What's happening in the Networking Landscape?

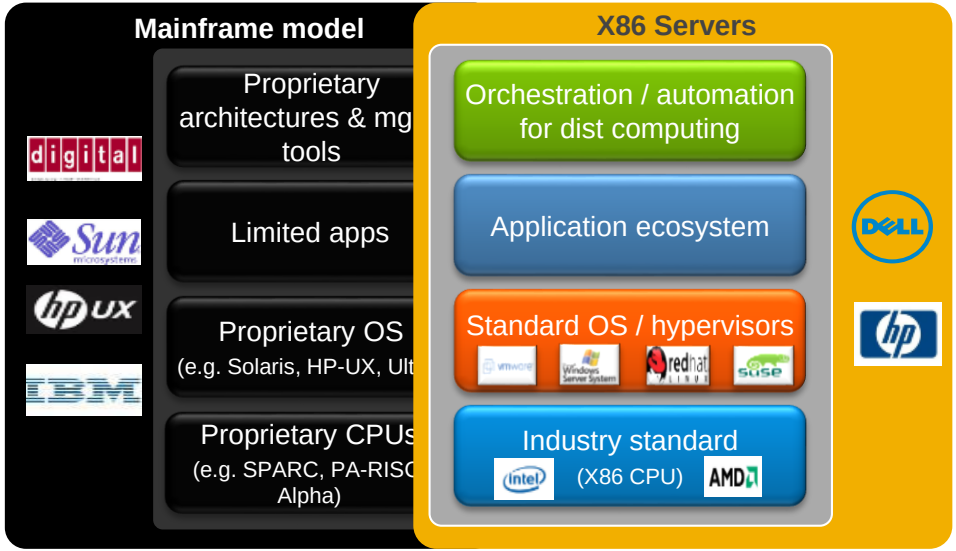
An overview on contemporary merchant switching silicon and SDN landscape

