Network Performance with Power 11

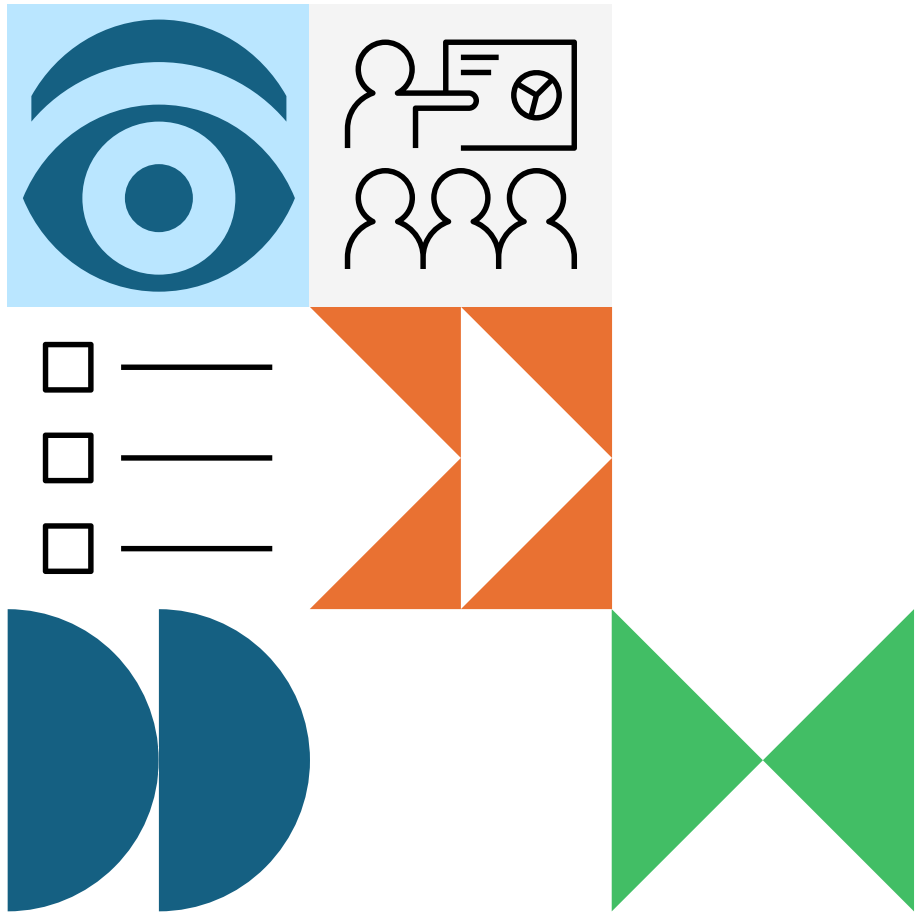


COMMON EUROPE CONGRESS 2026

14. – 17. June 2026 – Lyon, France

Alexander Paul (Alex)
Senior Technical Consultant
paulalex@de.ibm.com

Agenda



PowerVM Network Virtualization Options

Hybrid Network Virtualization (HNV) Architecture

Power11 and SR-IOV Capable Adapters Update

Network Performance Results

SR-IOV Adapter Configuration Options

NVIDIA OFED for SR-IOV Adapters on Linux

PowerVM Network Options

- **PowerVM Virtual Ethernet and Shared Ethernet Adapter**
  - Well known and established technology (since POWER5)
  - Build-in failover capabilities (SEA failover on Virtual I/O Server)
  - Supports Live Partition Mobility and Simplified Remote Restart

- **PowerVM (native) SR-IOV**
  - Hardware assisted network virtualization
  - Excellent performance
  - Client OS needs to care about redundancy as logical adapter ports are directly hardware related.
  - Does not allow the use of features like Live Partition Mobility (LPM).

- **PowerVM vNIC**
  - Based on PowerVM SR-IOV technology but using indirect assignment of adapter resources (via VIOS).
  - Allows that one vNIC can map to multiple SR-IOV backing devices.
  - Provides hypervisor-based failover and support for LPM and SRR.

- **Hybrid Network Virtualization (HNV)**
  - Combination of native SR-IOV and either Virtual Ethernet or vNIC.
  - Performance characteristics during normal operations similar to native SR-IOV.
  - Build-in failover capabilities within the OS and LPM support, due to HMC controlled automation.

Hybrid Network Virtualization (HNV) PoC at Bosch - Overview

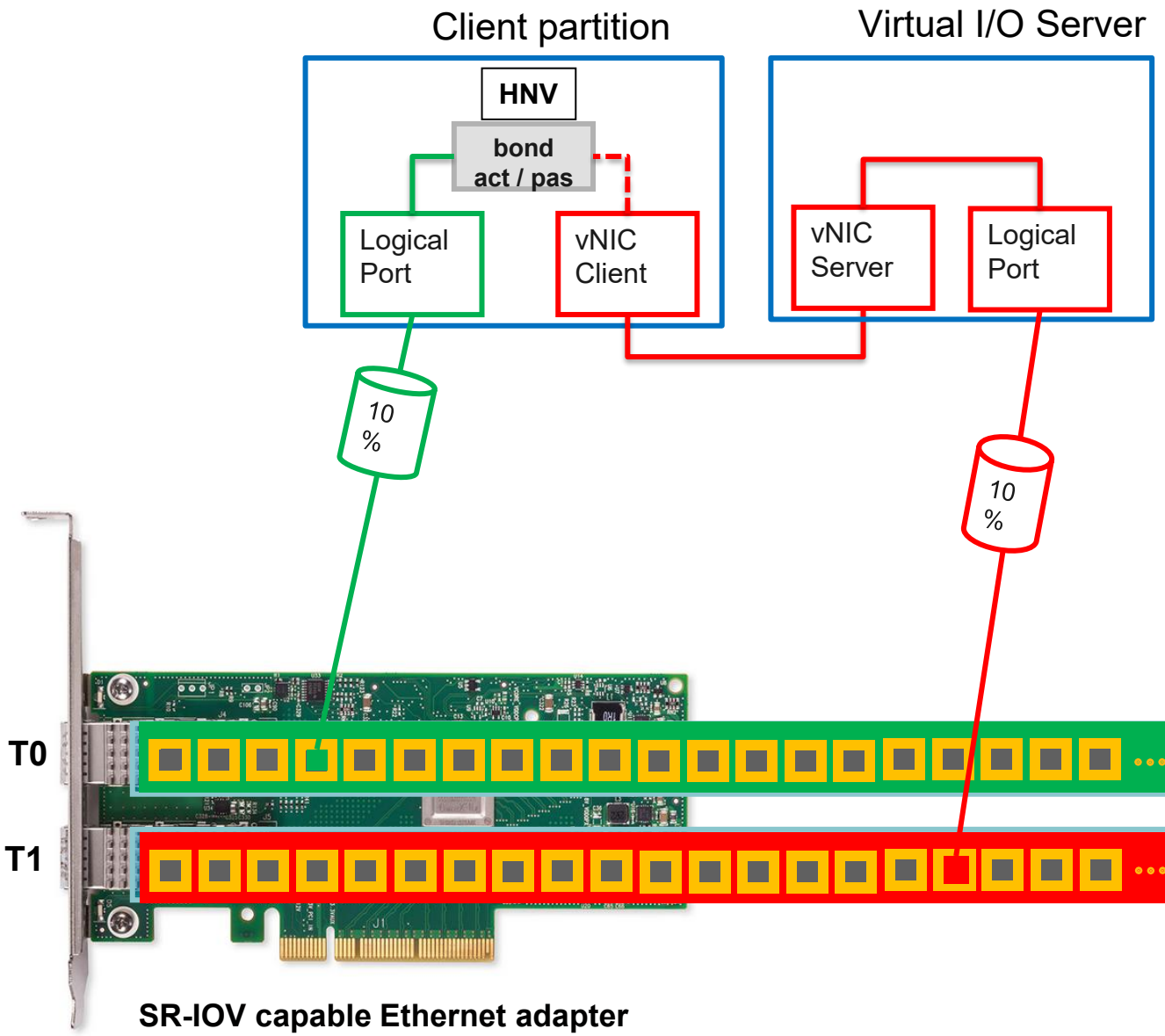
Goal:

- Evaluate network performance to achieve **maximum transactional throughput** on **Power10** and **Power11**.
- Highlight performance characteristics of **HNV** versus **vNIC** and **SEA**.
- Validate **PCIe5** and **PCIe4** network adapter capabilities.
- Target Operating Systems: SLES 15 SP7 (HANA DB) and AIX 7.3 TL3 (App. Servers and SAP classic).

