Acceleration of NFV - Integrated NFV Nodes and Intelligent NICs

IEEE ComSoc SCV

Bharat Mota, Director of Engineering, Networking Applications
bharat.mota@freescale.com

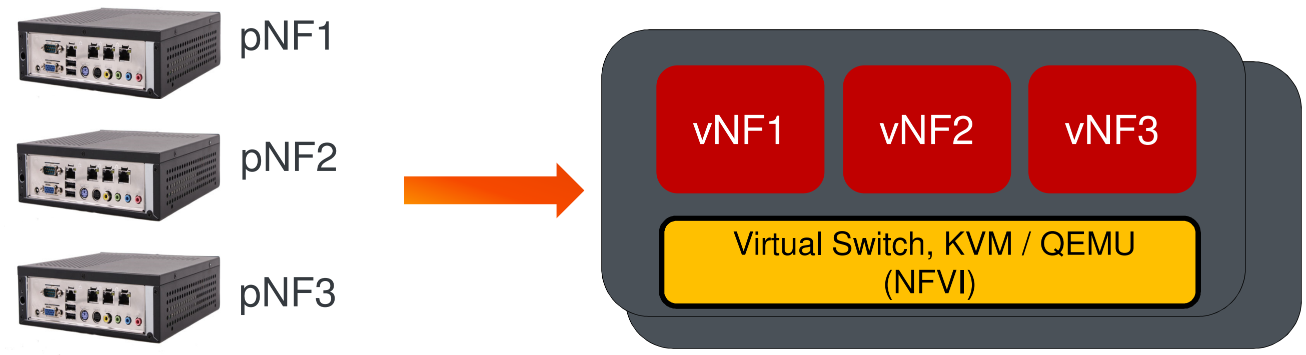
May 13, 2015

External Use

Agenda

- NFV Challenges
  - Comparison with Physical Network Functions
  - Performance Challenges
- Solutions to Mitigate Performance Challenges
  - New and innovative servers targeting NFV (NFV Servers)
  - Intelligent NICs in standard servers
- Freescale Solutions

NFV – Network Function Virtualization



Each network function as VM (vNF) - Benefits

- **Agility**: Scale-Out/In (bring more VMs on demand, based on load)
- **Efficiency**: Multiple network functions can share an NFV node (compute node)
- **Cost**: OPEX/CAPEX reductions - One common hardware for all types of network functions

Amount of traffic processed by vNFs is typically a lot higher than IT application VMs.

pNFs Versus Current vNFs

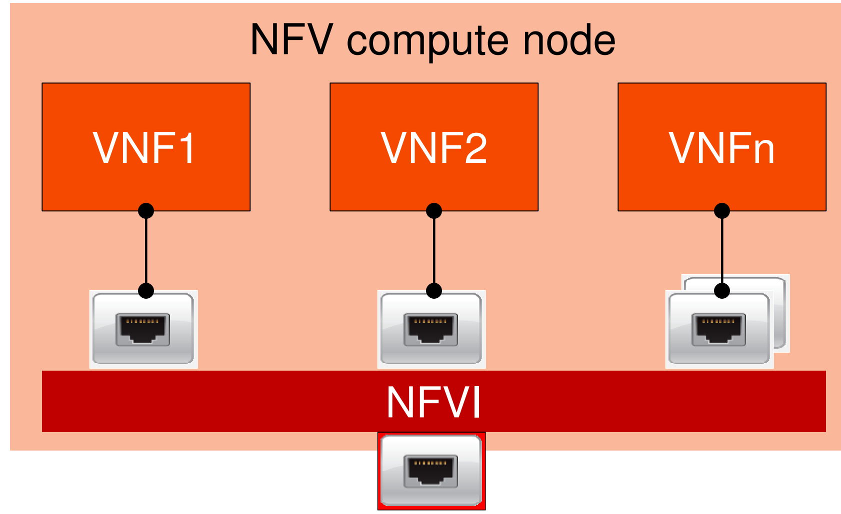
- vNFs runs on Virtualization Layer (NFVI)
  - Performance of vNFs on similar hardware is around 50-60% of pNFs
  - ***New innovative NFV servers are being created to mitigate this performance drop. Another solution is specialized iNICs to Servers to take up this load***
- pNFs take advantage of HW Look-Aside Accelerators
  - To make vNFs work on many platforms, accelerators are not used
  - ***Standardization of accelerator interfaces from vNFs is being considered in both ETSI and OPNFV***
- pNFs with control and data paths tend to implement data paths in special types of processing systems (FPGAs, multicore processors, ASICs, network processors etc..)
  - Packet processor nodes in the network
  - ***Openflow based packet processor nodes, keeping ETSI NFV principle in mind – one common hardware for multiple types of datapaths. Another solution again is specialized iNICs to Servers to take up this load.***

