Accelerating Telco NFV Deployments with DPDK and SmartNICs

Kalimani Venkatesan G, Aricent
Kalimani.Venkatesan@aricent.com

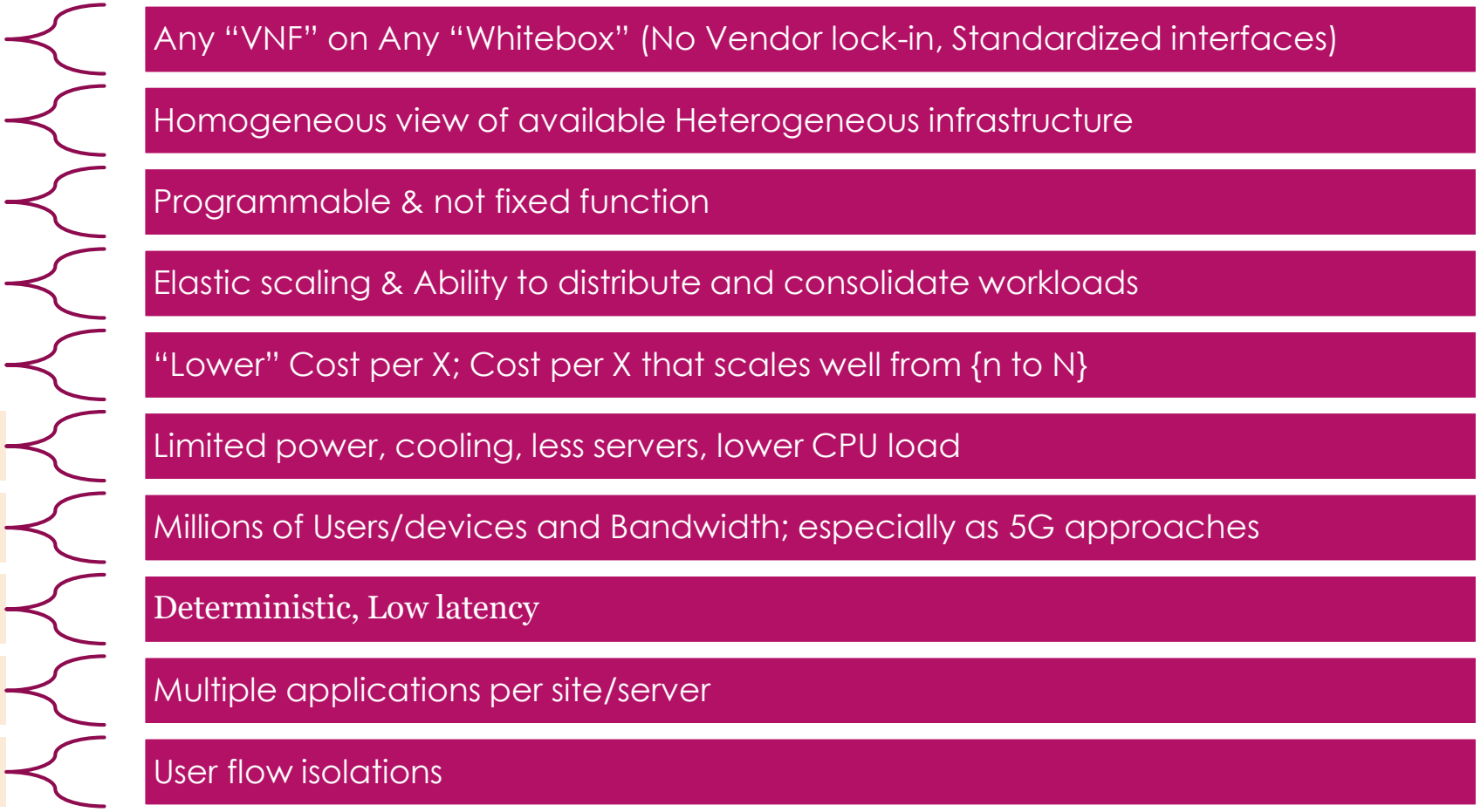
Barak Perlman, Ethernity Networks
Barak@Ethernitynet.com

DPDK Summit North America 2018
Dec 3rd 2018

Outline

- Telco requirements for NFV
- SmartNIC Acceleration
- DPDK and SmartNICs
- Software architectures for Telco VNFs with SmartNICs
- Telco SmartNIC use cases with DPDK