Test tools:

- Alex's Network Performance Toolkit
- Load generator: **netperf 2.7.0**

Hybrid Network Virtualization (HNV) Architecture



HNV

- Native SR-IOV for performance
- Virtual backup device for LPM (vNIC or vEth)
- HMC dynamically removes/adds native SR-IOV logical port as part of LPM operation

PowerVM Network Virtualization Comparison

Technology	Live Partition Mobility	Quality of service (QoS)	Direct access perf.	Redundancy Options	Server Side Failover	Requires VIOS
SR-IOV	No	Yes	Yes	Yes[1]	No	No
vNIC	Yes	Yes	No[3]	Yes[2]	vNIC failover	Yes
SEA/vEth	Yes	No	No	Yes	SEA failover	Yes
Hybrid Network Virtualization	Yes	Yes	Yes	Yes	No	Yes

Notes:

1. SR-IOV can optionally be combined with VIOS and virtual Ethernet to use higher-level virtualization functions like Live Partition Mobility (LPM); however, client partition will not receive the performance or QoS benefit.
2. Some limitations apply. For SR-IOV and vNIC see SR-IOV link aggregation support
3. Generally better performance and requires fewer system resources when compared to SEA/virtual Ethernet

SAP HANA Hardware and Cloud Measurement Tools (HCMT)

- HCMT is the replacement for SAP HANA Hardware Configuration Check Tool (HWCCT)
- SAP Note: 2493172 - SAP HANA Hardware and Cloud Measurement Tools.
  https://launchpad.support.sap.com/#/notes/2493172
- Official KPI certification tool for SAP HANA.
- HCMT performs various tests from an execution plan (JSON file) and collects results.
- Test results are in JSON format and packed together with a manifest file in a ZIP archive.
- Data must be uploaded to SAP Support Portal for analysis.

```
sles15sp1-2:~/HCMT # ./hcmt -v -p ./config/executionplan.json
hcmt-2.00.052.00.1599235305 (2020-09-04 18:21:41)
Loading executionplan
LogVolume (/hana/log):
DataVolume (/hana/data):
Hosts:
NvmBasePath:
Start execution of plan
Executing Test C9C9F832-854F-492D-8E7EFB4609AC435C
Note: CPU Micro Benchmark
        Plan Variant: CPU Performance
        Repeat Count: 5
        .....\
```

HCMT Network Results

- HCMT performs two network related test series
  - Network Topology Test
  - Network Loopback Adapter
- With IBM PowerVM, HCMT tests can be passed with either network virtualization (SEA or SR-IOV based technology).
- Note: HCMT results do not show significant difference between vNIC and SR-IOV.

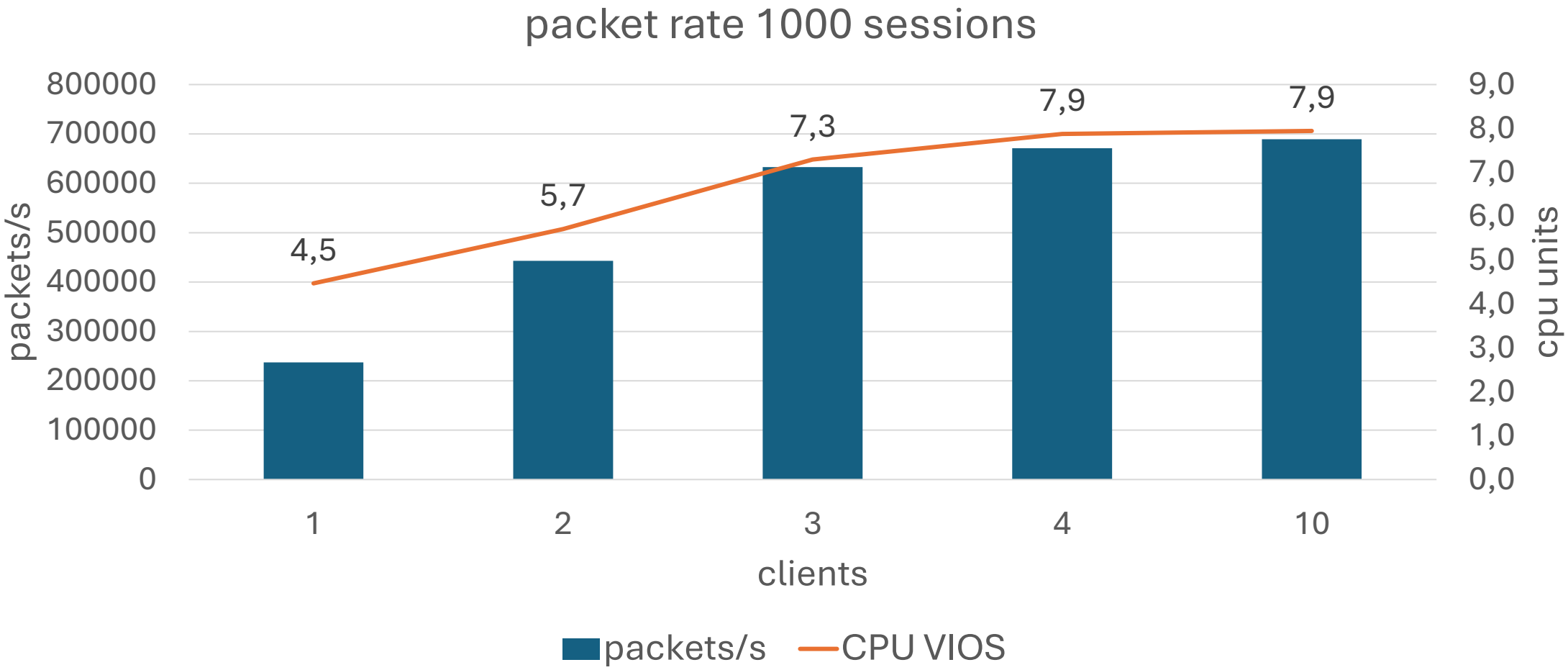
	Client DNS name.	Time required to receive the data, measured on client. [ms]	Received data size. [B]	Roundtrip time reduced by servers dataproce... [ms]	Complete roundtrip time, includes server processing time. [ms]	Time required to send the data, measured on client. [ms]	Sent data size. [B]	Server DNS name.	Throughput. [B/s] : ✓
SEA	sles15sp1-1	265	100,000,000	88	306	40	100,000,000	100.64.0.142	2,270,224,... ✓
vNIC (SR-IOV)	sles15sp1-1	245	100,000,000	86	304	59	100,000,000	100.64.0.142	2,323,474,... ✓

Virtual Ethernet Adapter Multi Queue Support

- Virtual Ethernet bandwidth and transactional performance is enhanced for networking between AIX LPARs on the same physical server and across different physical servers
- AIX 7.3 TL3 support performance enhancements on the client site.
- VIOS 4.1.1 adds multi queue support for Virtual Ethernet Trunk Adapter.