NFVI Challenges

- NFVI (VMM) enables virtualization of hardware and exposes each virtual hardware to VMs

- NFVI consists of multiple SW modules
  - Orchestration agent
  - Libvirt
  - Hypervisor such as KVM
  - QEMU for emulating hardware

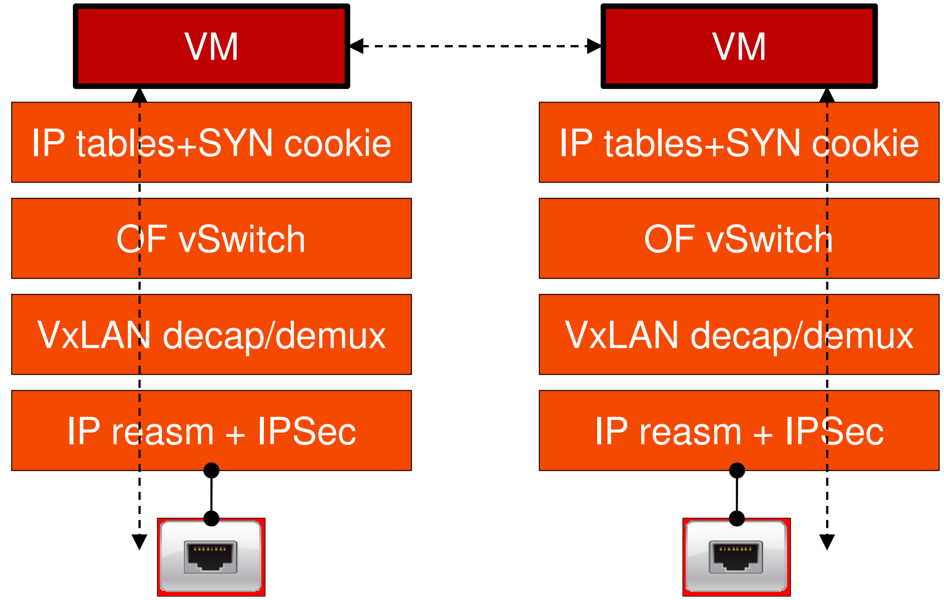
- Networking
  - VxLAN – Overlay based virtualization
  - OVS – Virtual Switching
  - Firewall – Filtering traffic going to/from VMs.
  - Traffic Control (Police & Shaping)
  - DDoS prevention (Syn flood and simple flood prevention)
  - IPSec for security-on-wire
  - IP Fragmentation & Reassembly

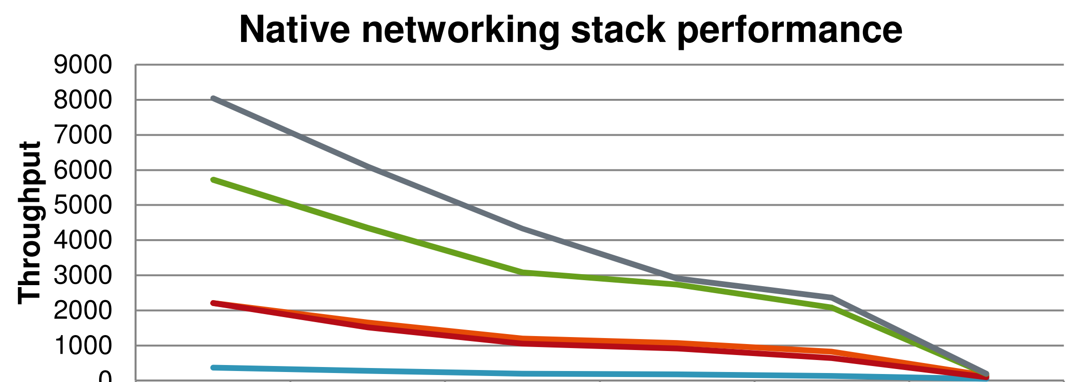


- Challenges
  - More intelligence is being added to VMM – Intelligence is pushed to the edge
  - Amount of traffic processed by vNFs is much higher than typical IT applications, therefore networking performance is important
  - For a similar hardware, vNFs provide only 50-60% of the performance over pNFs

