



The bridge to possible

# Building High-Performance Network Services with Cisco Smart Switch: Microsoft Success Story

Kristina Moore – Microsoft Principal TPM

Robert Murphy – Cisco Principal Architect

Shane Corban – AMD Technical Product Management

CSSSPG-1015

CISCO *Live!*

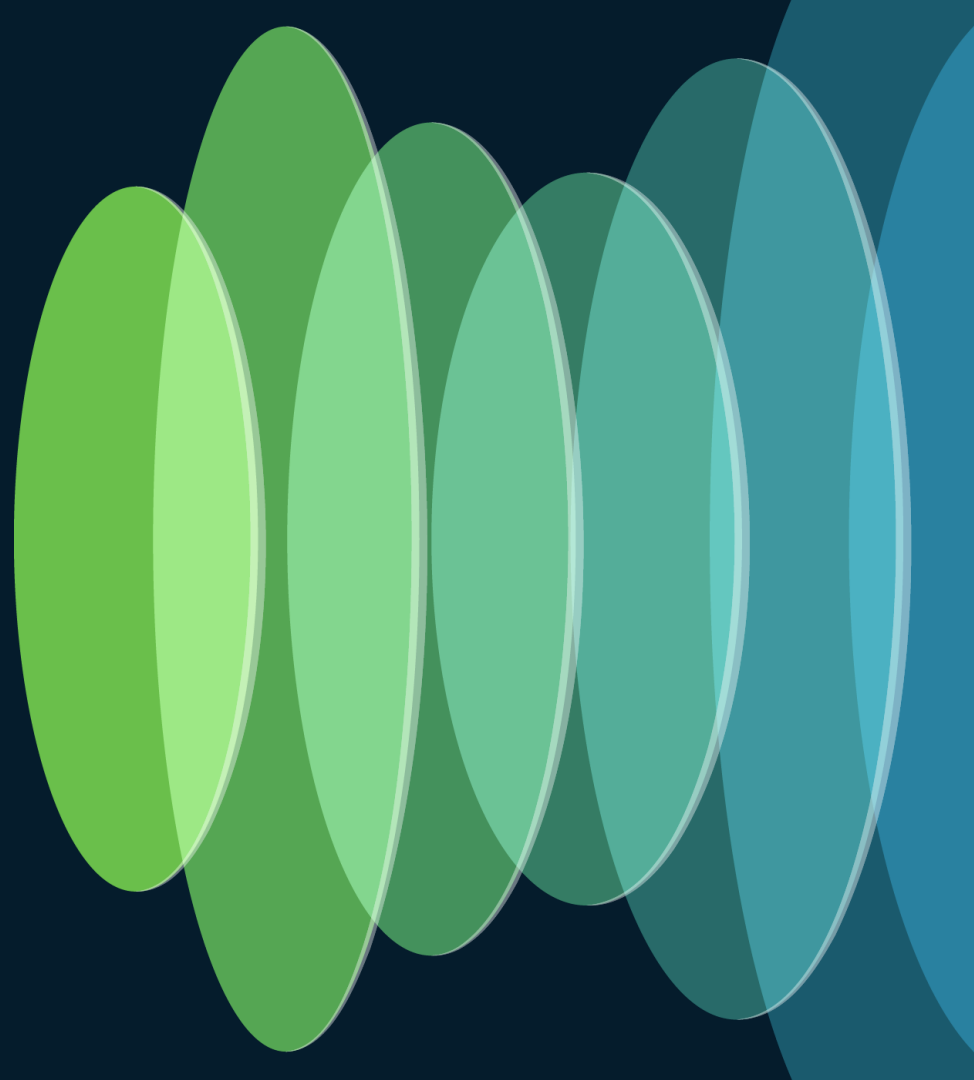
#CiscoLive



# Agenda

- The SDN Challenge
- Platform Introduction
- DPUs Accelerating SDN
- Pulling it All Together

# The SDN Challenge





- Available regions
- Regions coming soon
- Edge zones
- Network PoPs
- WAN links
- ★ Ground stations

**65+**  
Azure  
regions

**200+**  
datacenters  
worldwide

**175k+**  
miles of fiber

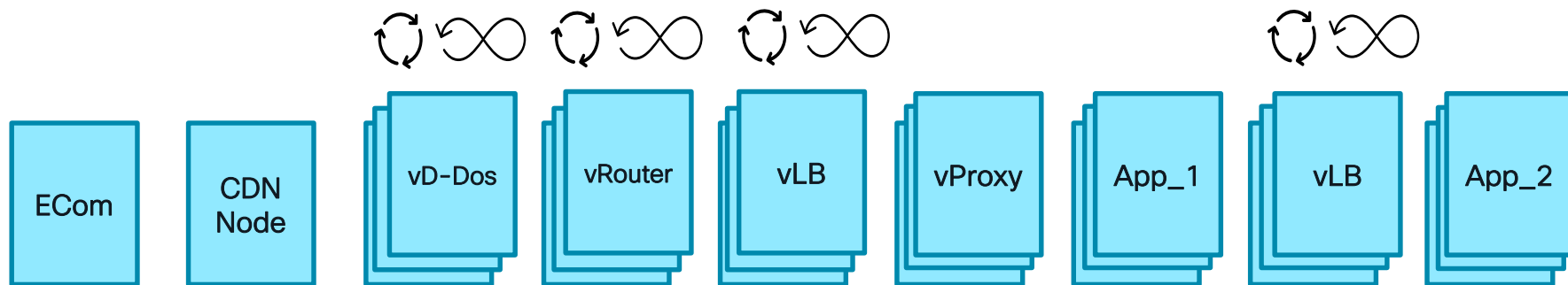
**190+**  
network PoPs

Network extends to space with Azure Orbital ground stations

**CISCO** *Live!*

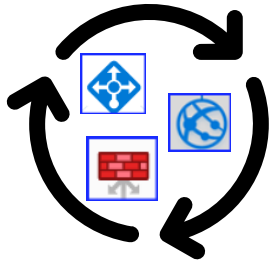
# Cloud Example of Virtual Appliances

Processing Millions of Simultaneous Users



- Millions of users are directed through layers of NVAs.
- Type of Services: Content Caching, Load Balancers, Firewalls, DDoS, Authenticators, Billing Systems, and Applications

# High-Scale, High-Performance, Connection-Oriented

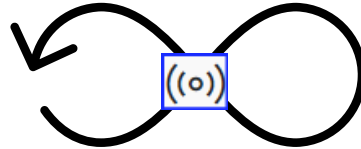


## Connections per Second

95% require < 10s of k/sec

<5% require 100s of k/sec

<1% require 1M+/sec



## Active Connections

10% of connections active

30% of connections active

100% of connections active



## Total Connections/Flows

100k connections

1M connections

10M connections

# Redefining SDN Services – Cloud Based Connection Mgmt



## Increased Performance

VM bottlenecks; Network Virtual Appliances can now rival bare-metal performance



## Decreased Costs

Customers application scale is decoupled from network service scale allowing smaller VMs



## Increased Control

Enables cloud flexibility by allowing/*selecting* performance levels at vNIC level



## Leader in Flexibility

Allows customers to enhance existing VM network capabilities without a new VM SKU



## Robust Ecosystem

Service definition through open-source DASH project with community contributions



## Programmability

Services definitions defined through DASH and consumed through open-source APIs

# SDN Service - Today

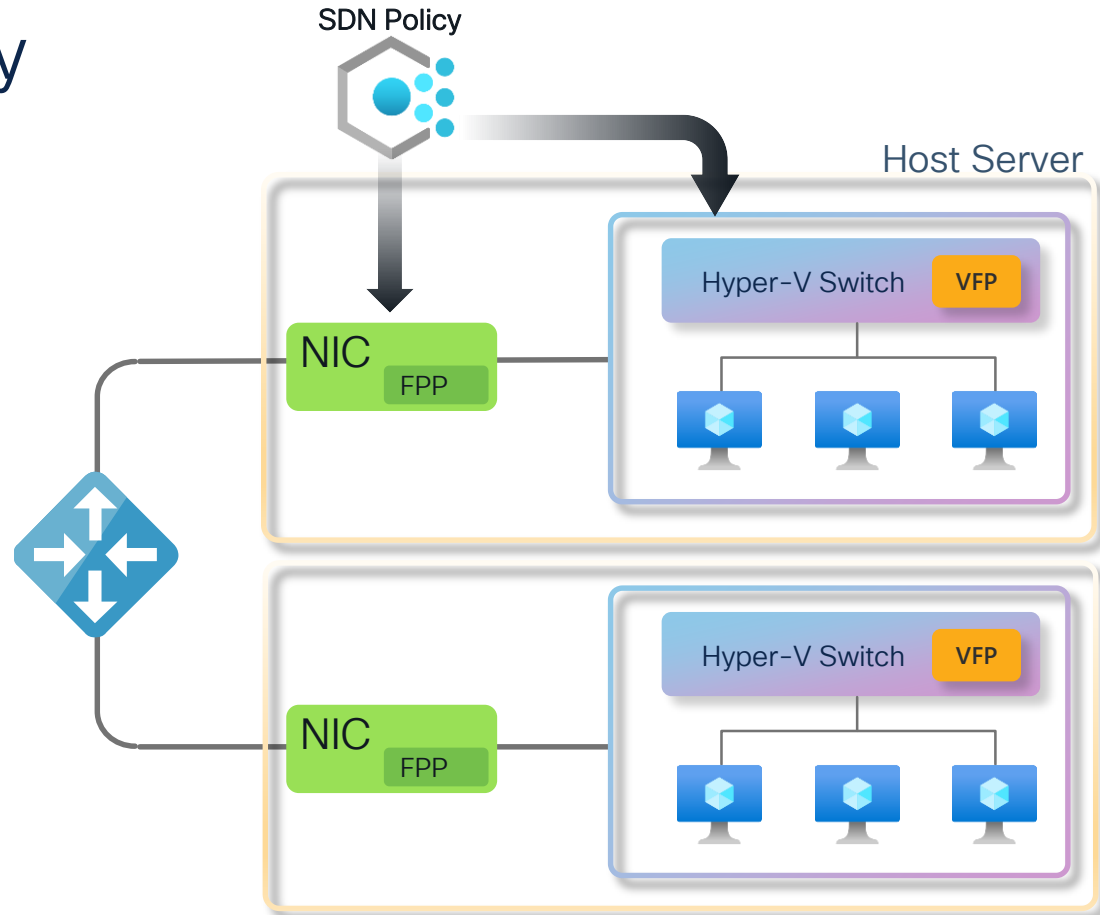
Network policy processing today is performed on a host node in a virtual switch (limits)