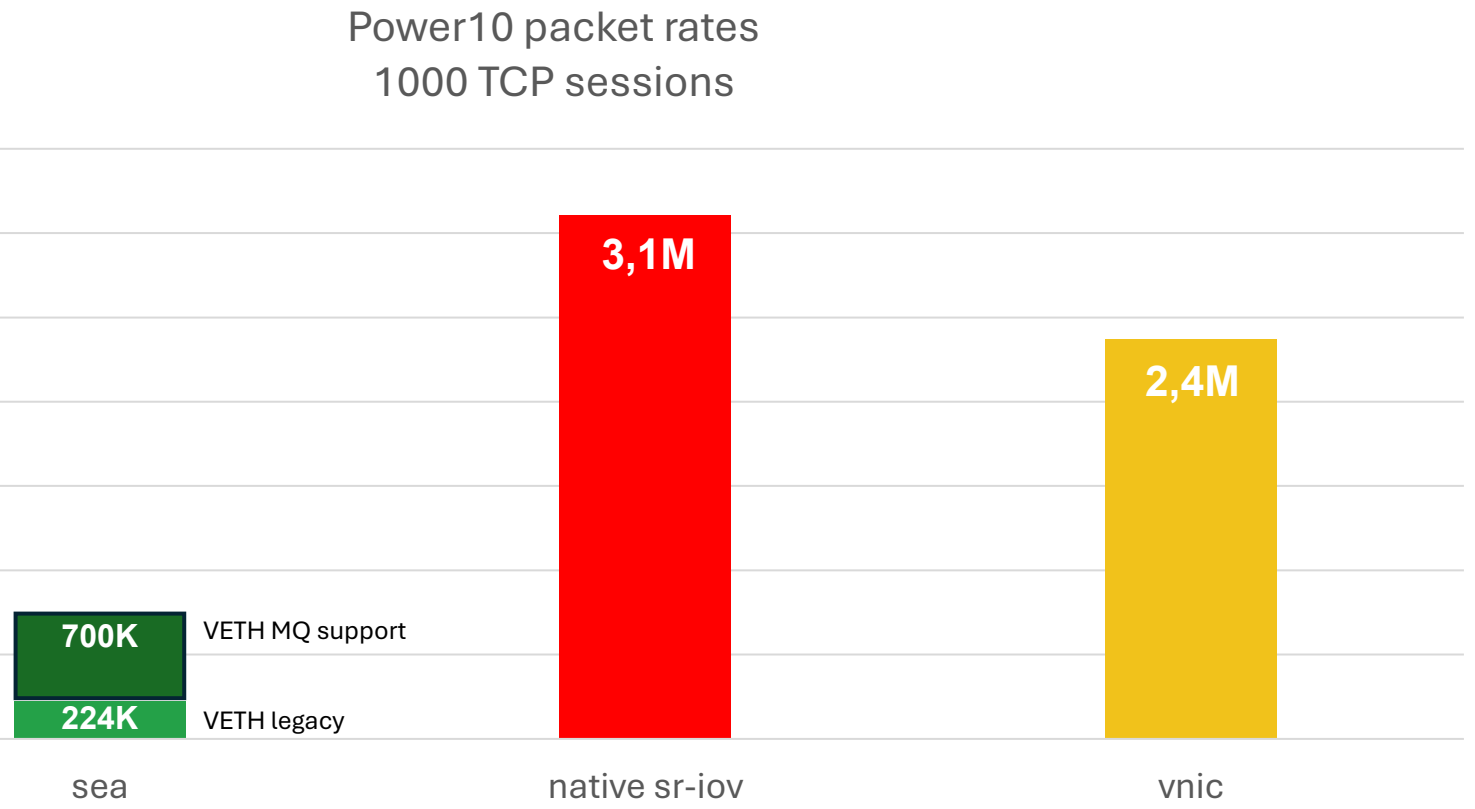
```
# lsattr -El ent5
...
flip_n_run      no              Send packets in receive context                 True
...
queues_rx       0               Requested number of Receive Queues              True
queues_tx       12              Requested number of Transmit Queues             True
rx_pool_policy  new             Receive Buffer Pools Policy                     True
rxq_depth       4000            Software Rx queue maximum packet count          True
...
tx_limit        32              Tx packets sent per transmit thread             True
```

Virtual Ethernet Adapter Multi Queue Support



Power10 Performance Reminder

- Power E1080 (9080-HEX) Systems Firmware MH1060.10
- PCIe4 x16 2-port 100/40GbE NIC & RoCE QSFP28 Adapter (#EC75/#EC76)
- **AIX 7.3 TL 2** and VIOS 4.1.0.21 (for vNIC)



SR-IOV Capable Adapters

	SR-IOV Capable Network I/O Adapters	Low profile FC	Full high FC
P8	PCIe3 4-port (2x10GbE+2x1GbE) LR Optical fiber and RJ45 [1,3]	EN0N	EN0M
	PCIe3 4-port 10GbE copper twinax [5]	EN18	EN17
	PCIe3 4-port (2x10GbE+2x1GbE) SR Optical fiber and RJ45 [1]	EN0J, EL38	EN0H, EL56
	PCIe3 4-port (2x10GbE+2x1GbE) copper twinax and RJ45 [1]	EN0L, EL3C	EN0K, EL57
	PCIe3 4-port 10GbE SR optical fiber	EN16	EN15
P9	PCIe3x16 2-port 100/40GbE NIC & **RoCE** QSFP28 Adapter [1,2,4]	EC3L	EC3M
P9 & P10	PCIe3 2-Port 10GbE NIC & **RoCE** SR/Cu Adapter [1,2]	EC2R	EC2S
	PCIe3 2-Port 25/10GbE NIC & **RoCE** SR/Cu Adapter [2]	EC2T	EC2U
	PCIe4x16 2-port 100/40GbE NIC & **RoCE** QSFP28 Adapter [6]	EC67	EC66
P10 + P11	PCIe4 x16 2-port 100/40GbE NIC & **RoCE** QSFP28 Adapter[7]	EC75	EC76
	PCIe4 2-Port 25/10/1GbE **RoCE** SFP28 Adapter[8]	EC71	EC72
	PCIe4x16 4-PORT 25/10/1GbE NIC&RoCE SR/Cu Adapter[9]	EN24	EN26
P11	PCIe5x16 2-PORT 200/100/40GbE NIC&RoCE SR/Cu Adapter[10]	**EC85**	**EC86**

Notes:

1. Withdrawn
2. SR-IOV support for NIC function prior to FW930. SR-IOV support for both NIC and RoCE with FW930
3. SR-IOV support available on POWER9 E950 & E980 servers with FW920.30
4. SR-IOV support available on POWER9 Scale-out servers with FW910 and POWER9 E950 & E980 servers with FW930
5. SR-IOV support available on POWER9 E950 & E980 servers with FW920.20
6. SR-IOV support available on POWER9 Scale-out, E950 & E980 servers with FW930
7. SR-IOV support available on Power10 Servers with FW1030
8. SR-IOV support available on Power10 Servers with FW1050
9. SR-IOV support available on Power10 Servers with FW1060.10
10. IBM does not currently offer 200 Gb cable or transceiver options. Contact your IBM sales representative for 200 GbE applications.

