TREX Workshop 2017

Ralf Korschner, Systems Engineer EMEA
ralf@arista.com

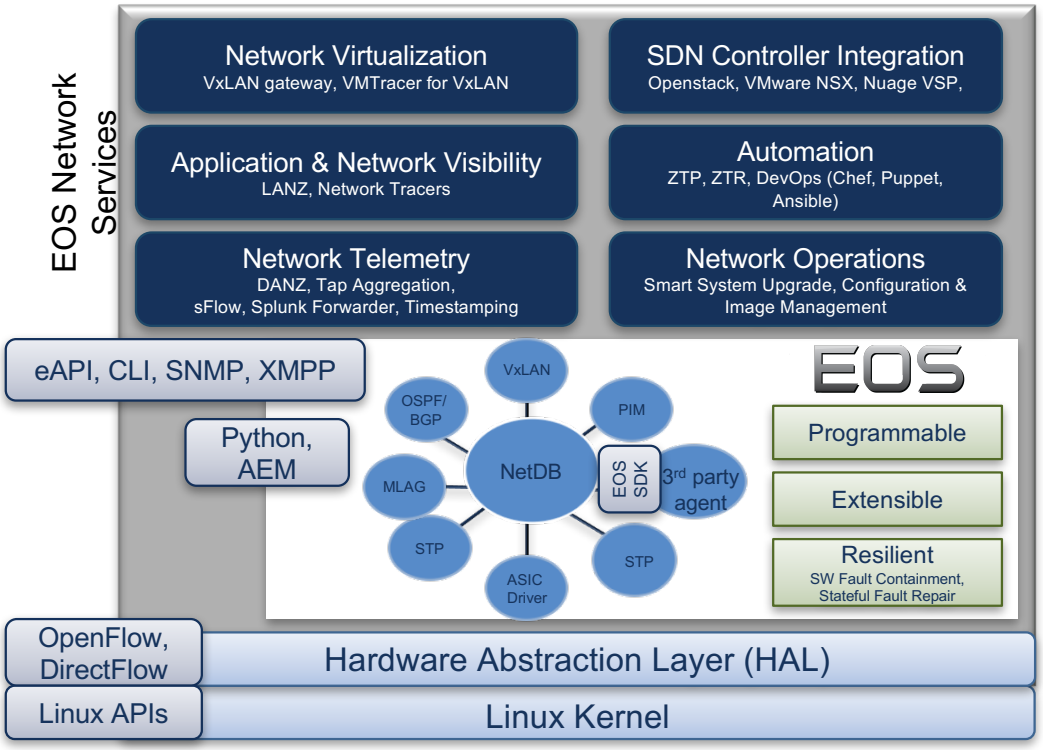
ARISTA

Quick Introduction

- Founded in 2004 to build a new class of networking platform
  - First product to market 2008 / Profitable since 2010
- Experienced Management and World Class Engineering Team
- Laser focus on key verticals: Cloud, SP, Web, Big Data, Finance
  - Platform of choice for Tier 1 financials and Cloud Titans
- 4500+ customers (>7 Million Ports) in 60+ countries
- The leading breakaway DC vendor: >14% share
- Gartner Data Center Magic Quadrant Leader
- Key innovators in:

Merchant Silicon	Open and Extensible Operating Systems
Network Scaling and Virtualization	SDN and Cloud Orchestration

