

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Redes de Nueva Generación
Área de Ingeniería Telemática



NFV



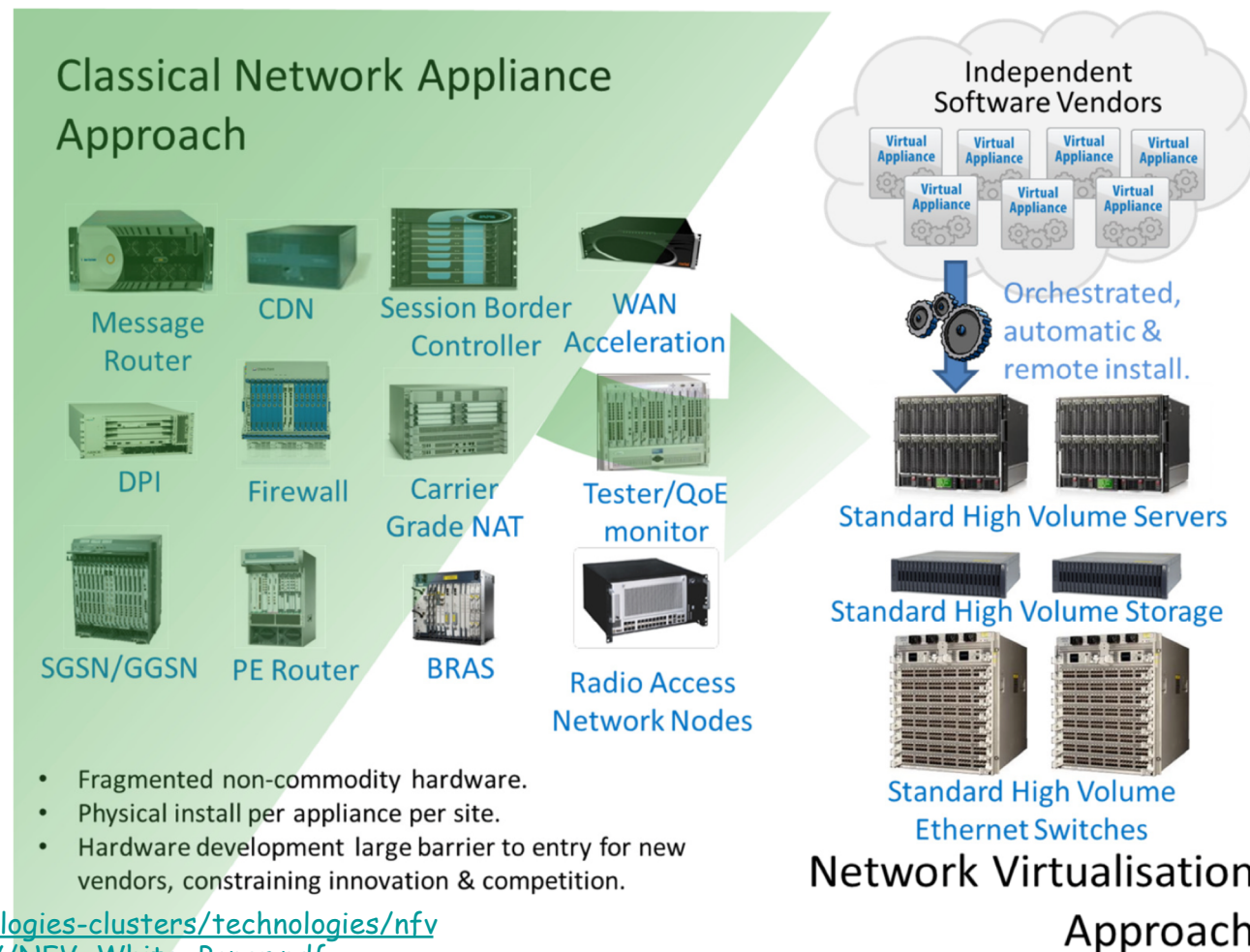
El problema

- Problema de las operadoras
- Gran cantidad de *appliances*
- Desplegar un nuevo servicio requiere espacio y alimentación para ese nuevo hardware
- Nuevas habilidades de la gente para diseñar, integrar y operar el servicio con ese nuevo hardware
- Ese hardware alcanza su límite de vida con rapidez, lo cual requiere políticas de remplazo que no crean nuevo beneficio
- Los operadores declaran no estar incrementando sus beneficios pero aumentan sus costes (más tráfico, más servicios)



