

Jul 24 '14

## POWER8: A Great Leap Forward in Performance

by [Jaqui Lynch](#) in [Enterprise Tech Journal](#)

On April 28, 2014, IBM once again delivered on its Power roadmap with the announcement of the first of the new POWER8 servers. Altogether, five new models were announced: three that run AIX, Linux and iOS and two that are Linux-only. With POWER, IBM offers a technology that can run multiple operating systems on the same hardware simultaneously. This lets users choose operating systems depending on the application while still using the same server. Additionally, within the Linux-only line of servers, users can now choose between using the traditional PowerVM for virtualization on POWER or they can use POWERKVM, the POWER implementation of Kernel-based Virtual Machine (KVM) for Linux. This lets users utilize the virtualization technology on Linux-only boxes they're most familiar with. Additionally, IBM, through the OpenPOWER Foundation, released detailed technical specifications for its POWER8 processor, making it one of the rare operating systems that's open.

The new systems are designed for performance, especially for the optimization of analytics performance and cloud applications. The initial servers are scale-out servers that also provide great reliability in a small package. These should be followed later by scale-up servers.

Power Systems are designed for Big Data with innovation that optimizes analytics performance. They provide superior cloud economics and the ability to scale-out or scale-up with security built into the foundation. As an open server platform at the heart of the OpenPOWER Foundation, Power Systems are revolutionizing the development and delivery of new technology.

### What Does It All Mean?

So what was actually announced? IBM announced three servers that can run AIX, iOS and/or PowerLinux (SUSE or Red Hat). These consist of the S814, which is a one-socket 4U server; the S822, which is a one- or two-socket 2U server; and the S824, which is a one- or two-socket 4U server. Additionally, IBM announced two Linux-only servers—the one-socket 2U S812L and the two-socket 2U S822L. The Linux-only servers support Red Hat and SUSE as well as Canonical Ubuntu Linux, and SUSE and Ubuntu will both run a little endian, which solves a major translation issue for MS Windows systems. The three operating system servers became generally available June 10, 2014.

All five servers take advantage of the POWER8 chip, which provides up to 12 cores per socket and allows for SMT8, which is the ability to run eight threads at a time on the core. The chip has huge caches to improve performance with a per-core allocation of a 64KB L1 data cache, a 32KB L1 instruction cache and a 512KB L2 private cache. The L3 cache is now 8MB per core and there are two integrated memory controllers. Memory is installed as memory cards and the intelligence for the memory is on those cards, which act as a form of L4 cache. The servers all use dual chip modules (DCMs), where each DCM equates to one socket and has two chips on it. A DCM can contain six, eight, 10 or 12 cores and each DCM has up to eight paths to the memory cards. The memory cards are available in 16GB, 32GB and 64GB, and it's cheaper per gigabyte to use the 32GB and 64GB cards. Active memory expansion (AME) is still provided as an option to further extend memory. To get the best performance, it's recommended that at least half the memory slots be filled.

The servers all use integrated PCIe Gen3 slots, which significantly improve I/O bandwidth; those slots are hotplug, which makes card replacement much simpler. The servers come with three different backplane options that impact the number of disk bays. The base backplane isn't available yet but it has one SAS controller and 12 SFF3 (small form factor) disk bays. The default backplane is the split backplane, which has two SAS controllers and 12 SFF3 disk bays.

The bays are split evenly between the two controllers (six and six). Finally, there's the expanded function backplane that provides 18 SFF3 disk bays on the one-socket S814 and 18 SFF3 disk bays plus 6 x 1.8-inch solid-state disk (SSD) bays on the two-socket S822 and S824. When using the expanded function backplane, it takes an additional slot but on the two-socket servers it allows you to take advantage of the SSDs using Easy Tier. Additionally, the expanded function backplane has write cache on it, which can improve performance for write-sensitive applications. Keep in mind that these are SFF3 disks so they use a different carrier from disks in the POWER6 and POWER7 servers and the expansion drawers.

With respect to I/O, initially there's no I/O drawer support, although an EXP24 (5887) SAS disk drawer can be attached. The number of slots varies, depending on the server:

- 4U1S – 7
- 4U2S – 11
- 2U1S – 6
- 2U2S – 9.