---

Demanding customers use a high number of SDN rules (ACL scale)

---

SDN policy evaluation can be resource intensive and complex. (billing, metering, etc... before session setup)





# SDN Service - Enhanced

Offload network policy processing to DC network HW

Complex policies are rendered in switch DPU

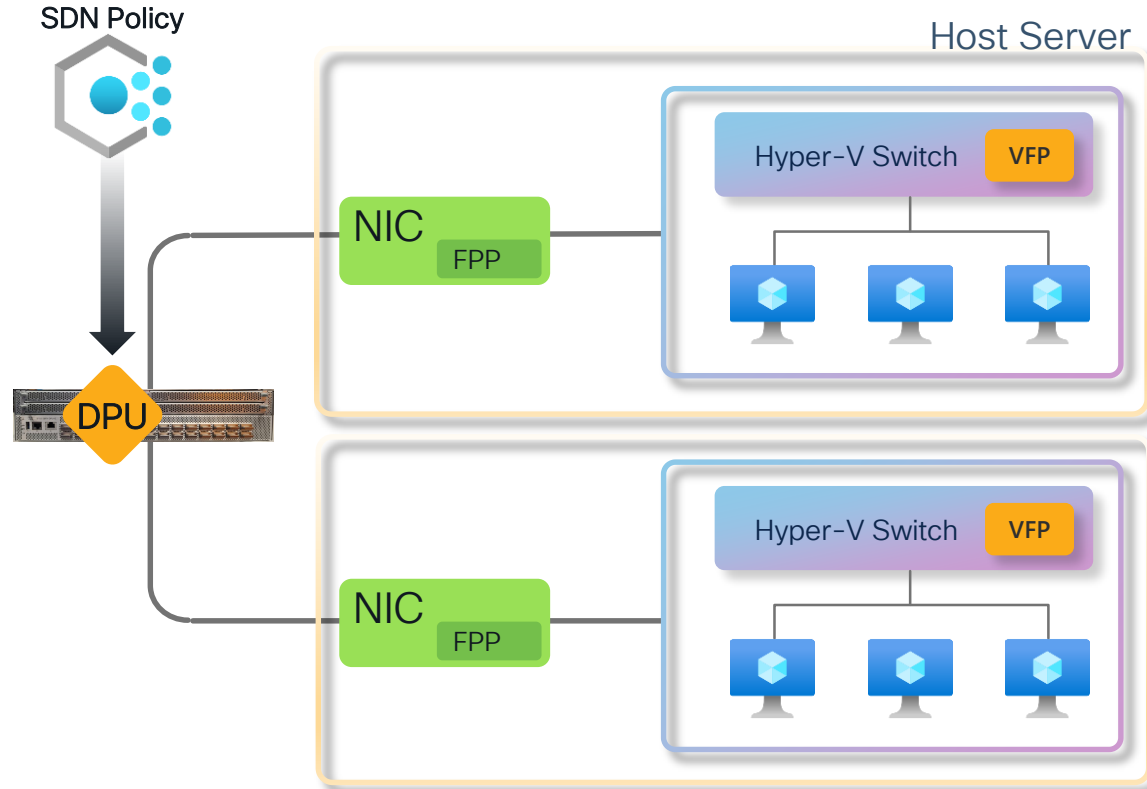
Service growth enabled through hardware scale in tier

**24+**

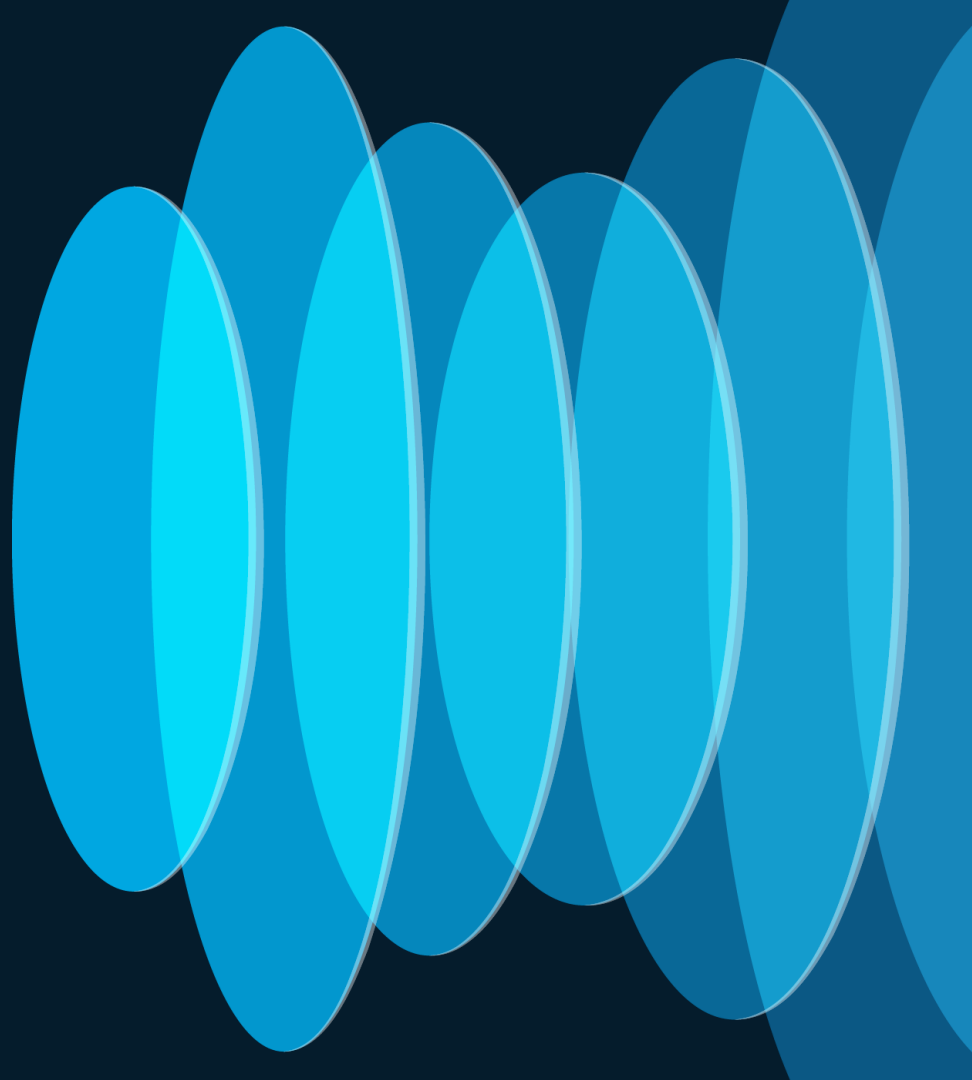
Million CPS

**400+**

Million PPS



# New Cisco Platform Introduction



# Smart Switch - New Software & Hardware Model

## DASH

Open Source Project for connection based services delivering both the model & API



## Cisco Silicon One

High performance and power efficient routing silicon with a P4 programmable pipeline



## Smart Switch

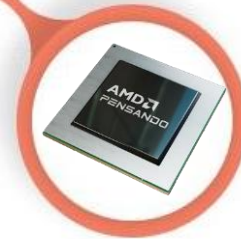
## SONiC

Open Source Project delivering a Network Operating System hosted in the Linux Foundation