NFV

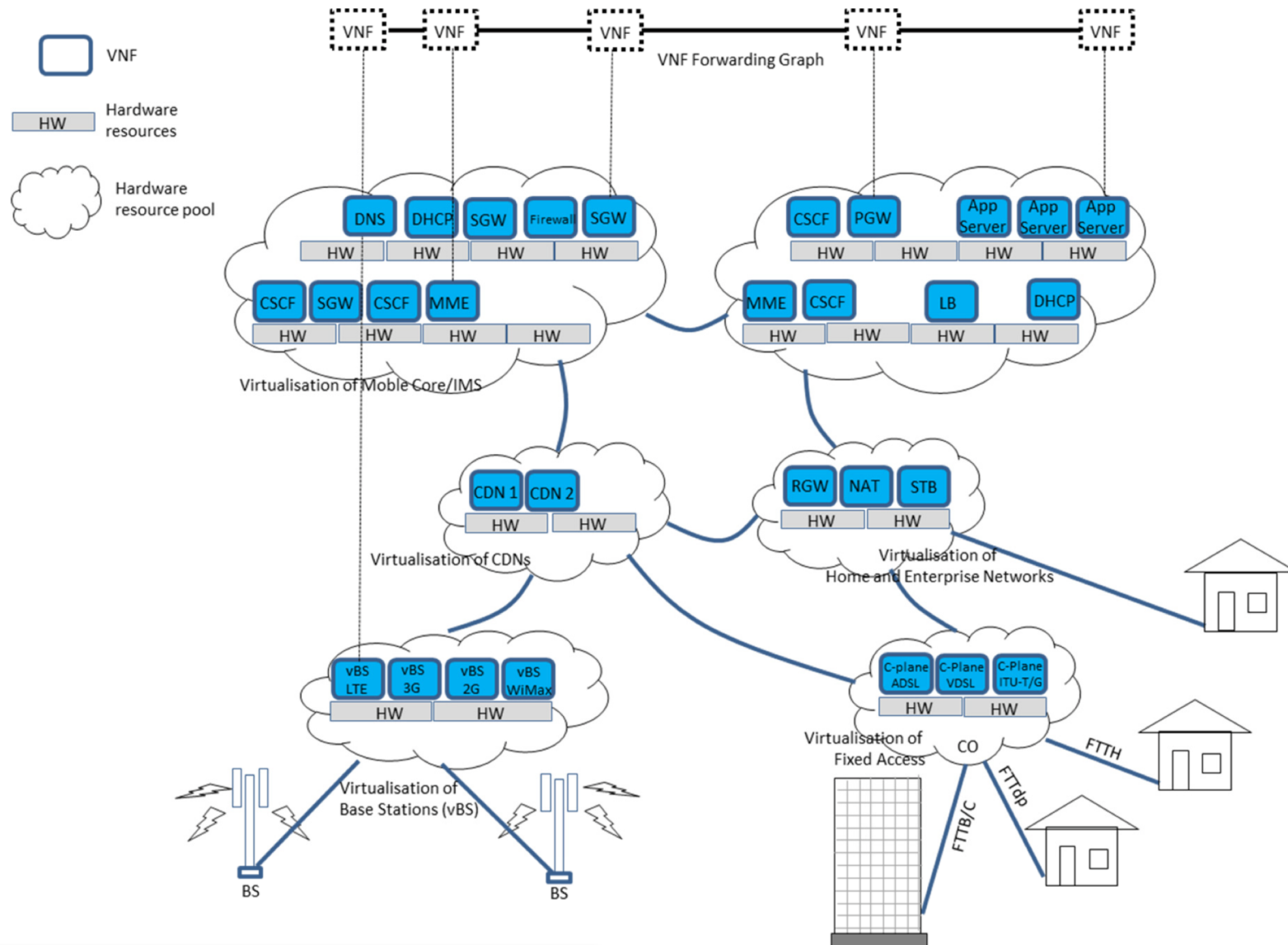
- *Network Functions Virtualisation* (complementario a SDN)
- Se busca mover de hardware dedicado a máquinas virtuales
- Un ISG (*Industry Specification Group*) de ETSI desde finales de 2012
- Hoy más de 200 compañías



Use cases

- Switching elements: BNG, CG-NAT, routers.
- Mobile network nodes: HLR/HSS, MME, SGSN, GGSN/PDN-GW, RNC, Node B, eNode B.
- Functions contained in home routers and set top boxes to create virtualised home environments.
- Tunnelling gateway elements: IPSec/SSL VPN gateways.
- Traffic analysis: DPI, QoE measurement.
- Service Assurance, SLA monitoring, Test and Diagnostics.
- NGN signalling: SBCs, IMS.
- Converged and network-wide functions: AAA servers, policy control and charging platforms.
- Application-level optimisation: CDNs, Cache Servers, Load Balancers, Application Accelerators.
- Security functions: Firewalls, virus scanners, intrusion detection systems, spam protection.