EOS Software Architecture Foundations

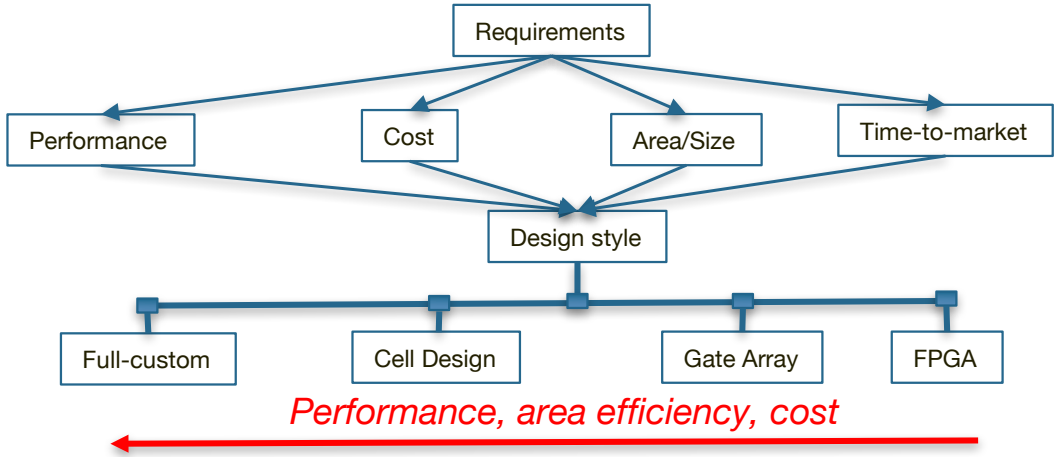


- Unmodified Linux Kernel
- State Separation - All processes in own user space
- State in NetDB, publish-subscribe state sharing
- No death-by-slow-memory-leak
- Programmable via JSON, Linux, EOS SDK
- User scripts for event changes
- Integrates with NetOps & DevOps
- Network services – Applications built on EOS
- Superior quality – proven over 5 years

What do we mean by Merchant silicon ?

- **Semi-Custom Design (Cell based ASIC)**
  - Top down approach
  - Focused on the application rather than the performance.
  - Functionality achieved using pre-defined building block and libraries
  - Fabricated carried out by by a third-party silicon manufacture

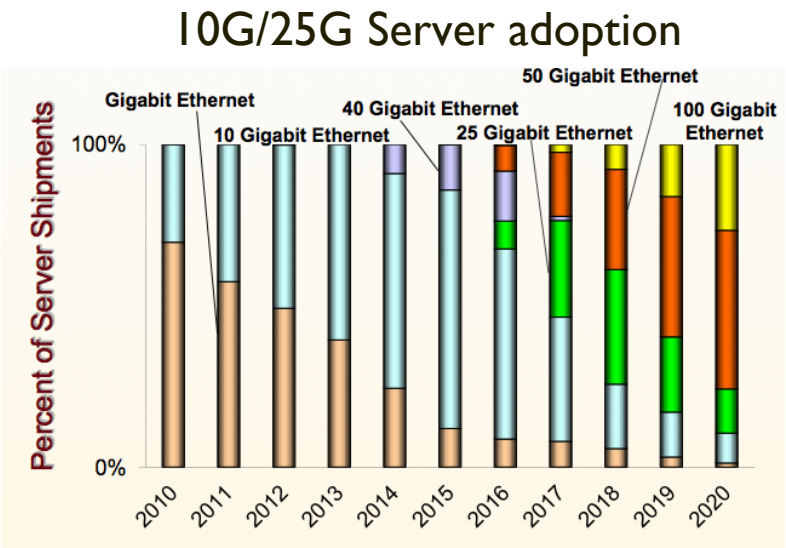
- **Full-Custom Design (Merchant Silicon)**
  - Bottom-up approach
  - Designed at the low-level component (transistor) layer
  - Allows focus on optimal clock rate/performance, density
  - Dev cycle and cost reduced, all development done by silicon vendors
    - Broadcom, Intel, Dune, Fulcrum etc
  - Hence the term Full-custom Merchant silicon



Moore's Law in the Data Center - Server

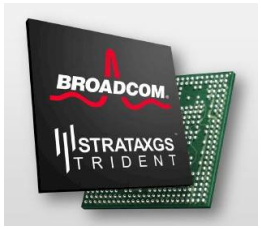
Moved to Cloud - Increased CPU power and Virtualisation

- **Faster CPUs need Faster Networks**
  - 50% attach rate 2014, 80% by 2015
  - From 5M ports 2010 to 67M ports 2016
  - 25G/50G projected to 10G by 2019
  - Faster End nodes need faster Backbones
  - 10/40/100G Market growing rapidly



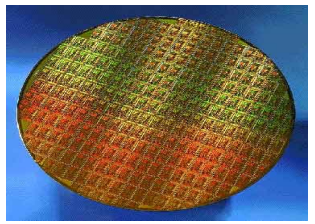
Growth of 10G/25G drove adoption of Merchant Silicon switches in Data Centers