Telco requirements for NFV

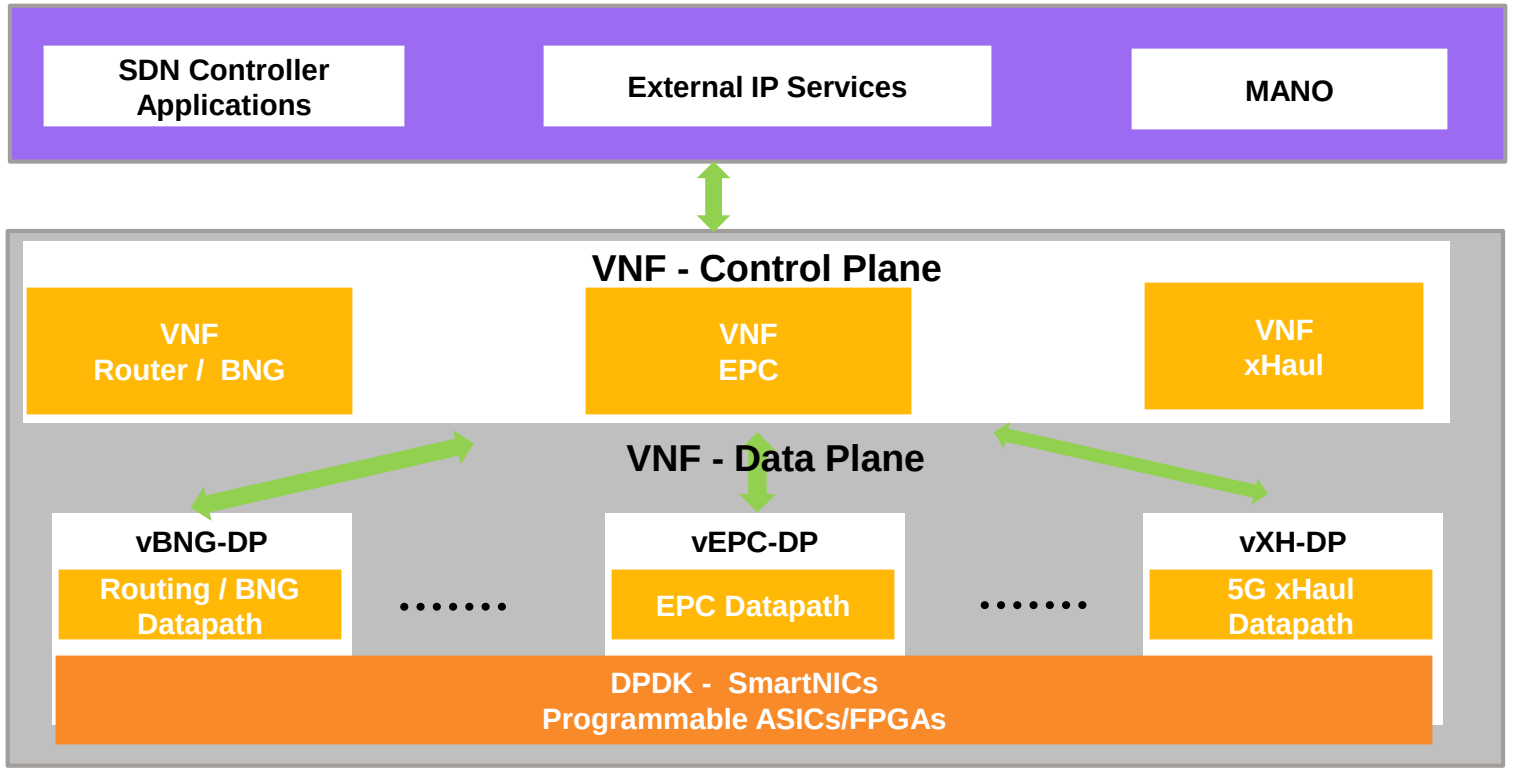


Telco Edge – Unique needs

Disaggregating Control plane and Forwarding plane

Control plane - User (data) plane separated VNFs **("CUPS" model)**

- Control Plane VNF: Control plane functions of 1 or multiple related data planes (VNFs and PNFs)
- Data Plane VNF: Routing / BNF Data Path, EPC Data Path, 5G xHaul Data Path, VPN GW Data Path
- Control Plane VNF and Data Plane VNFs can be on same compute node or distributed across compute nodes depending on the use cases

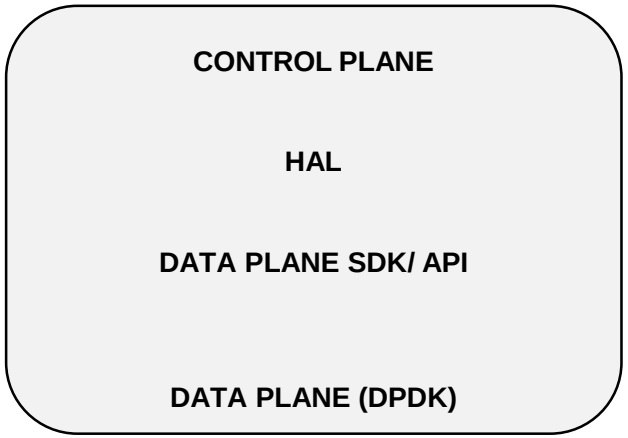


Dataplane

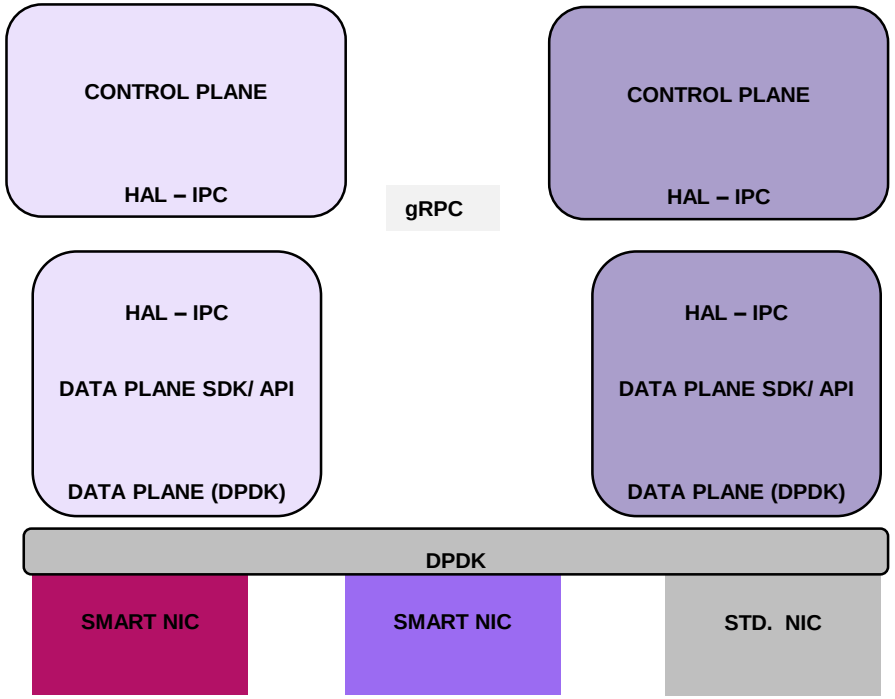
- **VNFs – x86/DPDK, SmartNIC**
- **PNFs – Switch-based data plane**