Ejemplos



VNF = Virtualised Network Function

Algunos beneficios

- Reducción de coste de equipos
- Reducción de consumo eléctrico
- Reducción de tiempo de despliegue de un nuevo servicio
- Posibilidad de tener servicios en producción, prueba y desarrollo en la misma infraestructura
- Escalado rápido del servicio
- Abre el mercado a desarrolladores de soft (no necesitan desarrollar hardware)
- Multi-tenancy
- Mejores habilidades existentes para la gestión de infraestructura IT de gran escala que de equipos de red
- Reducción de tiempos de reparación
- Reducción de tiempos de actualización de software
- Etc etc



Facilitadores

- *Cloud Computing*

- Virtualización (hypervisores, vSwitch, smart NICs)
- *Orchestration*
- Open APIs



- Grandes volúmenes de servidores

- Componentes estándar (por ejemplo x86), vendidos por millones (escala) e intercambiables (competencia)
- En lugar de *appliances* que dependen de ASICs



Ejemplo: B4N CG-NAT

B4N CG-NAT SPECIFICATIONS

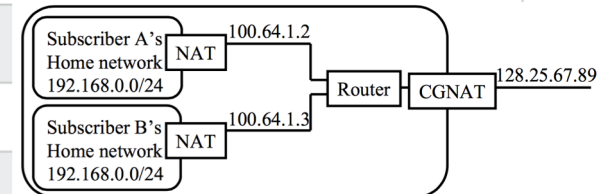
WHAT IS B4N CG-NAT?

B4N CG-NAT is an NFV-based solution designed to provide high performance and transparent address and protocol translation. B4N CG-NAT helps Service Providers to mitigate IPv4 address exhaustion by using address and port translation in large scale and provides native integration within existent operators' infrastructure.

B4N CG-NAT is a fully virtualized and SDN ready solution that utilizes commodity x86 servers and provides carrier grade performance by using Intel® Data Plane Development Kit libraries.

CG-NAT solution provides maximum **500Gbps** throughput performance and fully compliant with **ETSI NFV ISG architecture**

	CONF.10	CONF.50	CONF.500	DISTRIBUTED
Max Throughput	10 Gbps	50 Gbps	500 Gbps	Unlimited
Connections per Second	200K per 10Gbps			Depends on OpenFlow switches performance, but not less than BOXED
Two-way sessions	10M per 10Gbps			
Resiliency	N+1. Active-Active, Active-Standby			
Supported protocols	NAT44 PCP			
Interfaces	REST API NETCONF			
Management	WEB CLI			
Supported hypervisors	LXC (Linux Containers) KVM VMware			
Logging	Local or external SYSLOG Server			



SCALABILITY

Simple extend capacity and performance by adding new B4N NAT VNFs and Distributed Switches, while maintaining existent network architecture



COMMODITY HARDWARE

Using commodity x86 servers instead of dedicated hardware devices



UNIFIED MANAGEMENT

Single point of management through powerful WEB-interface



AUTOMATION

B4N CG-NAT provides set of tools for automate service management



CONFORMANCE WITH REFERENCE ARCHITECTURE

Fully compliant with MANO Framework. B4N CG-NAT includes VNF-manager that can be integrated with Customer orchestration and management system.



