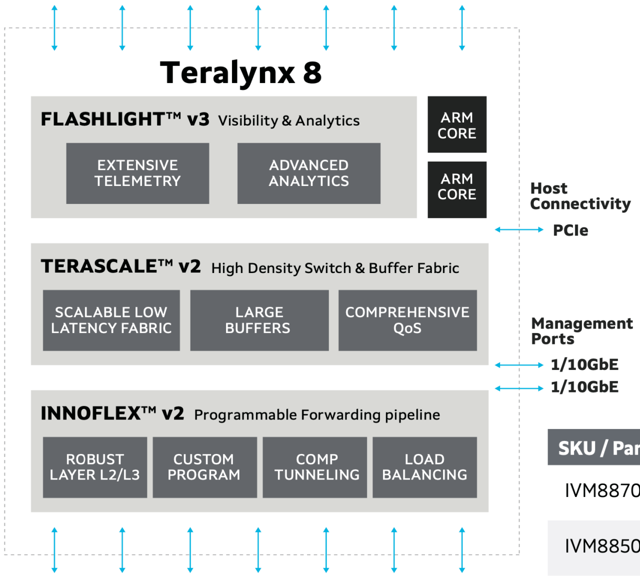
https://www.arista.com/assets/data/pdf/Whitepapers/7280R3-Platform-Architecture-WP.pdf

Jericho3AI

- 28.8 Tb/s
- 144 SerDes @ 106-Gb/s PAM4
- Up to 18 × 800GbE, 36 × 400GbE, or 72 × 200GbE
- Support for 25GE, 50GE, 100GE, 200GE, 400GE, 800GE Ethernet port interfaces
- *“Hierarchical traffic manager, scalable packet buffer memory and low latency forwarding”*
- *“The BCM88890, combined with the BCM88920 (Ramon3), is designed to meet the unique requirements for next-generation Artificial Intelligence / Machine Learning (AI/ML) routed networks.”*
- *“... using the BCM88920 two-stage fabric to create a scalable core platform that delivers up to 25,000 ports of 800GbE”*

Marvell Teralynx8

Features	Benefits
· 25.6 Tbps throughput with 256x 112 Gbps SerDes	· Industry leading performance and scale enables customers to deploy fewer network switches and tiers dramatically reducing cost, power, latency & management
· Comprehensive IP forwarding and highly scalable/flexible layer 2 and 3 tables for IPv4, IPv6 and hybrid networks	· Proven, innovative & highly scalable architecture delivers 64 x 400Gbe, 128 x 200G and 256 x 100GbE ports
· Line-rate programmability to accommodate future networking protocols with software upgrades	· 100G LR SerDes enables higher scale IO with backward compatibility to 50G PAM4 & 10/25G NRZ
· Extensive tunneling capabilities such as IP-in-IP, GRE, MPLS, VXLAN and Geneve	· Breakthrough visibility and analytics capabilities enable predictive, faster & more accurate issue resolution, higher automation and self-healing autonomous networks
· Very low latencies - cut-through and store-and-forward	· Superior power efficiency enables customers to design 1RU 32 x 800G switches for best power and cost per bit
· Advanced QoS/traffic management feature set such as DCB, RDMA/RoCE	· InnoFlex™ programmable forwarding pipeline enables support of custom & new standard protocols without requiring ASIC spins to future proof the network
· FLASHLIGHT™ v3 innovations delivers breakthrough visibility and telemetry addressing Cloud customer requirements	



