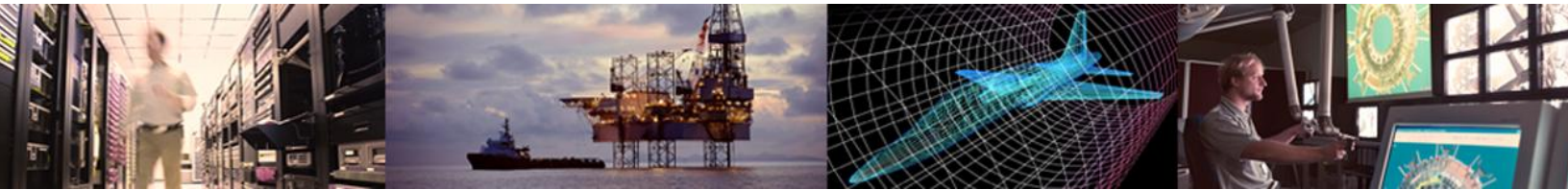


Strategic Direction

Unified SR-IOV for Citrix XenServer “Project Boston”



Solarflare 10G Ethernet Server Adapters Deliver Unified Single-Root I/O Virtualization (SR-IOV) for Citrix XenServer “Project Boston”



Solarflare 10 Gigabit Ethernet Server Adapters enable IT managers to consolidate more virtualized operating systems over fewer physical servers by providing scalable network I/O bandwidth and removing I/O bottlenecks in virtual server environments. Solarflare has released support for a complete single-root I/O virtualization, or SR-IOV solution, which delivers native network performance on virtualized servers. Furthermore, Solarflare’s unique implementation allows users to leverage hypervisor-based services and management tools, such as virtual machine live migration, while benefitting from dramatically improved I/O performance.

In addition to native performance with full hypervisor functionality, Solarflare’s SR-IOV solution provides the highest workload consolidation scalability in the market. Solarflare server adapters support more VFs and DMA channels accelerating and isolating more VMs than other solutions, up to 254 VFs and 2048 DMA channels (up to 8 times more than other adapters).

Traditionally, virtualized servers have relied on the hypervisor to fully manage data flow, which provides powerful services and management capabilities, but at the same time introduces overheads and bottlenecks that can dramatically impact network I/O performance. To address these performance issues, Citrix XenServer’s “Project Boston” implements a new emerging standard called single-root I/O virtualization, or SR-IOV, that promises to significantly improve network I/O performance on virtualized servers. SR-IOV provides a mechanism for guest operating systems to bypass the XenServer dom0 and interact directly with the network hardware. The objective of this PCI standard is to enable virtualized servers to achieve far greater network performance than is realized today with hypervisor-based network virtualization techniques. As servers become increasingly powerful, and the number of consolidated workloads on virtualized servers continues to scale, addressing this network I/O bottleneck is becoming increasingly important.

While promising dramatic improvements in I/O performance, typical SR-IOV implementations are incompatible with hypervisor-based management tools that enable such capabilities as live migration. This limitation imposes severe restrictions on the usability of many SR-IOV solutions in real data center environments.

Solarflare has worked jointly with Citrix to develop a unique unified approach to SR-IOV that delivers unprecedented cut-through I/O performance while maintaining full hypervisor-based management. As a result, virtualized servers using Solarflare’s 10G Ethernet adapters can benefit from accelerated network performance as well as hypervisor-based

services and management capabilities. This means that users can achieve higher I/O performance, scale to consolidate more workloads, and maintain efficient hypervisor-based services and management that makes server virtualization so attractive to users.

TRADITIONAL SERVER VIRTUALIZATION (WITHOUT SR-IOV)

Traditional Xen paravirtualized networking involves bridging (paravirtualized) Netback devices to physical network devices in dom0. The Netback driver interfaces with the Netfront driver running in the guest, facilitating guest to guest connectivity (see Diagram 1).

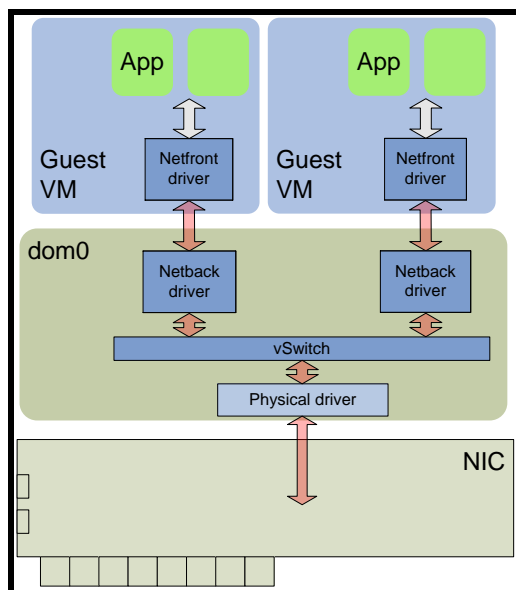


Diagram 1. Guest to Guest Connectivity

As a result of hardware virtualization, data copies between the Netfront driver in the guest and the Netback driver in dom0 within the XenServer host become the bottleneck to the performance of this implementation.

SR-IOV

In emerging SR-IOV solutions PCIe hardware virtual functions (VF) are used to allow a single network hardware device to appear as multiple virtualized network devices to dom0. These virtualized network devices operate independently of one another and present characteristics of actual physical devices to dom0. Using SR-IOV, a VF is passed-through to the guest operating system, and the guest's network driver binds directly to this PCIe VF, allowing the guest to bypass dom0 and providing direct access to the network adapter from the guest VM. As a result of this direct guest access to the network hardware, the overheads associated with dom0-based networking (virtualization, data copies, etc.) are eliminated, providing significantly improved performance (see Diagram 2).