Paolo Bianco
GCN Systems Engineer
paolo.bianco@dell.com

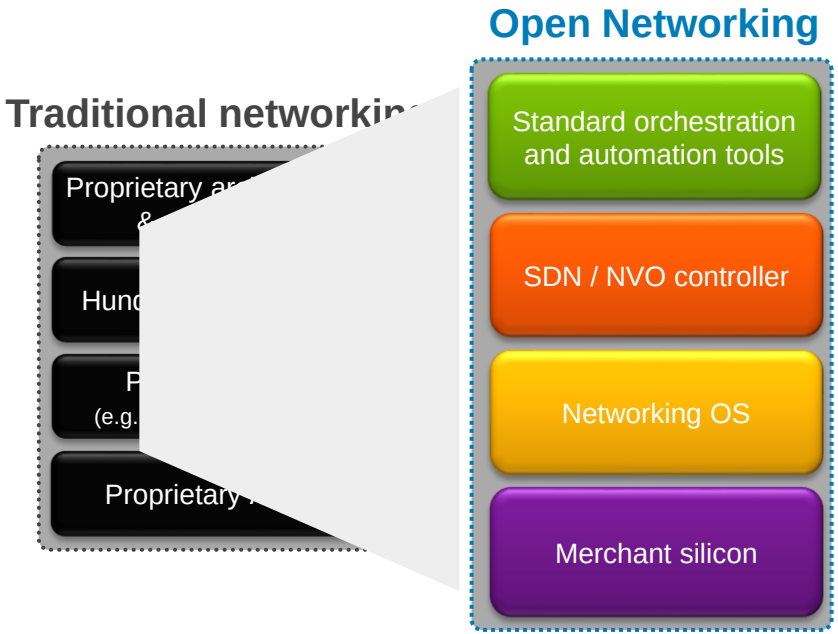
Microsoft

Windows Server: Power your business

25 years ago: Compute **paradigm shift**



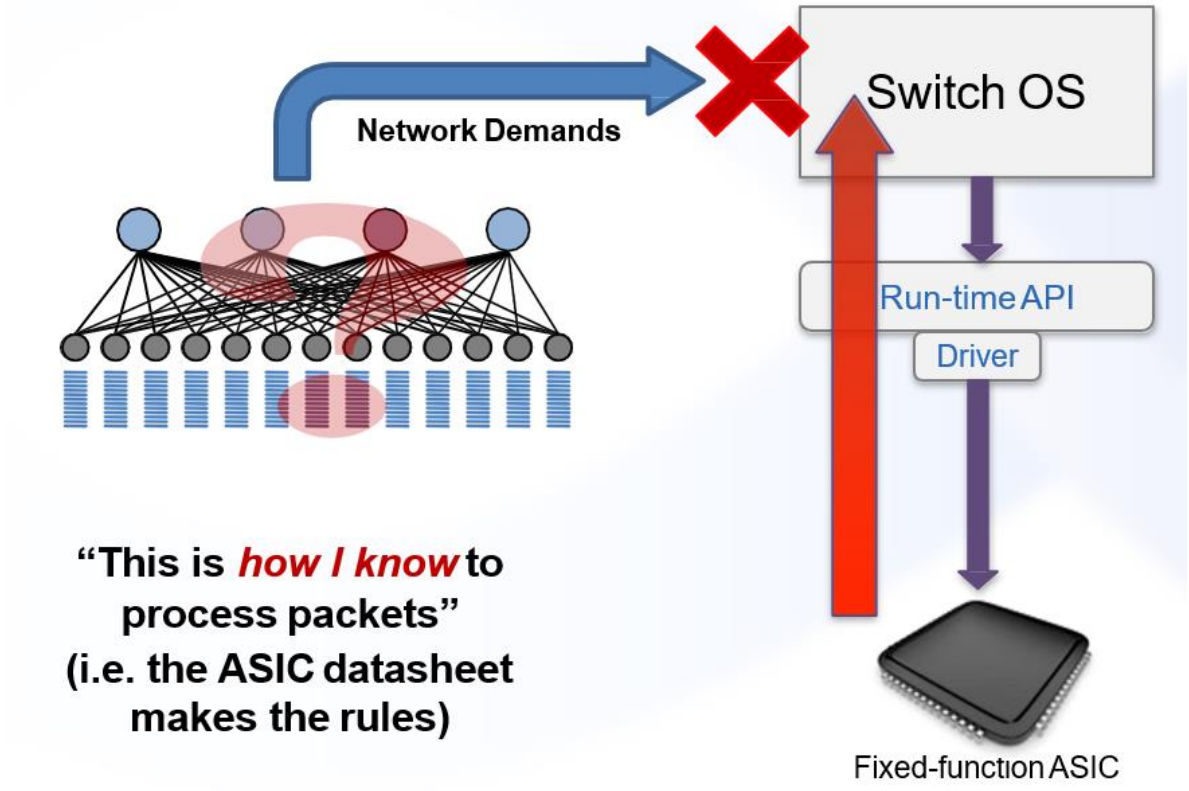
Now: Networking paradigm shift



DELL EMC | Microsoft

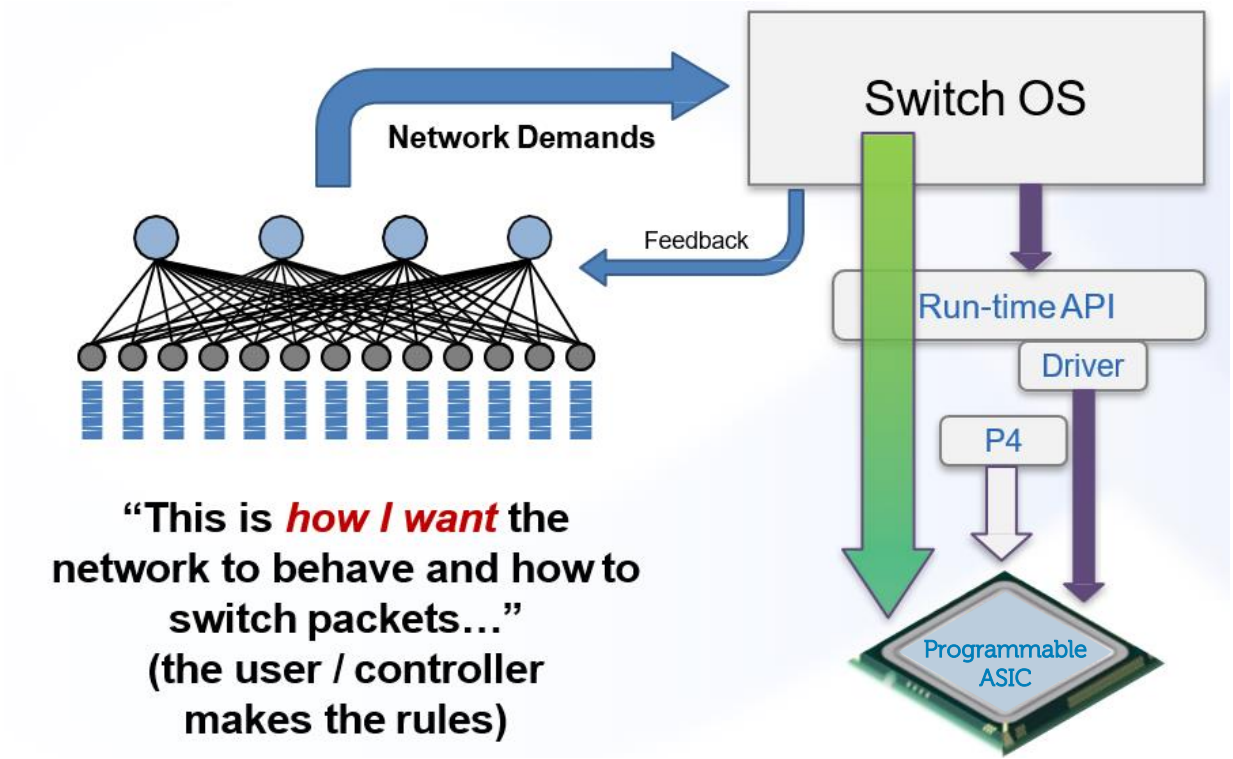

DELL EMC
Contemporary Merchant Silicon Landscape
Microsoft
Windows Server: Power your business

Traditional Approach: Bottom-Up network design



DELL EMC | Microsoft

New Approach: Top-Down network design



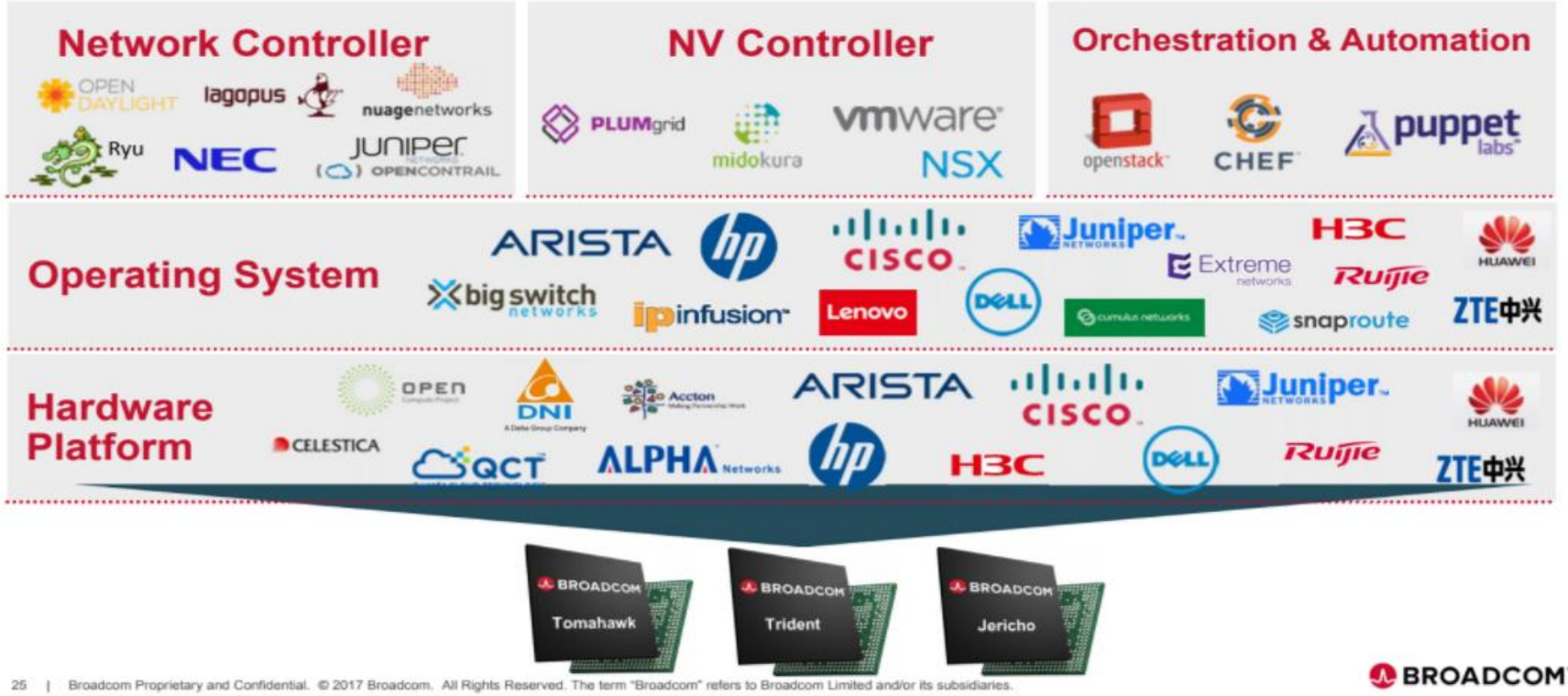
DELL EMC | Microsoft

The Programmable Switching Silicon Landscape

- One big Player: Broadcom
  - 95% of market share in 2015
  - Broader Ecosystem
  - First with fastest Programmable Switch Silicon (12.8TB/s)
- Some Emerging Contenders
  - Cavium
    - First with 25GbE
    - XPA programmable architecture on all DC chipset
  - Barefoot
    - Innovative approach
    - PISAprogrammable architecture
    - HW+SW «Weapon System» with P4
  - Innovium
    - Very New company
    - Now Sampling 12.8TB/s Teralynx chipset
    - Innoflex programmable architecture

DELL EMC | Microsoft

Broadcom Ecosystem



BROADCOM

DELL EMC | Microsoft

Broadcom Offering

RoboSwitch™

Broadcom's RoboSwitch solutions are highly integrated, cost effective, and smart-managed based on a field-proven, industry-leading architecture.

StrataConnect®

StrataConnect switches support 1 Gigabit, 2.5 Gigabit, 5 Gigabit, 10 Gigabit, and 25 Gigabit I/O speeds covering switch bandwidths from tens of gigabits to hundreds of gigabits.

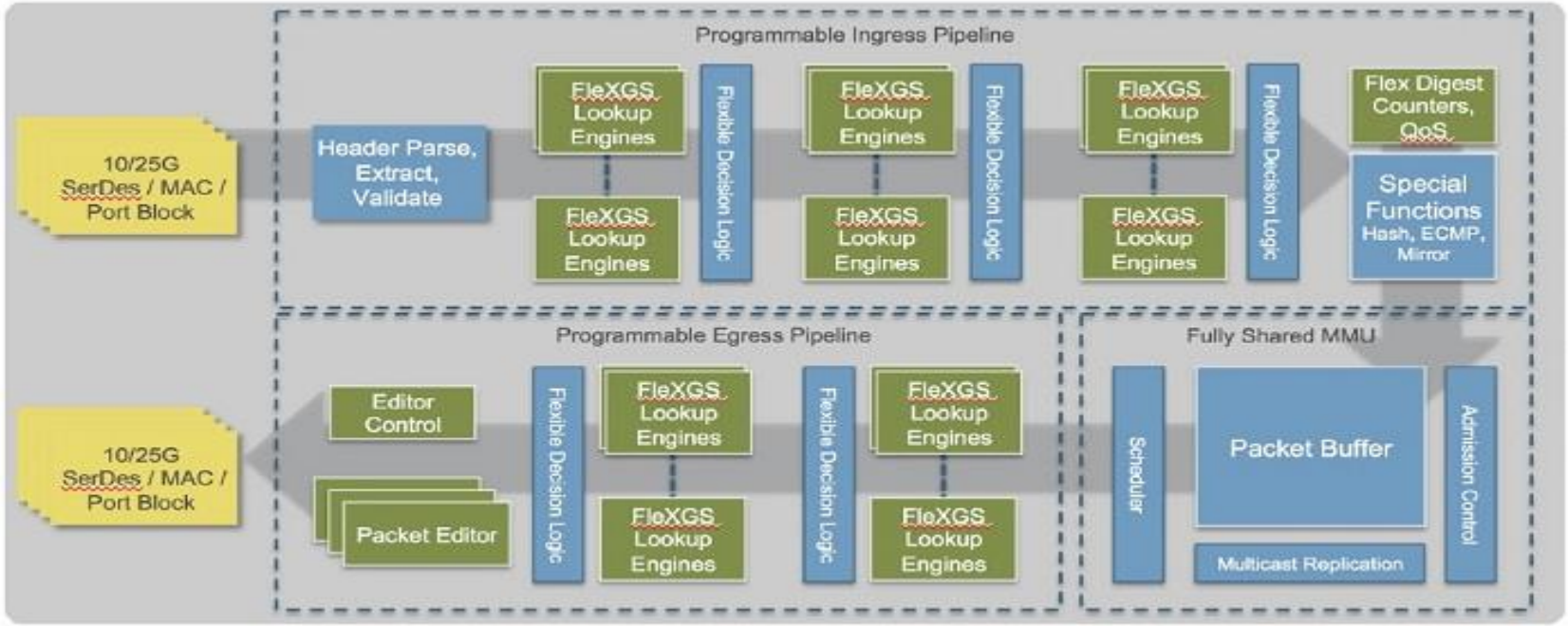
StrataDNX™

The StrataDNX product line offers the greatest extensibility and scalability of any merchant silicon switch in the industry, with the ability to scale both tables and buffering with external TCAMs or DRAMs.