Example Timeline of Merchant Silicon



T+ → T2 → T2+ → T3

T2 → TH → TH2



64x10G 32x40G 32x100G 64x100G

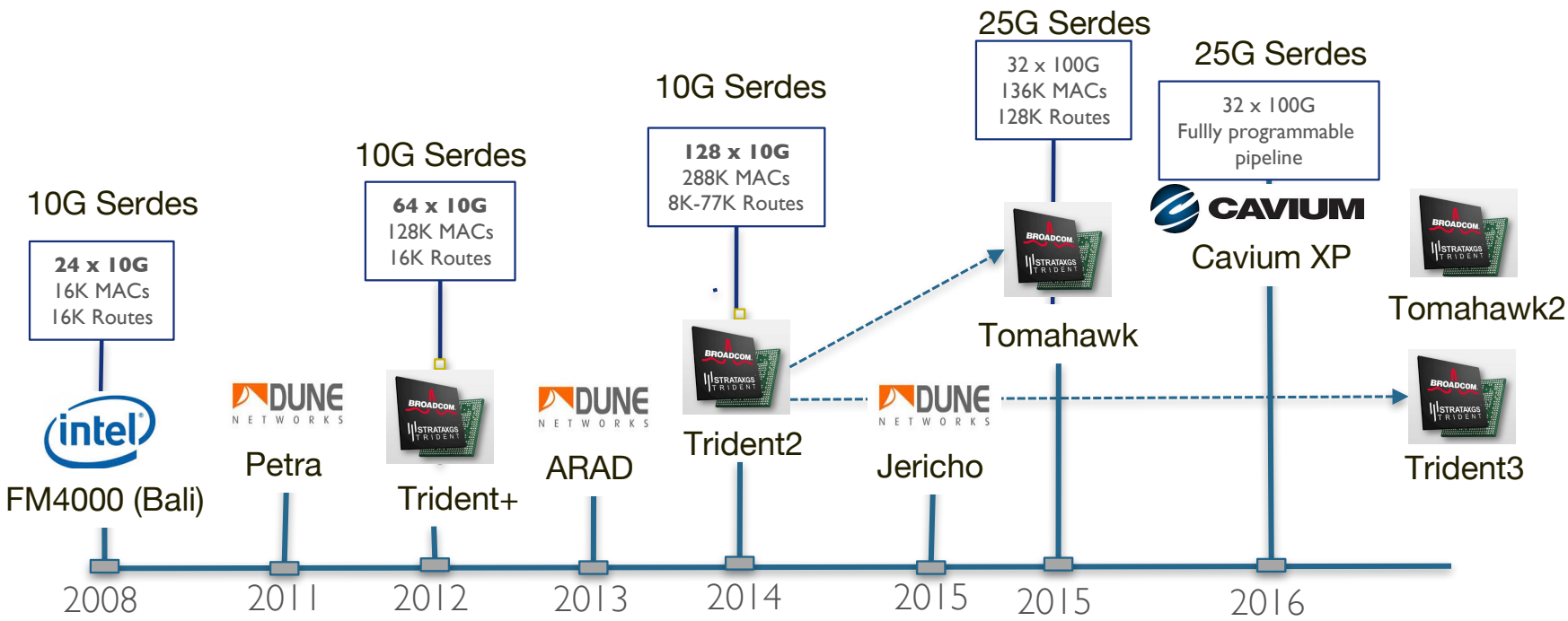
DUNE NETWORKS

Petra → Arad → Jericho → Jericho2

8x10G 6x40G 6x100G 18x100G

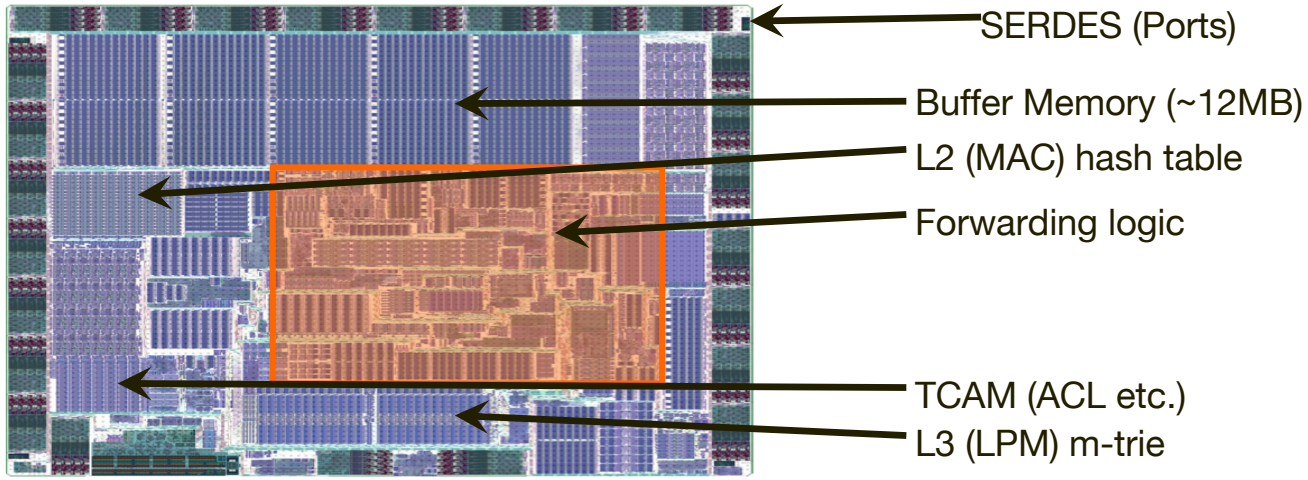
Evolution of Merchant Silicon

- To keep pace with CPU performance and bandwidth in the DC
- Merchant Silicon switches has allowed networking to follow Moore's Law
- Increasing competitive, with the introduce of new silicon vendors