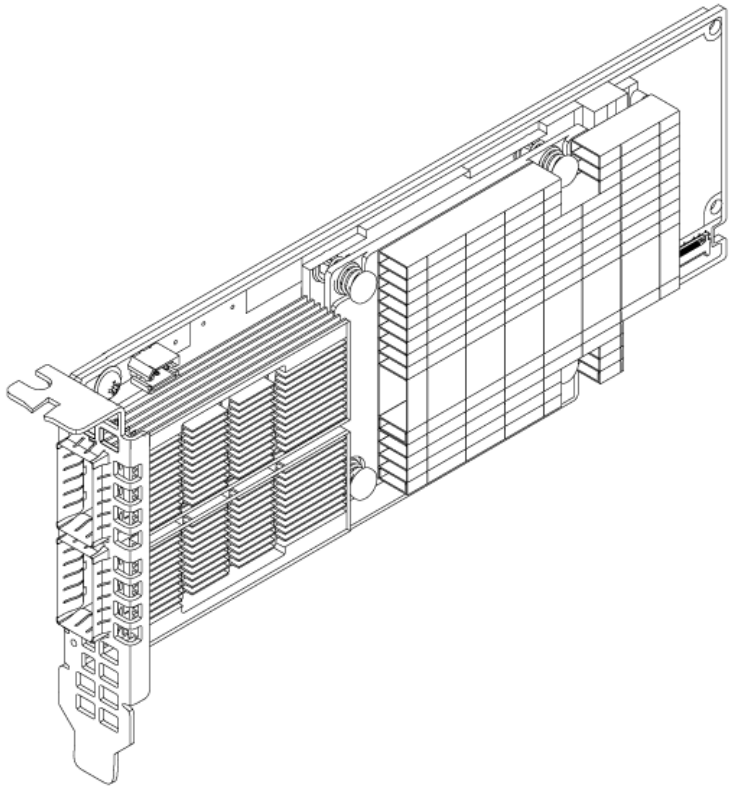
Maximum Logical Ports (VFs) per Adapter

SR-IOV Capable Network I/O Adapters	Feature codes	Physical port link speed	Max Logical ports per physical port	Max Logical ports per adapter
PCIe3 4-port (2x10GbE+2x1GbE) SR Optical fiber and RJ45	EN0J, EN0H, EL38, EL56	1Gb	4	48
		10Gb	20	
PCIe3 4-port (2x10GbE+2x1GbE) copper twinax and RJ45	EN0L, EN0K, EL3C, EL57	1Gb	4	48
		10Gb	20	
PCIe3 4-port (2x10GbE+2x1GbE) LR Optical fiber and RJ45	EN0N, EN0M	1Gb	4	48
		10Gb	20	
PCIe3 4-port 10GbE SR optical fiber	EN16, EN15	10Gb	16	64
PCIe3 4-port 10GbE copper twinax	EN18, EN17	10Gb	16	64
PCIe3x16 2-port 100/40GbE NIC&RoCE QSFP28 Adapter	EC3L, EC3M	100/40Gb	60	120
PCIe3 2-Port 10GbE NIC&RoCE SR/Cu Adapter	EC2R, EC2S	10Gb	40	80
PCIe3 2-Port 25/10GbE NIC&RoCE SR/Cu Adapter	EC2T, EC2U	25/10Gb	40	80
PCIe4x16 2-port 100/40GbE NIC&RoCE QSFP28 Adapter	EC67,EC66	100/40Gb	60	120
PCIe4 x16 2-port 100/40GbE NIC & **RoCE** QSFP28 Adapter	EC75,EC76	100/40Gb	80	160
PCIe4 2-Port 25/10/1 Gb **RoCE** SFP28 Adapter	EC71,EC72	25/10/1Gb	40	80
PCIe4x16 4-PORT 25/10/1GbE NIC&RoCE SR/Cu Adapter	EN24,EN26	25/10Gb	20	80
PCIe5x16 2-PORT 200/100/40GbE NIC&RoCE SR/Cu Adapter	EC85, EC86	200/100/40Gb	80	160