StrataXGS®

The StrataXGS product line offers the most highly integrated, highest-bandwidth switching solutions available in the market.

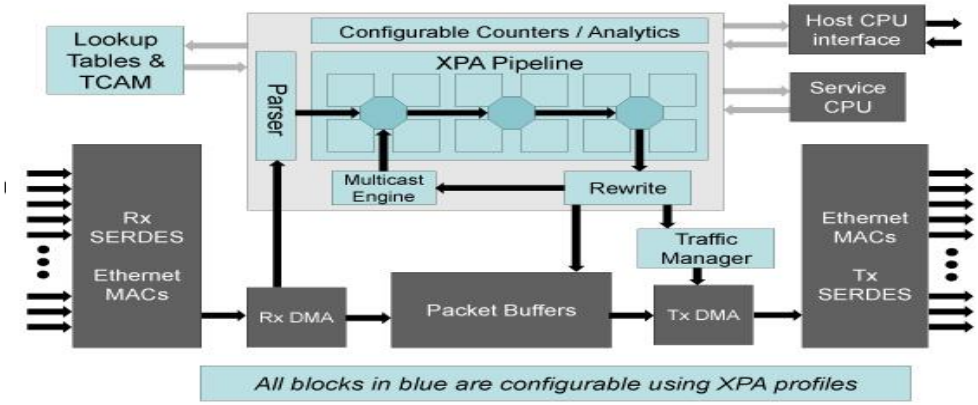
Trident 3 chipset Architecture



Cavium - XPliant

- **Scales up to 3.2Tbps**
  - Family of products from 880Gps
- **25G Ethernet Consortium support**
  - 25G / 50G specification
- **XPA architecture**
  - Flexible and Agile
  - Customized protocol implementation
- **Complete Protocol Feature Set**
  - L2, L3 Bridging / Routing
  - Tunnels and overlays
- **Standard API support**
- **Management CPU Support**
  - Host: OCTEON, x86 or PowerPC via PCIe & 10G connection
  - Internal Service CPU

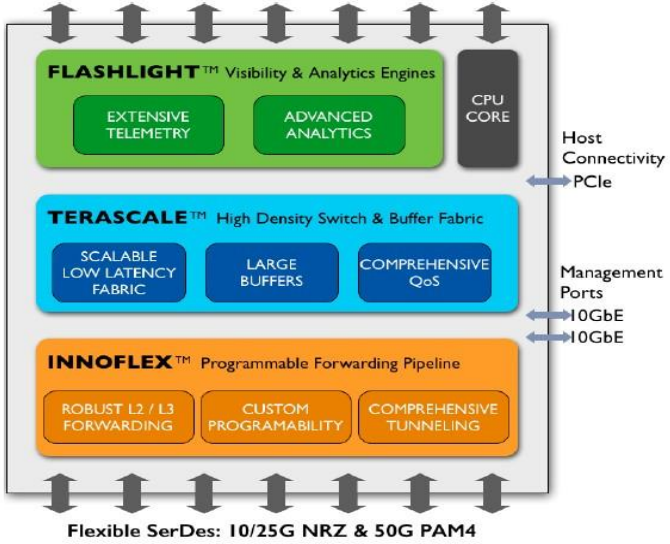
Flexible Port Configuration
32 x 100GE Ports
64 x 50GE Ports
64 x 40GE Ports
128 x 25GE Ports
128 x 10GE Ports

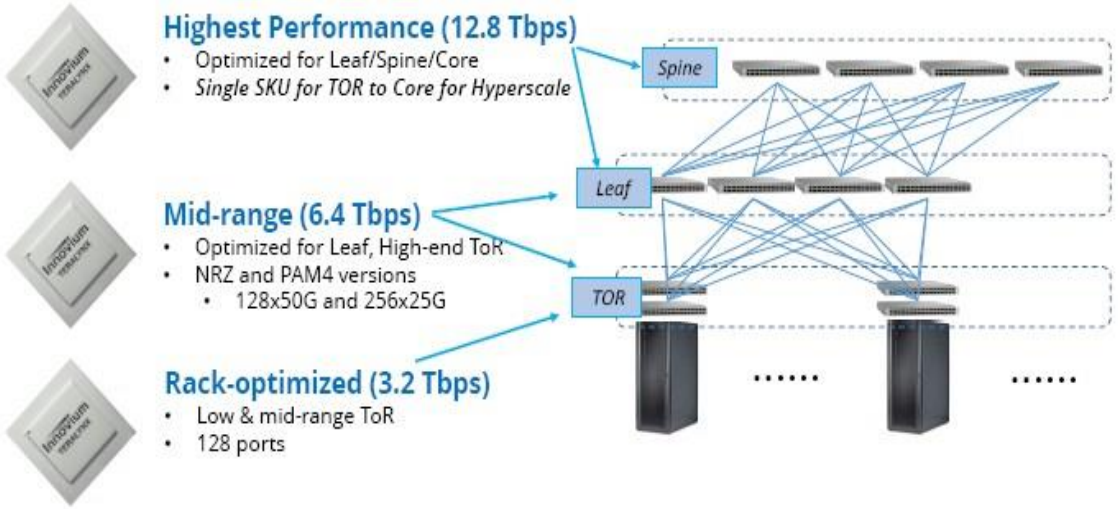


Broadcom Vs XPliant

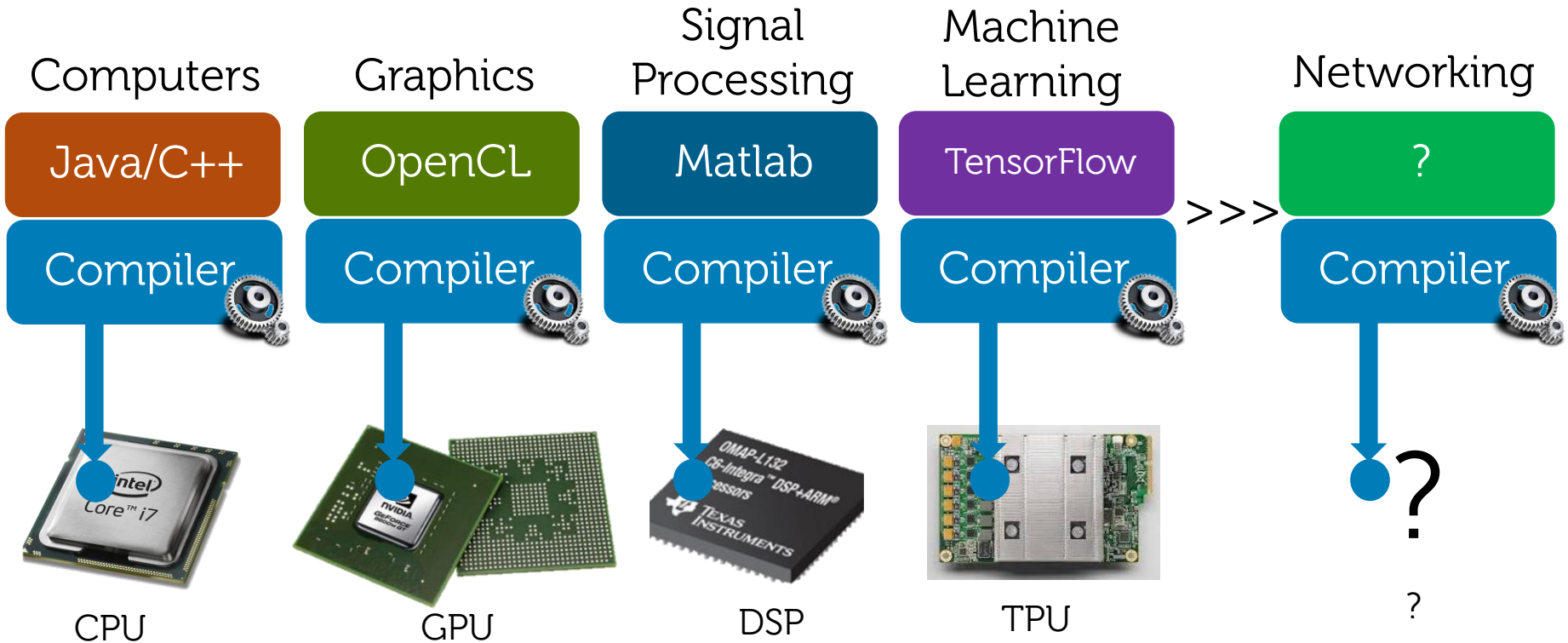
Chipset	TCAM	L2, L3 tables
Cavium / Xpliant	Up to 2,000,000	
Broadcom Tomahawk	Up to 256,000	
Apollo 2, Helix 4, Firebolt 3, Triumph 2, Trident+, Trident 2	1,000 to 2,000	~10,000s