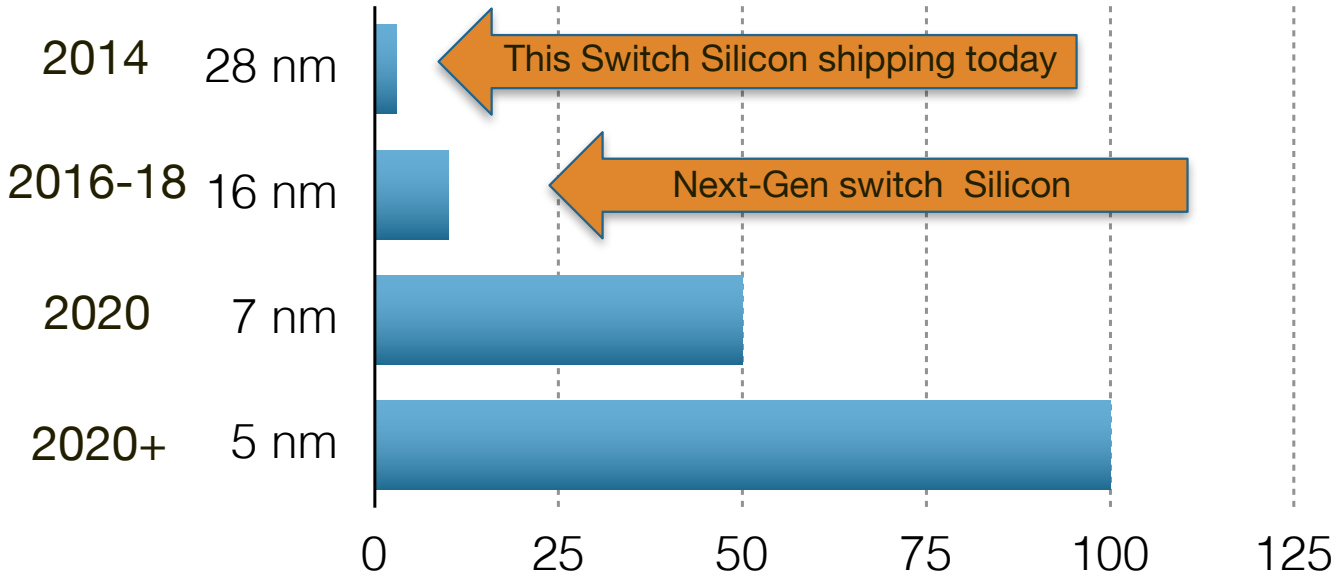


Merchant Silicon – Design trade-offs

- **Designing merchant silicon – multi-dimension challenge, silicon finite size**
  - Number of Pins and their speed (10G/25G) – defining the throughput requirements
  - Pins for physical interfaces, fabric links, access to off- chip resources?
  - Layer 2/3 forwarding logic and tunnel support (GRE, MPLS, VXLAN etc.)
  - Table sizes, MAC table (Exact match table), ARP tables (LEM), LPM for Layer 3 tables