PCIe5 x16 2-Port 200 GbE RoCE Adapter FC EC85 und FC EC86; CCIN EC2C

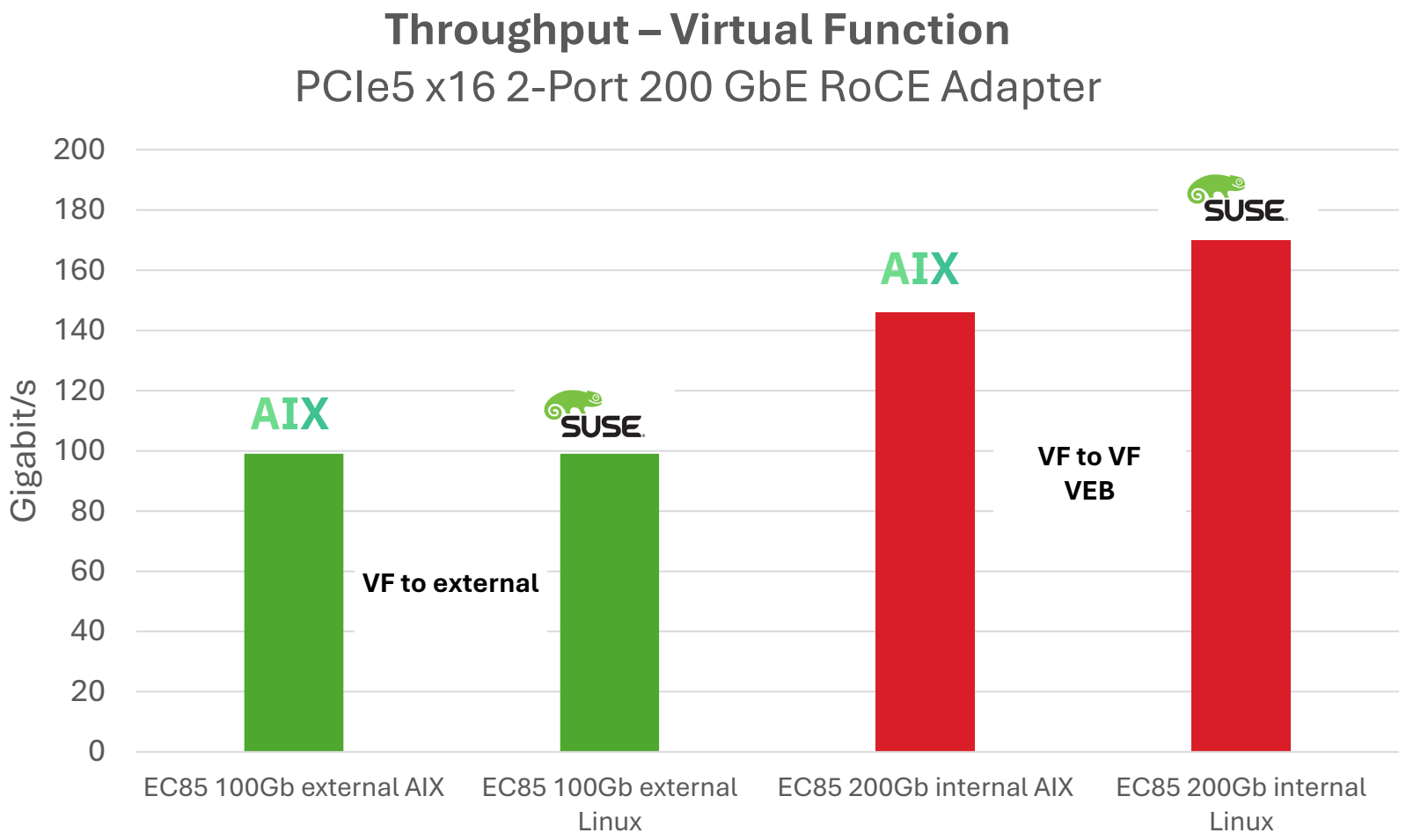


- PowerVM SR-IOV supported with Power11
- **Minimum Requirements**
  - **AIX 7.3 TL3 SP2, AIX 7.2 TL2 SP5 or above**
  - **Red Hat Enterprise Linux 9.4, SuSE Enterprise Server 15 Service Pack 6 or above**
  - **IBM i 7.6 or above**
  - VIOS 4.1.1.10, 3.1.4.60 or above
- Two 200 GbE QSFP56 ports
- Optical transceiver options
  - QSFP28 100Gbase-SR4 (FC EB59)
  - QSFP+ 40G BASE-SR4 (EB57)
- Copper options
  - QSFP28 Passive Copper 100 Gb[1)]
  - QSFP28 AOC 100 Gb[2)]
  - Passive QSFP+ 40 Gb[3)]
- IBM does not currently offer 200 Gb cable or transceiver options. Contact your IBM sales representative for 200 GbE applications.



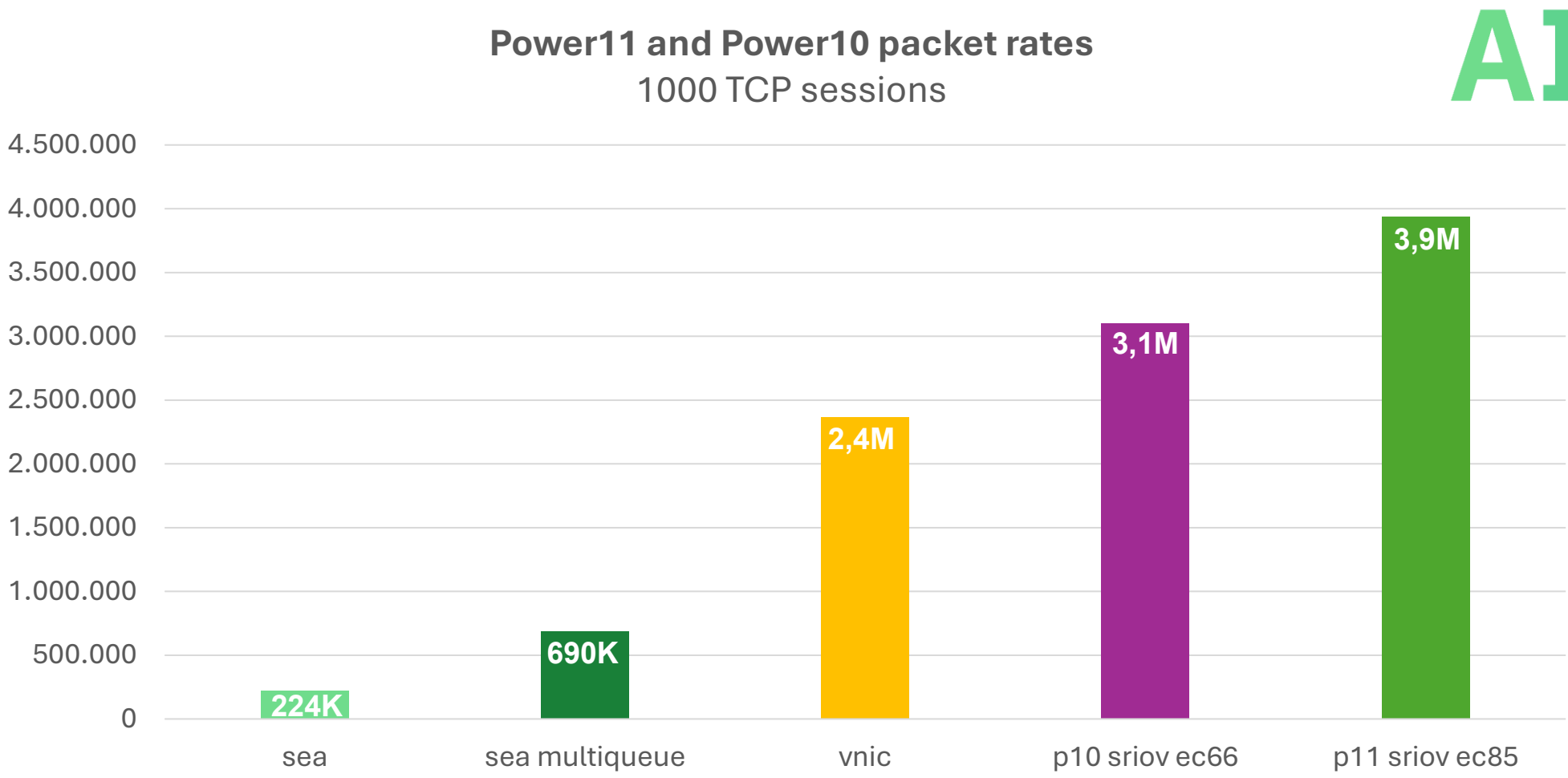
1) EB5J, EB5K, EB5L, EB5M (0.5 – 2 meter)
2) EB5R, EB5S, EB5T, EB5U, EB5V, EB5W, EB5X, EB5Y (5 – 100 meter)
3) EB2B, EB2H

PCIe5 x16 2-Port 200 GbE RoCE Adapter
SR-IOV Throughput



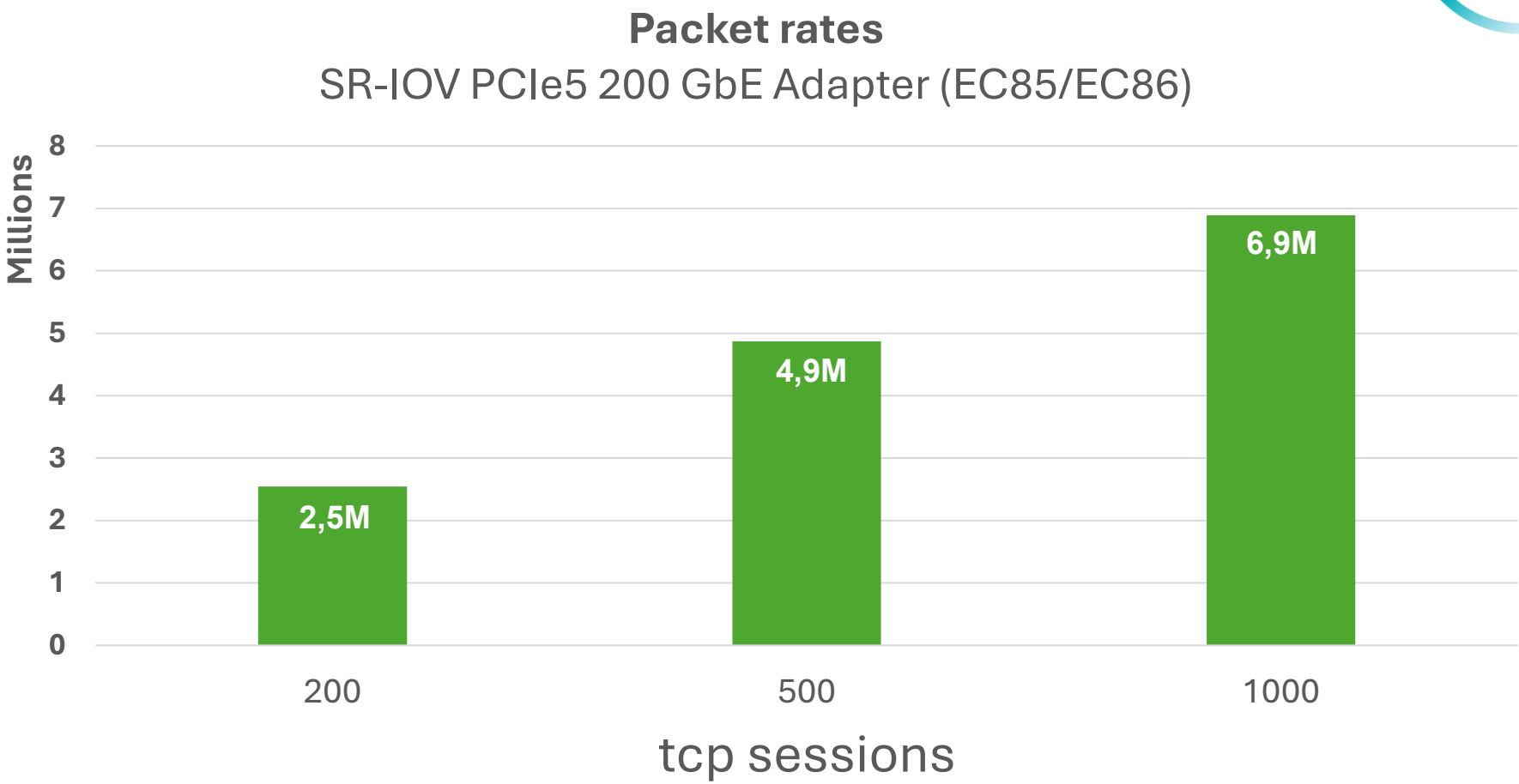
Tested with pre-GA code – real world results may vary.