Innovium Teralynx



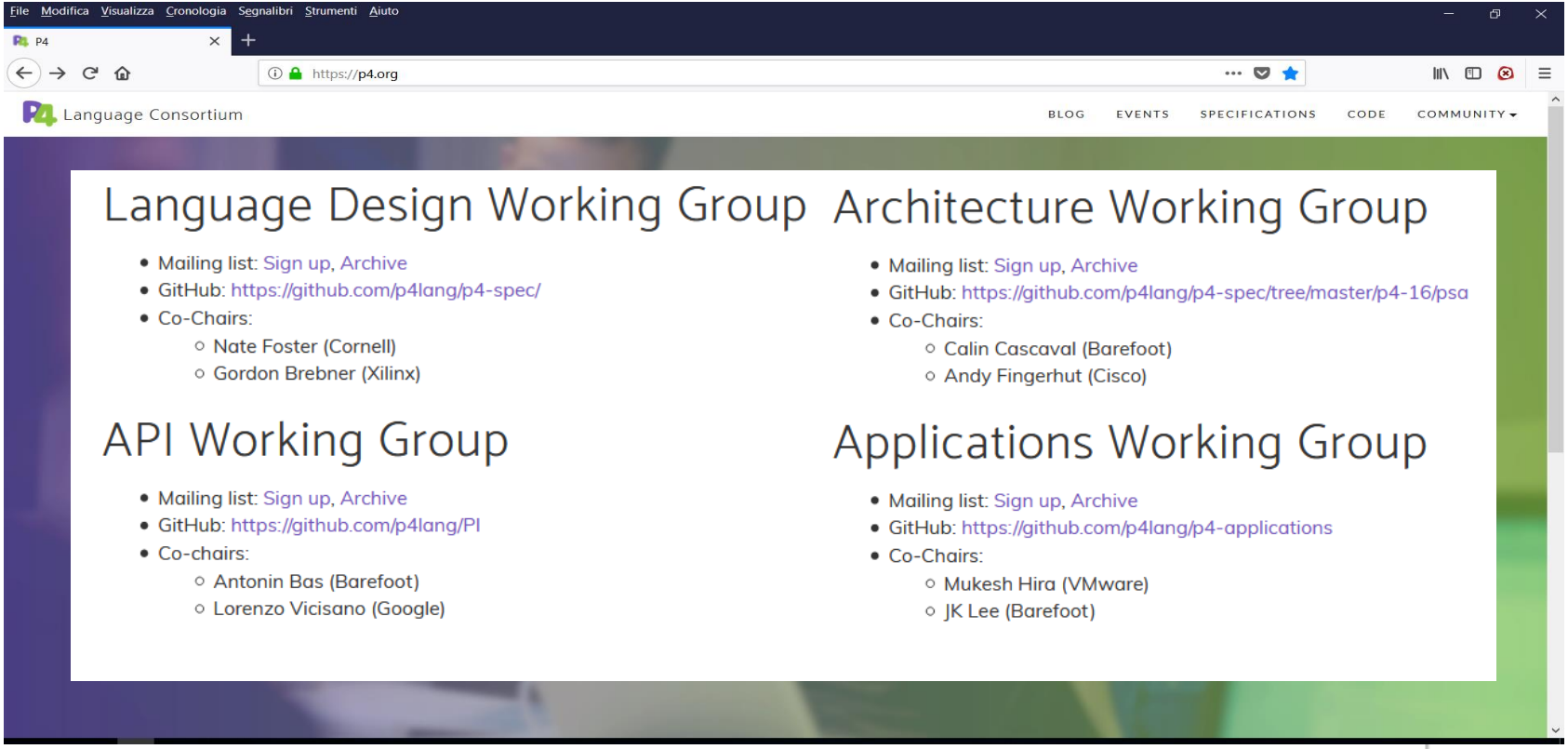


Domain Specific «Weapon Systems»



DELL EMC | Microsoft

P4.org – P4 Language Consortium



DELL EMC Microsoft

Barefoot Tofino

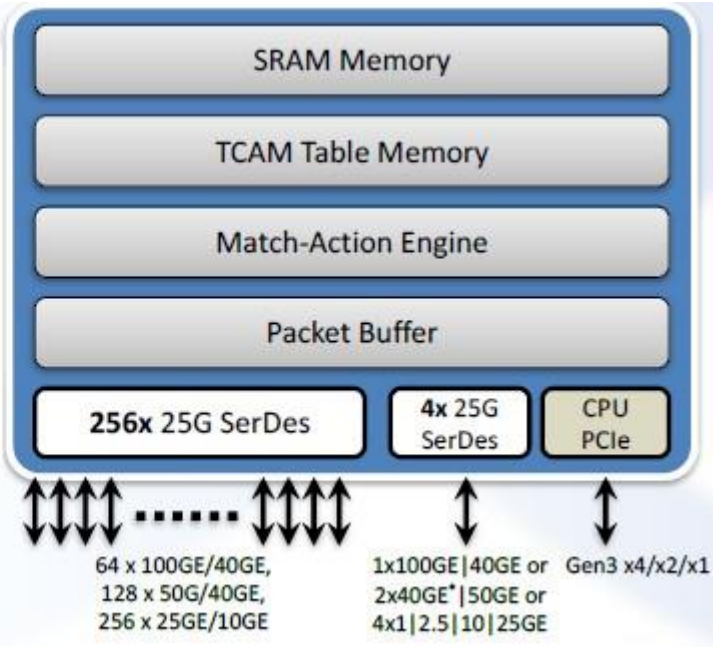
6.5Tbps single chip Ethernet switch
- Programmable Compiler Target
- All features operate at line rate

Port Configurations
- 65 x 100GE/40GE, 130 x 50GE
- 130 x 40GE (w/ gearbox), 260 x 25GE/10GE