However, a significant downside of this approach is that XenServer would block migration of a VM when a VF is present, which would significantly reduce the functionality of server virtualization.

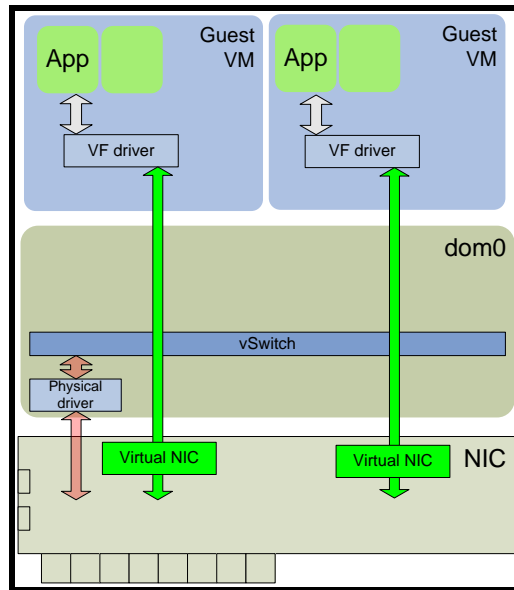


Diagram 2. Direct Guest Access to NIC Hardware

It is also worth noting that the networking interface in this model is now a vendor supplied driver and creation/configuration of networking interfaces must be done via the vendor's driver and not the standard XenServer network drivers.

SOLARFLARE UNIFIED SR-IOV MODEL

Solarflare, in a collaborated development with Citrix, has implemented a unique unified approach to SR-IOV that enables accelerated cut-through performance while maintaining full compatibility with hypervisor-based services and management tools. Furthermore, Solarflare's 254 VFs enable highly scalable workload consolidation for large data center environments and virtual desktop infrastructure (VDI) deployments. Unlike other adapters, each Solarflare VF can utilize multiple DMA channels that enable the VF to scale over vCPUs using RSS and/or RFS. Solarflare's 2048 DMA channels provide the most scalable virtualized network I/O solution available in the market.

Solarflare's SR-IOV implementation uses a plug-in approach that maintains the traditional (software) data path through the XenServer paravirtualized drivers, and then through the vSwitch to the physical device network driver. In addition, there is also an alternative (accelerated) data path through the VF driver directly to the network adapter from the guest. Packets are received on either data path transparently to the guest VM's network stack. For transmitted data, the plug-in (when loaded and enabled) makes the decision on whether or not to use the accelerated path (see Diagram 3).

With this approach a VM can be created/cloned using traditional tools, and networking to/from the VM can initially utilize the standard software network path. Citrix XenServer's "Project Boston" will pass-through a VF from the network adapter into the guest, the guest sees new hardware has been hot plugged and binds the Solarflare plug-in driver to this VF. This plug-in driver automatically registers with the Netfront driver as an accelerated network plug-in. Once the VF driver has registered, subsequent traffic to/from the guest uses the accelerated data path accessing the adapter directly from the guest. If the VF is hot unplugged (i.e. removed from the guest), the plug-in deregisters with the Netfront driver and the networking traffic reverts to the software data path.

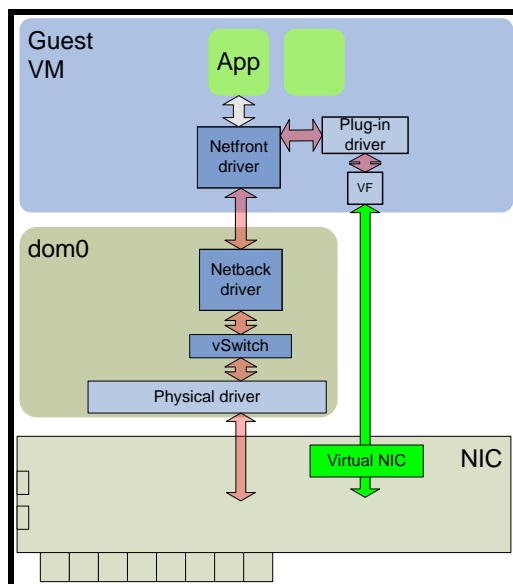


Diagram 3. Network Plug-in Model

This approach means there is no dependency on the VF or its driver for the networking data path of the VM. Acceleration can be disabled at any time if needed without losing network connectivity, and migration is fully supported in this model. When a VM is being migrated, XenServer hot unplugs the VF from the VM and the network path reverts to the software path, and after migration is complete XenServer hot plugs a new VF from an adapter on the destination machine back to the VM and the networking path is again accelerated. Solarflare's unique approach also removes the limitations imposed by typical SR-IOV implementations, enabling VMs to be moved between non-identical hosts. For example, if a guest is migrated to a host containing a different vendor's NIC, then the hot plug event provides the opportunity to bind a new VF driver within the guest and re-establish accelerated networking. Equally, if the new host does not contain suitable hardware, then no accelerated path is established, yet networking will continue using the original (or traditional) paravirtualized architecture.

SUMMARY

Solarflare's unique approach to SR-IOV combines the benefits of accelerated cut-through performance while maintaining full compatibility with hypervisor-based services and management tools. In addition to improved performance and superior manageability, Solarflare adapters support up to 254 VFs, each of which can support multiple DMA channels (up to a total of 2048), providing far more scalable workload consolidation for large data center environments and virtual desktop (VDI) deployments than is otherwise available. To further improve performance, Solarflare adapters utilize the available DMA channels within a VF to support Receive Side Scaling (RSS) for guest VMs, spreading VM workloads across many CPU cores enabling performance to scale with the number of CPU cores.