All the slots are hot swap and PCIe3. The Gen3 cards will have twice the bandwidth as the older Gen2 cards and three times the bandwidth of the Gen1 cards. If you use the expanded function backplane, there's one less slot available and, typically, you lose one slot for the LAN adapter. Default hardware maintenance on the servers is three years, 9 to 5, next business day.

## Firmware

Starting with POWER8, firmware will require a license key to be installed and activated. The POWER8 servers require machine code update entitlement at activation via an update access key (UAK) that's checked at every activation and installation of firmware code. If it fails entitlement, the update won't proceed. Replacement UAKs can be obtained at [www.ibm.com/servers/eserver/ess](http://www.ibm.com/servers/eserver/ess). The initial UAK is placed on the box by IBM and is based on machine type and serial number.

## New I/O Cards

To get the maximum value out of the I/O slots, four new network cards were announced. They consist of a two-port 10GbE RJ45, a four-port 10GbE+1GbE SR optical, a four-port 10GbE+1GbE copper twinax and a two-port 40GbE QSFP+. Currently, the first three cards aren't supported for network install manager (NIM) installations, so if NIM is to be used to install the LPARs, it will be necessary to use one of the other LAN adapters for the network. The specific feature codes are: #enow, enox, enos, enot, enov and enou. Note that the DVD drive is connected to one of the SAS controllers and doesn't float. It will go to whichever LPAR has the disk bays associated with that controller.

## New HMC

Along with the new POWER8 servers, IBM also announced a new HMC and monitor. The new HMC is the 7042-CR8 and has dual 500GB disk drives in a RAID-1 configuration with a base of 8GB. I recommend adding a second 8GB dimm to ensure the best performance. To support POWER8, the HMC will need to run the new HMC code 8R810, which became generally available June 10, 2014. This code is supported on the 7042-CR5 and higher. The 7316-TF3 monitor has been withdrawn and a new monitor (7316-TF4) is now available.

## SMT8 Support and Performance

To get SMT8 support, the LPAR will need to run AIX v7.1 TL3 SP3. Prior versions will provide SMT4 support. IBM

provided rPerf (relative performance) ratings for the new servers for SMT1, SMT2, SMT4 and SMT8. Even if you can't upgrade to the latest version of AIX, there are performance advantages to moving to POWER8. As an example, a POWER7+ 3.5GHz 740 with 32 cores is rated at 354.90 rPerf when running in SMT4 mode. A new S824 with 24 x 3.5GHz cores is rated at 394.1 rPerf in SMT4 mode and 421.8 rPerf if running SMT8. So, even if you can only run in SMT4 mode, there are significant savings due to only needing 24 licenses on the S824 vs. 32 on the 740. To put this in perspective, a 40-core 5GHz p6-595 is rated at 368.82 rPerf (only does SMT2) and requires significantly more power, cooling and space as well as more licenses than the 24 core S824. Additionally, it has significant requirements for the physical environment (such as floor loading, etc.). A 32-core 4GHz p7-795 is rated at 372.27 rPerf (SMT4), which is less than the 24-core S824 when it's in SMT4 mode.

## Summary

IBM's latest evolution in the Power line of servers is a great leap forward in terms of performance. These low-end, scale-out servers provide the performance of the higher-end servers from previous Power versions in the size and power and cooling envelope of a low-end server. They still have all the built-in reliability functions such as memory sparing, redundant and hot swap power and fans and blowers, hot swap adapters and processor instruction retry. They also include support for concurrent firmware updates, thus maximizing uptime for the servers.

## Resources

- POWER8 home page at [www.ibm.com/power](http://www.ibm.com/power)
- IBM PowerVM Live Partition Mobility at [www.redbooks.ibm.com/redbooks/pdfs/sg247460.pdf](http://www.redbooks.ibm.com/redbooks/pdfs/sg247460.pdf)
- POWER8 Facts and Features at [www-03.ibm.com/systems/power/hardware/reports/factsfeatures.html](http://www-03.ibm.com/systems/power/hardware/reports/factsfeatures.html)
- AIX Virtual User Group (lots of topics including POWER8) at [www.tinyurl.com/ibmaixvug](http://www.tinyurl.com/ibmaixvug). **ETI**