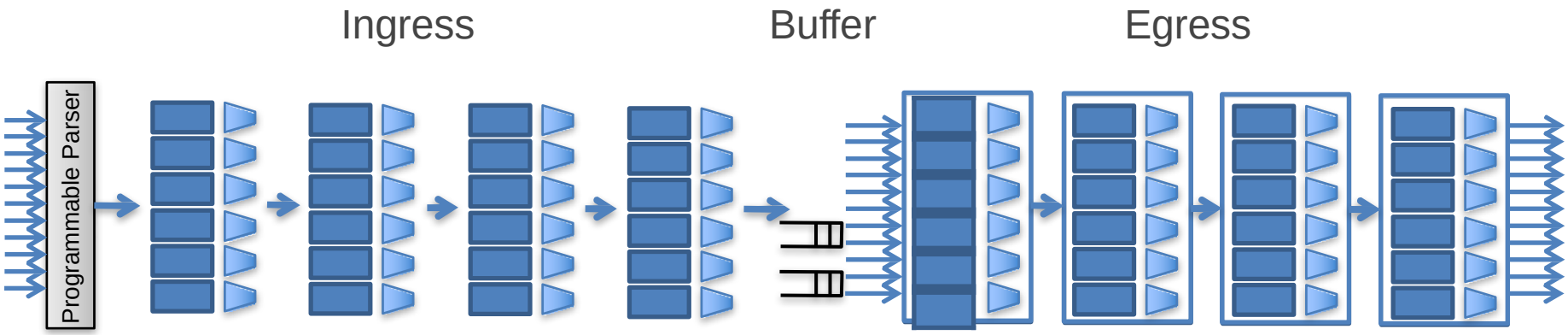
SerDes
- 25G with Integrated FEC

CPU Interfaces
- PCIe: Gen3 x4/x2/x1

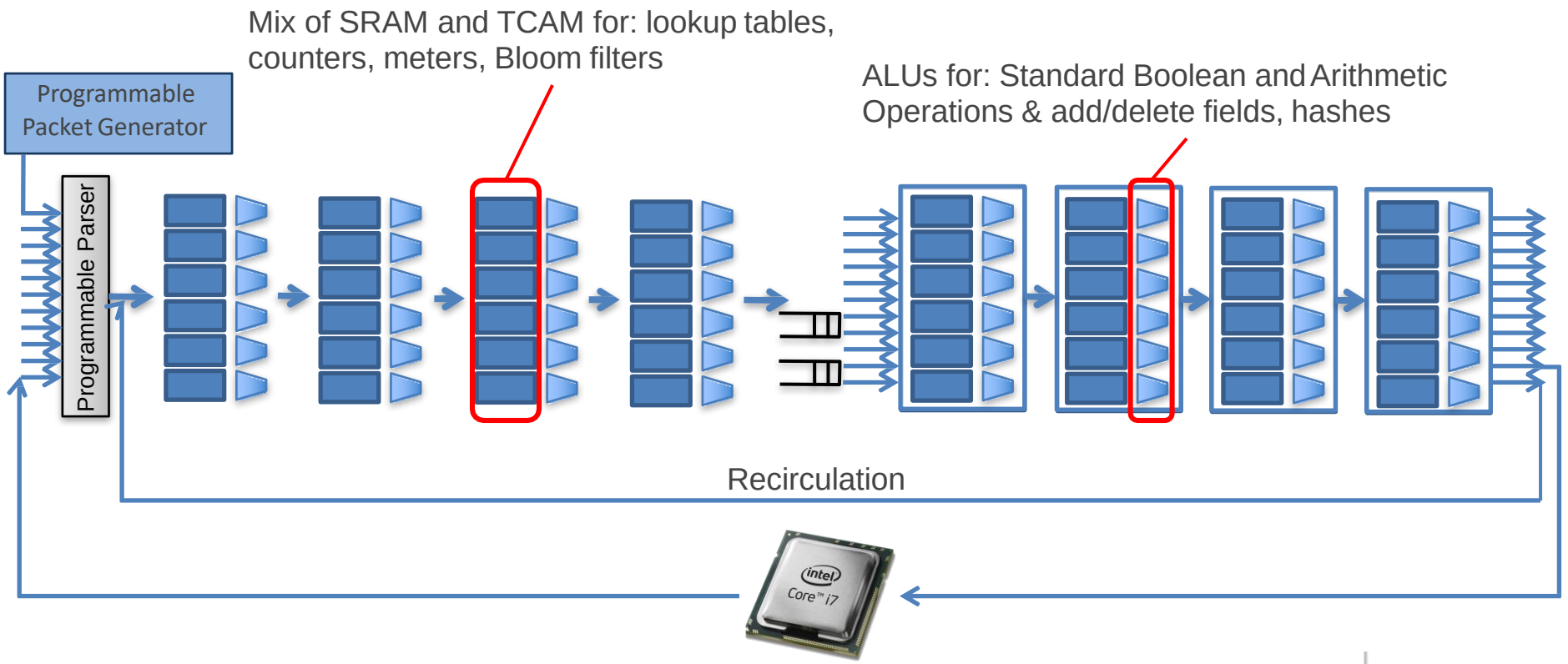
Programmable Pipeline (Embodiment of PISA)
- Create your own forwarding pipeline in P4 OR
- Use pre-built switch.p4 program



PISA: Protocol Independent Switch Architecture

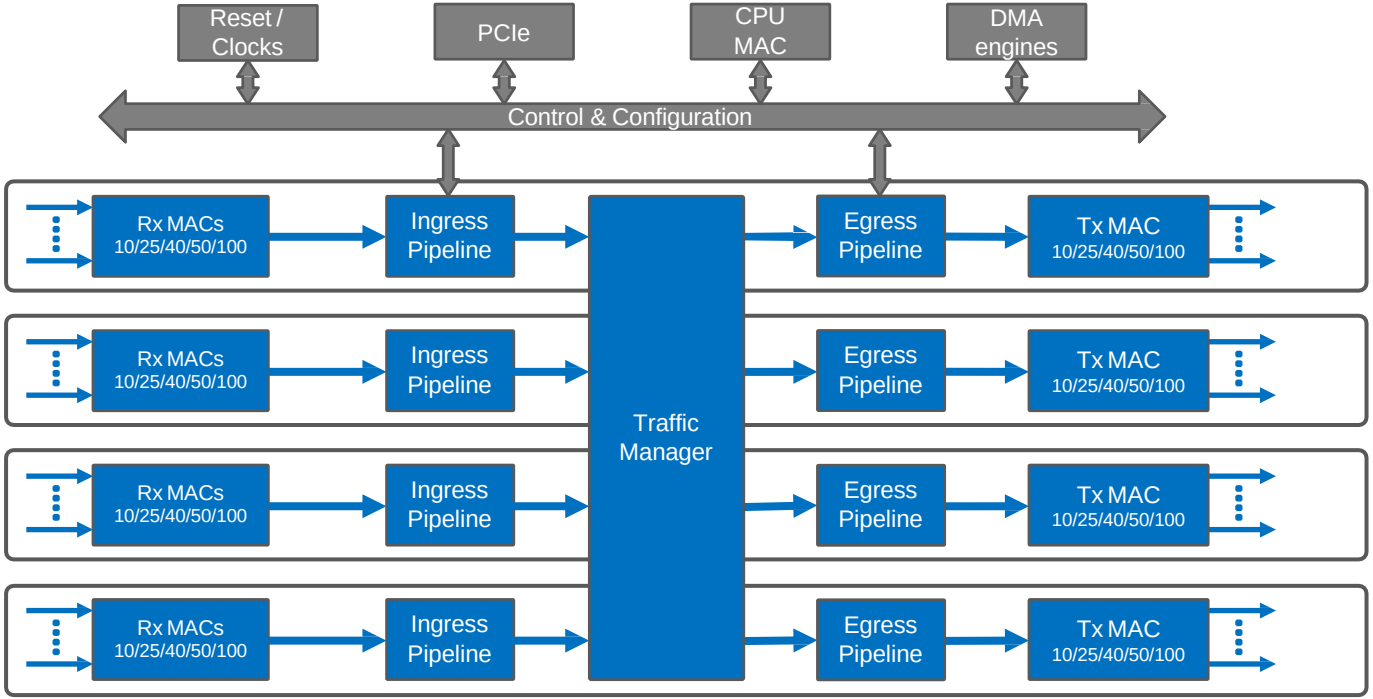


PISA: Protocol Independent Switch Architecture



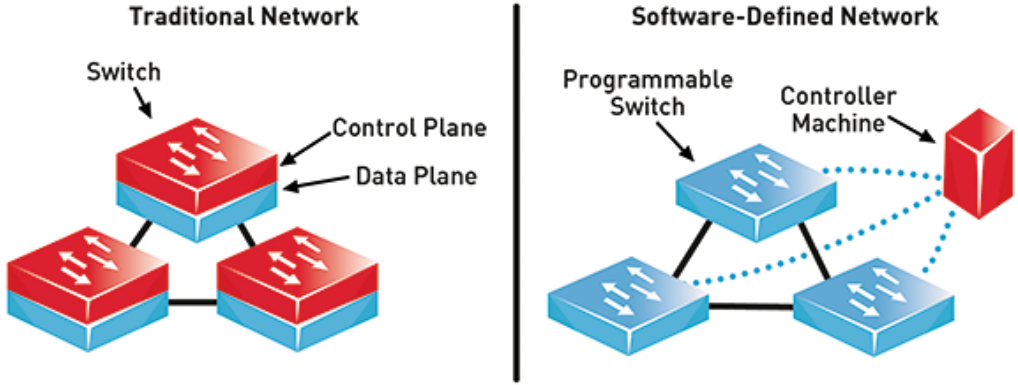
DELL EMC | Microsoft

Tofino Block Diagram




DELL EMC
SDN for Merchant Silicon Landscape
Microsoft
Windows Server: Power your business

What is Software Defined Networking?