Teralynx 8

FLASHLIGHT™ v3 Visibility & Analytics

EXTENSIVE TELEMETRY

ADVANCED ANALYTICS

ARM CORE

ARM CORE

Host Connectivity

PCIe

TERASCALE™ v2 High Density Switch & Buffer Fabric

SCALABLE LOW LATENCY FABRIC

LARGE BUFFERS

COMPREHENSIVE QoS

Management Ports

1/10GbE

1/10GbE

INNOFLEX™ v2 Programmable Forwarding pipeline

ROBUST LAYER L2/L3

CUSTOM PROGRAM

COMP TUNNELING

LOAD BALANCING

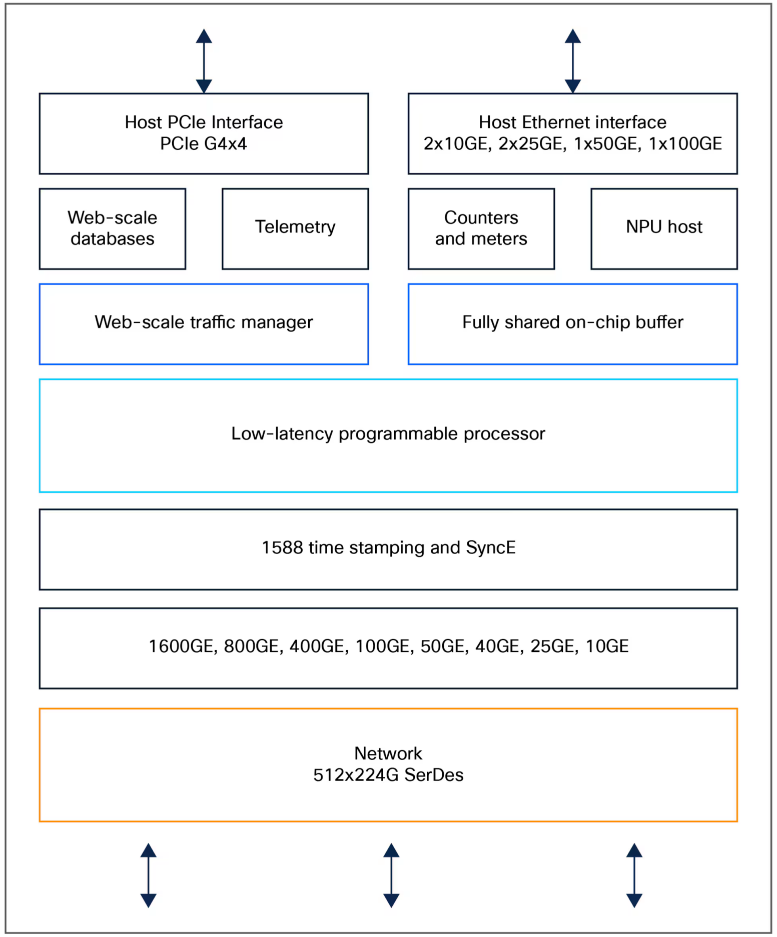
Flexible SerDes: 25G NRZ & 50/100G PAM4

SKU / Part Number	Capacity (Tbps)	# of SerDes @ Gbps	10/25/50G ports	100G ports	200G ports	400G ports
IVM88700	25.6	256@ 10/25/50/100	256	256	128	64
IVM88500	12.8	256 @ 10/25/50	128	128	64	32
IVM88300	8.0	160 @ 10/25/50	80	80	40	20

https://www.marvell.com/content/dam/marvell/en/public-collateral/switching/marvell-teralynx-8-product-brief.pdf

Cisco Silicon One G300

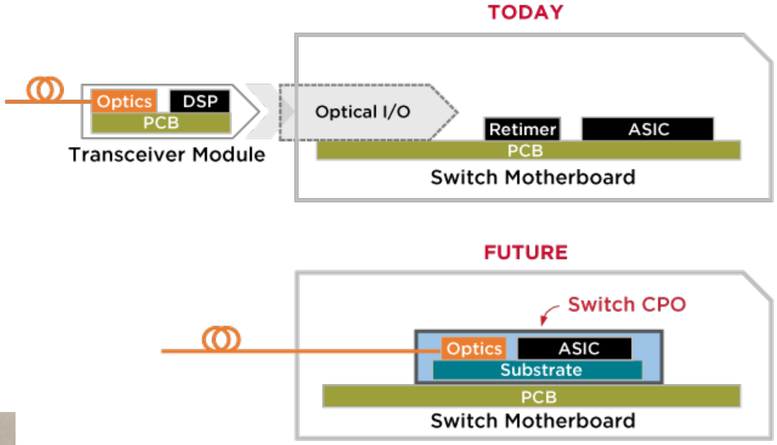
- 102.4 Tb/s
- 512 x 224Gb/s SerDes (hasta 1.6 Tb/s Ethernet)
- 802.1Qbb PFC, ECN, WRED
- Dynamic Load Balancing
- P4-INT (in-band telemetry)

