I/O Virtualization
Increase Adapter Utilization by 80% and Cut I/O Management by 50%

Micron Technology, Inc.
May, 2012

Server Evolution

Volume servers today trace their ancestry (and, unfortunately, their architecture) back to the humble PC of the early 1980s. The PC was designed to provide low cost, commodity home computing. Applications included simple word processing, spreadsheets, and games. Input and output beyond the keyboard and display were optional and rare, and networking was irrelevant.

In contrast, business computers (or enterprise servers, as they are now known) were designed with a completely different aim: The storage and manipulation of large amounts of business-critical data. Their architectures were much more I/O-centric and had surprisingly few processing capabilities.

A quarter of a century later, the ubiquity of the PC has changed the face of enterprise servers. Volume servers today are effectively PCs—albeit with far more powerful CPUs, memory, and I/O devices. We gained acquisition cost and scalability advantages that came with the high-volume PC market, but the business demands on enterprise servers remain much the same as they always were in terms of reliability, storage capacity, bandwidth, networking, and connectivity—demands that the PC was never intended to address.

Over the last decade, PCs have addressed these demands by simply providing more hardware, but that trend is no longer sustainable. In particular, power, floor space, and management costs have become dominant costs in the data center. Mainframe-like capabilities have been making their way into both data centers and CPUs for some years. Networked storage, optical networks, and server virtualization have all sought to address the growth and manageability requirements, but the architecture of the PC server has resolutely remained the same as it was 20 years ago, with CPU and memory and some dedicated optional I/O.

What is I/O in Server Architecture?

I/O can be defined as all components and capabilities that provide the CPU, and ultimately the business application, with data from the outside world and enable it to communicate with other computers, storage, and clients.

The I/O in a typical server consists of:
- Ethernet network adapters (NICs), which enable it to communicate with clients and other computers
- Networked storage adapters (HBAs), which provide connectivity into shared storage pools
- Local disk storage (DAS) for nonvolatile storage of local data, operating system (OS), and server “state”

I/O also includes all the cables and networking infrastructure required to interconnect the many servers in a typical data center. Each server has its own private set of I/O components.

In 1965, Intel’s Gordon Moore observed that transistor densities (and by implication, CPU capabilities) were growing exponentially, doubling approximately every two years. This trend has continued to this day. Unfortunately for servers, this trend has been more relevant to CPUs than to I/O. I/O performance is determined by more than transistor density, and it has lagged behind, particularly on a measure of cost versus performance. Today, I/O can account for more than half the cost of server hardware.
I/O Virtualization

In recent years, data centers have turned to a variety of virtualization technologies to ensure that their capital assets are used efficiently. Virtualization is the concept of separating a function, such as storage, memory, or processing, from the underlying physical hardware (disks, DRAM, and CPUs). This enables the physical hardware to be pooled and shared across multiple applications, increasing its utilization and capital efficiency while maintaining the standard execution model for applications. Virtualization also makes the deployment of new applications much quicker since it does not involve the manual installation and configuration of new servers, networks, and storage.

Virtualization consists of three distinct steps:

- **Separation of resources**: Providing management independence
- **Consolidation into pools**: Increasing utilization and saving cost, power, and space
- **Virtualization**: Emulating the original I/O functions as virtual functions to avoid software disruption

I/O virtualization (IOV) follows the same concept. Servers, networked storage, and local direct attached storage (DAS). Instead of providing each server with dedicated adapters, cables, network ports, and disks, IOV separates the physical I/O from the server. Servers are transformed into highly compact, space-efficient pure compute resources (such as 1U servers), and the I/O from multiple servers can be consolidated into an IOV appliance.

Because the I/O components are shared across many servers, they can be better utilized. Also, the number of components can be significantly reduced compared to a nonvirtualized system, making the system more cost-, space-, and power-efficient. The system is also more reliable and easier to manage due to architectural advantages and fewer components.

The final step is to create virtual I/O devices that look the same to the server software as the original physical I/O devices. This functional transparency preserves the end-users’ significant investment in software applications, operating systems, drivers, and management tools. The server hardware appears to the software to have the same I/O architecture, the same I/O devices, and the same drivers—all managed with the same tools—but with the cost, space, power and dynamic configuration, and manageability advantages that come with I/O consolidation.
Micron’s I/O Virtualization Technology

The benefits of IOV have been available to high-end systems for some time, either as proprietary implementations in mainframes or based on InfiniBand. Until now, there has not been a solution that was either fully transparent (as previously described) or available at a suitable cost point for the volume server market.

The default I/O interconnect in volume servers is PCI. All server chipset and I/O devices have used PCI as their common interface since the early 1990s. The introduction of PCI Express® (PCIe) in 2004 has enabled this interface to reach beyond the confines of the server motherboard, and it can now be run over backplane and cables. Today, all volume servers and I/O devices support PCIe. The PCI-SIG has also recently defined a number of extensions to PCIe to support IOV capabilities—within a single server (SR-IOV) and across multiple servers (MR-IOV). However, these extensions have not been fully transparent with respect to standard PCIe and require new modified I/O devices.