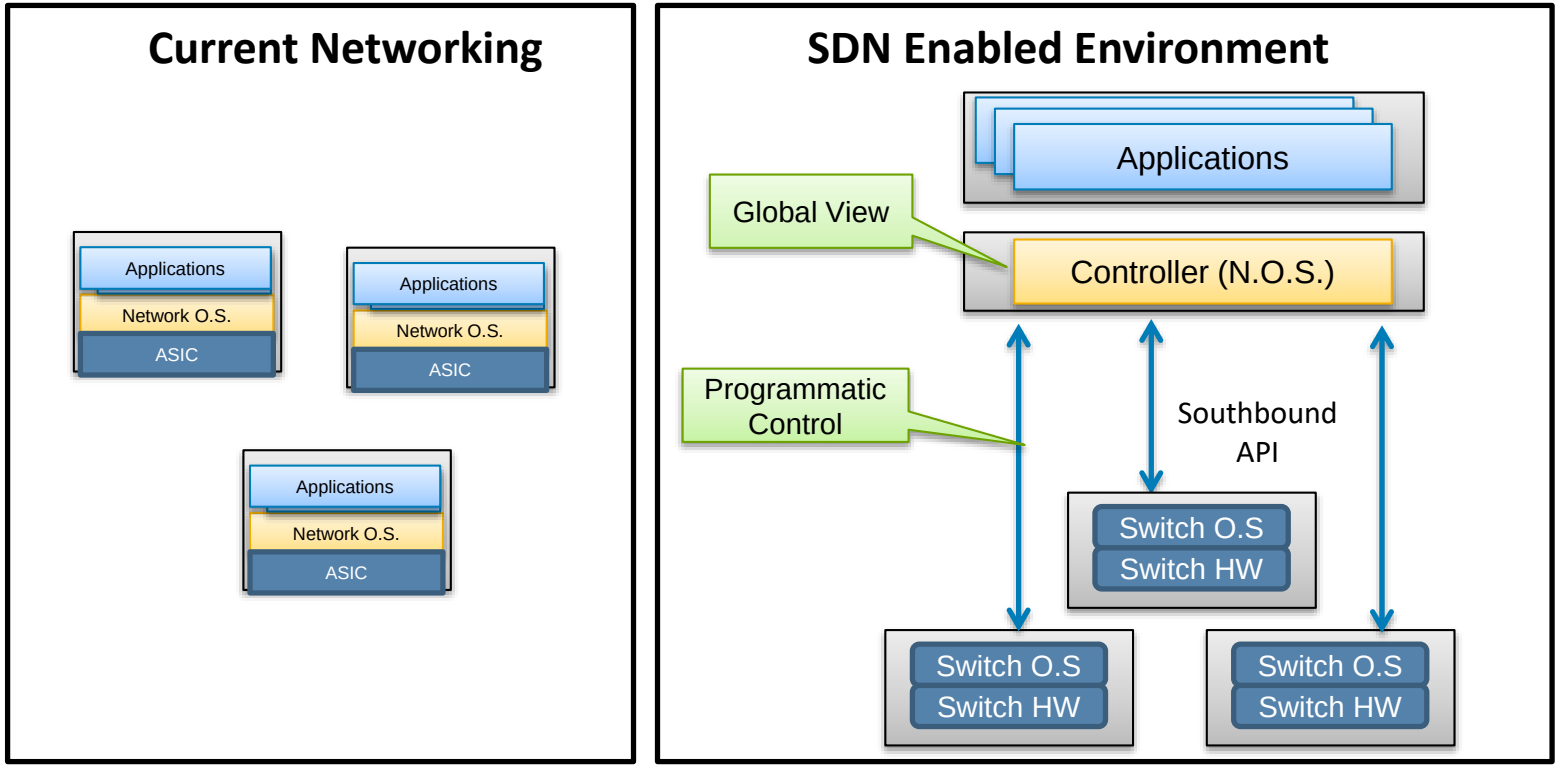


- A software-defined networking (SDN) architecture defines how a networking and computing system can be built using a combination of **open**, software-based technologies and commodity networking hardware that separate the control layer and the data layer of the networking stack.
- **Open** Networking is the foundation for SDN (but is not SDN!)
- In the SDN architecture, the splitting of the control and data forwarding functions is referred to as "**disaggregation**" because these pieces can be sourced separately, rather than deployed as one integrated system.

In turns: **SDN focuses on network stack disaggregation**

DELL EMC | Microsoft

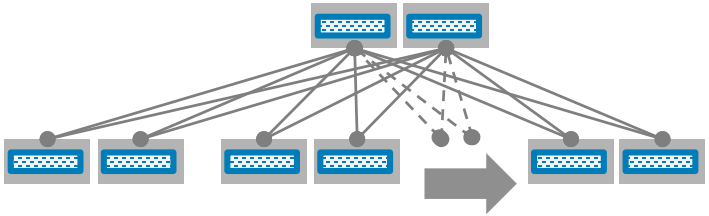
Implications of SDN



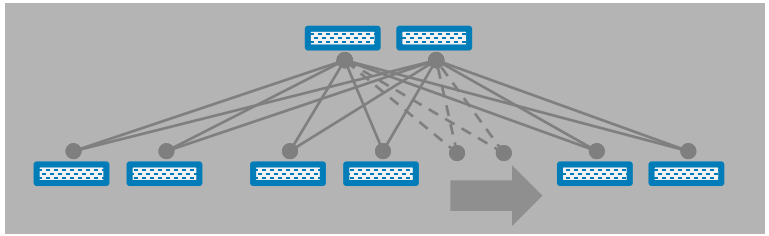
DELL EMC | Microsoft

From Network of boxes to Networked fabric

Traditional Approach
"Network of boxes"



Modern Approach
"Networked fabric"



- Deployment
  - Configuration prone to error
  - Time consuming configuration
  - Complex topology validation
- Management
  - Box by box visibility
  - Device level troubleshooting

Deployed, managed, optimized, and automated as a single entity

DELL EMC | Microsoft

Contemporary SDN Ecosystem (not necessarily complete!)

	Data plane (Elements used for traffic handling)		Controller solutions (Decoupled control plane)		Fabric (Combined data and control plane)	Management (Extensible mgmt software and API)
	PSwitch	VSwitch	OpenFlow	Overlay	Fabric	Mgmt
L2-L4 routing	NEC, IBM, Brocade, ipinfusion, big switch networks, Dell, Pica, cumulus networks, HP, Juniper Networks	Cisco, VMware NICira, VMware PLUMgrid, midokura, IBM, VMware	big switch networks, NEC, IBM, HP, OpenDaylight, Juniper Networks, ADARA	Cisco, VMware NICira, PLUMgrid, VMware, nuagenetworks, midokura, VMware	Dell EMC, Pluribus Networks, Plexxi, Insieme Networks, Cisco, ConteXtream, HP, xsigo systems, Oracle, vello systems, Ribbon Communications, ~~Infinite networks~~	cariden, CYAN, Infoblox, openstack, tail-f, anuta networks delivering a better cloud

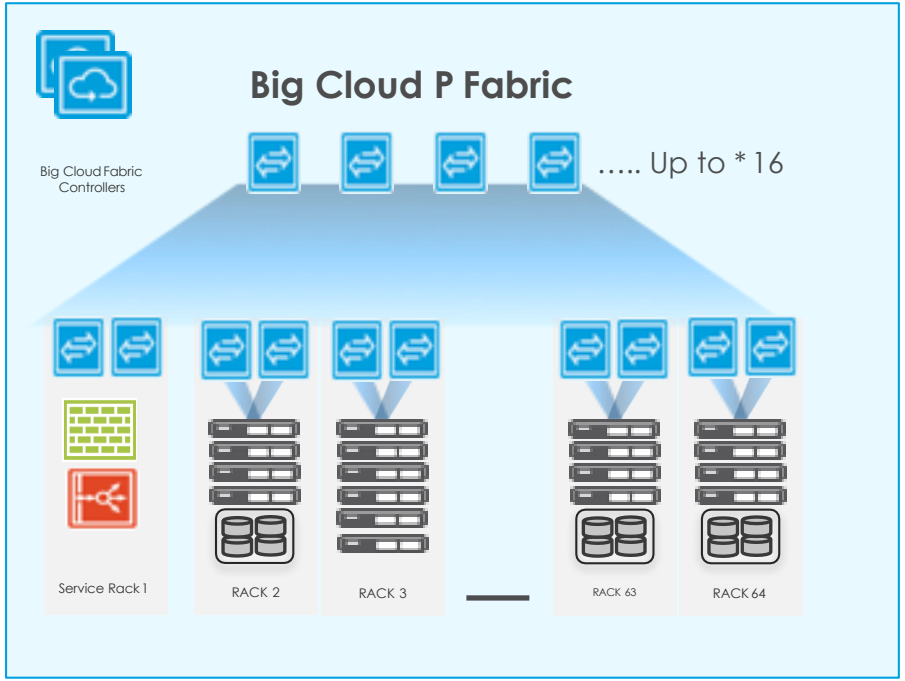
Big Switch Networks Portfolio

Big Switch Networks	***The Next-Generation Data Center Networking Company***
Company Mission	**Next-Generation DC switching** **Next-Generation DC security and monitoring**
Products/ Solutions	**Big Cloud Fabric** VMware SDDC (vSphere, NSX, vSAN) OpenStack (NFV/Private Cloud) Containers (Docker, Kub, Mesos, RH) **Big Monitoring Fabric** Pervasive Visibility DMZ Security Cloud Monitoring
Networking Architecture (Hyperscale Inspired)	Google facebook Microsoft SDN Controller Software Open Networking Hardware Scale-out Fabric Architecture