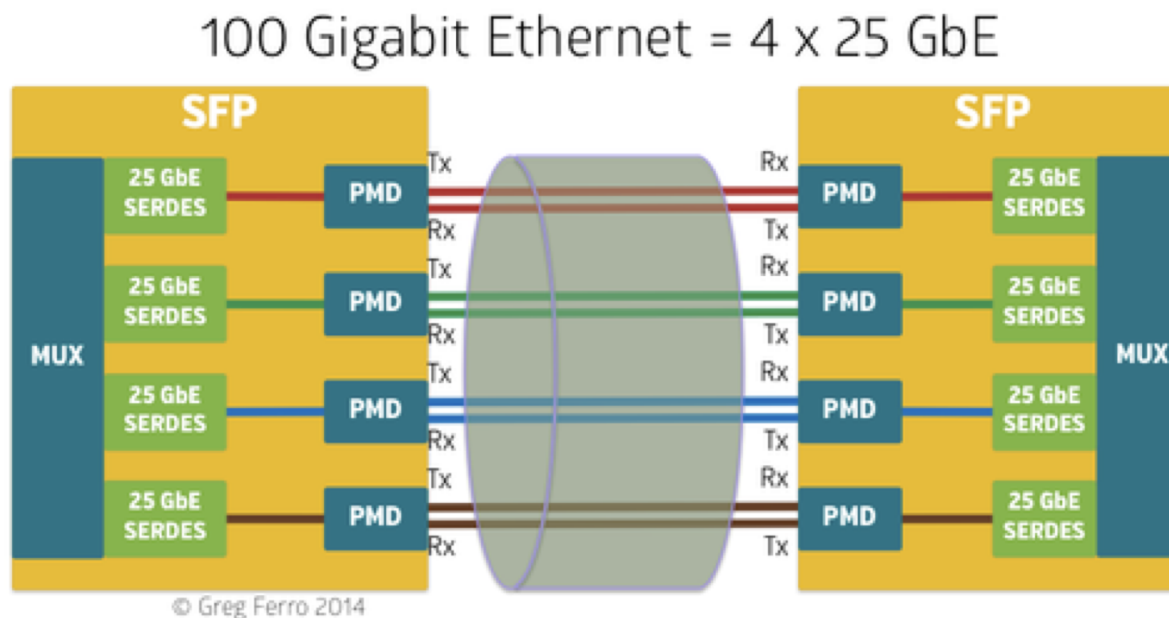
SDN READY

B4N CG-NAT designed to be easy integrated with Customer SDN infrastructure

Networking hardware y el software

Evolución

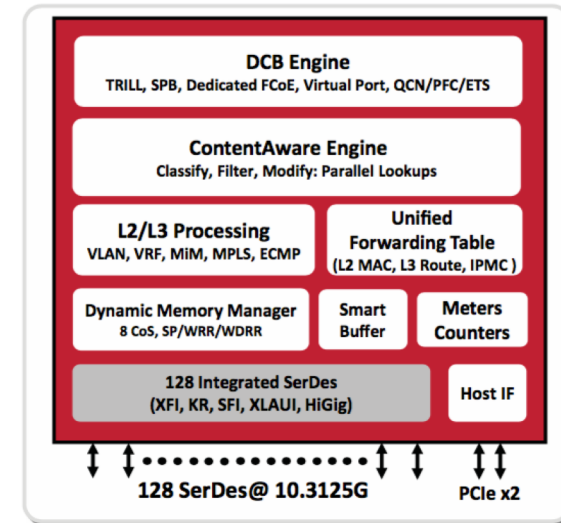
- Los fabricantes de equipos de red están adoptando los ritmos de producción de electrónica
- Empujados por pocos grandes clientes
- Por ejemplo: donde teníamos SerDes a 10Gbps los tendremos este año a 25Gbps, al mismo coste
- Esto permite interfaces 100GE donde antes teníamos 40GE, al mismo precio
- A día de hoy SoC (Switch on Chip) a 3.2Tbps



