

SR-IOV

By Jaqui Lynch

Introduction

The ability to share physical resources is the fundamental underpinning of the concept of virtualization, something IBM has been focusing on for a very long time. In the case of network technologies, IBM has multiple offerings and, as of 2014) they added an additional offering. That offering is SR-IOV or single root I/O virtualization. The initial SR-IOV offering was for some of the Power7 servers (9117-MMD, 9179-MHD and 8412-EAD) in April 2014. As of March 2015 the high end Power8 servers (E870 and E880) also have SR-IOV options. Finally, in June 2015 and December 2015 there were additional updates providing support for the POWER8 scale-out servers, the Power E850 and the PCIe Gen3 I/O expansion drawer.

What are my options?

There are multiple options for network adapters. The first option is dedicated adapters which offer the best possible performance but no resource sharing. This means the adapter is bound to the specific LPAR and is not available for sharing. Any unused bandwidth on that adapter is not available to any other LPAR.

The second option is the virtual Ethernet adapter. This allows you to set up a virtual adapter that allows multiple LPARs on the same server to communicate using hypervisor internal switching. This is also the underpinning of Shared Ethernet Adapters.

The third option is Shared Ethernet Adapters (SEA). These require the use of a VIO server. LPAR to LPAR on the same server uses the virtual Ethernet that underpins the SEA but all communication outside the server goes through the physical adapter that is associated with the SEA. The SEA offers multiple high availability options including SEA failover, SEA failover with load sharing, and NIB (network interface backup).

On Power7 (not Power7+ or Power8) there was also the option of a Host Ethernet Adapter (HEA or IVE) that could be shared between LPARs. This is no longer available on the more recent technology.

The final and newer option is SR-IOV. While SR-IOV has been around for well over a year now it is only now starting to be used for virtualization.

What is SR-IOV?

SR-IOV is an extension to the PCIe specification and it allows multiple LPARs or operating systems to share the same PCIe adapter with minimal intervention by the hypervisor. It also does not require the use of a VIO server, however it can be used with a VIO server should you wish to take advantage of technologies like LPM (live partition mobility).

SR-IOV allows you to share a PCIe adapter between multiple LPARs, each using direct access I/O to the adapter. Each LPAR gets assigned a virtual slice of the adapter through logical ports. Initially up to 48 logical ports are supported per adapter although this varies by adapter. To support more LPARs you can assign the SR-IOV adapter to a VIO server and use the logical port as the physical device for the SEA, allowing VIO server clients to then use the adapter over the SEA. SR-IOV provides direct access performance, high scalability and Quality of service (QOS) options. Additionally, SR-IOV ports can be aggregated into an etherchannel to provide for redundancy and bandwidth. This is done using LACP (link aggregation control protocol) aggregation. For AIX and IBM i, all of the ports in the aggregate must be SR-IOV – you cannot aggregate SR-IOV ports with other adapters.

By itself SR-IOV does not support LPM, however, when assigned to a VIO server SR-IOV adapters can be used in an LPM scenario.

Why SR-IOV over SEA?

SR-IOV is designed for high performance solutions where bandwidth is an issue. The use of direct access to the hardware reduces latency and path lengths and helps with performance. There is significant overhead involved in passing client instructions through a VIO to the adapter and back – the direct access provided by SR-IOV reduces that significantly. Additionally, SR-IOV provides QOS options that allow you to control bandwidth allocations per logical port.

SR-IOV Architecture

SR-IOV introduces two new concepts – Physical Functions (PFs) and Virtual Functions (VFs). A Physical Function is a PCIe function that supports SR-IOV capabilities. A Virtual Function is associated with a PF and is accessed by an operating system. The VF shares physical resources with the PF and any other VFs associated with the same PF. Power systems implement VFs as logical ports and each logical port is associated with a physical port on the adapter.

SR-IOV adapters can be used as regular dedicated adapters or they can be used in shared mode. Shared mode is where the hypervisor owns the adapter and shares the ports and functions out to the LPARs.

With PowerVM there is a new adapter type called a vNIC (virtual network interface controller). This is a new virtual adapter that leverages SR-IOV to enable advanced virtualization features such as LPM and QOS. This feature was announced in December 2015 as part of VIOS 2.2.4 and provides an alternative to using an SEA as well as a one to one relationship between the vNIC client and the SR-IOV logical port allowing for logical port QOS to be extended through the VIO server to the VIO client.