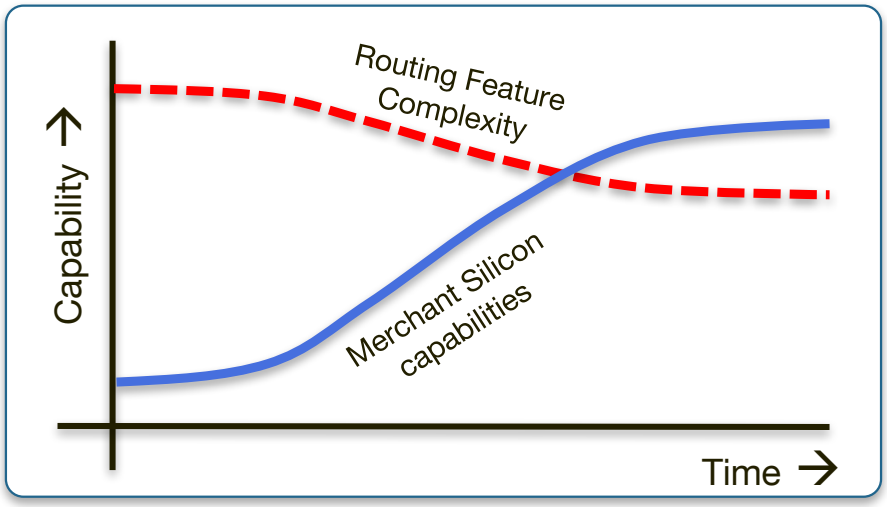
Evolution of Merchant Silicon



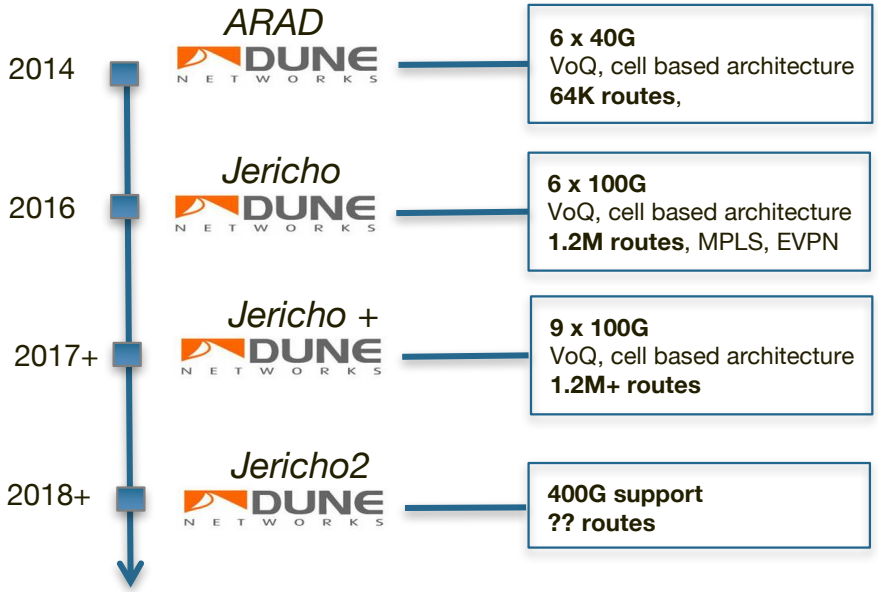
5 nm in 2X denser than 7, 10X denser than 16, 30X denser than 28nm

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 the latest merchant silicon.
  - Dune chipset, VoQ, cell-based lossless architecture