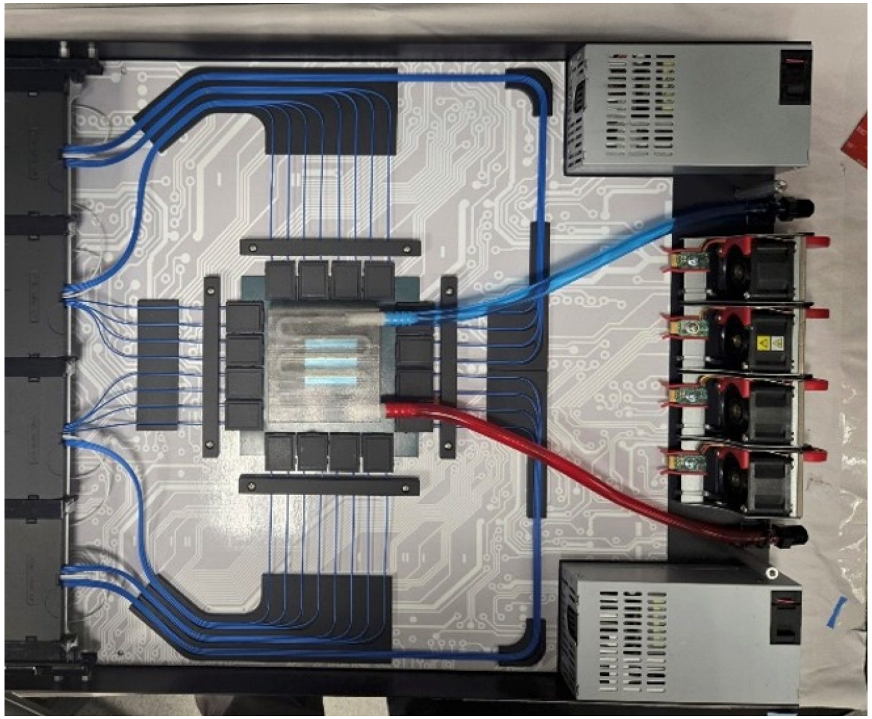

Host PCIe Interface
PCIe G4x4
Host Ethernet interface
2x10GE, 2x25GE, 1x50GE, 1x100GE
Web-scale databases
Telemetry
Counters and meters
NPU host
Web-scale traffic manager
Fully shared on-chip buffer
Low-latency programmable processor
1588 time stamping and SyncE
1600GE, 800GE, 400GE, 100GE, 50GE, 40GE, 25GE, 10GE
Network
512x224G SerDes

https://www.cisco.com/c/en/us/solutions/collateral/silicon-one/silicon-one-g300-ds.html

CPO

- Co-Packaged Optics
- Optical transceiver en el ASIC





https://www.corning.com/optical-communications/worldwide/en/home/the-signal-network-blog/what-is-co-packaged-optics.html