## AMD DPU

Pensando DPU delivering service offload to dedicated silicon at 200G per unit



# Introducing the New Cisco 400G Service Accelerated Switch

Q4 CY24

Cisco Silicon One Q200L

Supports 28x400G QSFP-DD

1.6T of DPU Services

Using 8 AMD Elba DPU

Four field replaceable sleds

SONiC Operating System

SONiC instance on switch & DPU SLED

Upgraded System CPU & Memory

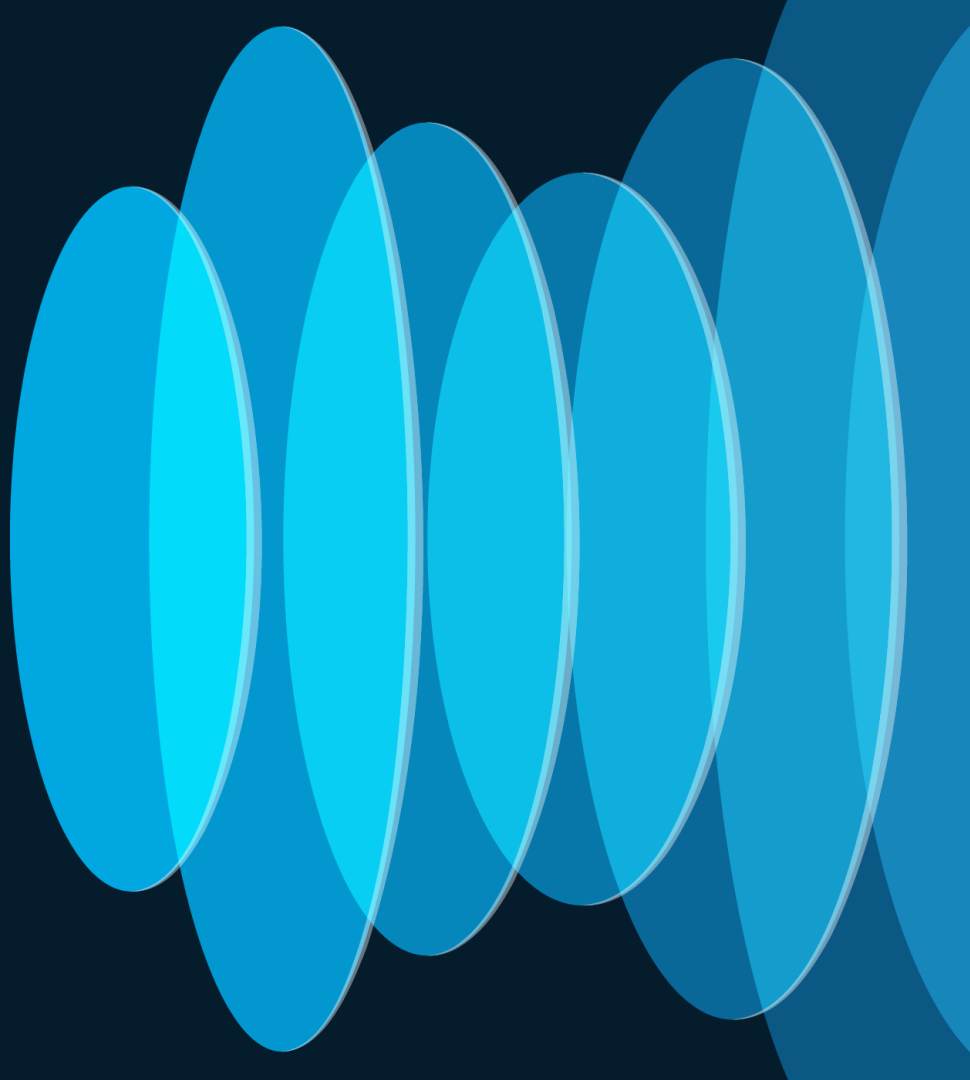
Cisco 8102-28FH-DPU-O



8K-DPU400-2A



# DASH Service Stack

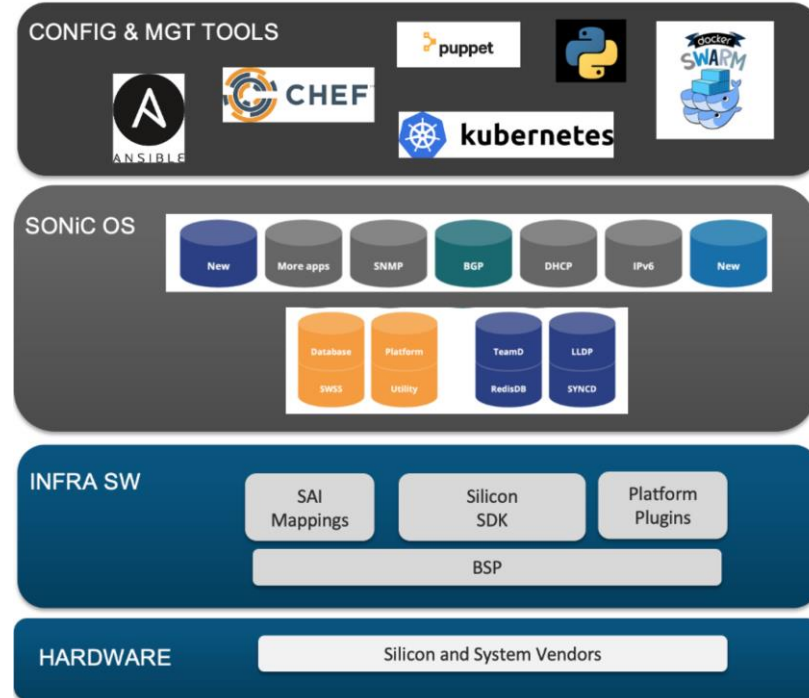


# What is SONiC?

Software for Open Network in the Cloud



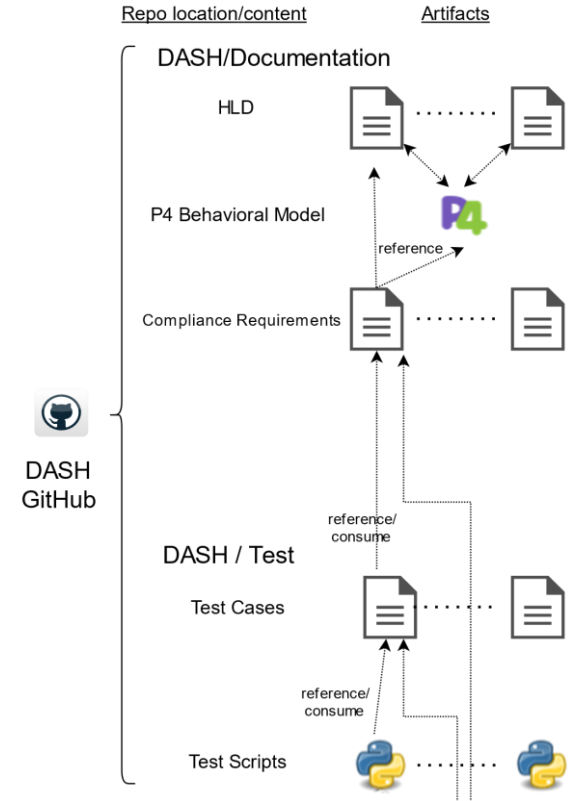
SONiC is an open source network operating system (NOS) that is supported on multiple vendors switches and ASICs.



# What is SONiC-DASH?

## Disaggregated APIs for SONiC Hosts

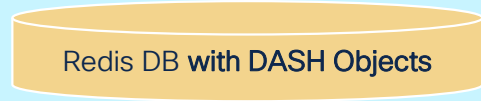
- Built upon a SONiC Underlay dataplane
- DASH services are P4 based and define the behavioral model
- Designed to achieve 10-100X performance over todays cloud offerings
- IPU/DPU vendor agnostic
- DASH services defined in public GitHub repo
- Test-suite are written against SAI-DASH



# SONiC DASH

## SONiC Underlay + DASH Overlay for Services

SDN North Bound Interface  
(gNMI)



South Bound Interface

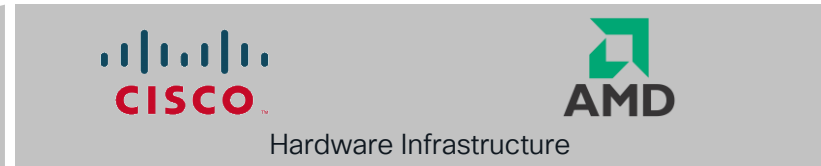
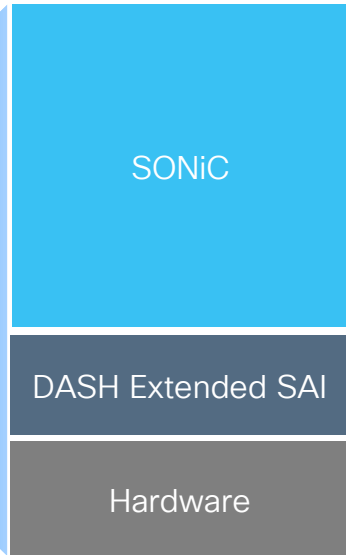


NB APIs

SB APIs



Switch Abstraction Interface - (SAI) DASH Extensions

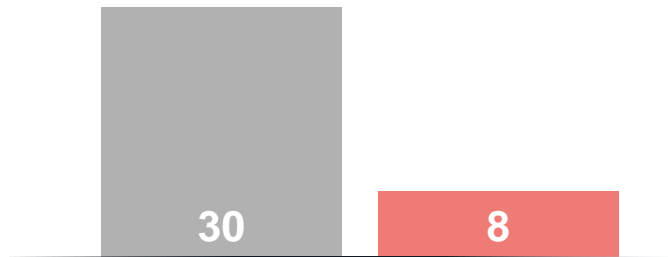




# Azure Benefits Accelerated NVA

## Decreased CPU Utilization

**22**  
Core savings



3rd Party SLB  
Standard Azure  
Offering

3rd Party SLB with  
Azure Accelerated  
Offering

**CPU Cores**

Consumed on x86

## Customer Benefits

- Frictionless Insertion
- >10x performance versus previous thresholds
- Total active connections increasing up to 2x
- Increased connections per second (CPS)
- Consistent active connections
- Increased CPU stability for high traffic network optimized VM
- Reduced jitter