Software Architectures for Telco VNFs

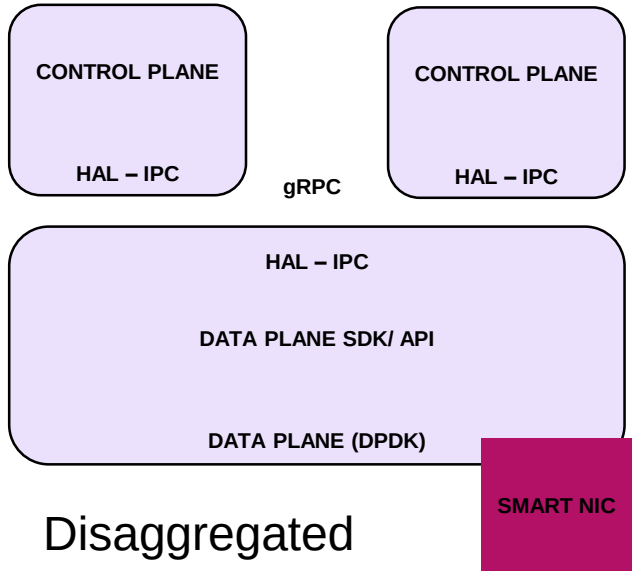
CP-DP Disaggregated Models



Traditional Aggregated



Disaggregated & Decoupled

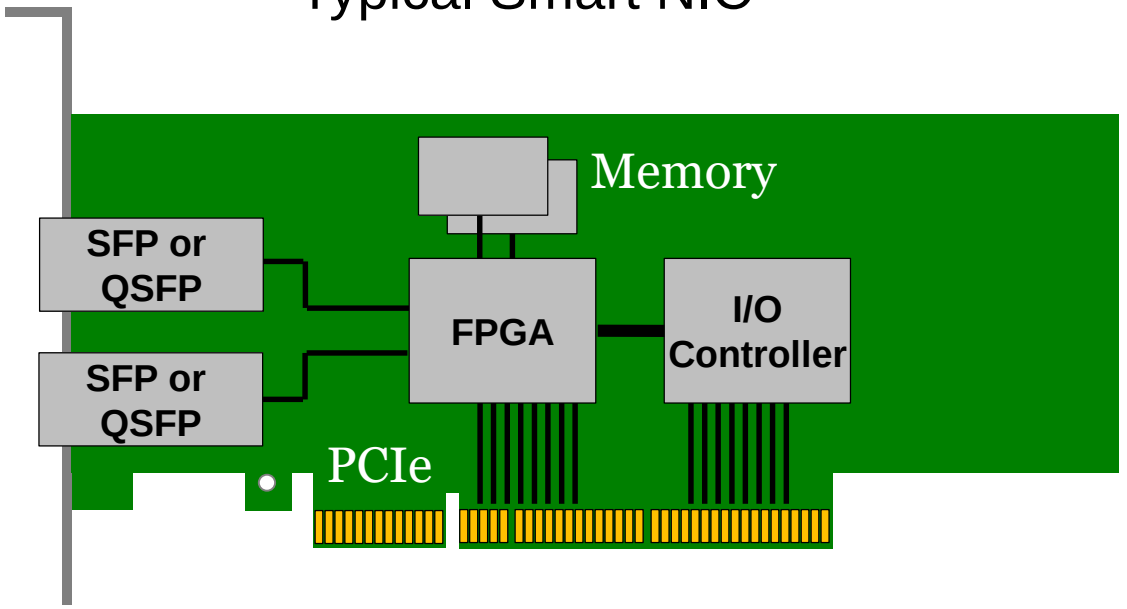


Disaggregated

Smart NIC Acceleration

- Smart NICs accelerate application performance
- Replacing standard NICs
  - Hyperscale data centers
  - Edge computing
- Multi-host CPU offload
  - Applications
  - Network functions
- FPGA or processor based
- I/O controller integrated or separate