DELL EMC | Microsoft

Big Cloud Fabric Maximum Deployment

Maximum deployment for a single BCF Pod



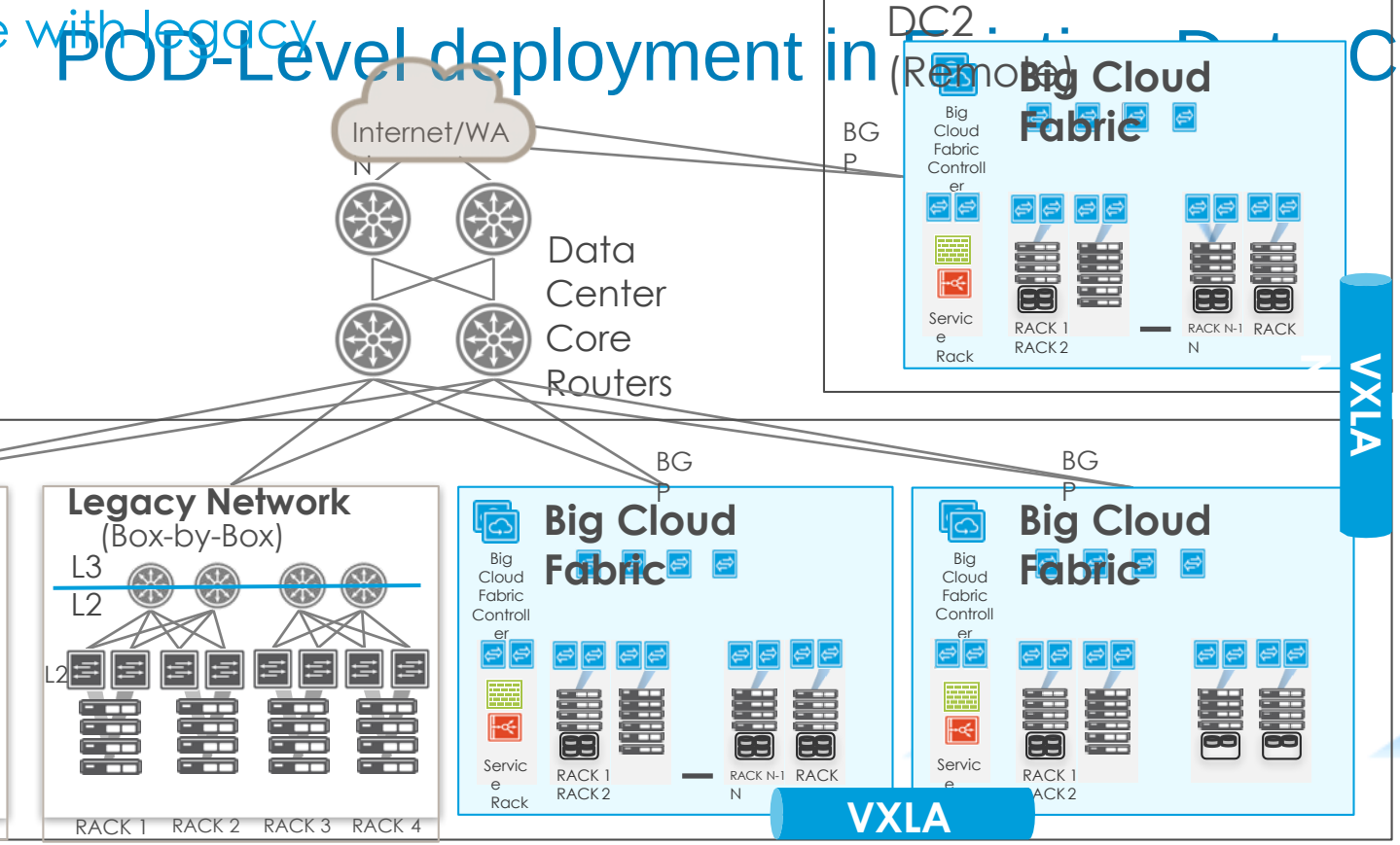
Total 10G Leaf Compute/Service connectivity:
48 * 64 = 3072 (MLAG)

- Pair of Hardware Controllers
- 64 racks
- Maximum 16 Spine Switches
  - 10G spine connections to Leaf Switches
  - 40G breakout to 4 * 10G on Dell Z9100-ON (due to breakout capabilities available on all ports)
    - 32 * 4 = 128
- 128 Leaf Switches
  - 2 per rack, 64 racks: 2 * 64 = 128
- Whilst 16 Spine Switches can be used, you must remember that each Spine Switch must connect to every Leaf Switch. Therefore each Leaf Switch would require 16 Spine to Leaf connections