**Future Services:**

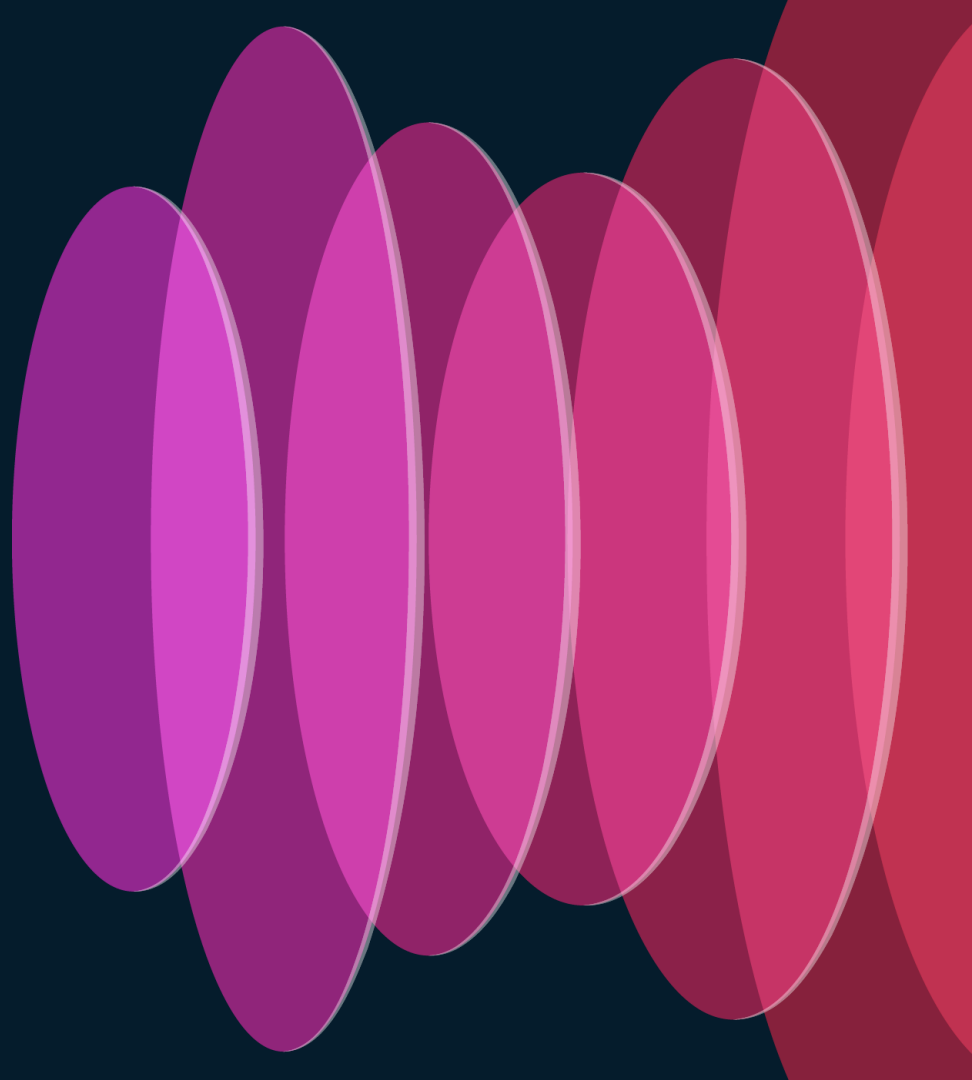
DPU Accelerating Floating NIC & Key Exchange

*“We have seen a 40x improvement in overall cloud based connection related performance... This partnership has enabled a transformational limitless networking”*

Deepak Bansal  
(Azure Core Engineering, CVP)



# DPU Accelerating SDN



# Industry Dynamics – Relentless Increase in Speeds & Feeds

Architectures are at an Inflection Point, how to deliver performance at scale

## Tug-of-War between

### ASIC-like Fixed Function Logic

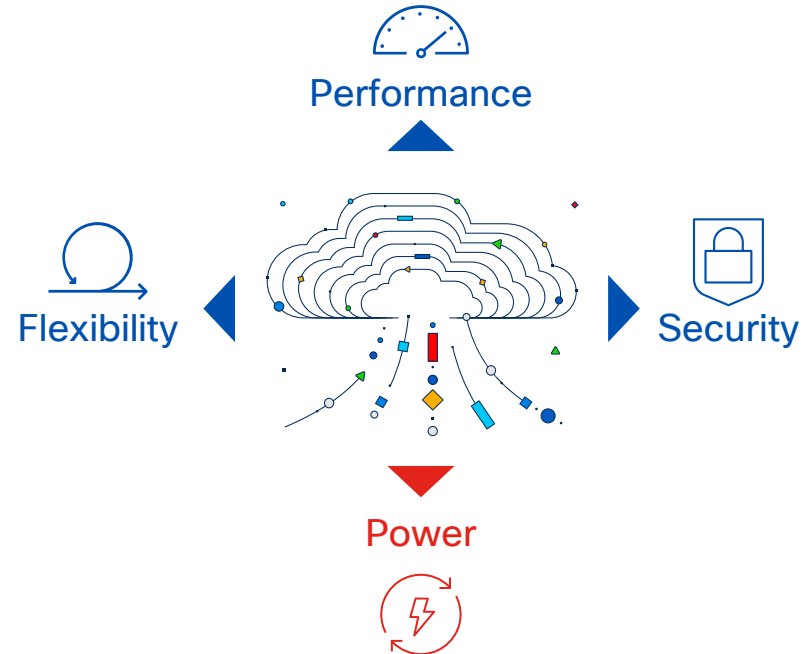
Frequent vs occasional functions

### Embedded Processor Cores

Workload flexibility vs processing offload

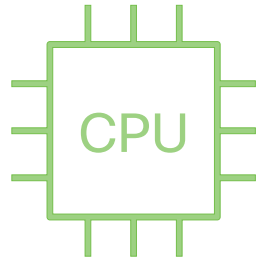
### FPGAs or Programmable Logic

Adapting new and existing functions

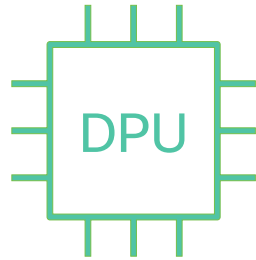


# Where does a DPU fit in the ecosystem?

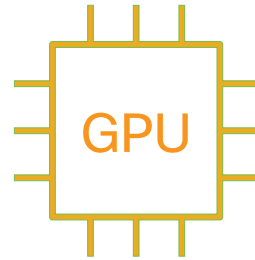
## Understanding the Fuzzy Marketing Terminology



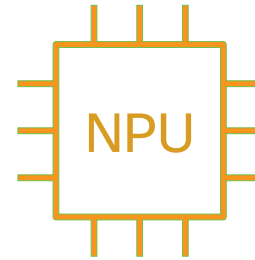
- **Strength** in serial processing tasks
- Small number of cores
- Highly versatile through software



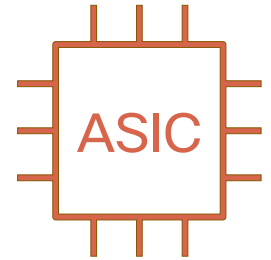
- **Strength** in data processing tasks
- A few dozen number of cores.
- Optimized for Network, Security, and Storage tasks



- **Strength** in Massive Parallel Processing
- Many 1000s of cores of different types.
- Heavy mathematical tasks such as graphic processing or AI



- **Strength** efficiency + reasonable flexibility
- Small number of cores
- High performance network processing

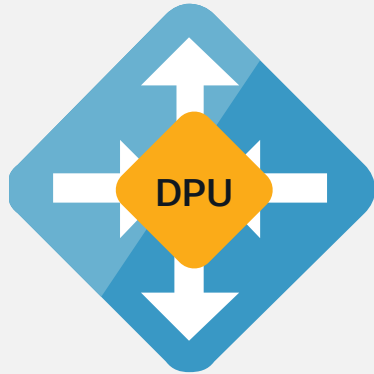


- **Strength** optimized for an application
- None or 3rd party cores
- Application specific

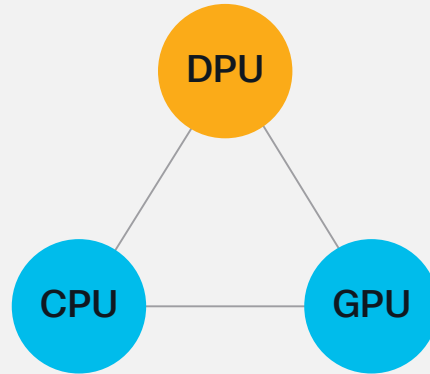


# What is a DPU?

Solution to scale-out services at 400G, without impacting performance



New Generation of Domain Specific Programmable Processor for Infrastructure Services



Third Pillar / Socket in the Datacenter Alongside CPU's, GPU's



Networking  
(SDN, Routing/Overlays, NAT, SLB...)