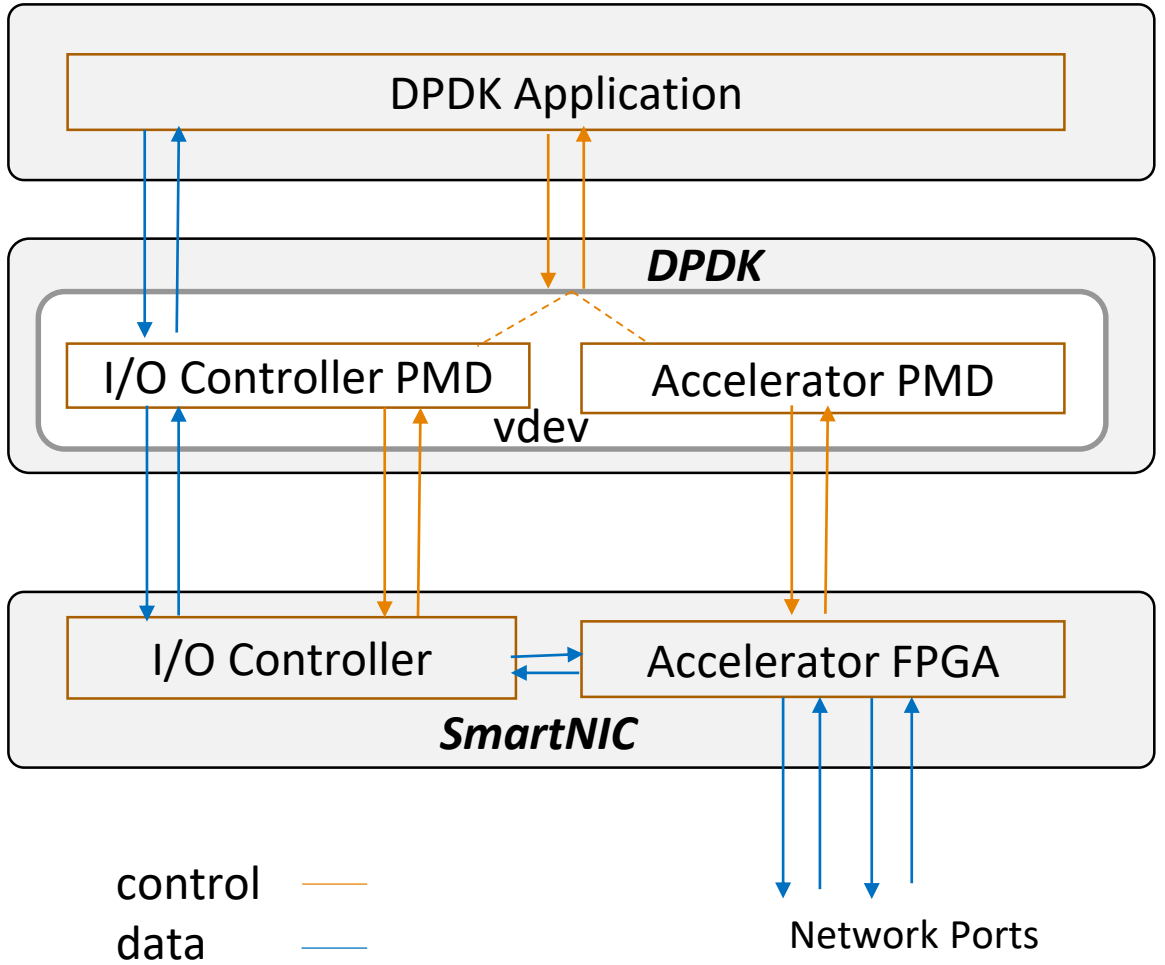
Typical Smart NIC



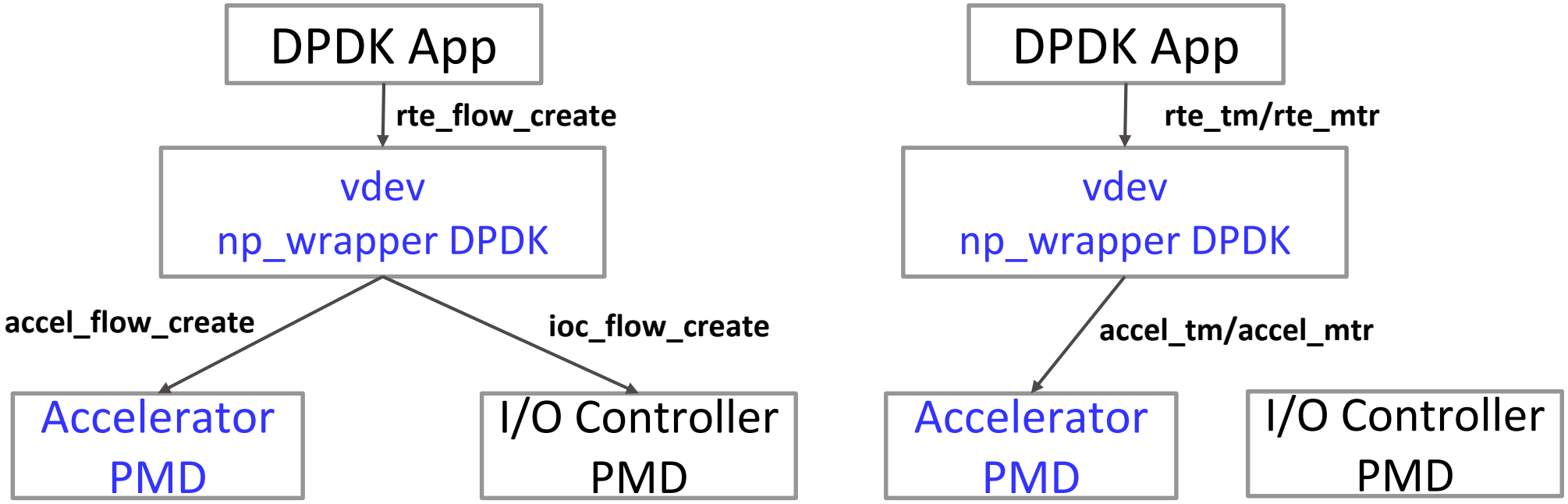
Source: HeavyReading and Earlswood Marketing

DPDK and Smart NICs

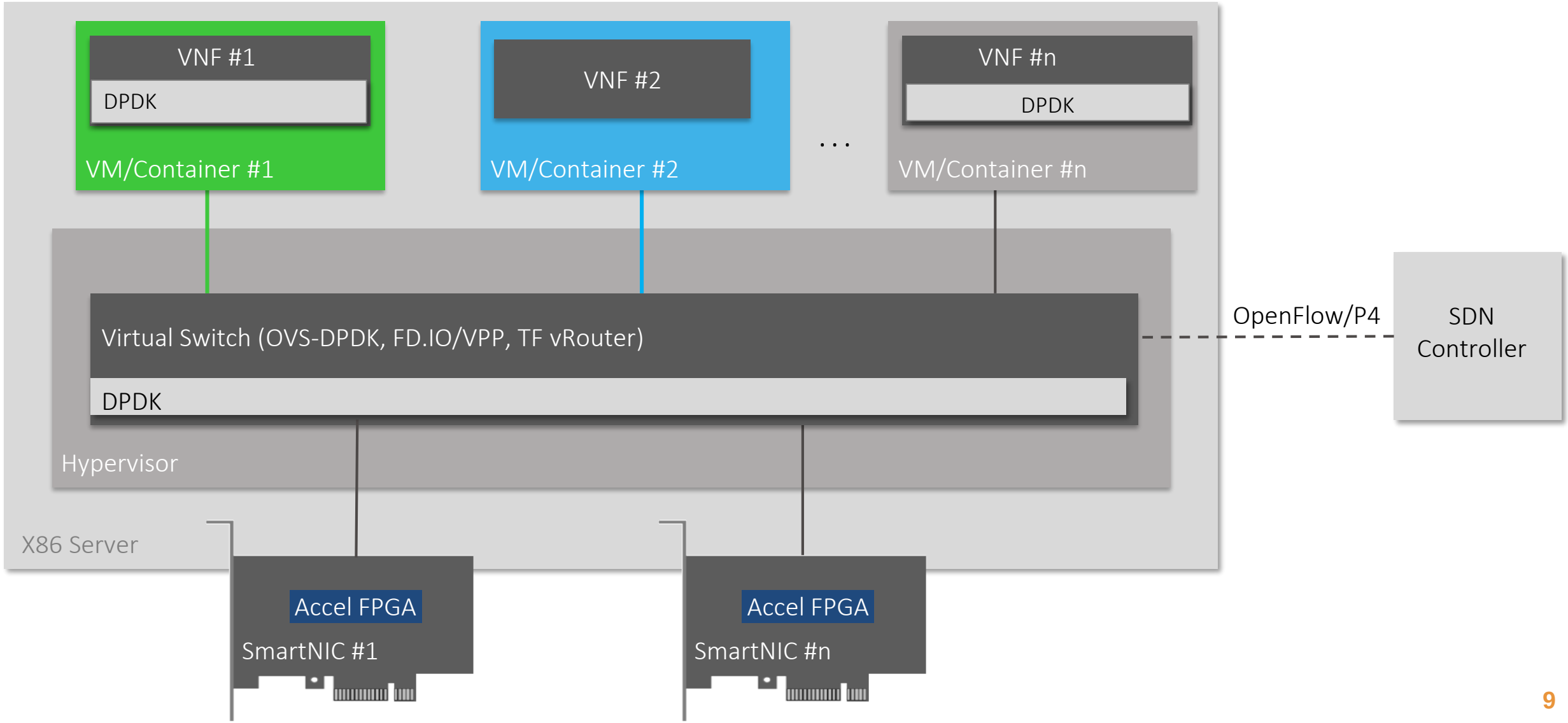
- Wrapper vdev encapsulates Accelerator PMD and I/O Controller PMD
- Any I/O Controller & any accelerator can be used
- Active/Active mode
- Application sees a single entity
- Split logic is inside wrapper
- Application transparency