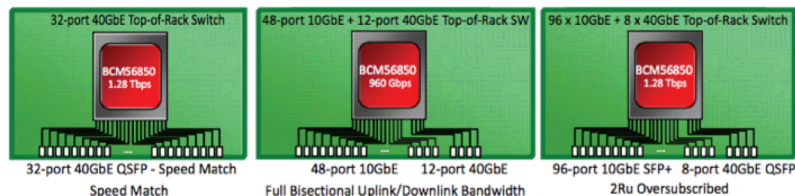
Broadcom Trident 2



- 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



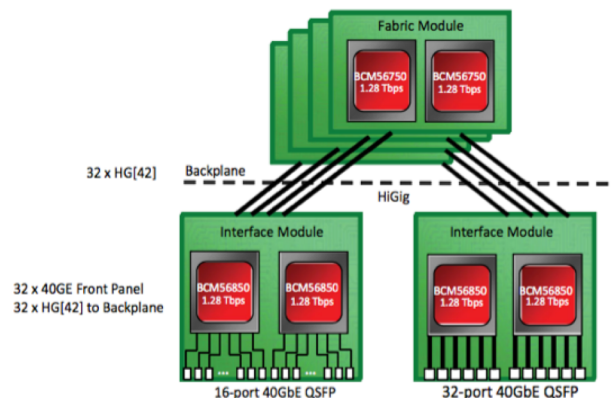
Fixed/Top-of-Rack Switches



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



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

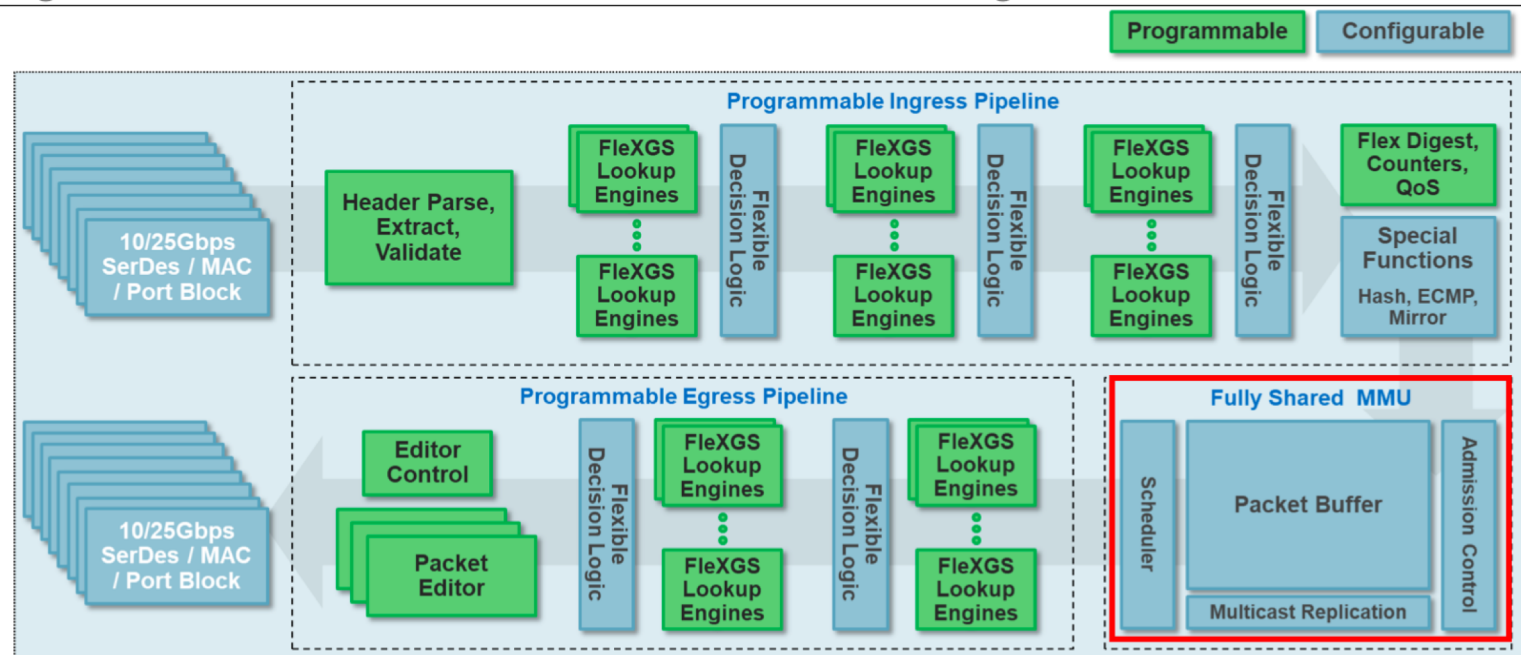
<https://www.broadcom.com/collateral/pb/56850-PB03-R.pdf>

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



Trident 2 y Trident 3

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

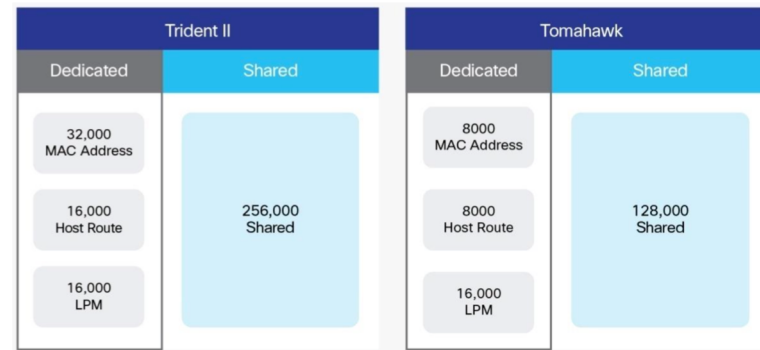


Table 1. Broadcom Trident 2 Forwarding Tables

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	8,000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (8,000)	16,000
Mode 1	8,000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (8,000)	16,000
Mode 2	8,000	Layer 2 (32,000)	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (8,000)	16,000
Mode 3	8,000	Layer 2 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (32,000)	Layer 3 (8,000)	16,000
Mode 4	8,000	LPM (32,000)	LPM (32,000)	LPM (32,000)	LPM (32,000)	Layer 3 (8,000)	16,000