NFVI Performance Challenge

- Increasing complexity of infrastructure stack
  - Trending to more intelligent networking stacks. Netflow, BFD, monitoring, replication etc..
- Performance of NFVI is proportional to number of functions enabled in host Linux.



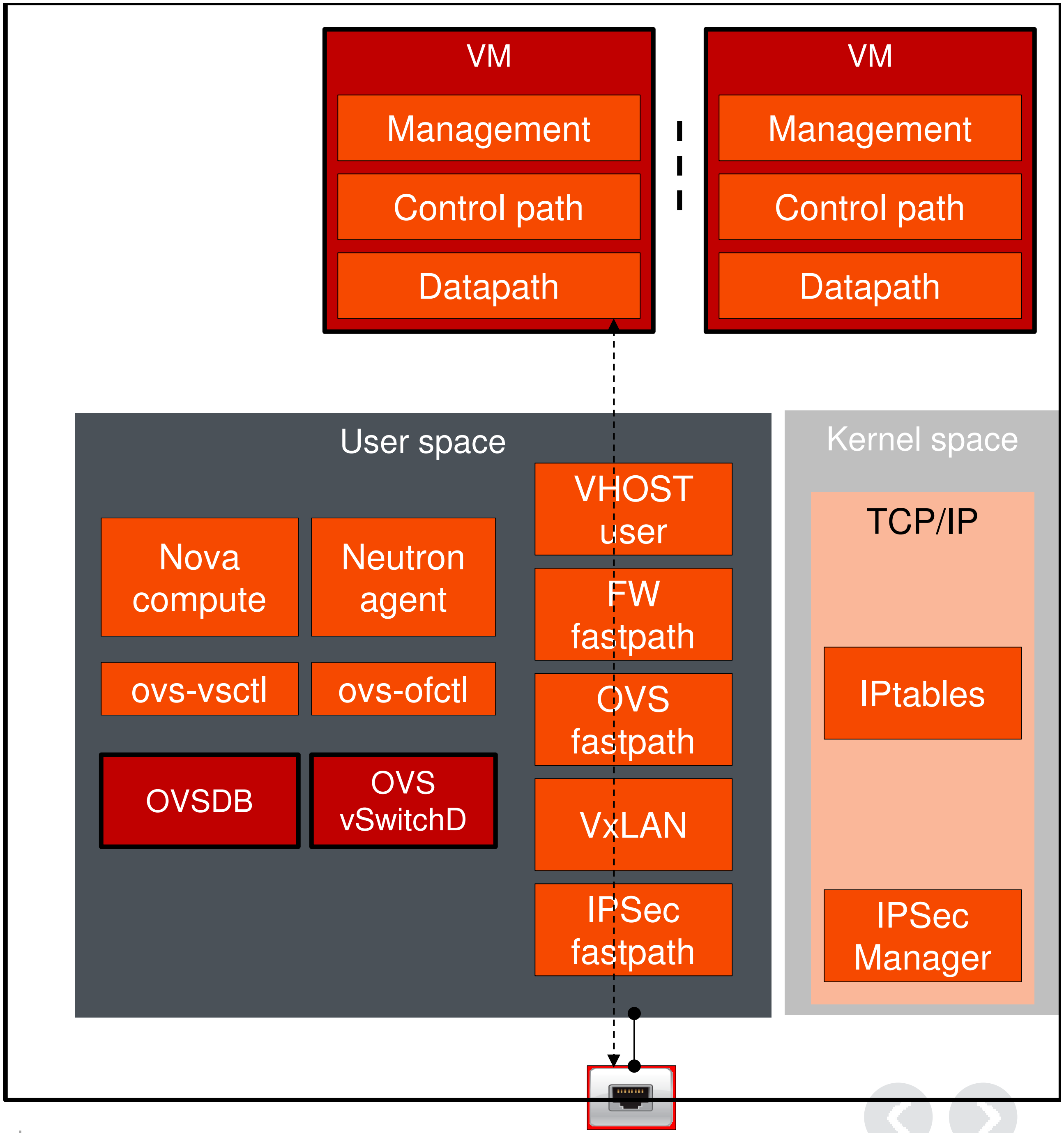