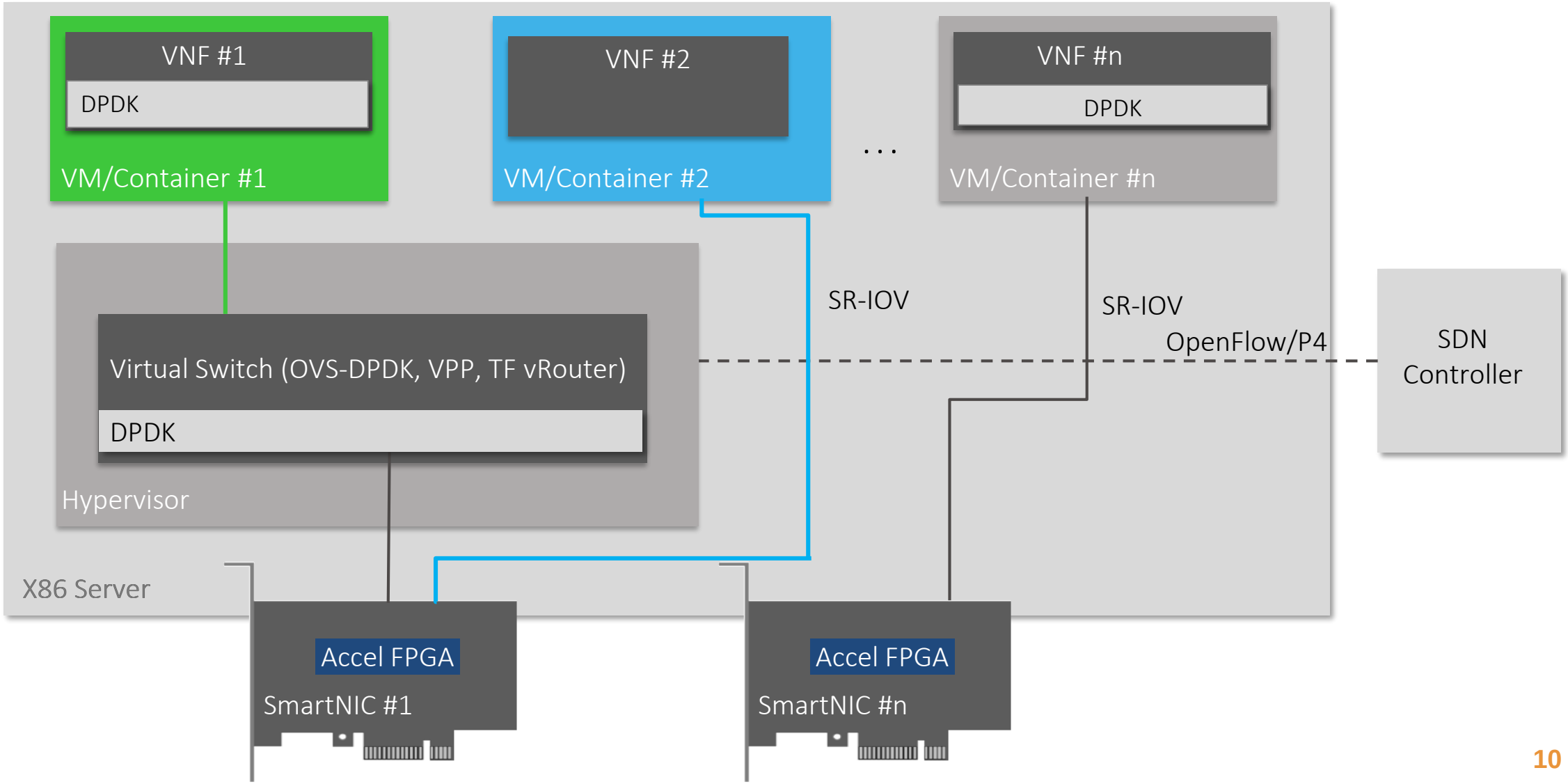
Generic Wrapper Virtual Device



NFVi Offload

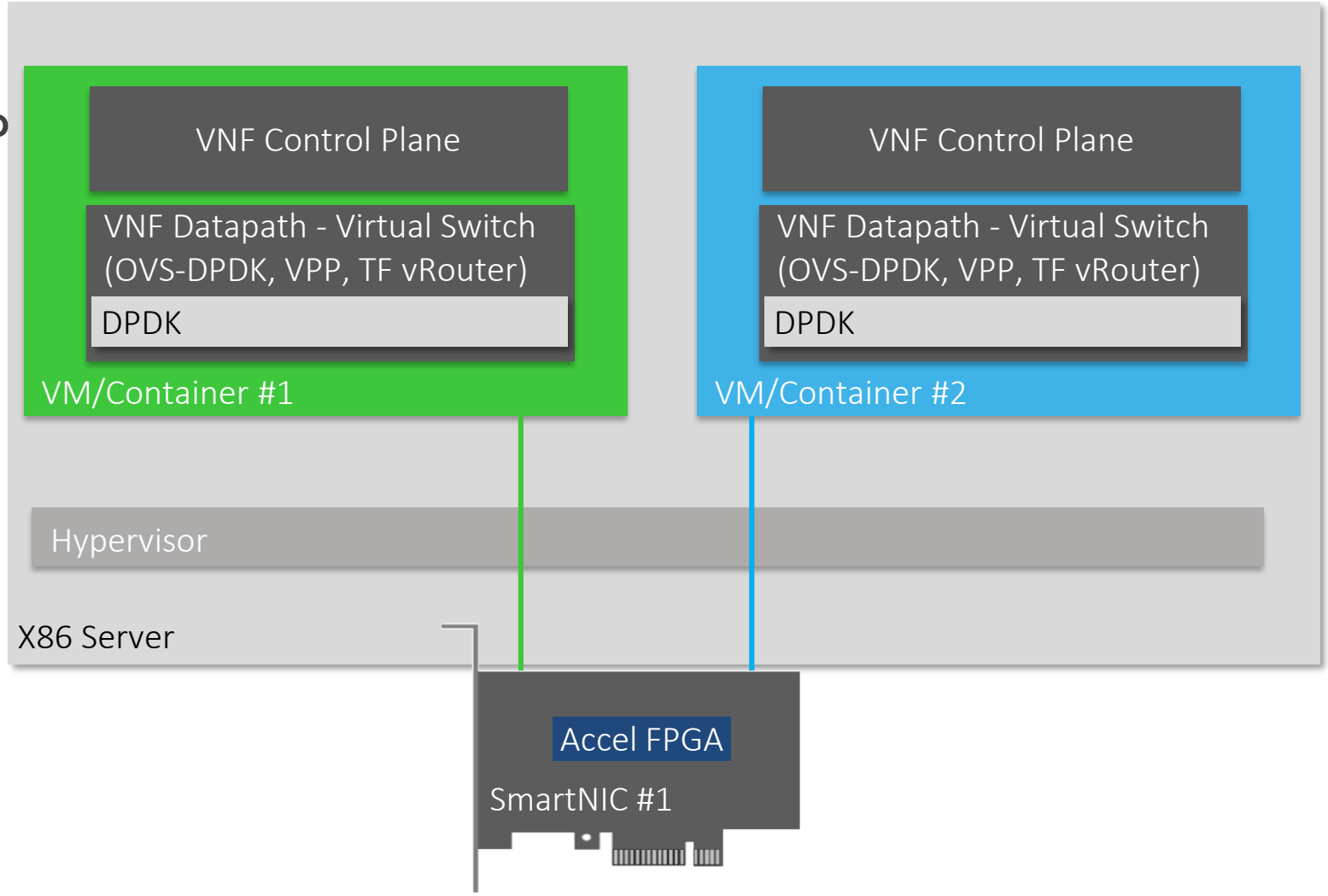


VNF Offload Using SR-IOV



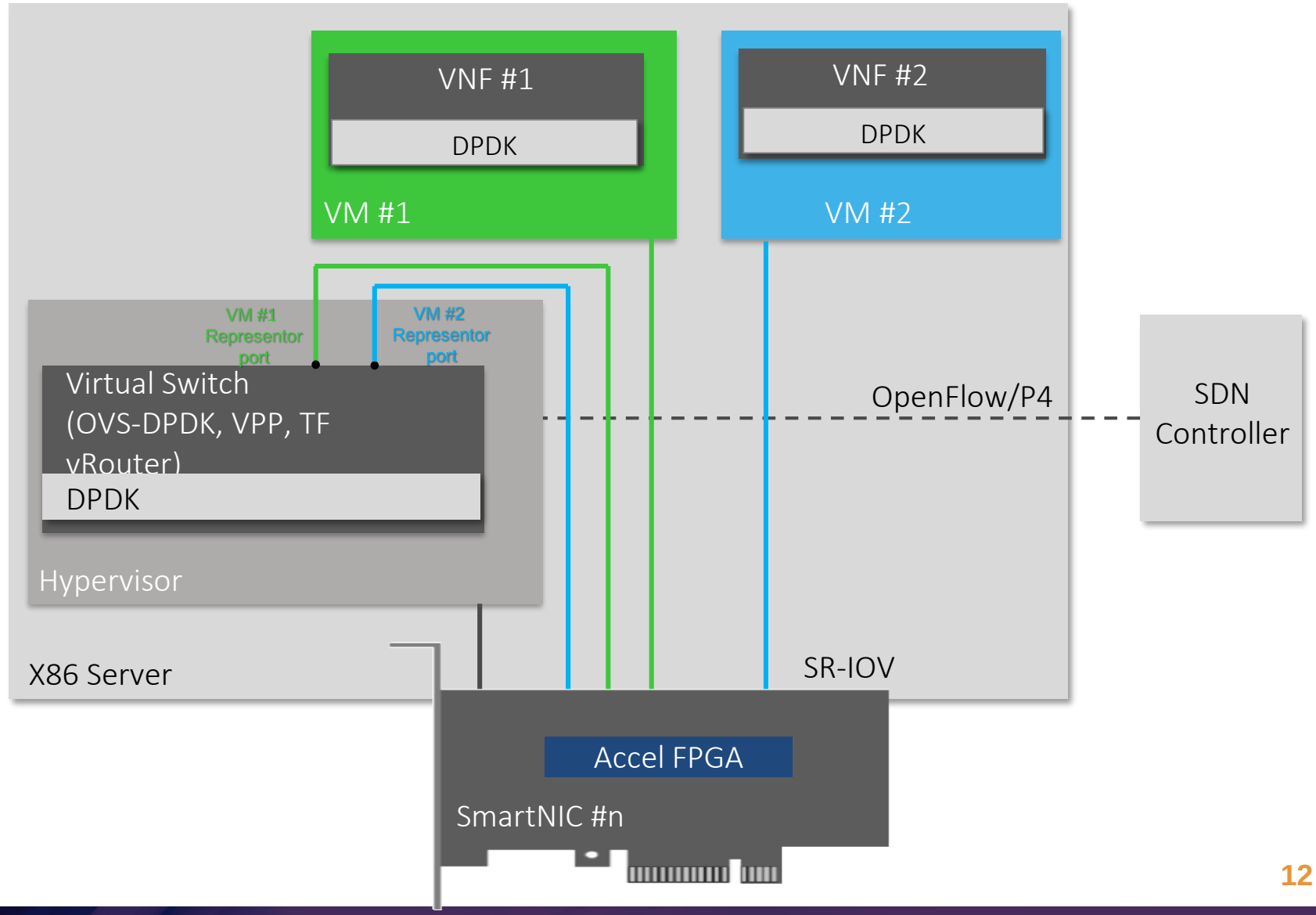
VNF Offload Using Common vSwitches & DPDK

- VNFs evolve to use common elements for datapath: OVS-DPDK, VPP & TF vRouter
- DPDK is used as the baseline for all these elements
- DPDK acceleration method used for a wide range of VNFs
- Open source vRouters need to add more features