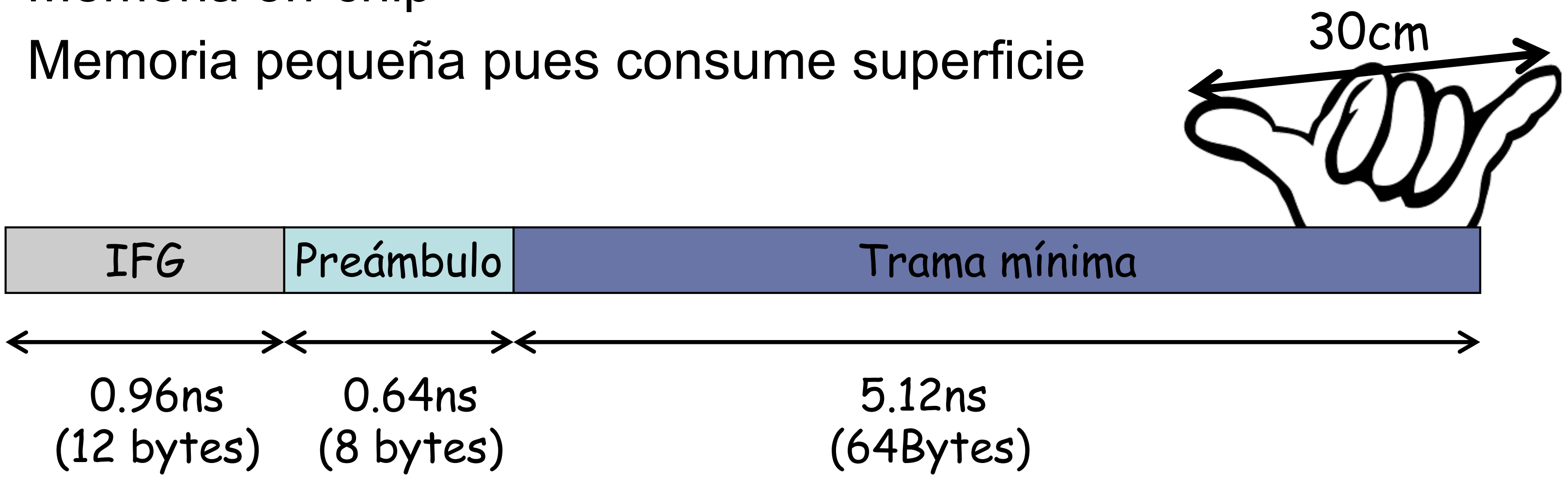
Buffering

Shallow vs Deep buffers

- Alternativas tradicionales:
  - High bandwidth switches + small buffers
  - Low bandwidth routers + large buffers
- Arquitecturas que siempre almacenan los paquetes en buffer
- Pero en realidad buffers necesarios solo ante congestión
- Congestión debería ser de breve duración
- ¿Quieres un buffer de 1GByte?
- Ejemplo:
  - Llegan flujos por 4 puertos a 1Gb/s hacia el mismo puerto a 1Gb/s
  - Llega hasta 4x lo que puede salir
  - Si se llena el buffer, el último paquete debe esperar a ser transmitido lo que tarde en transmitirse 1GByte a 1Gb/s ...
  - ¡¡ 8 segundos !!
  - Si era un flujo TCP es probable que ya haya caducado el timer de retx
  - Pero tal vez 1GByte de buffer compartido para 128 puertos a 800Gb/s no es tanto problema
  - 8Gbits / 128 puertos / 800Gb/s = 78 μs

Shallow vs Deep buffers

- 40Gb/s, 100Gb/s, 400Gb/s
- A 100Gb/s la trama de 64bytes tarda 6.4ns
- Las memorias más rápidas responden en el rango de 1-2ns
- Pero de nuevo eso es solo con un puerto, con varios puertos hay que poder atender a varias peticiones
- La limitación es más seria pues la luz recorre en un 1ns…
- $3x10^{8}$m/s x 10^{-9}s/ns = 0.3m/ns = 30cm/ns
- ¡ Si tenemos la memoria a 15cm del procesador tarda ya 1ns una señal en viajar de uno al otro y volver !
- Memoria on-chip
- Memoria pequeña pues consume superficie



Deep buffers

- Hoy en día existe la opción de “High Bandwitdh Memory” (HBM)