	TCP/IP	OVS	IPtables + OVS	OVS + VxLAN	IPtables + OVS + VxLAN	IPtables + OVS + VxLAN + IPsec
64	370	279	195	181	136	49
390	2205	1652	1194	1072	824	146
390 (1K conn)	2205	1514	1051	914	639	98
1024	5722	4346	3085	2737	2080	190
1472	8042	6097	4334	2906	2365	197

freescale™

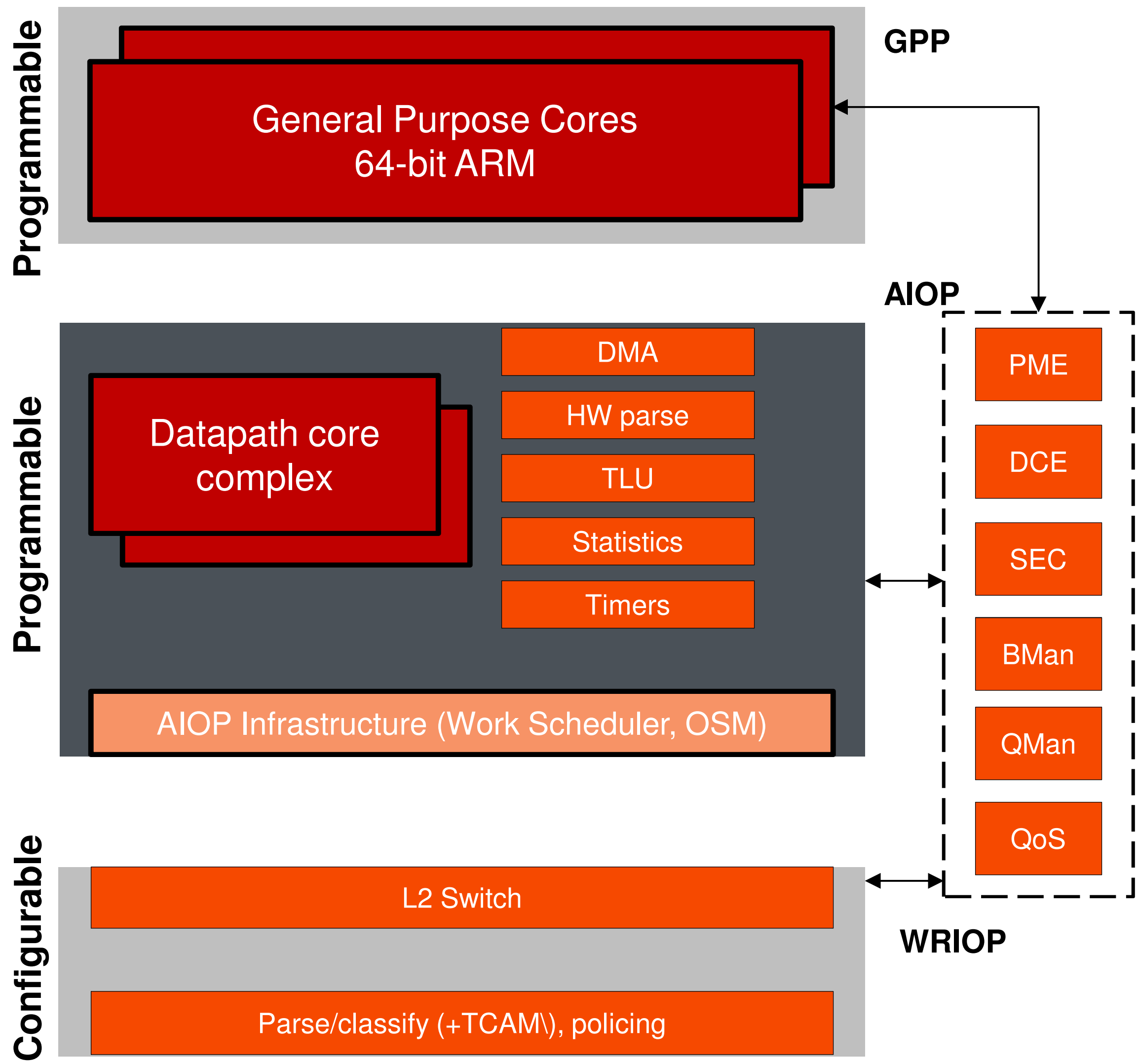
DPDK/ODP based NFVI-Networking Fast Path