Cisco ASE-2

- ACI Spine Engine 2 (ACI = Aplicacion Centric Infrastructure)
- 3.6Tbps (todos los tamaños de paquetes)
- 36x100GE, 72x40GE, 144x25GE
- 16K VRF, 32 SPAN, 64 mcast, 4K NAT
- Push/swap 5 etiquetas VPN
- DWRR con 16 colas por puerto
- WRED, ACN, AFD (Approximate Fair Dropping)

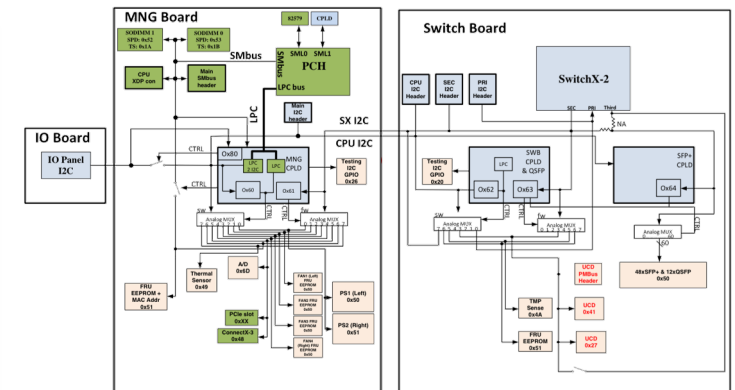
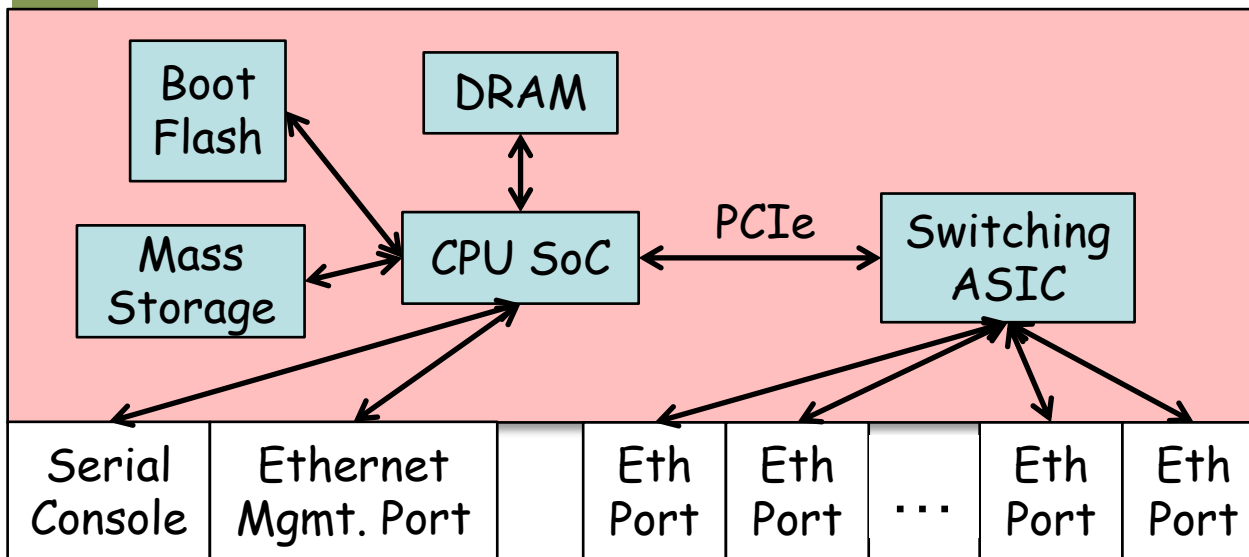
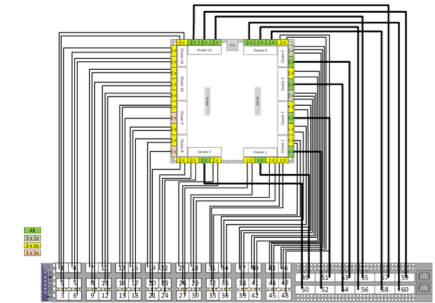
¿Evolución?