Security  
(Firewall, DoS, IPSec Encryption)



Telemetry  
(XDR, NPM...)



Storage  
(NVMeoTCP, RDMA, Compression, Dedupe, Encryption, Checksum)

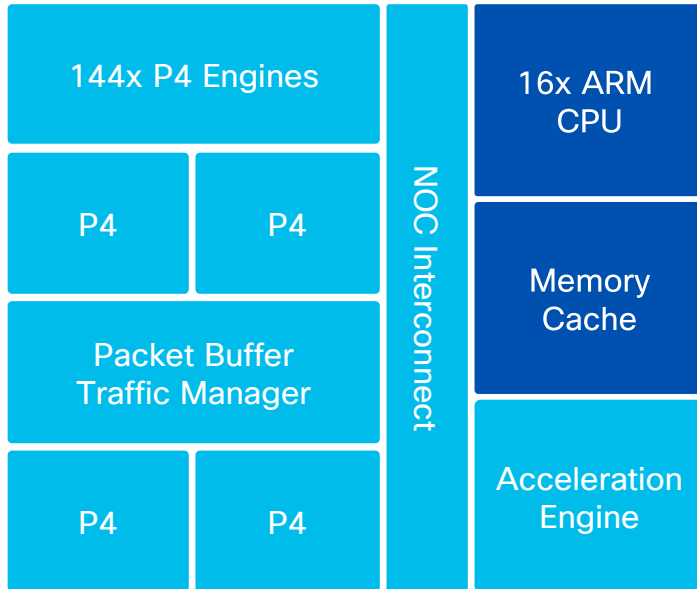
Infrastructure for Wire-Rate Acceleration of Software Defined Stateful Services

# AMD Multi-Generational DPU Architecture

High performance and Scale with the Flexibility of a FULLY P4 Programmable System



## DPU Architecture

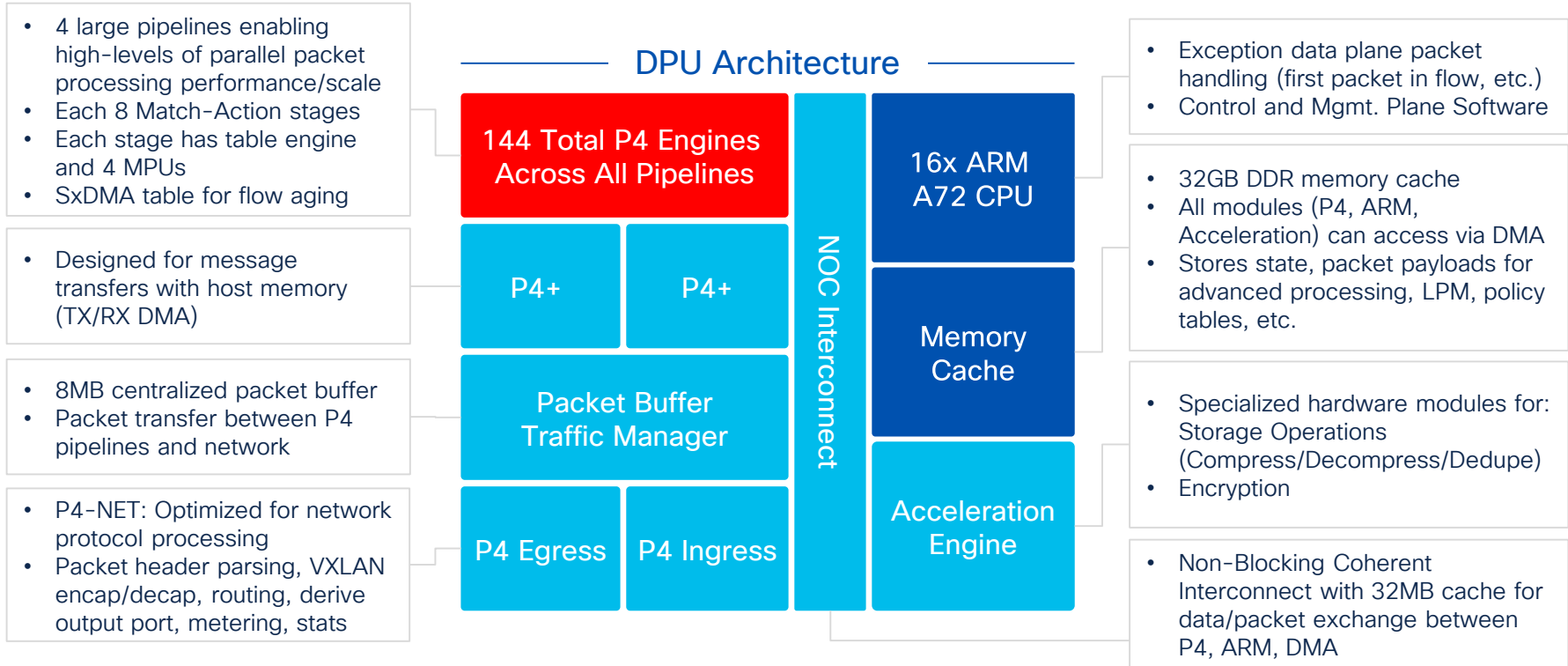


- Fully programmable
  - Control, data plane and management plane
- Order of magnitude performance and scale
  - Up to 100x session setup rate
  - Up to 2-3x packet rate
  - Tens of millions of active sessions
- Concurrent stateful services at line rate

# AMD Multi-Generational DPU Architecture



## Functional Components & Operations





# Performance Benefits for Cloud Providers with AMD Elba

## Orders of Magnitude Ahead of Competition and Traditional Approaches

	AWS NITROv5	AWS NITROv4	AMD DPU ELBA
AVG Latency (64 byte UDP)	28 $\mu$ sec	38 $\mu$ sec	<b>3 <math>\mu</math>sec</b>
Packets per Sec (96 bytes)	24M	2.7M	<b>60MPPS</b>
Throughput (1500 bytes)	100Gbps	15Gbps	<b>200Gbps</b>
Average Connections per sec	500K	64K	<b>5M</b>

BM-to-BM      VM-to-VM      BM-to-BM

Increased Revenue per Rack


Better Cloud Performance


More Density per Rack


Delivering Performance of Hardware with Programmability of Software



Full Stack Solution


# Enterprise DPU Use-Cases: DC East-West & Edge Security



 **Native Security**

 **Datacenter**

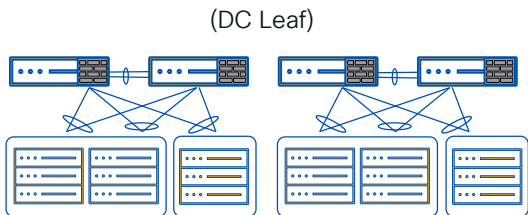
 **Hybrid Multi-Cloud Security**

  **DCI / Colocation Edge**

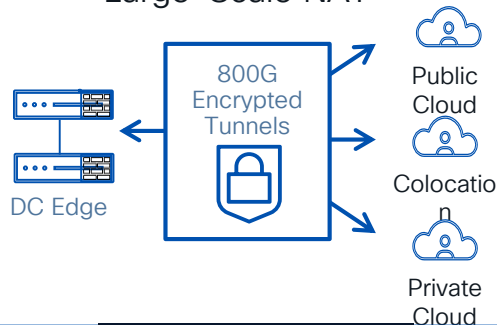
 **AI Ready Telemetry**

  **Datacenter**

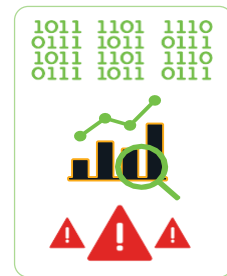
Zero-Trust, Micro Segmentation, Firewall, DoS Protection