- Ethernet Ports are attached to the fast path user space process
  - Cores are typically dedicated to the process
- Higher throughput
  - Based on Lagopus and extrapolated with common (except IPSec) NFVI virtual networking functions enabled - around 1 Gbps per core for 64 bytes and around 5 Gbps for IMIX traffic.
  - 25Gbps at IMIX requires 6 to 7 cores (Almost 30-40% of a 10-core dual CPU is used by NFVI, leaving 60% of CPU for VMs)



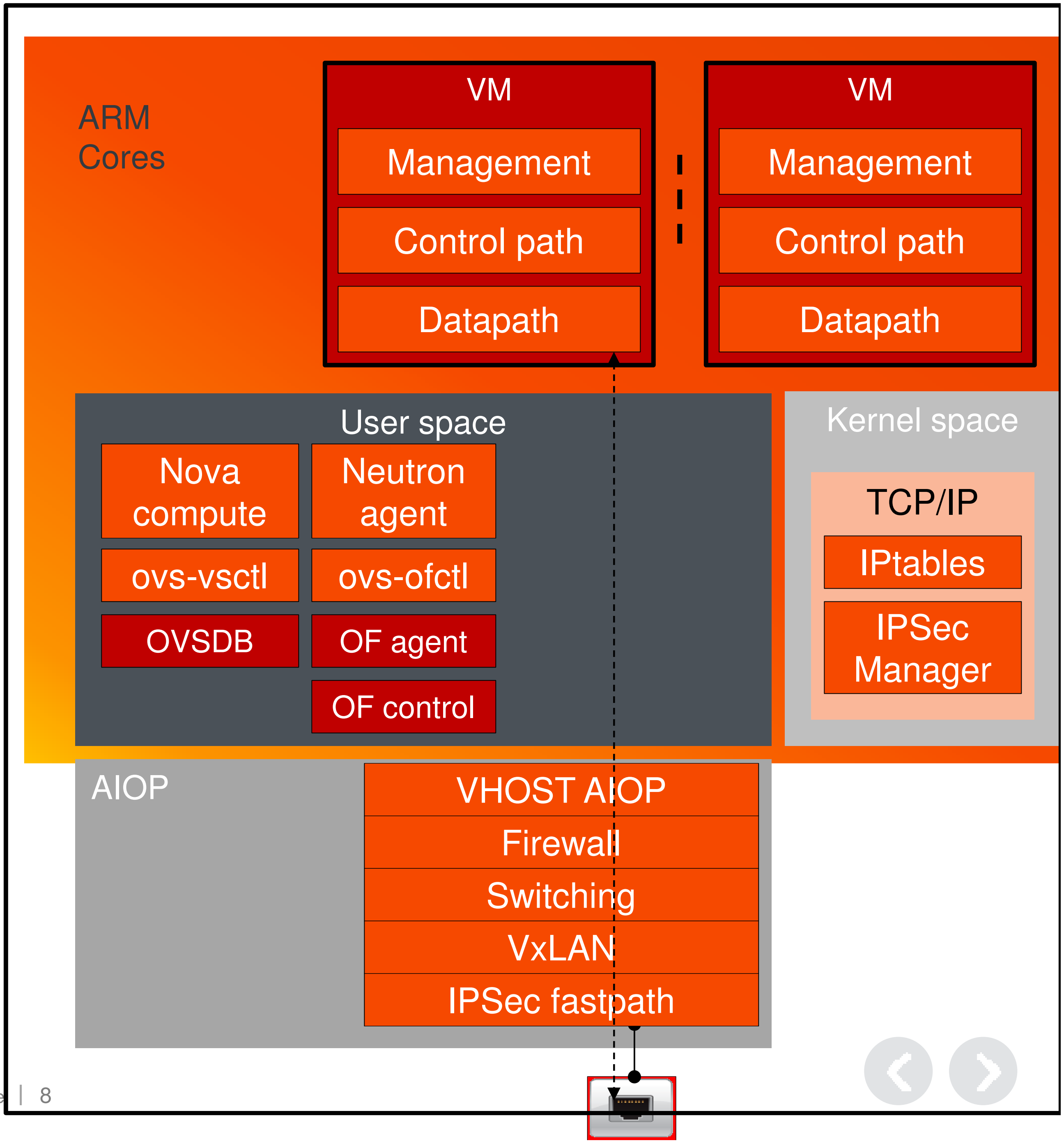
Layerscape Architecture

AIOP
Fast path data plane/packet processor
Hardware task scheduler
No interrupt overheads
C programmable
Packet processing accelerators • Table lookup (EM/LPM/ACL) • Packet infrastructure (BQMan, DMA,…) • Parser, SEC, timer etc..
SG Buffer Management in hardware
Packet order maintenance & synchronization in hardware
Synchronous programming model
Deterministic performance (maintained for large number of flows)