The lines between a Merchant silicon switch and router will continue to blur

Bringing Merchant Silicon to the Routing Market

- 7500E – ARAD based architecture

36 x 40G per line card
96 x 100GE for 8-slot chassis

64 x 40GE in 2RU

Multi ARAD Design
48x10GE – 4x40GE

2014

Evolution of 100GE Densities

- 7500R – Jericho based architecture

36 x 100G per line card
288 x 100GE, 3W per 10G port

48 x 100G + 8 x 40G in 2RU

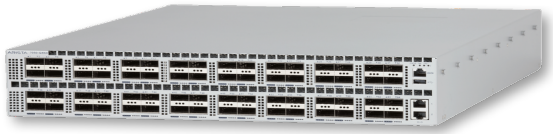
Single Jericho chip
48x10GE – 6 x100GE

2016

Bringing Merchant Silicon to the Routing Market

- 7500E – ARAD based architecture

36 x 40G per line card
96 x 100GE for 8-slot chassis



64 x 40GE in 2RU

Multi ARAD Design
48x10GE – 4x40GE

2014

Evolution of 100GE Densities

- 7500R – Jericho based architecture

36 x 100G per line card
288 x 100GE, 3W per 10G port

48 x 100G + 8 x 40G in 2RU

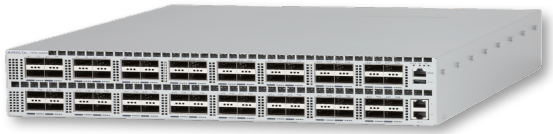
Single Jericho chip
48x10GE – 6 x100GE

2016

Bringing Merchant Silicon to the Routing Market

- 7500E – ARAD based architecture
- 7500R – Jericho based architecture

36 x 40G per line card
96 x 100GE for 8-slot chassis



64 x 40GE in 2RU

Multi ARAD Design
48x10GE – 4x40GE

2014

Evolution of 100GE Densities

- Full Internet Table – 1.2M
- MPLS , EVPN, VXLAN, API

36 x 100G per line card
288 x 100GE, 3W per 10G port

- Full Internet Table – 1.2M
- MPLS , EVPN, VXLAN, API

48 x 100G + 8 x 40G in 2RU

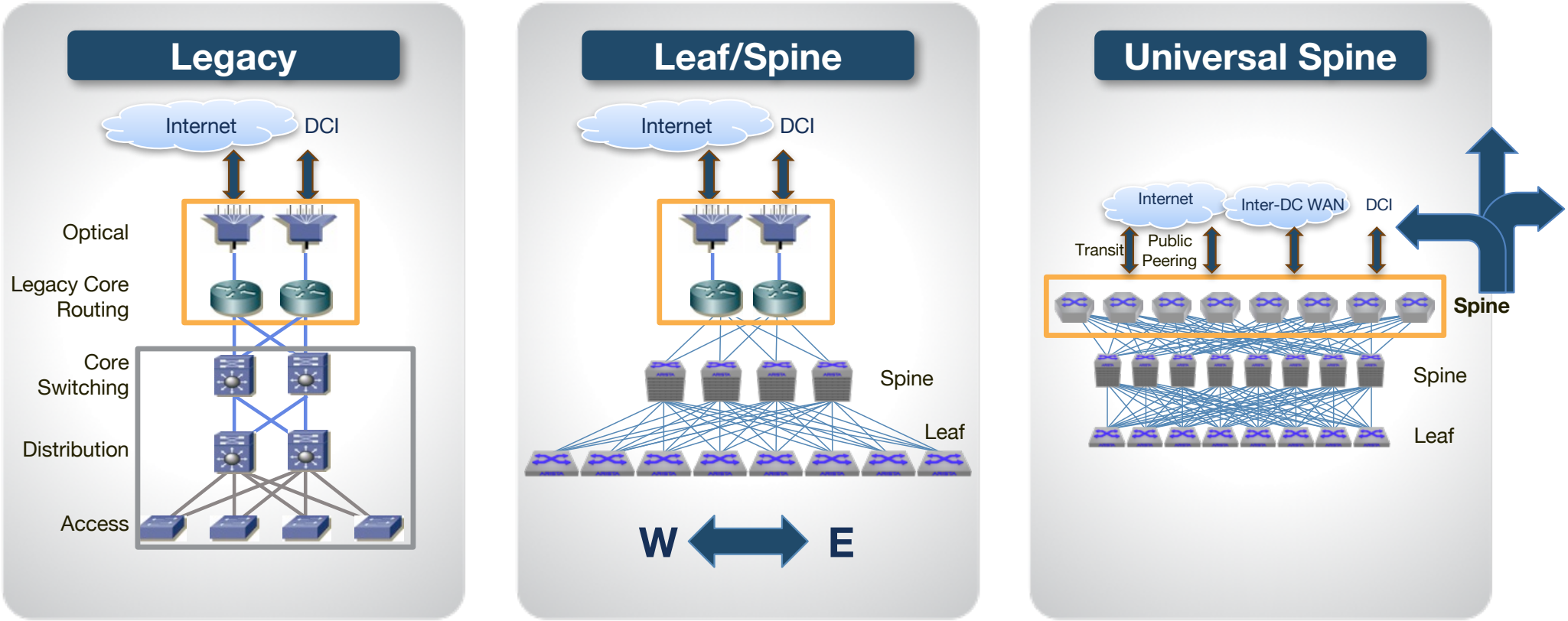
- Full Internet Table – 1.2M FIB
- MPLS , EVPN, VXLAN, API