DELL EMC | Microsoft

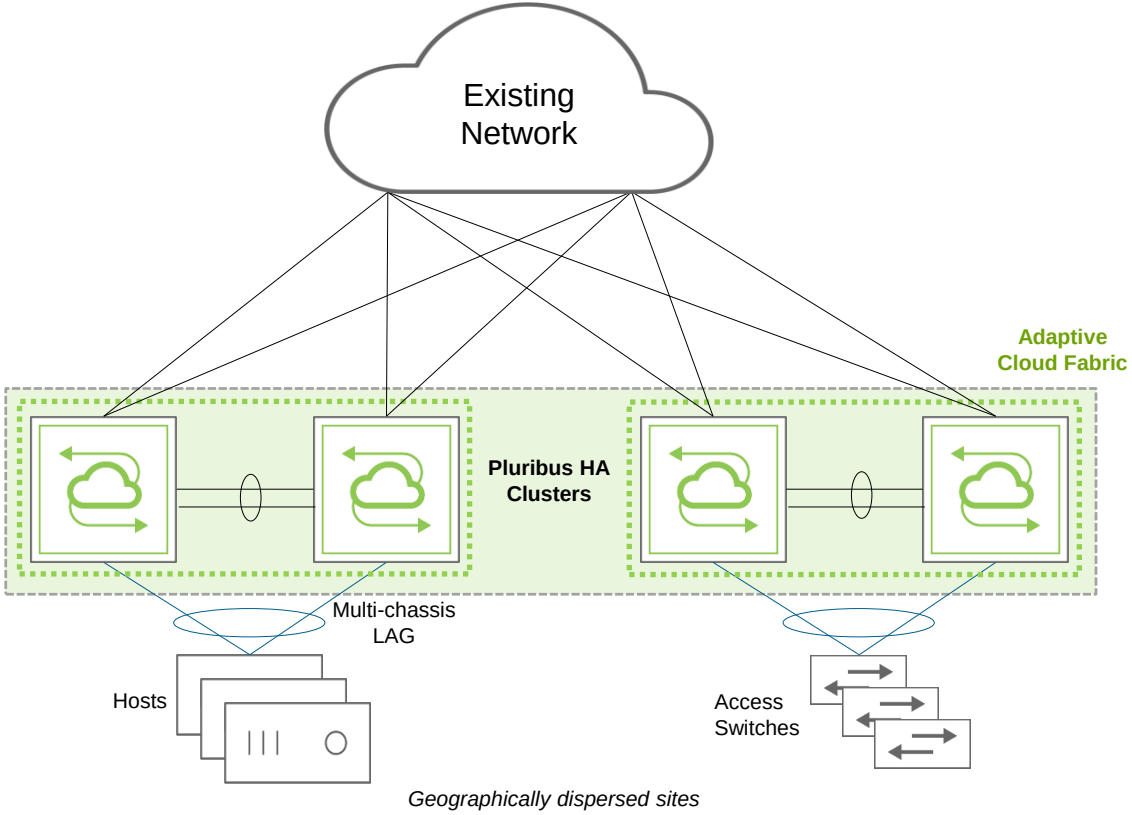
e with legacy

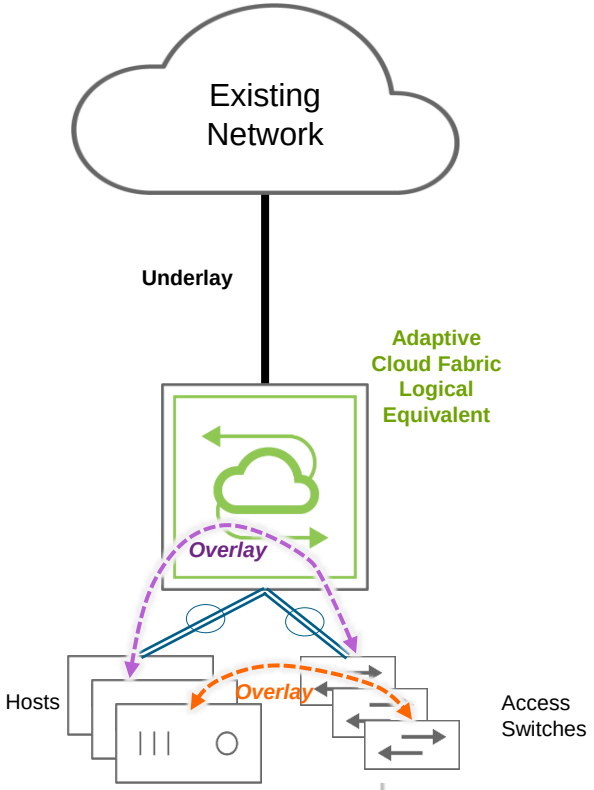
POD-Level deployment in (Remote) Data Center



DELL EMC | Microsoft

Pluribus Adaptive Cloud Fabric

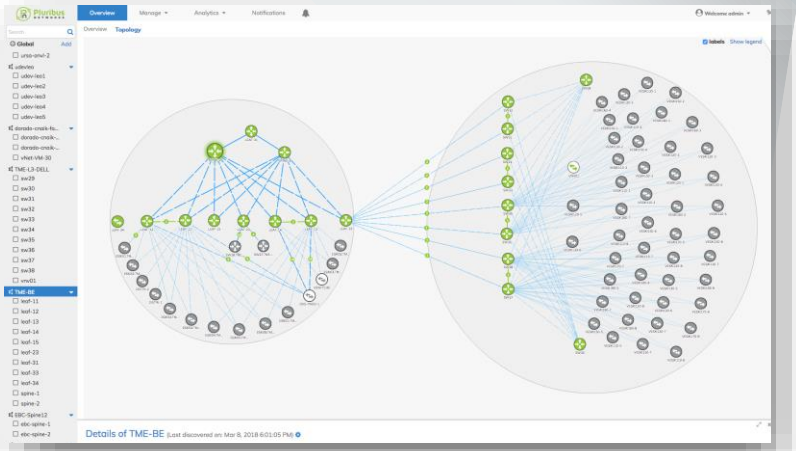




Pluribus UNUM™ Unified Management, Automation and Analytics

Virtual appliance

Fabric Manager

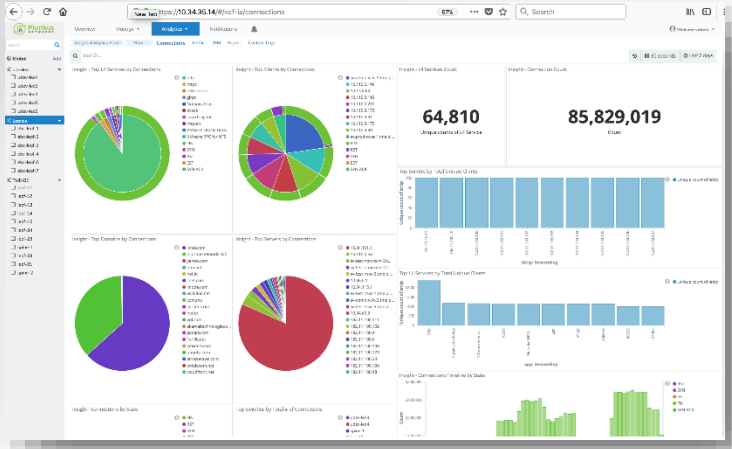


Device & Fabric Management

Automated Provisioning

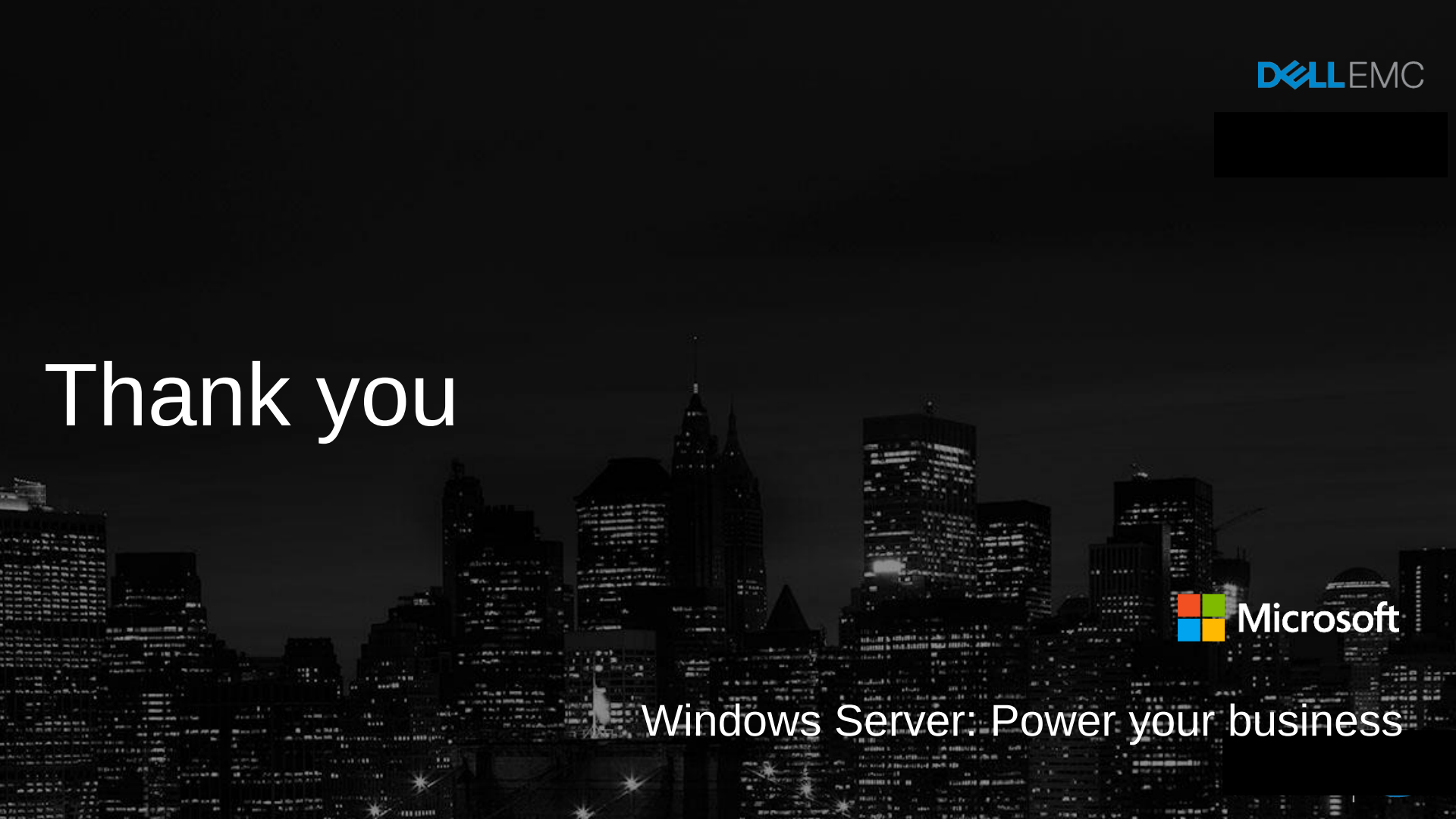
Topology Visualization

Insight Analytics



Analytics & Performance Management

DELL EMC | Microsoft


DELL EMC
Thank you
Microsoft
Windows Server: Power your business