- La infraestructura se está simplificando
- Principalmente el hardware, controlable por software
- *White boxes* no solo servidores sino también switches
- También se venden ya switches "*Bare metal*" = solo el hardware



Bare metal switches

- Menores costes
- El mismo equipo un día es un switch, otro un firewall, otro un balanceador... dentro de las limitaciones del ASIC
- Ejemplo:
 - Open Compute Networking Project
 - <http://www.opencompute.org/wiki/Networking>
 - Especificaciones completas de conmutadores
- Fabricantes: Mellanox, Quanta, Penguin Computing, Edge-core, Acton, Dell, etc



¿Evolución?

- Para estos equipos sistemas operativos y gran cantidad de software, generalmente basados en linux, muchos de código abierto
- Ejemplos:
 - Open Network Install Environment (ONIE): <http://onie.opencompute.org>
 - Open Network Linux: <http://opennetlinux.org>
 - Big Switch's Switch Light OS
 - Pica8 PicOS
 - Cumulus Linux
- Es decir, igual que en el entorno de servidor, puedes cambiar el hardware, instalar el sistema operativo que quieras y desarrollar tus aplicaciones (...)



¿Evolución?

- Para diferenciarse, los proveedores desarrollan software propietario para ofrecer sus servicios
- Porque hoy en día ya es el software por lo que principalmente están cobrando los fabricantes “no-open”
- Muchos modelos ToR de fabricantes conocidos son switches bare-metal que han comprado, cambiado el software y el frontal



Quanta Bare Metal Switch

A Powerful 10GBASE-T Top-of-Rack Switch for Data Center and Cloud Computing

- 10GBT
- ONIE Pre-loaded
- x86 CPU Board Support
- Ubuntu Compliant
- SDN Ready

Datacenter networks is facing a major paradigm shift toward the disaggregation of software and hardware. This move, combined with the benefits of software-defined networking (SDN) allows network administrators respond quickly to changing business requirements at a lower capital cost as well as reducing the network operations complexity.

QuantaMesh BMS products offer higher performance, increased availability, and better serviceability to meet datacenter installation environment. QuantaMesh T3048-LY9 supports 48 100/1000/10G Base-T ports and 6 40G QSFP+ ports in a compact rack unit size. By leveraging the new generation merchant silicon chips, T3048-LY9 is a high performance high density Ethernet switch with an affordable price for the deployment of data center infrastructure. Moreover, the CPU board design of T3048-LY9 provides option for 3rd party software choose. Finally, with ONIE (Open Network Installation Environment) pre-loaded, it provides the flexibility and allows choice of network operating system supported by different software vendors. This provides agile installation process and faster response for the changing business demand.

About QCT

QCT (Quanta Cloud Technology) is a global datacenter solution provider extending the power of hyperscale datacenter design in standard and open SKUs to all datacenter customers. Product lines include servers, storage, network switches, integrated rack systems and cloud solutions, all delivering hyperscale efficiency, scalability, reliability, manageability, serviceability and optimized performance for each workload. QCT offers a full spectrum of datacenter products and services from engineering, integration and optimization to global supply chain support, all under one roof. The parent of QCT is Quanta Computer Inc., a Fortune Global 500 technology engineering and manufacturing company.

Physical ports

- **Port configuration:** 48 100/1000/10GBASE-T and 6 QSFP+ ports
- **Management Port:** Out-of-band management port (RJ-45, 10/100/1000Base-T)
- **Console Port:** 1 (RJ-45)
- **USB:** 1 (v 2.0)

CPU Board 1

- **CPU:** Freescale P2020
- **Memory:** 2GB DDR3/ECC
- **Flash:** 128MB
- **Storage:** 8GB Micro SD

CPU Board 2

- **CPU:** Intel Rangeley
- **Memory:** 4GB DDR3/ECC
- **Storage:** 32GB SSD

Performance

- **MAC:** Unified Forwarding Table to dynamically allocate the L2/ L3 tables
- **Switching capacity:** 1.44Tbps
- **Maximum forwarding rate:** 1071Mpps
- **Latency:** <3us

High Availability

- **Redundant power supply:** 1+1
- **Hot-swappable fan tray:** N+1



Dell S4810-ON

The Dell Networking S4810-ON switch is the industry's first disaggregated hardware and software data center networking solution that empowers organizations to deploy modern workloads and applications designed for the open networking era.

Organizations that benefited from utilizing the disaggregation model with their data center server platforms can now leverage even greater benefits from Dell open networking platforms. Organizations can take advantage of this disaggregated networking model using industry-leading hardware and a choice of leading network operating systems to simplify data center fabric orchestration and automation and accelerate innovation.

These new offerings provide organizations the needed flexibility to transform their data centers and offer high-capacity network fabrics that are cost-effective, easy to deploy and provide a clear path to a software-defined data center.