Transactional Performance Power 11 and Power10 with AIX



Tested with pre-GA code – real world results may vary.

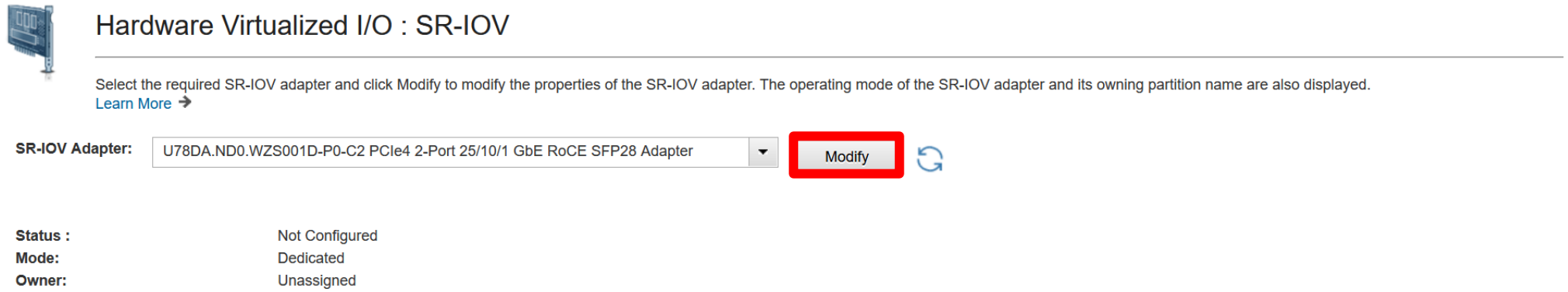
PCIe5 x16 2-Port 200 GbE RoCE Adapter
SR-IOV / HNV Performance on Linux



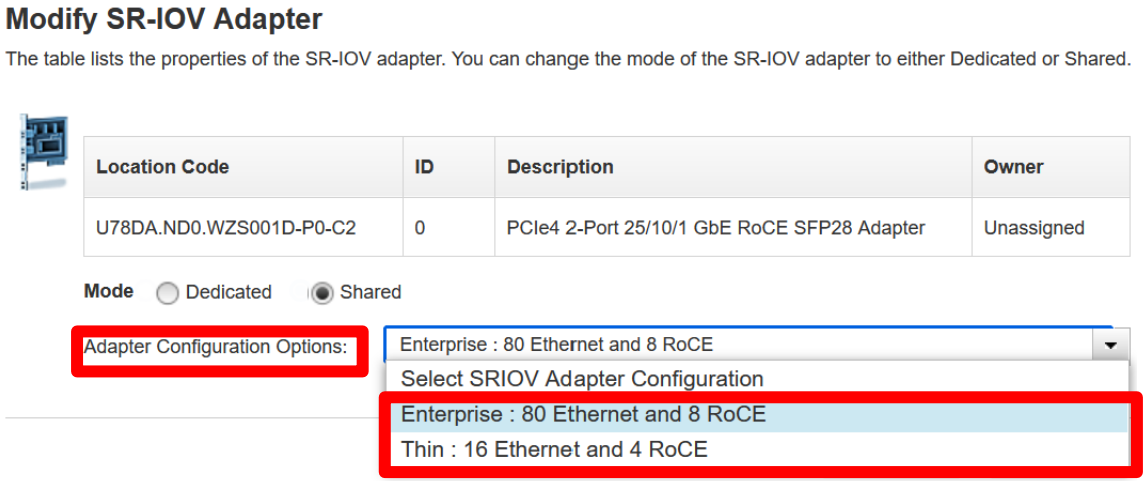
Tested with pre-GA code – real world results may vary.

Adapter Configuration Options (ACO) Configuration

- To determine if Adapter Configuration Options are available for an SR-IOV capable adapter, select Modify on the Hardware Virtualized I/O: SR-IOV screen when the adapter is in Dedicated Mode.



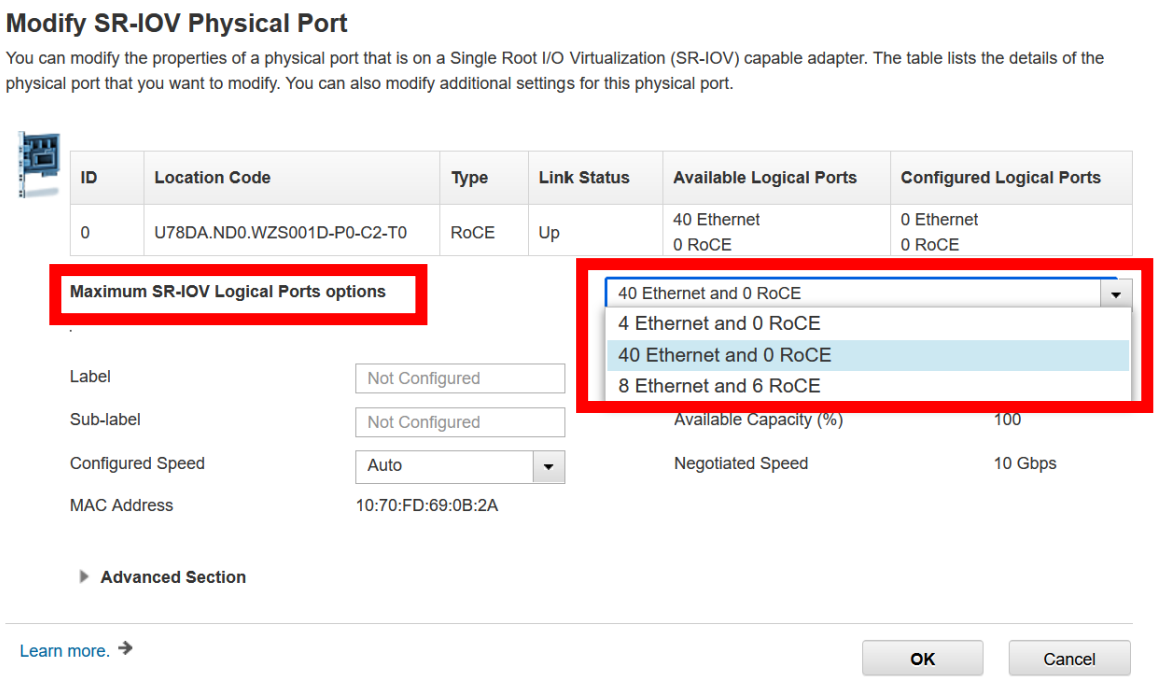
- Then select Shared for the Mode. This will display the Adapter Configuration Options.