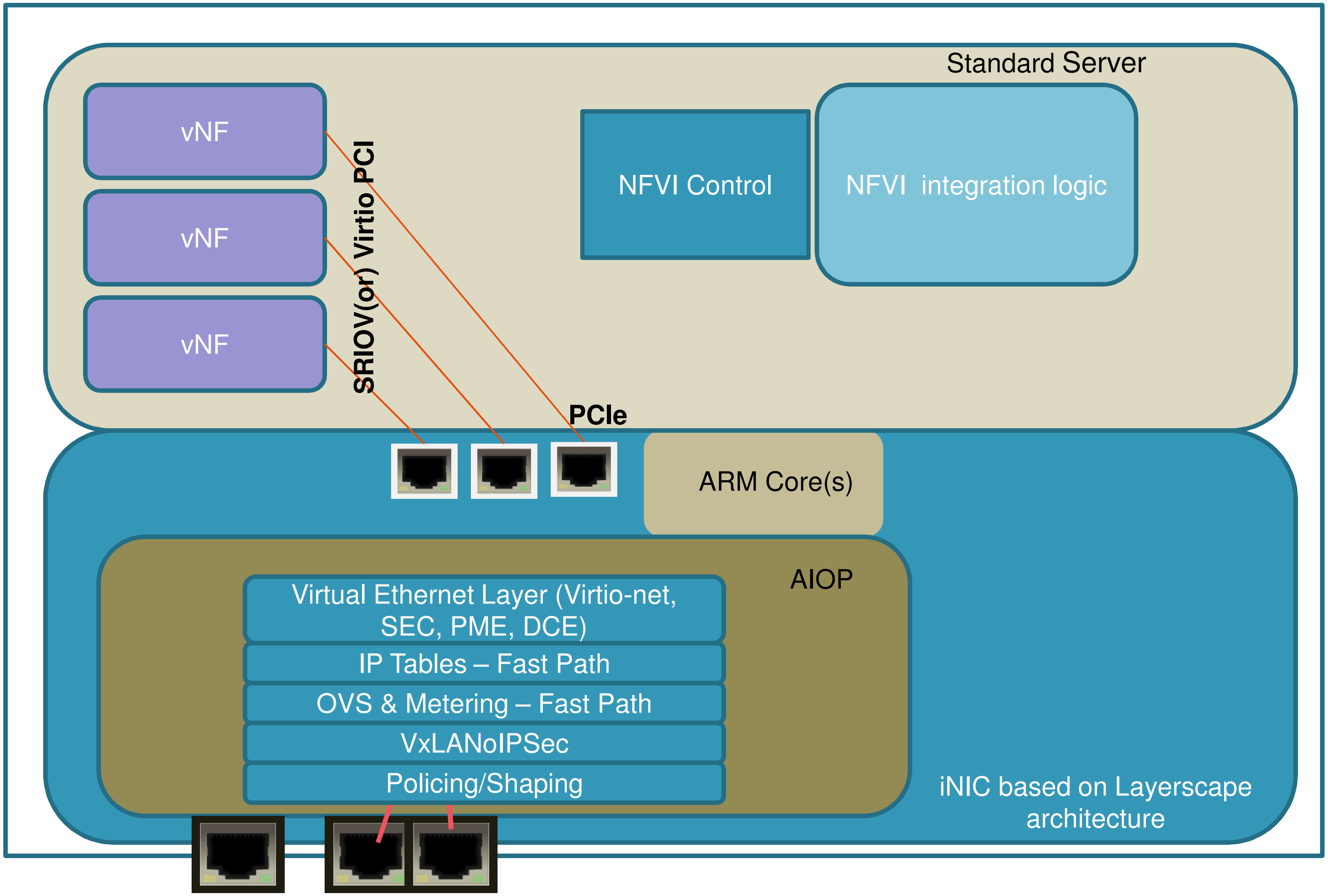


NFV Server with NFVIxl in AIOP

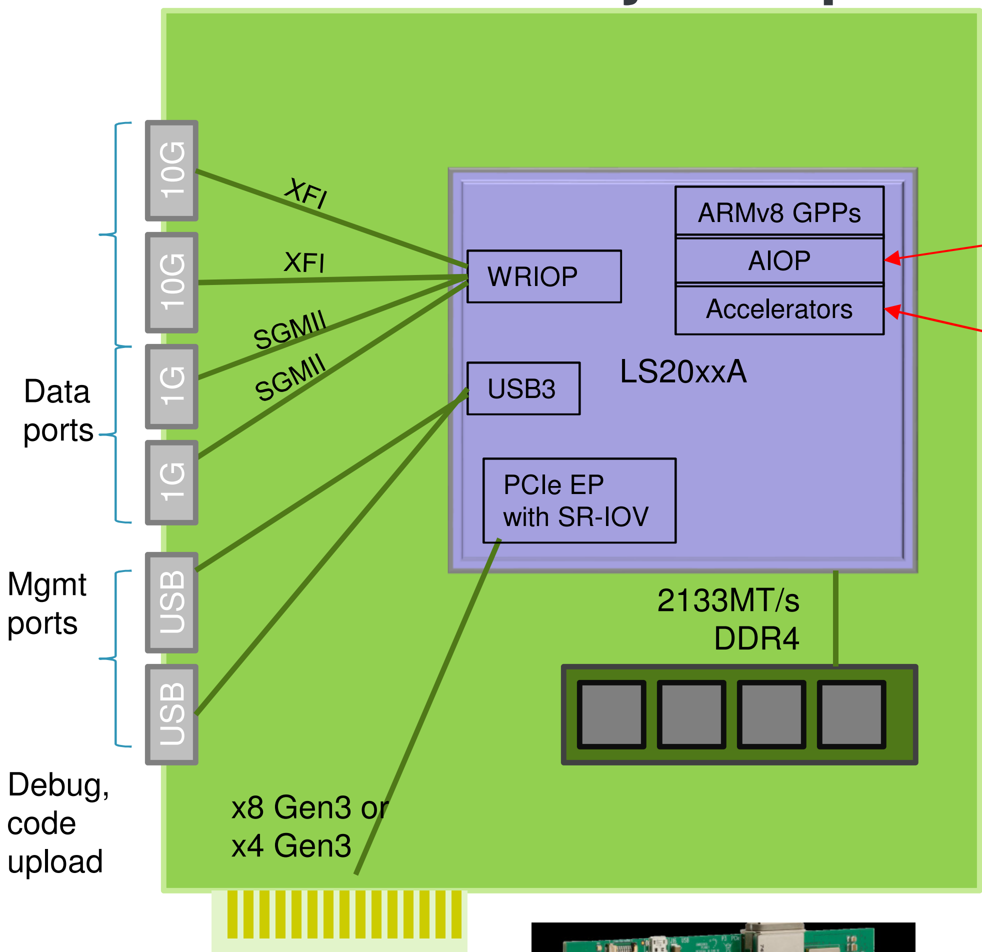
- Limited GPP involvement (management only)
- Offload as much packet processing to AIOP
  - AIOP implementing all fast paths
  - Direct connectivity to VM
  - AIOP acting as virtio backend
- Faster Connection rate
  - IP Table Policy Caching
  - Entire OF pipeline processing for switching
  - All OF based data paths (Easy to add new features with no changes to AIOP)
- Software Enablement
  - No changes to Openstack controller (North bound interface is kept intact).
  - Full software enablement in both GPPs and AIOP
- Goal: 25Gbps+ for IMIX traffic