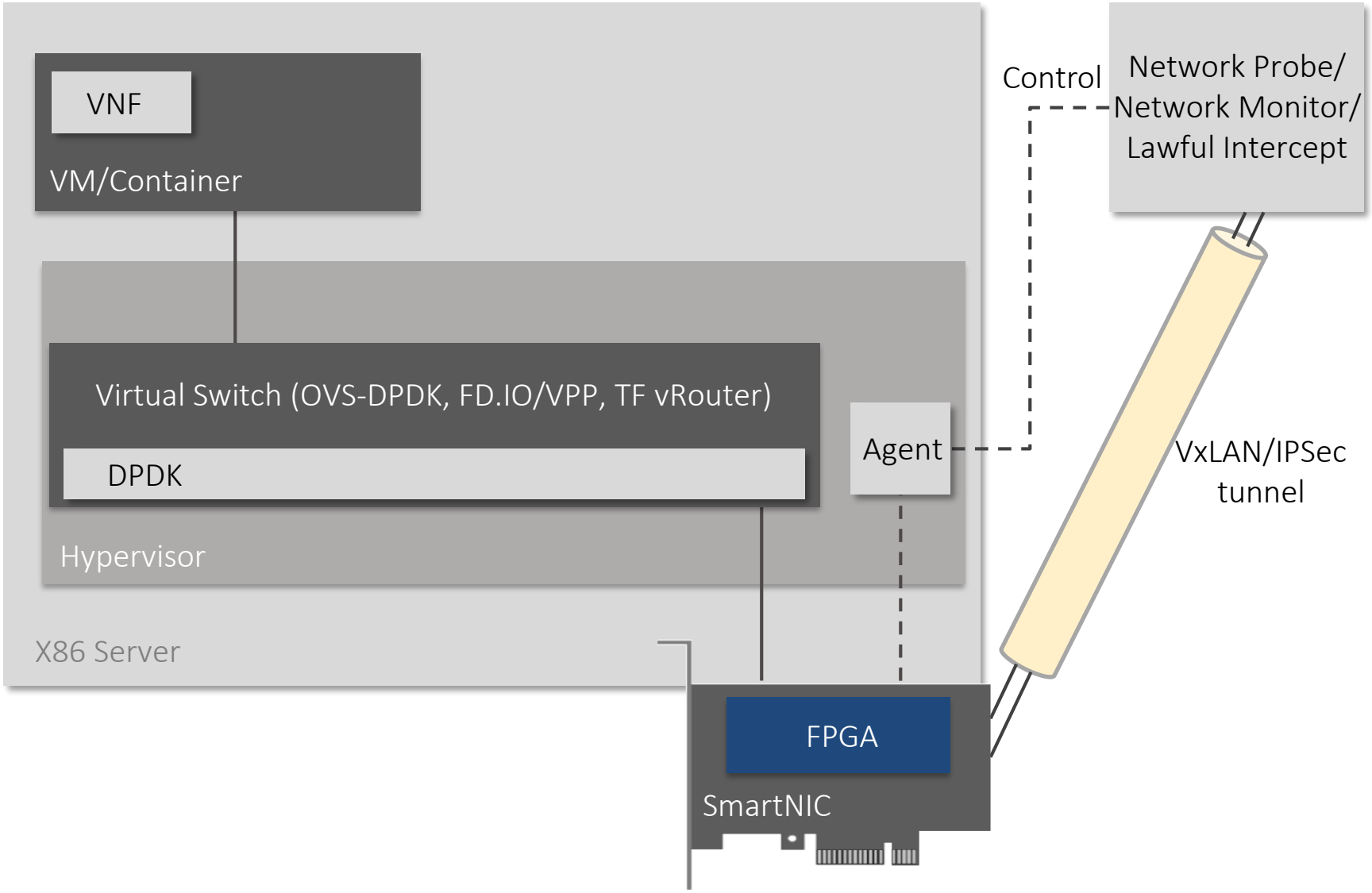
VNF Offload Using Representor Ports

- Adds support for VM migration
- SmartNIC needs to support vSwitch offload & VNF offload at the same time