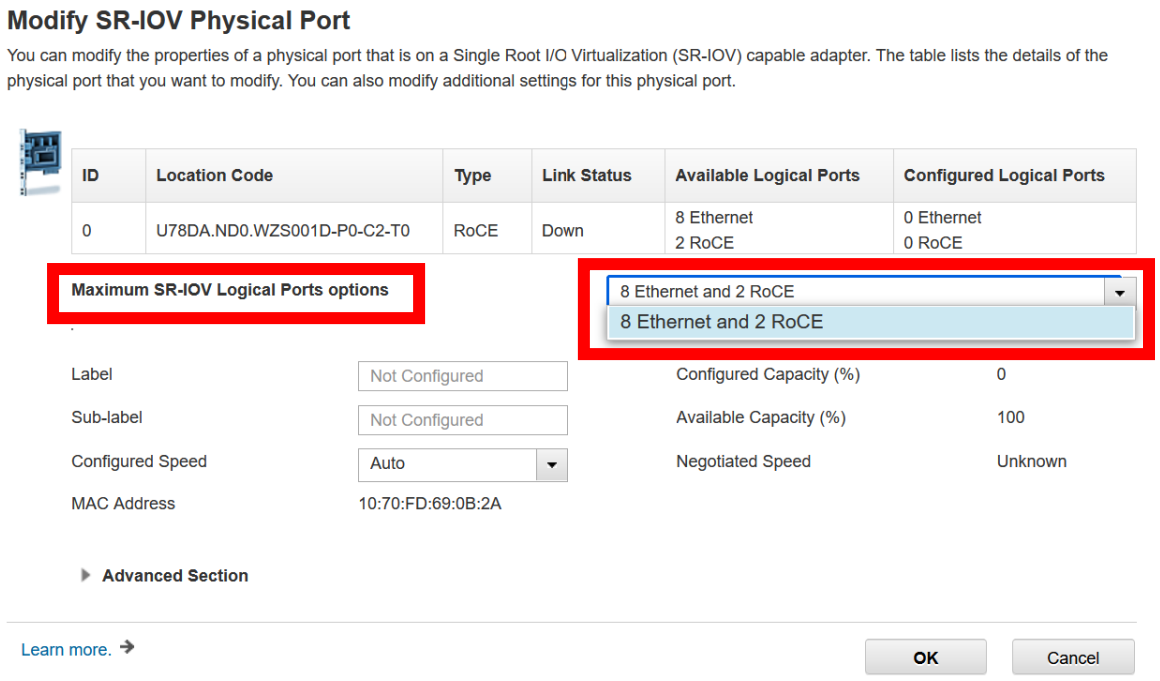
ACO Configuration Physical Port Options

- The ACO option selected determines the Physical Port's Logical Port options. The screen shots below shows the EC71/EC72 adapter Physical Port's Logical Port options for the two ACO options:

Enterprise Physical Port options



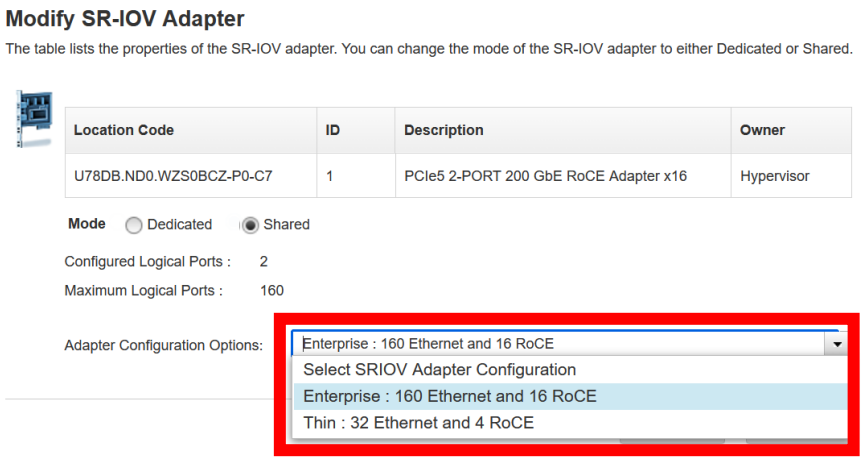
Thin Physical Port option

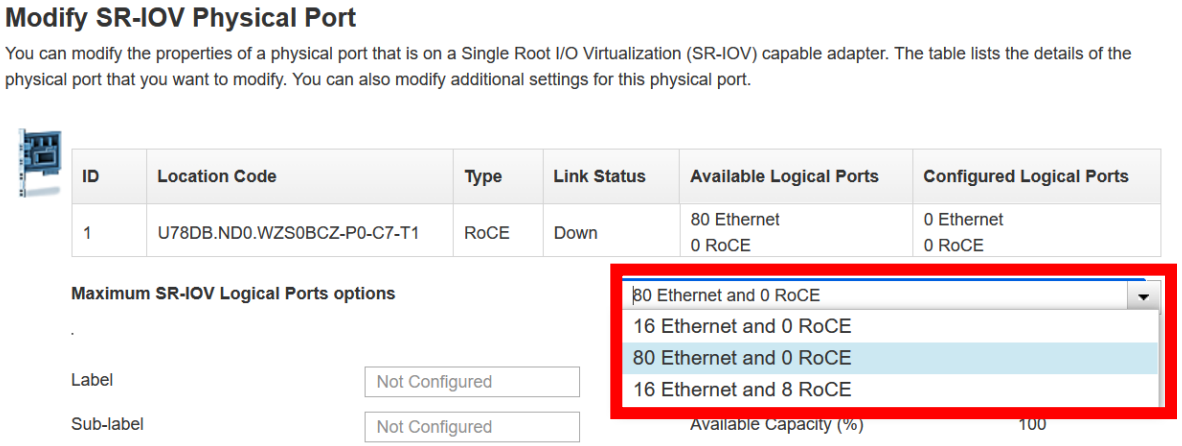


ACO vs. Physical Port Configration Mode

- Adapter Configuration Option (ACO) is **configurable per adapter**.
- Available for selected adapter features.
- Thin mode reduces hypervisor memory usage.
- Enterprise mode enables maximum VFs.