Planning for SR-IOV

There are a limited number of adapters that support SR-IOV so it is important to get the correct adapter. There are also specific hardware, software and firmware requirements, including the requirement for an HMC (hardware management console) at a minimum level of 7/7/9/0. Although a VIO server is not required, SR-IOV requires PowerVM standard or Enterprise edition as these are needed for vNIC support. The PCIe adapter guide also provides information on which I/O slots support SR-IOV. Section 2.2 of the SR-IOV Technical introduction and overview manual goes into details on the hardware and software required to support SR-IOV. It also provides a list of the supported adapters and their feature codes for ordering purposes. Section 2.3 covers operating system requirements to support SR-IOV.

Adapter Options

There are 6 options for the 2 port 10Gbe CAN and 2 port 1Gbe adapter. These are feature codes EN0J, EN0H, EN0L, EN0K, EN0N and EN0M. All of these provide for 2 x 10Ge ports and 2 x 1Gbe ports. The 10Gbe ports can each support up to 20 VFs each and the 1Gbe ports can support up to 4 VFs each for a total of 48 VFs for the adapter. Each logical port can be assigned to one or more LPARs and can have QOS controls applied.

The PCIe3 4 port 10Gb adapter (feature codes EN15, EN16, EN17 or EN18) provides for 4 x 10Gbe ports, each of which can support up to 16 VFs per port for a total of 64 VFs for the adapter.

Deployment Options

There are multiple ways to implement SR-IOV. Two of them require no VIO servers and two of them make use of VIO servers.

The first option is simply a single partition that owns the adapter and treats it as a regular Ethernet adapter. The adapter is in dedicated mode, uses direct access I/O, and does not provide any sharing options. All of the resources on the adapter are available to the owning LPAR.

The second option also uses Direct access I/O and the adapter is in sharing mode. There is no VIO server but the adapter allocates different logical ports to different LPARs and allows them to share the adapter. Capacity can be controlled using QOS settings per logical port.

The third option allocates the SR-IOV adapter to a VIO server and the VIO server treats it like a regular adapter using it for the SEA. This allows the VIO clients to be eligible for LPM (live partition mobility) actions.

The fourth and final option is a combination of Direct access I/O and VIO server access. The adapter runs in sharing mode allocating logical ports to LPARs and also logical ports to the VIO server.

For options three and four, LPARs accessing the adapter through the VIO server are eligible for LPM but do not receive the performance benefits that LPARs using direct access I/O receive. Additionally, while the logical port assigned to the VIO server can have QOS desired bandwidth assigned to it, there is no way to assign that at the more granular level of the VIO client LPAR.

Summary

SR-IOV is a new networking technology that provides enterprise capabilities such as adapter sharing and enhanced performance along with the ability to provide QOS options for the adapter so that it can share out different slices to clients depending on their bandwidth needs. SR-IOV adapters can be used as regular network adapters or can be run in shared mode supporting multiple LPARs from one card. Additionally, they can be combined into a VIO server environment to provide backing for the SEA or they can be configured as vNICs to provide support for enhanced functions such as LPM and client level QOS.

SR-IOV has the ability to reduce costs, to provide sharing with multiple deployment options, to provide for adapter bandwidth controls (QOS) and to provide for direct access I/O and performance. With the reduction in latency and overhead provided by SR-IOV adapters it is well worth investigating how they may improve your current environment as long as you are current on firmware, HMC, and operating system levels.

References

SR-IOV Technical Introduction and Overview

<http://www.redbooks.ibm.com/abstracts/redp5065.html?Open>

P011895 - SR-IOV Fundamentals – presentation given at pTechU comes to you in Chicago (May 2016)

PowerVM SR-IOV FAQ

https://www.ibm.com/developerworks/community/wikis/home?lang=en_us#!/wiki/Power%20Systems/page/SR-IOV%20Frequently%20Asked%20Questions

vNIC – Introducing a new PowerVM Virtual Networking Technology

https://www.ibm.com/developerworks/community/wikis/home?lang=en_us#!/wiki/Power%20Systems/page/vNIC%20-%20Introducing%20a%20New%20PowerVM%20Virtual%20Networking%20Technology

IBm PowerVM vNIC FAQ

https://www.ibm.com/developerworks/community/wikis/home?lang=en_us#!/wiki/Power%20Systems/page/vNIC%20Frequently%20Asked%20Questions

Blog – Introduction to vNIC FAQs

https://www.ibm.com/developerworks/community/wikis/home?lang=en_us#!/wiki/Power%20Systems/page/Introduction%20to%20vNIC%20FAQs