IPsec VPN, Firewall, Large-Scale NAT



Event Based Telemetry for NPM and XDR



**TCO Savings**

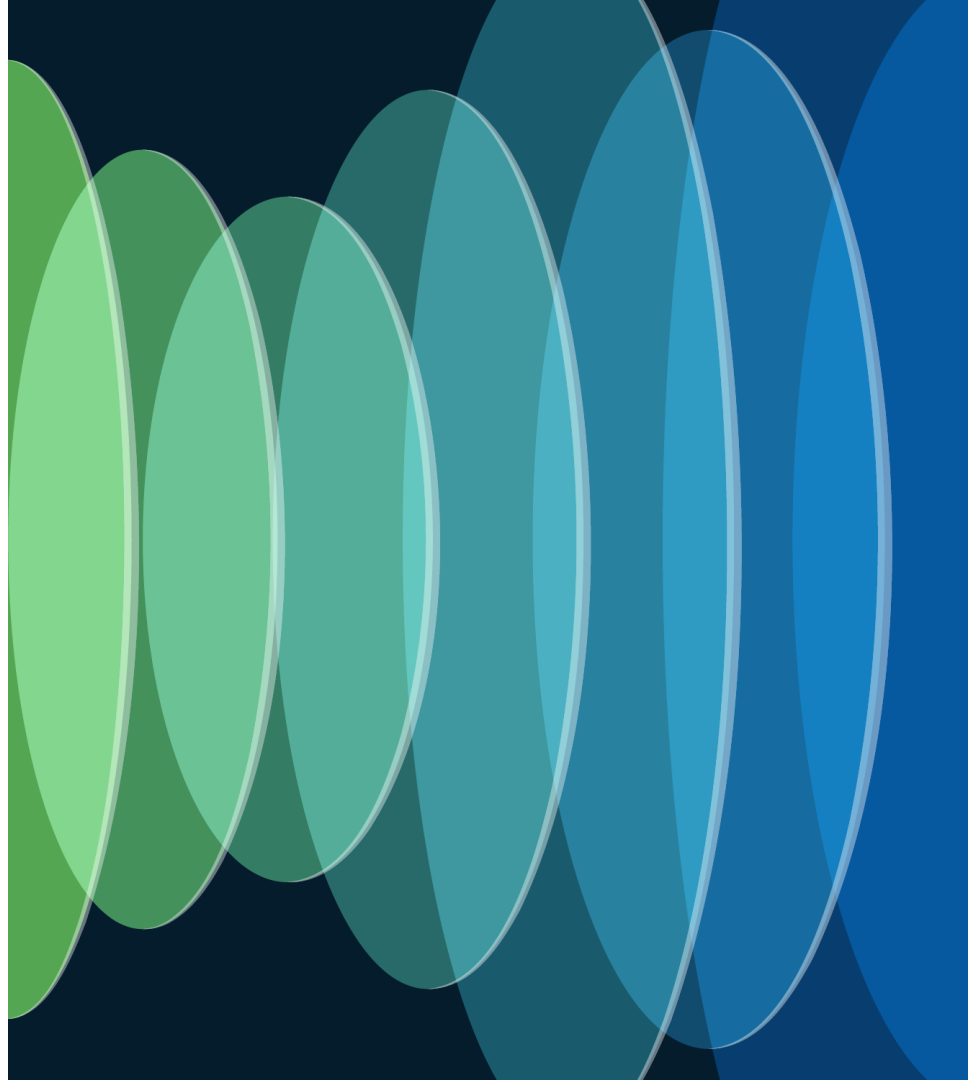
~70%

~89%

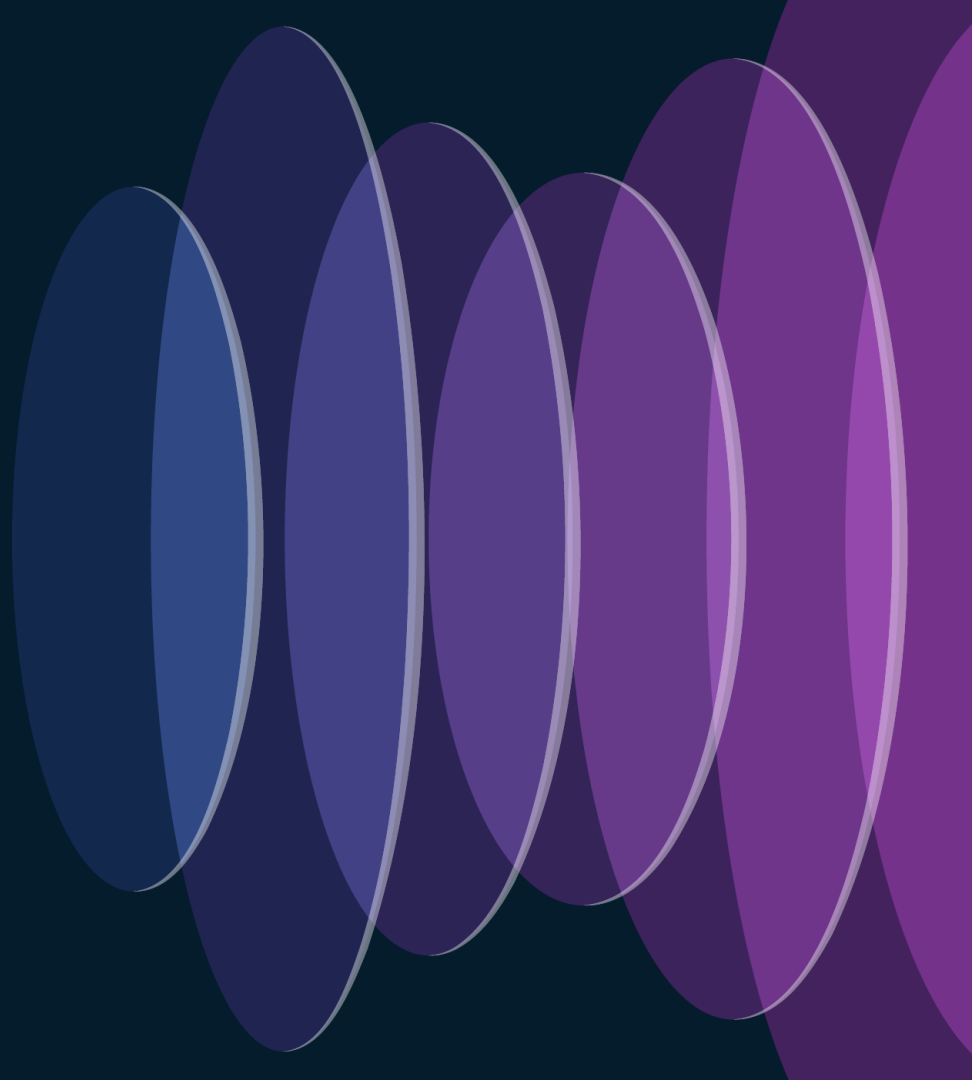
~50%

# Demo

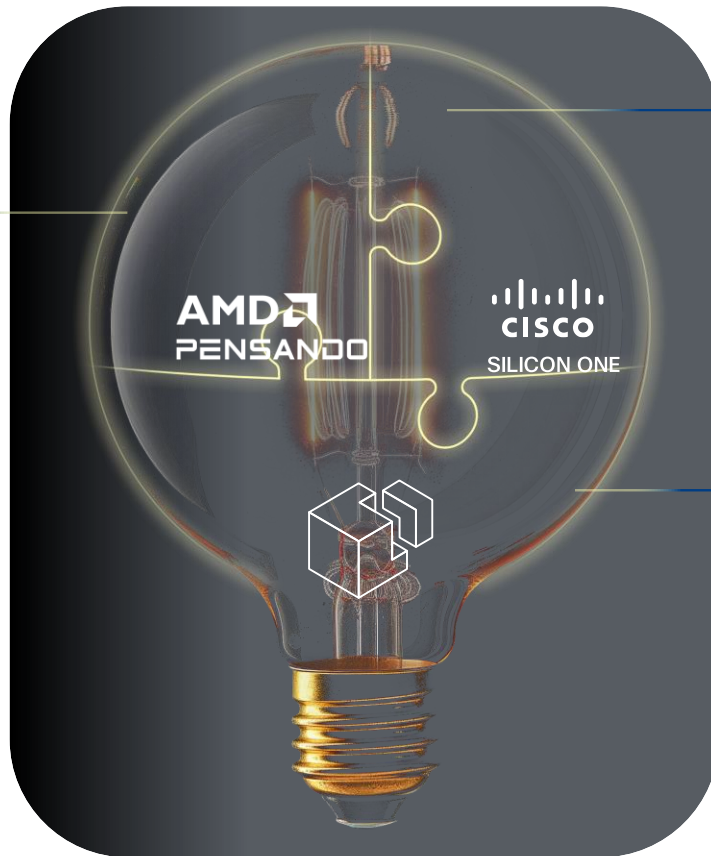
CISCO *Live!*



# Pulling It All Together



# Technology Partners Solving Complex Challenges



## AMD DPU

Provides service acceleration at wire-rate performance



SDN POLICY



NAT



FIREWALL



IPSEC



TELEMETRY

## Cisco Silicon One

12.8T routing silicon with P4 programmability  
(Public/Private Cloud - Committed)

## Network OS

DASH+SONiC (Cloud)  
Enables SDN control of both services and networking policy

# Cisco 8102-28FH-DPU-O



The Cisco 8102-28FH-DPU-O and are fixed form factor (2RU) switch with 28xQSFP-DD56 interfaces capable of 12.8T of throughput. The switch has four, port facing, field replaceable DPU sleds.