Micron’s IOV approach, although compatible with the PCI MR-IOV standard, is based on virtualizing the high-volume standard PCIe I/O adapters available today. Micron’s IOV is fully transparent to existing servers and I/O adapters and requires no changes to server or I/O hardware, drivers, or management tools. Based on the native server I/O interconnect (PCIe), it does not require any extra adapters or interfaces to other interconnects, such as InfiniBand or 10Gb Ethernet. Architecturally, Micron’s IOV technology looks very much like PCI today.

The server chipset has a number of PCIe interfaces that connect to I/O devices, typically via a PCIe switch. With standard PCIe, only a single server can connect to a switch and, hence, to the I/O devices.

With Micron’s IOV (and PCI MR-IOV), the PCIe switch (IOVE) has been enhanced to enable multiple servers to connect to the switch and for the I/O devices to be shared across many servers. However, the sharing in MR-IOV has required the I/O devices to be modified. With Micron’s IOV, the sharing functionality has been built into the switch fabric in the form of a virtualiza-
tion proxy controller (VPC), which is a Micron hardware device that works with the switch to virtualize multiple standard PCIe cards.

The final component of Micron’s IOV technology is the VIO management software (VMS). This provides out-of-band management of the PCIe switching and virtualization capabilities and an execution platform for an I/O vendor’s standard management tools.

Key Features and Benefits

Hardware Cost Reduction Through Consolidation

The I/O subsystem can account for half of the acquisition cost of server hardware, but it is typically poorly utilized (about 15%) due to the variable and unpredictable I/O requirements of different applications. By consolidating the I/O from multiple servers, utilization can be increased and the total amount of hardware can be reduced for a given workload.

Power Reduction

As with acquisition costs, increasing the I/O utilization through consolidation minimizes the amount of I/O hardware required and, thus, power consumption.

Floor Space Reduction

Eliminating the need for servers to support physical I/O adapters enables the increased use of high-density 1U form factors, saving valuable floor space and extending the life of a data center.

Management Simplification

IOV transforms server configuration from a hands-on, lights-on manual operation involving installation of adapters, cables, and switches to a software operation suitable for remote or automated management. It is estimated that 40% of data center outages are due to human error. Removing humans from the data center and providing automated validation of configuration changes enhances data center availability.

Dynamic Configuration Agility

Businesses today need to adapt quickly to change if they want to succeed. Their IT infrastructure needs to be agile to support rapidly changing workloads and new applications. IOV enables servers to be dynamically configured to meet the processing, storage, and I/O requirements of new applications in seconds rather than days.

Ease of Deployment/Nondisruptive Integration

Micron’s PCIe-based IOV technology has been designed specifically to avoid any disruption to existing software, hardware, or operational models in data centers. Micron’s IOV works with and is invisible to existing volume servers, I/O adapters, management tools, operating systems, and drivers, making its deployment in data centers extremely transparent.

Figure 5. Today’s Server I/O Architecture

Figure 6. Micron’s I/O Virtualization Architecture
Compatibility with Server Architectures

By basing its IOV technology on PCIe, Micron has ensured compatibility with the architectures and strict space and power constraints of server chassis.

This technology also simplifies the management of servers by effectively removing the networking aggregation switches from the chassis and restoring a clean separation between server, storage, and network management operations.

No Performance Cost-of-Virtualization

Micron’s technology virtualizes I/O adapters directly at the PCIe hardware level, ensuring that there is no “cost-of-virtualization,” which has been seen in older software virtualization approaches. Software is used only in nonperformance-critical management functions.

Rapid Adoption of New CPU and I/O Technologies

CPU and I/O technologies have been evolving at different rates. New, more powerful and cost/power effective CPUs typically appear every nine months, while new I/O technology generations come only every three to five years. The separation of I/O from the compute resources in servers (CPU and memory) enables new CPUs to be introduced quickly without disrupting the I/O subsystems. Similarly, new I/O technologies can be introduced as soon as they are available. Since these new high-cost and high-performance I/O adapters are shared across multiple servers, their introduction cost can be significantly reduced compared with today’s deployment model.

Conclusion

I/O virtualization is a breakthrough innovation that enables I/O adapters to be separated, consolidated, and virtualized away from the physical confines of a server enclosure. IOV can increase adapter utilization by 80% and cut the cost of I/O management by 50% while delivering better space, power, and cooling benefits to data centers.

The ideal approach to IOV comes from leveraging ubiquitous PCIe standards that enable server I/O bus to reach beyond the confines of the server motherboard and run over backplane and cables. Micron is the world’s leader in implementation of IOV technologies based on PCIe standards.