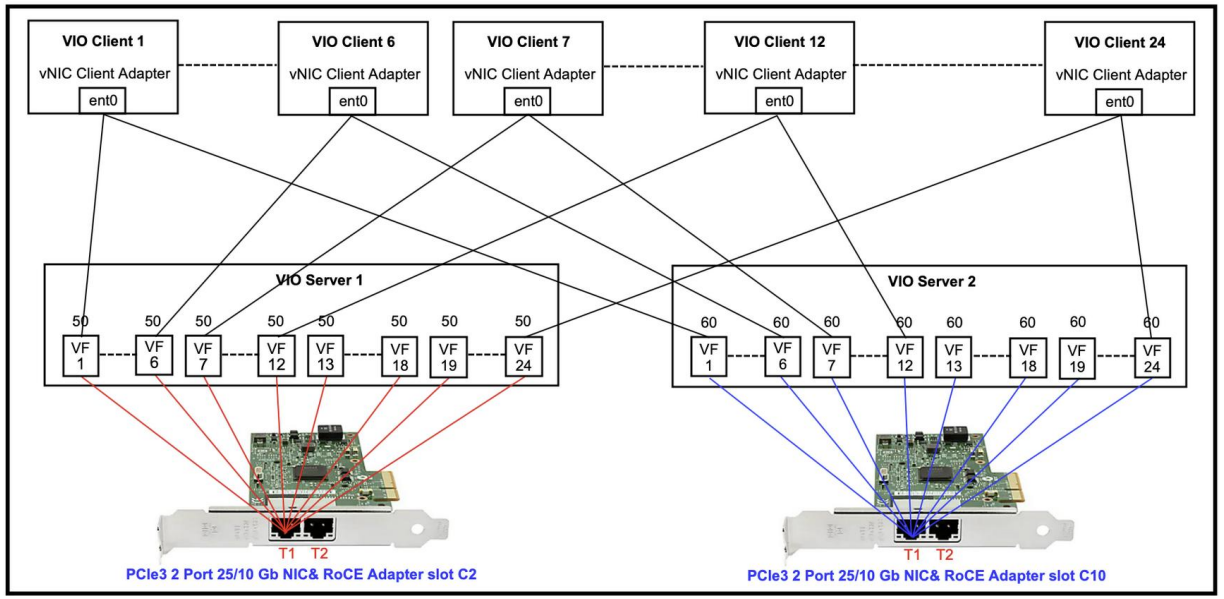
- Port Configuration Mode is configured **per physical port** if adapter is in SR-IOV shared mode.
- Reducing VFs enables more interrupt sources the client OS can use.
- Can bring ~30% better transaction rate (effective results depend on adapter feature).





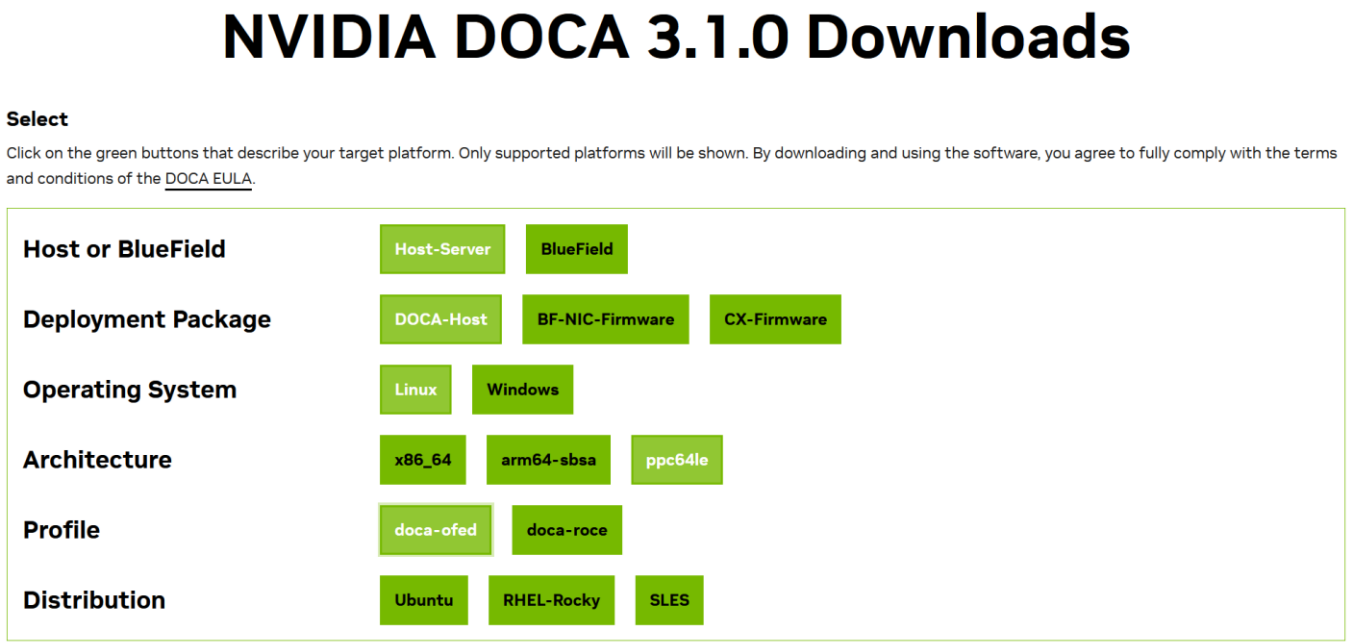
vNIC Failover Enhancements

- In the event of a failure, all affected active backing devices must fail over to a backing device with the next highest failover priority.
- Effective recovery time depends on available CPU capacity on the target VIOS.
- Using alternating failover priorities can help reduce the number of failover events.
- With VIOS 4.1.1.10 and later, vnic failover enhancements for planned disruptions (VIOS shutdown) are reduced.
- New fo_heartbeat attibute for viosvnic0 pseudo device was introduced.



NVIDIA OFED for SR-IOV Capable Adapters on Linux

- NVIDIA OFED can be used when assigning native SR-IOV Logical Ports or HNV to a Linux on Power LPAR.
- New: The OFED driver is now available via NVIDIA DOCA (Data Center Infrastructure-on-a-Chip Architecture) framework.
- OFED via DOCA can be installed online or offline.
- https://developer.nvidia.com/doca-downloads



NVIDIA OFED installation via DOCA

- **Red Hat Enterprise Linux**
  - **Local installation:**
    - `wget https://www.mellanox.com/downloads/DOCA/DOCA_v3.1.0/host/doca-host-3.1.0-091000_25.07_rhel96.ppc64le.rpm`
    - `sudo rpm -i doca-host-3.1.0-091000_25.07_rhel96.ppc64le.rpm`
    - `sudo dnf clean all`
    - `sudo dnf -y install doca-ofed`

- **SuSE Enterprise Server**
  - **Local installation:**
    - `wget https://www.mellanox.com/downloads/DOCA/DOCA_v3.1.0/host/doca-host-3.1.0-091000_25.07_sles15sp7.ppc64le.rpm`
    - `sudo rpm -i doca-host-3.1.0-091000_25.07_sles15sp7.ppc64le.rpm`
    - `sudo zypper refresh`
    - `sudo zypper install -y doca-ofed`

IRQ binding for ibmvnic

- Manual IRQ binding may further increase vNIC performance.
- IRQs need to be banned from Irqbalancer.
- Example for a LPAR with 8 Virtual CPUs:

```
/etc/sysconfig/irqbalance:
IRQBALANCE_ARGS="--banirq=55 --banirq=56 --banirq=57..."

systemctl restart irqbalance
ethtool -L $eth rx 8 tx 8
TXQs=$(cat /proc/interrupts | grep ibmvnic | grep tx | cut –d: -f2 | tr -d ' ')
RXQs=$(cat /proc/interrupts | grep ibmvnic | grep rx | cut –d: -f2 | tr -d ' ')
i=0 ; for rxq in $RXQs
do
          echo $i > /proc/irq/${rxq}/smp_affinity_list
          let i=i+4
done

i=2 ; for txq in $TXQs
do
          echo $i > /proc/irq/${txq}/smp_affinity_list
          let i=i+4
done
```



Network Performance with Power 11

Alexander Paul
Senior Technical Consultant
Technology Expert Labs
paulalex@de.ibm.com

COMMON EUROPE CONGRESS 2026
14. – 17. June 2026 – Lyon, France

IBM