Physical Summary
NPU: Single 12.8T Q200L
CPU: 16 Core x86 CPU 128GB DRAM
DPU: AMD Elba 16 ARM CPU Cores 32G Memory
4 x DPU Sleds Field Replaceable Units
Software Operating System SONiC
3+1 Fan, 1+1 PSU Redundancy Port-side Intake Airflow
Interface Support: 400/200 /100/50/40/25 GE

## Highlights:

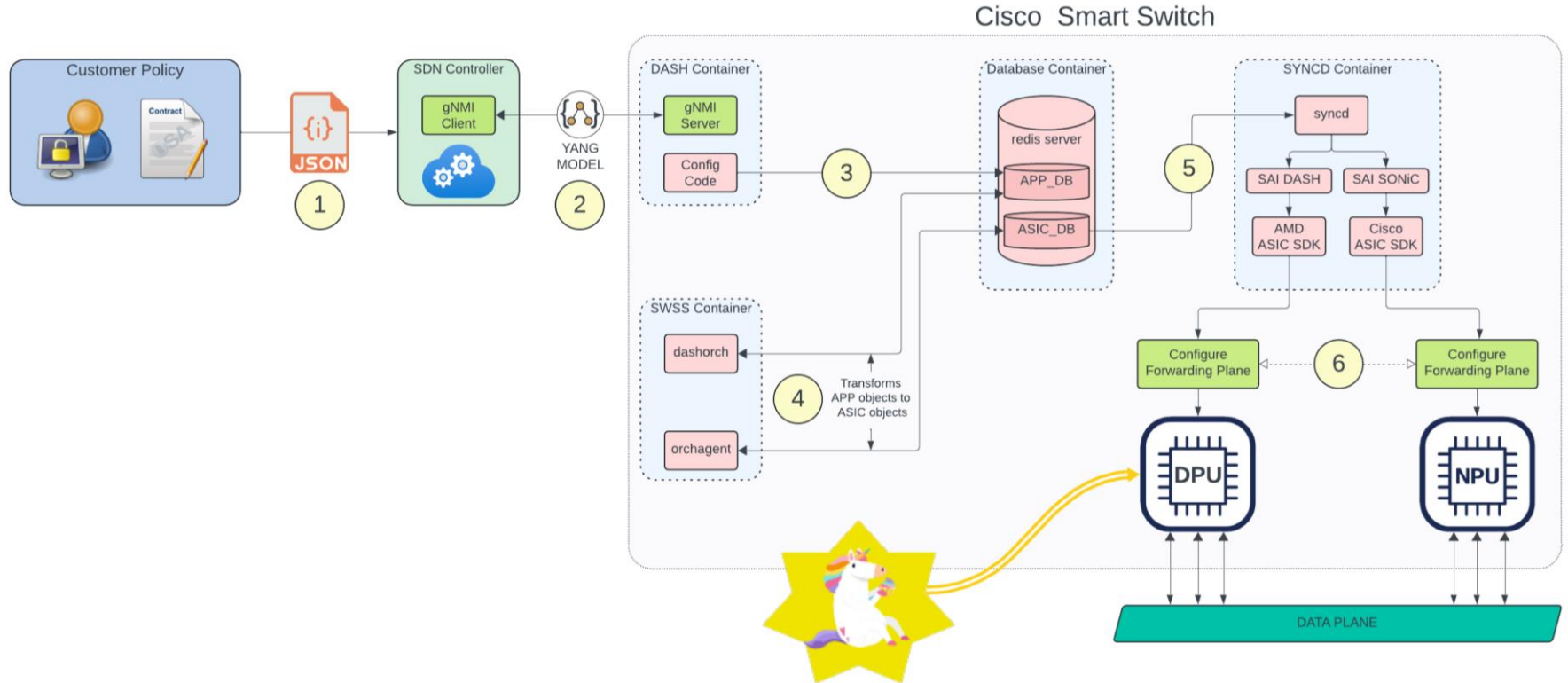
- Increased CPU cores and memory for future workloads
- SAI/SONiC support (202405)
- 28x400G QSFP-DD56 switch optimized for power efficiency
- Support up to 1.6T of DPU services
- RDMA support for AI and storage workloads