The Dell S4810-ON supports the open source Open Network Install Environment (ONIE) for zero-touch installation of alternate network operating systems.

Ultra-low-latency, data center optimized

The Dell Networking S-Series S4810-ON is an ultra-low-latency 10/40GbE top-of-rack (ToR) switch purpose-built for applications in high-performance data center and computing environments. Leveraging a non-blocking switching architecture, the S4810 delivers line-rate L2 and L3 forwarding capacity with ultra-low-latency to maximize network performance. The compact S4810 design provides industry-leading density of 48 dual-speed 1/10GbE (SFP+) ports as well as four 40GbE QSFP+ uplinks to conserve valuable rack space and simplify the migration to 40Gbps in the data center core (Each 40GbE QSFP+ uplink can support four 10GbE ports with a breakout cable). In addition, the S4810 incorporates multiple architectural features that optimize data center network flexibility, efficiency and availability, including I/O panel to PSU airflow or PSU to I/O panel airflow for hot/cold aisle environments, and redundant, hot-swappable power supplies and fans.

Key applications

- Ultra-low-latency 10GbE switching in HPC, high-speed trading or other business-sensitive deployments that require the highest bandwidth and lowest latency
- High-density 10GbE ToR server aggregation in high-performance data center environments

Key features

- 1RU high-density 10/40GbE ToR switch with 48 dual-speed 1/10GbE (SFP+) ports and four 40GbE (QSFP+) uplinks (totaling 64 10GbE ports with breakout cables) with OS support
- 1.28Tbps (full-duplex) non-blocking switching fabric delivers line-rate performance under full load with sub 700ns latency
- I/O panel to PSU airflow or PSU to I/O panel airflow
- Supports the open source ONIE for zero-touch installation of alternate network operating systems
- Redundant, hot-swappable power supplies and fans
- Low power consumption



Performance

Switch fabric capacity:	1.28Tbps (full-duplex) 640Gbps (half-duplex)
Forwarding capacity:	960Mpps
Latency:	Sub 700ns
Packet buffer memory:	9MB
CPU memory:	2GB

Cisco Nexus 9300

The Cisco Nexus 9300 platform consists of fixed-port switches designed for top-of-rack (ToR) and middle-of-row (MoR) deployment in data centers that support enterprise applications, service provider hosting, and cloud computing environments. They are Layer 2 and 3 nonblocking 10 and 40 Gigabit Ethernet switches with up to 2.56 terabits per second (Tbps) of internal bandwidth.

Models

Table 1 summarizes the Cisco Nexus 9300 platform switch models.

Table 1. Cisco Nexus 9300 Platform Switches

Model	Description
Cisco Nexus 9332PQ Switch	32 x 40-Gbps Quad Enhanced Small Form-Factor Pluggable (QSFP+) ports
Cisco Nexus 9372PX-E Switch	48 x 1/10-Gbps SFP+ and 6 x 40-Gbps fixed QSFP+ ports
Cisco Nexus 9372TX-E Switch	48 x 1/10GBASE-T and 6 x 40-Gbps fixed QSFP+ ports
Cisco Nexus 9372PX Switch	48 x 1/10-Gbps SFP+ and 6 x 40-Gbps fixed QSFP+ ports
Cisco Nexus 9372TX Switch	48 x 1/10GBASE-T and 6 x 40-Gbps fixed QSFP+ ports
Cisco Nexus 9396PX Switch	48 x 1/10-Gbps SFP+ and up to 12 x 40-Gbps QSFP+ ports
Cisco Nexus 9396TX Switch	48 x 1/10GBASE-T and up to 12 x 40-Gbps QSFP+ ports
Cisco Nexus 93120TX Switch	96 x 1/10GBASE-T and 6 x 40-Gbps fixed QSFP+ ports
Cisco Nexus 93128TX Switch	96 x 1/10GBASE-T and up to 8 x 40-Gbps QSFP+ ports



HPE FlexFabric 5930 Switch

Key features

- Cut-through with ultra-low-latency and wire speed
- VXLAN VTEP OVSDDB support for virtualized environments
- High-density 10GbE and 40GbE spine/ToR connectivity
- IPv6 support with full L2 and L3 features
- Convergence-ready with DCB, FCoE, and TRILL



Software Defined X

- Software Defined Networking (SDN)
- Software Defined Infrastructure (SDI)
- Software Defined Data Center (SDDC)
- Software Defined Storage (SDS)
- Software Defined Radio (SDR)
- Software Defined WAN (SD-WAN)
- Software Defined Power (SDP)
- Software Defined Internet of Things (SDIoT)
- etc

Networking hardware y el software