Single Jericho chip
24x40GE + 12 x100GE

- < $3,000 per 100G port

2016

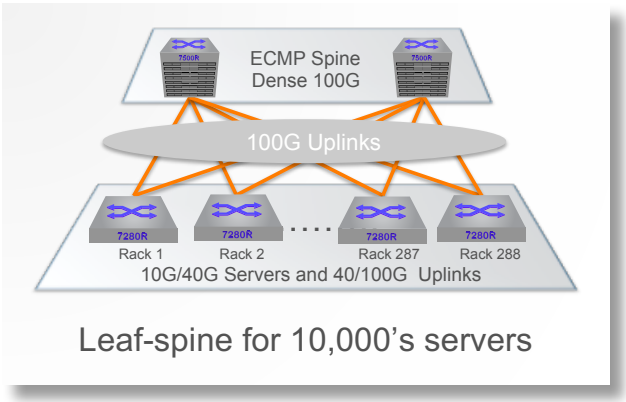
The Universal Spine Subsumes Core Router Functions



Improving Scale while Reducing Complexity

Merchant Silicon Use Cases for Service Providers

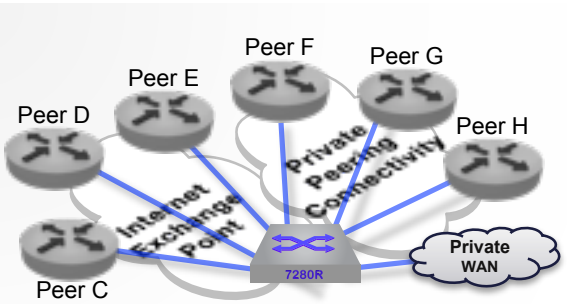
NFV / CDVR / DC/ Cloud



Leaf-spine for 10,000's servers

- Wire speed L2 & L3
- VXLAN and eVPN
- Flexible 40/100G uplinks
- Ease of migration
- Ultra deep buffers

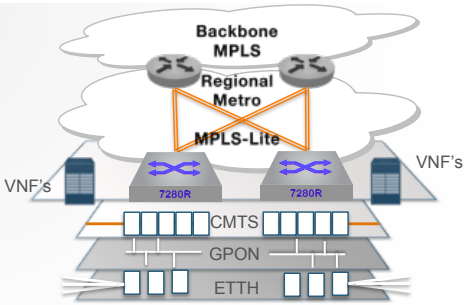
Internet / content peering



Internet Peering and Transit

- Non-blocking 10G & 100G
- Large forwarding tables
- Alternative to legacy
- Dense 100G Capacity
- Ultra deep buffers
- Power efficient

WAN / Aggregation



IP Services Edge Routing

- L2, L3 Services, MPLS-Lite
- Segment Routing
- 10/40/100G Aggregation
- Power efficient
- Very cost effective bandwidth
- Ultra deep buffer architecture
- DC and NEBS