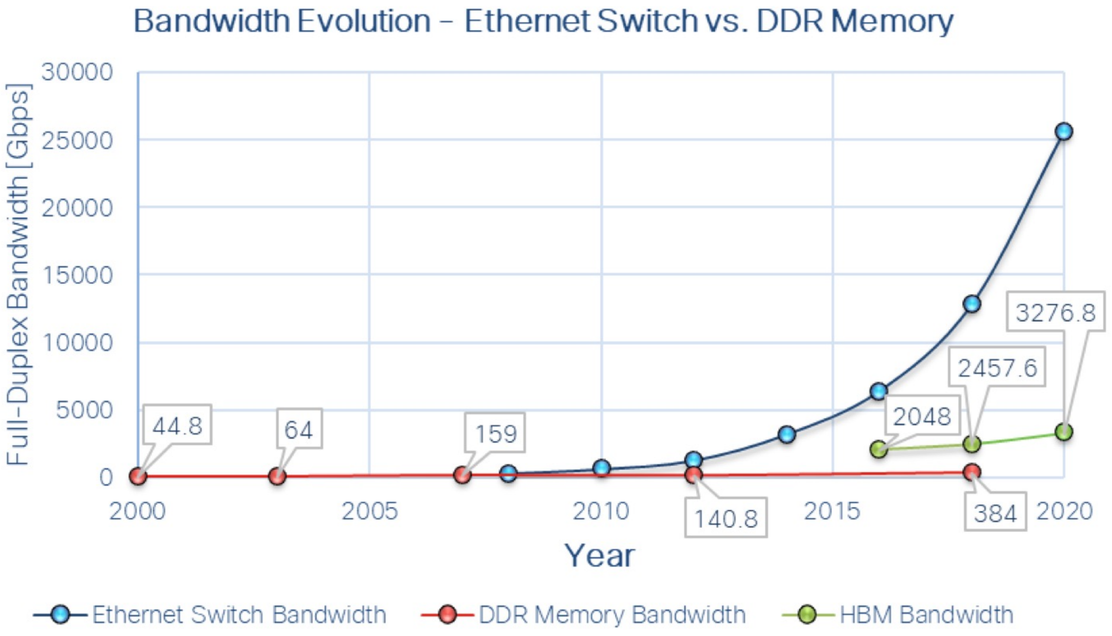


Figure 1. Bandwidth growth in DDR memories and ethernet switches

https://www.cisco.com/c/dam/en/us/solutions/collateral/silicon-one/white-paper-sp-hybrid-buffer-architecture.pdf

HBM

- Mejor interconexión con el ASIC de conmutación
- Permite GBs de packet buffer
- Sigue siendo más lenta que la on-chip

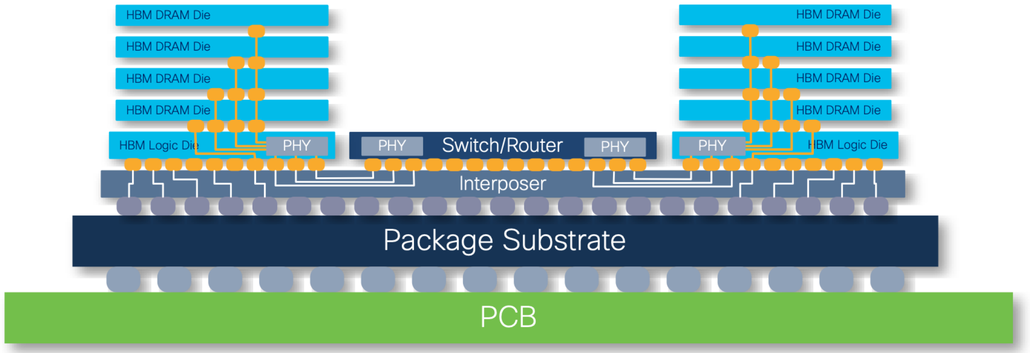


Figure 2. HBM - In-package memory

- El switch debe seleccionar los flujos que envía al buffer externo
- Puede haber API para controlarlo

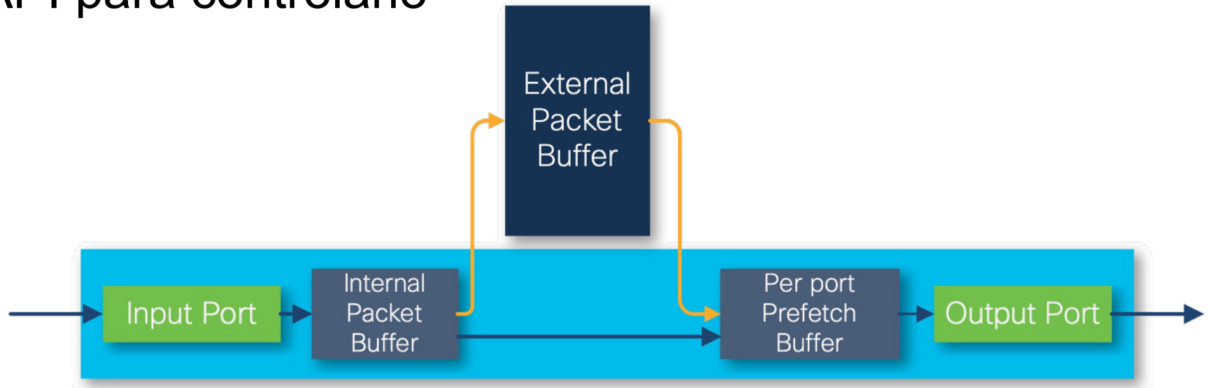


Figure 4. Hybrid memory architecture

https://www.cisco.com/c/dam/en/us/solutions/collateral/silicon-one/white-paper-sp-hybrid-buffer-architecture.pdf

Arquitectura de conmutadores Switching ASICs

Switches modulares

- Hemos comentado lo que se puede hacer con un ASIC
- Tiene limitaciones
- ¿Y switches más grandes? Por encima del límite de un ASIC...