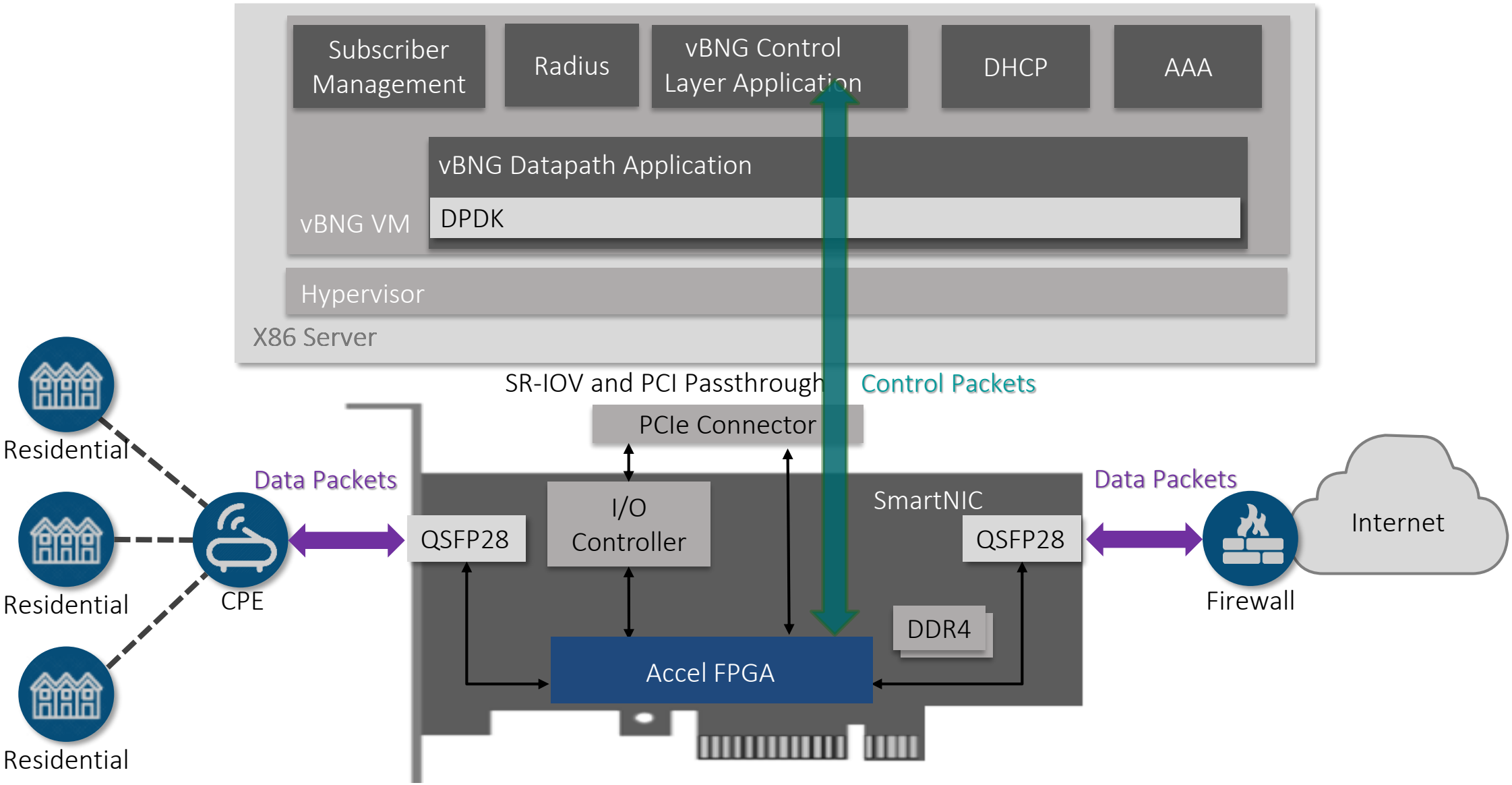


Telco VNF example: TAP-as-a-Service

- Flow-based TAPing, monitor/probe
- Millions of flows
- Flexible flow classification
- Can tap any flow (n-tuple)
- Programmable tunnels



Telco VNF Example: vBNG with Full Offload



Telco VNF example: vBNG Full offload requirements

- Per-Subscriber scaling (Millions)
- Per Subscriber PM Counters & Statistics
- PPPoE
- NAT / CGNAT
- Hierarchical Traffic Management
  - Policing
  - Shaping
  - Scheduling
- Lawful Intercept
  - Per subscriber flow monitoring
- TR-101 VLAN management (Double tag)
- N-tuple ACLs
- LAG hash
- ECMP hash
- Multicast
- OAM
- Stateful load-balancing of Subscribers to Control VNFs

- Most of the offload functions are available in dpdk
- Some offload functions, even if available in DPDK, are not supported to scale in many NICs

Current DPDK offload support

Speed capabilities
Link status
Link status event
Removal event
Queue status event
Rx interrupt
Lock-free Tx queue
Fast mbuf free
Free Tx mbuf on demand
Queue start/stop
Runtime Rx queue setup
Runtime Tx queue setup
MTU update
Jumbo frame

Scattered Rx
LRO
TSO
Promiscuous mode
Allmulticast mode
Unicast MAC filter
Multicast MAC filter
RSS hash
RSS key update
RSS reta update
Inner RSS

VMDq
SR-IOV
DCB
VLAN filter
Ethertype filter
N-tuple filter
SYN filter
Tunnel filter
Flexible filter
Hash filter
Flow director
Flow control
Flow API

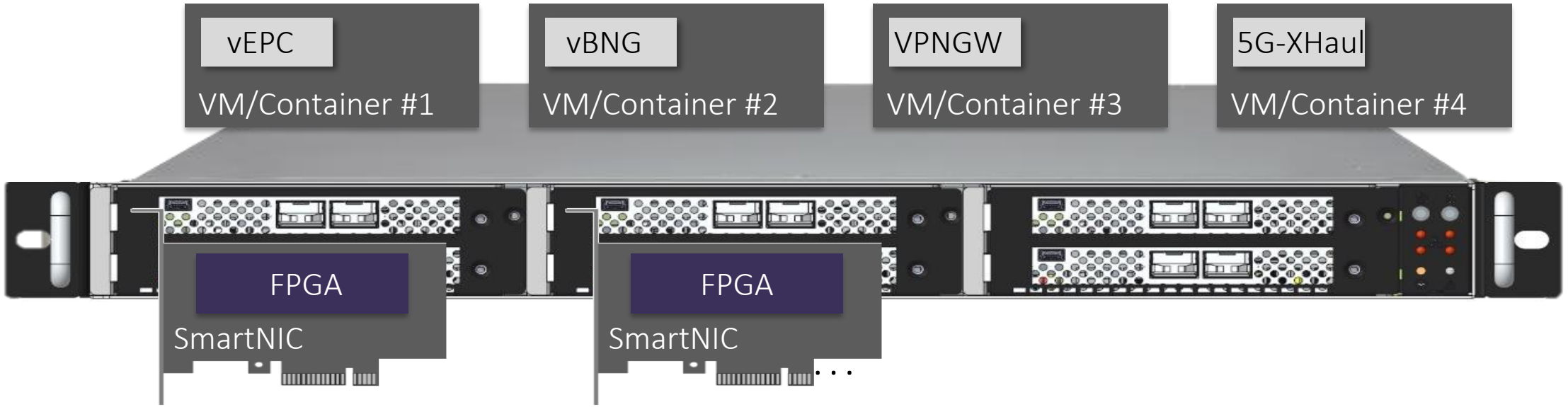
Rate limitation
Traffic mirroring
Inline crypto
CRC offload
VLAN offload
QinQ offload
L3 checksum offload
L4 checksum offload
Timestamp offload
MACsec offload

Inner L3 checksum
Inner L4 checksum
Packet type parsing
Timesync
Rx descriptor status
Tx descriptor status
Basic stats
Extended stats
Stats per queue

- http://doc.dpdk.org/guides/nics/overview.html#id1 - Table 1.2 Features availability in networking drivers
- Many more features not listed here, but available through rte_flow library

Telco Multi-Access Edge VNFs

- 1U server-based solution with HW acceleration
- Optimal for network edge deployment
- High performance, fully programmable, future-ready



System architectures for Telco VNFs

- Servers (1U upwards) with dual CPU socket, redundant power supplies
- Multiple (1 to 6) Smart NICs @ 100Gbps to 400Gbps each – 600Gbps to 2.4Tbps
  - 4 x 25GE, 2 x 50GE, 2 x 100GE, 2 x 200GE
- PCIe Gen3/Gen4 x 32
- High Availability with Clustering (N:1) or Redundancy (1:1)

- Design goals
  - Smart NICs with local processing / full offload
    - In/Out (Downstream / Upstream) traffic association inside a NIC itself
    - Minimum PCIe loading - PCIe and x86 limited to cross-traffic (across NICs) and Control plane traffic
  - Critical performance factors
    - New flow setup (setup latency, #sessions/sec)
    - Fault handling (fast reprogramming – protection switch/session-reestablish times, #sessions reestablish/sec);
    - Time critical – Detection/Analysis/Restoration – for large number of sessions/failures
  - Co-exist with multiple 3rd party VNFs

Thanks

Presenters:

Kalimani Venkatesan G
Kalimani.Venkatesan@aricent.com

Barak Perlman
Barak@Ethernitynet.com