## Port Side:

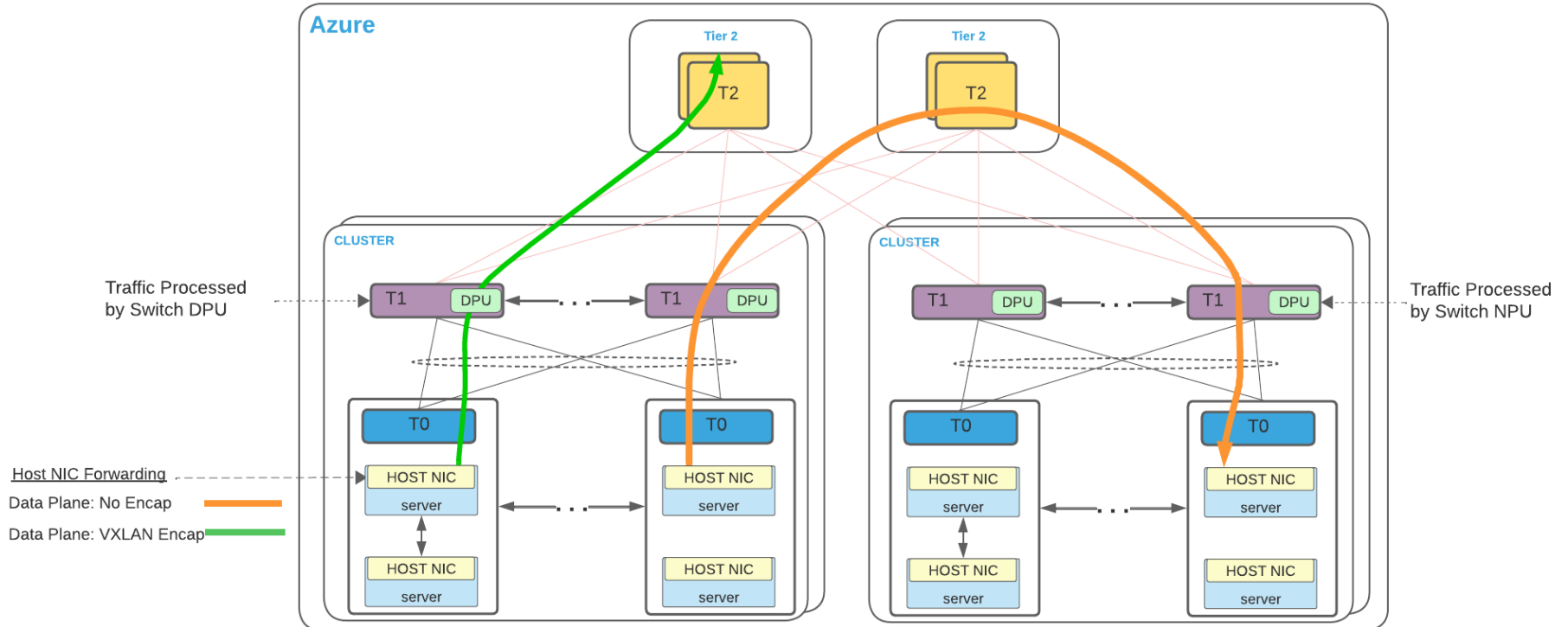


# Peeling Back the Onion

## Understanding the Flow of Policy



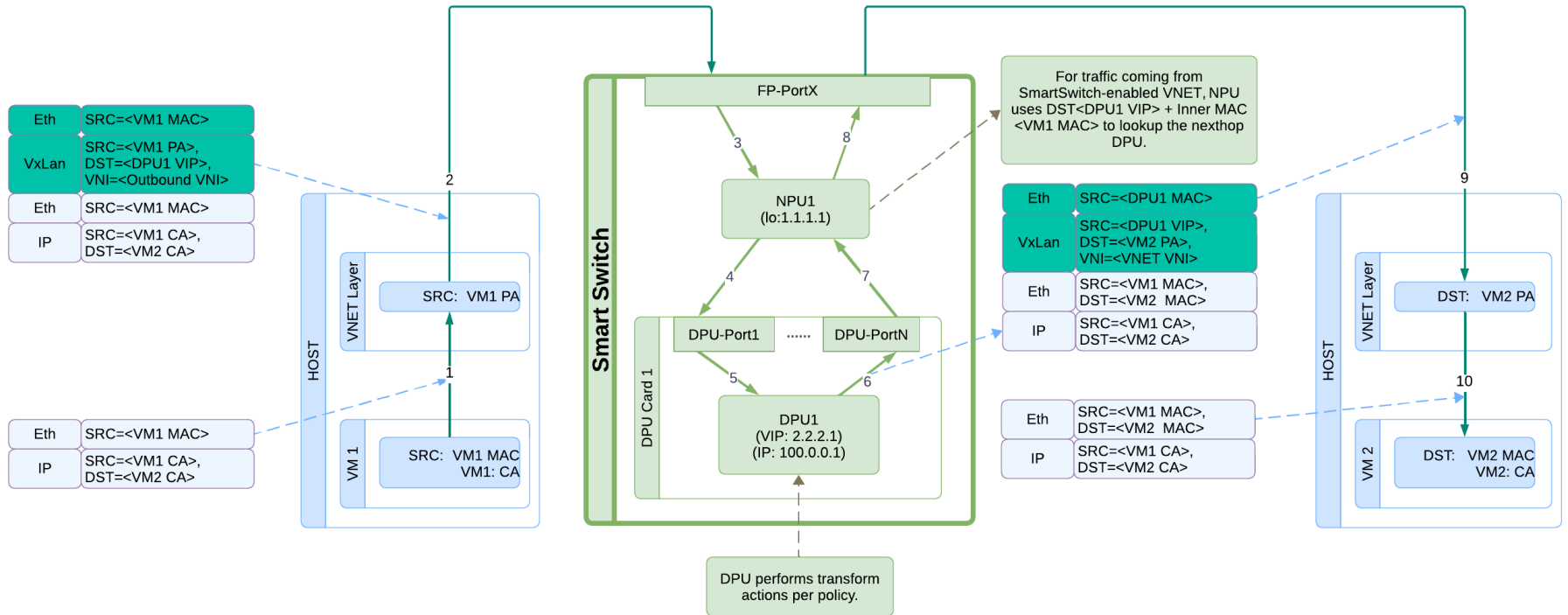
# Smartswitch in DC Fabric





# DPU Magic in Service Graph

## VM to VM in VNET Communication Example

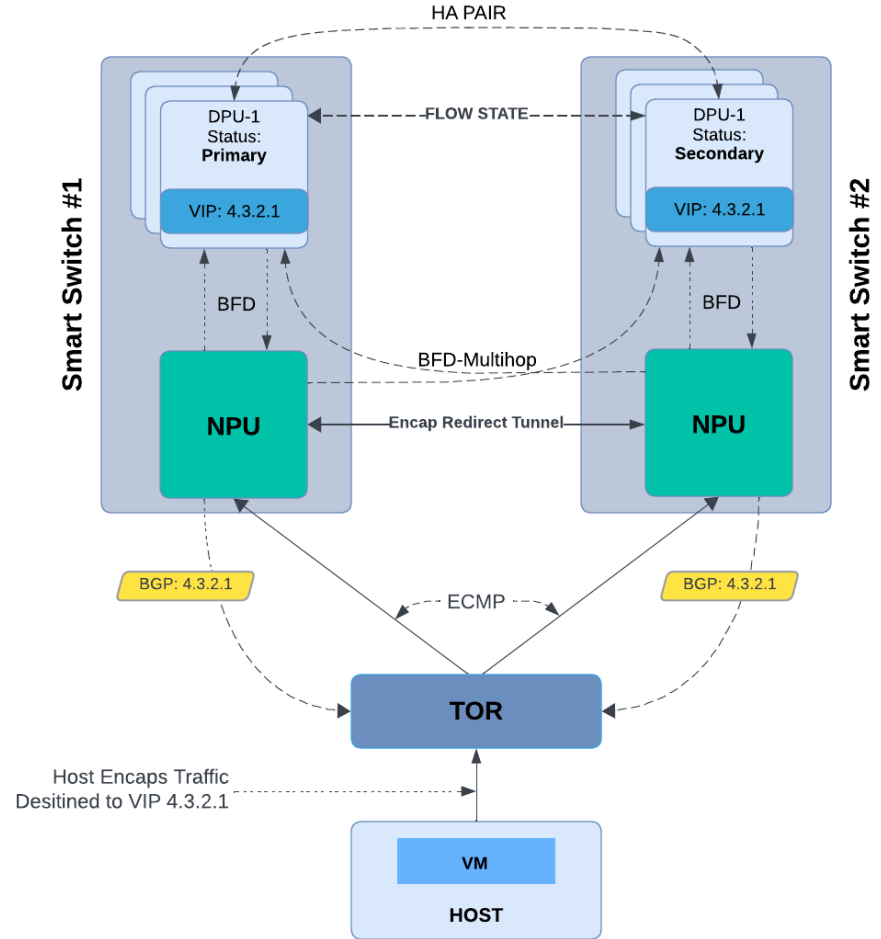


Encap: VXLAN, GRE, or NVGRE

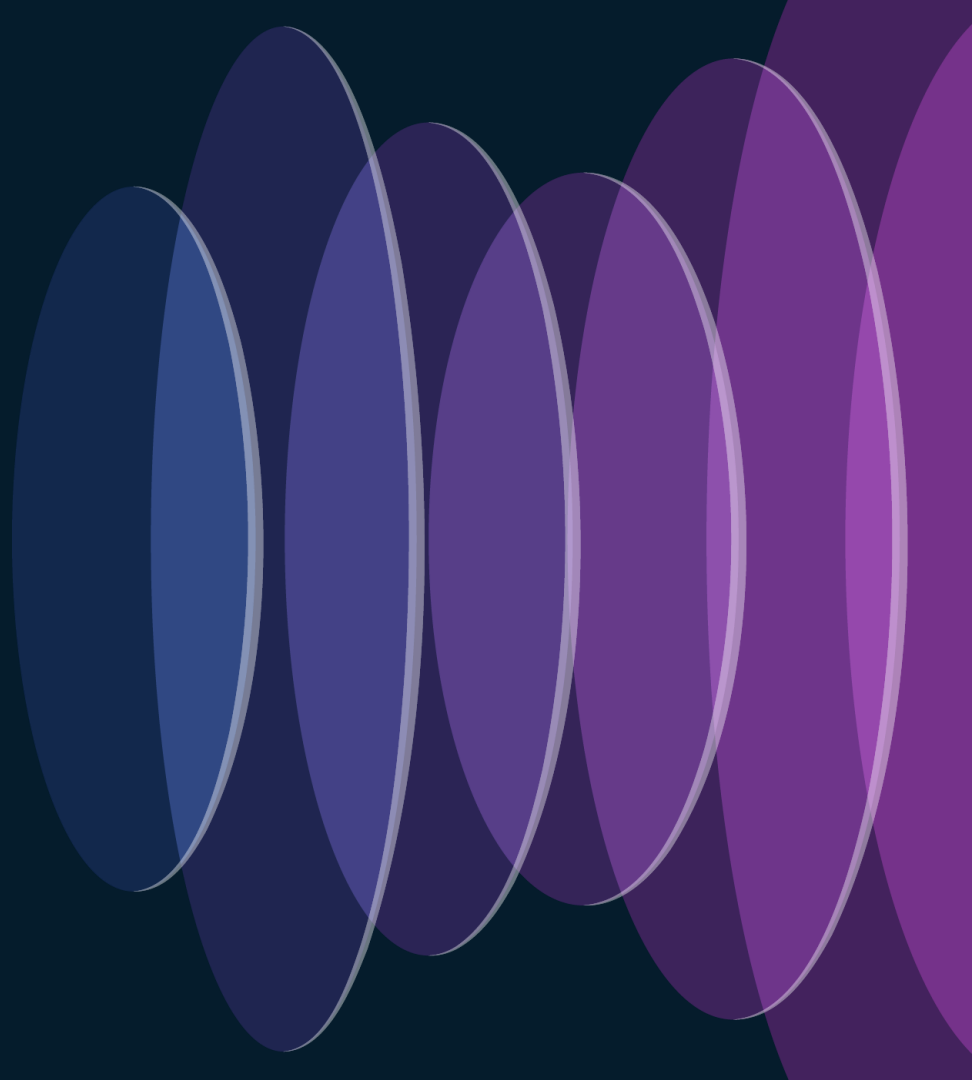
# High Availability

## Understanding HA in Context

- DPU Provisioned in HA pairs
- BFD used for DPU link state
- Flow state mirrored to Standby
- TOR uses ECMP for forwarding
- Encap tunnel failover / redirect



# Looking Forward



# What's the Big Picture

- The 8102 Smartswitch project will deliver an order of magnitude increase in NVA services for Azure.
- Creation of limitless networking enabling the migration of traditional network services to the 8102 Smartswitch.
- A TCO savings of 73% in just server cores, allowing Azure to repurpose cores for revenue generation.
- Contributions are welcome in the DASH open-source community



# Further Learning

- Disaggregation of Stateful Network Functions Research Paper:  
<https://www.usenix.org/system/files/nsdi23-bansal.pdf>
- DataCenter and Cloud Networking – AMD DC and AI Technology Premiere:  
<https://youtu.be/t43v1MvqV-E?si=QGbOSxf5m5W23iq->
- DASH Github: <https://github.com/Azure/DASH>
- SONiC Github : <http://sonicfoundation.dev/>
- Mailing lists: [sonic-dash@googlegroups.com](mailto:sonic-dash@googlegroups.com)  
[sonicproject@googlegroups.com](mailto:sonicproject@googlegroups.com)

# Cisco Webex App

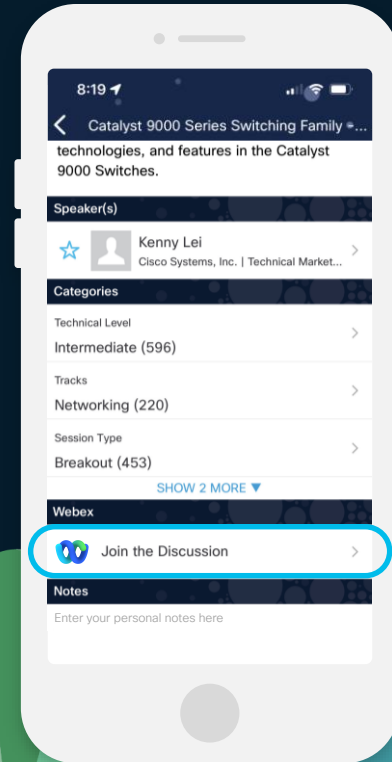
## Questions?

Use Cisco Webex App to chat with the speaker after the session

## How

- 1 Find this session in the Cisco Live Mobile App
- 2 Click “Join the Discussion”
- 3 Install the Webex App or go directly to the Webex space
- 4 Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 7, 2024.



# Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Engineer meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at [www.CiscoLive.com/on-demand](http://www.CiscoLive.com/on-demand)



The bridge to possible

# Thank you

CISCO *Live!*

#CiscoLive