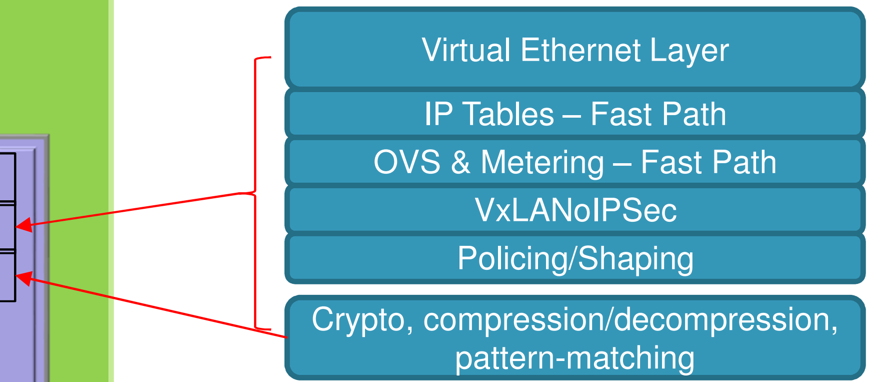


Standard Server with LSx based iNIC

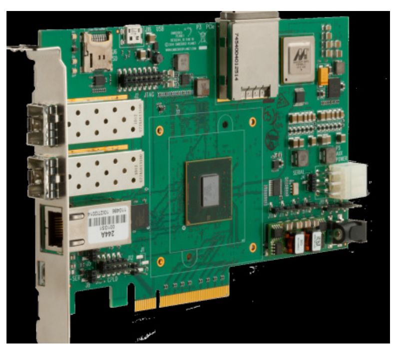


10G Dual Port Layerscape iNIC



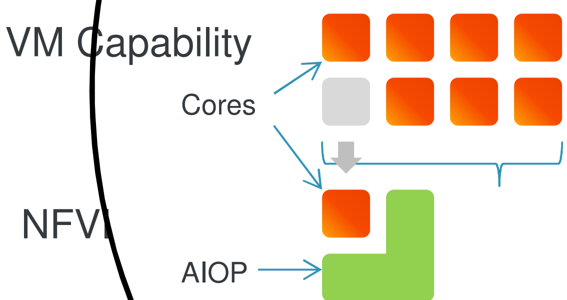


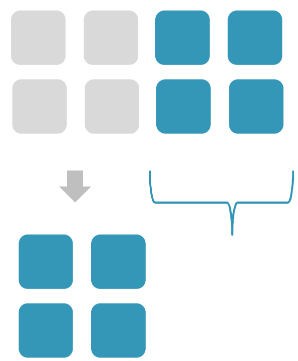
- Well-balanced device for 20Gb/s bi-directional application:
  - QDMA engines move about 20Gbps
  - x4 Gen3 or x8 Gen2 PCIe moves up to 64Gb/s
- SR-IOV allows virtual machines on host to see a private NIC. Also support virtio-net to allow generic vNFs to run.
- Offload accelerators for services cards: 20Gb/s IPSEC or Kasumi, 10Gb/s pattern matching, 20Gb/s data compression
- PCIe card reference board available 2H-2015



Layerscape Platform Efficiently Addresses the NFV Performance Challenge

	Layerscape Platform	Standard Server	Standard Server with iNIC based on Layerscape architecture
NFVI	Offloaded to AIOP	On GPP cores	Offloaded to iNIC
VM-to-VM datapath	Offloaded to AIOP	On GPP cores	Offloaded to iNIC





Minimum number of cores required to match AIOP performance - 4

Additional GPP cores can be freed up by vNFs that can utilize accelerators & offloads provided by iNIC.

iNIC based on Layerscape architecture provides perf/W and perf/$ advantage

Summary

- Layerscape Platform from Freescale alleviates performance limitations of vNFs running on standard servers that are designed for non-network workloads
- An iNIC based on Layerscape architecture offers a VMM (NFVI) networking offload mechanism that can co-exist in a standard server environment
- An NFV node / iNIC based on the Layerscape platform offers significant perf/W and perf/$ advantages

www.Freescale.com