close window

Web Exclusive

Print 🖨

My Hands-On Experience With POWER8

August 2014 | by Jaqui Lynch

In June, my company received its first two POWER8 servers to use for testing. One server was a single-socket S814 with PowerVM, RHEL and AIX, and the second was a Linux-only box running PowerKVM and RHEL Linux. I had the privilege of working with the AIX server, which I attached to an HMC and our V7000 disk subsystem.

The Specs

The S814 is configured with the high-function backplane so it wasn't a split backplane. It has 8 x 3.7ghz cores, 128GB of memory, 12 x 387GB SSDs, 4 x 300GB hard drives, 2 x dual port 16Gb Fibre cards, 2 x 4 port dual 10Gb and dual 1Gb FCoE cards as well as the 4 port 1Gb network card. I chose the FCoE cards rather than the new 10Gb network cards because they're supported by network install manager (NIM). The S822L Linux only box has 10 x 3.4ghz cores and 64GB of memory with a 4 port 1Gbe adapter and 2 x 300GB hard drives. Because the S822L is running PowerKVM, it's not connected to an HMC. This article will focus on the S814.

The S814's high-function backplane means all disks and SSDs were on the same controller, which went to one VIO server. The second VIO server will be booted from SAN and configured later. I installed two client LPARs—one running AIX v7.1 tl03 sp3 and the other running AIX v6.1 tl09 sp3. The HMC was a CR7 and was installed at the latest 8.8.1.0 MH01441, which is still the highest level.

The Setup

After connecting everything, powering up the server and flashing the microcode, I ensured the logical memory block or memory region size was set to the same as our other servers. This is required using live partition mobility and you must power cycle the box if you have to change it. The configuration and setup of the VIO and LPAR profiles was no different to any other server and very few differences were noted in the HMC GUI, so the setup was simple. Since the NIM server had been upgraded to AIX v7.1 tl03 sp3, the next step was to create the LPP source and SPOT for NIM for the new AIX v6 and v7 levels to be installed. Additionally, I used installios against the VIO server DVDs to create the necessary VIO resources on the NIM master.

The VIO images were also loaded to an FTP location on the NIM server. After creating the VIO profile, I booted the LPAR and checked the Yes box at the top of the activation pop-up that says, "Install Virtual I/O Server as part of activation process?" Once the VIO server was running, I went through all the normal checks. Ioslevel shows as 2.2.3.3 and oslevel –s shows the operating system at 6100-09-03-1415. This means the VIO will be running in SMT4 mode since SMT8 requires 7.1 tl03 sp3. The following shows this is the case:

```
# smtctl
This system is SMT capable.
This system supports up to 4 SMT threads per processor.
SMT is currently enabled.
```

After setting up the VIO on two of the four internal disks, I then set up FBO (file backed optical) and loaded a bunch of ISO images for various sets of software. I always set up FBO if I have a spare disk as I tend to lose the DVDs. By ripping them to ISO images and uploading them to FBO, I can always find them and recreate the DVDs as needed. And, of course, I can load them remotely to any of the LPARs controlled by that VIO server. After setting up FBO, it looked as follows:

```
$ lssp
Pool              Size(mb)    Free(mb)  Alloc Size(mb)    BDs Type
rootvg              540672      407552             512      2 LVPOOL
fbovg              1089024      577024             512      1 LVPOOL

$ lsrep
Size(mb) Free(mb) Parent Pool          Parent Size     Parent Free
  511414   490594 fbovg                     1089024         577024
```

From this, we can tell that the FBO pool is approximately 1TB in size with about half of that space unused.

```
$ lsvopt
VTD           Media                                        Size(mb)
vtopt0        No Media                                          n/a
vtopt1        No Media                                          n/a
vtopt2        No Media                                          n/a
vtopt3        No Media                                          n/a
```

Finally, with the VIO server up and running it was time to create the client LPARs. I created an AIX v7 LPAR and an AIX v6 LPAR, both with an entitlement of three and six virtual processors (VPs) and 32GB of memory. Both were provided SAN-based disks via vSCSI from the VIO server. Once everything was running, I ran smtctl to check the LPARs. It turns out that by default the AIX v7 LPAR comes up in SMT4 initially. I ran 'smtctl –t 8' to change the AIX v7 LPAR to SMT8. vmstat shows an entitlement of 3 and logical CPUs (LCPUs) of 48 on the AIX v7 LPAR and LCPUs of 24 on the AIX v6 LPAR.

SMT8 Support and Performance

To get SMT8 support on POWER8, the LPAR needs to run AIX v7.1 TL3 SP3, otherwise the LPAR will run in SMT4 mode. I decided to use nmem, ncpu and my own test program written in C to compare the three LPARs, which all had the same entitlement, VPs and memory. On the S814, one LPAR ran AIX v7 with SMT8 and a second ran AIX v6 with SMT4. The third ran AIX v7 with SMT4 on a 750B (which has 32 x 3.3ghz cores). Memory on the 750 was 1066mhz and on the S822L and S814 was 1600mhz, which puts the 750 memory at about 66 percent of the speed of the S814 memory. Comparing rPerfs for three cores on each shows the 750 LPAR to be estimated at 60.42 percent of the rPerf of the SMT8 AIXv7 POWER8 LPAR and 64.66 percent of the SMT4 AIXv6 POWER8 LPAR. The AIX v6 SMT4 POWER8 LPAR is estimated at 93.44 percent of the AIX v7 SMT8 POWER8 LPAR. So how did they stack up?

Using nmem I saw the following as an overall average (OPS/S K is thousands of operations per second):

```
              OPS/S K
b814aix1      29820.78              P8 AIX v7 SMT8
b814aix2      27483.75              P8 AIX v6 SMT4
b850nl1       21708.83              P7 AIX v7 SMT4

Comparatives  OPS/S K
nl1/aix1      73.89%
```

```
aix1/nl1        135.54%
nl1/aix2        81.94%
aix2/nl1        122.69%
aix2/aix1       90.43%
aix1/aix2       110.78%
```

Comparing the OPS/S K, we can see the 750 LPAR performed better than predicted by the difference in mhz in the memory. As we all know, mhz is not everything; variables like buffering, cache sizes and even instruction type all impact how quickly things move through memory. With that said, the memory performance in operations per second was still 22 percent to 35 percent better on POWER8 than it was on POWER7, which is a significant improvement.

Comparatives	rPerf	Results (128 processes)	C Program
nl1/aix1	60.42%	67.95%	59.43%
aix1/nl1	165.5%	147.17%	168.28%
nl1/aix2	64.66%	74.19%	71.31%
aix2/nl1	154.65%	134.78%	140.23%
aix2/aix1	93.44%	91.58%	83.33%
aix1/aix2	107.02%	109.19%	120%

When comparing rPerf, I'd expect the 750 LPAR to be between 60 and 65 percent of the POWER8 LPARs and the AIXv7 POWER8 LPAR to be around 7 to 8 percent faster than the AIX v6 POWER8 LPAR. Multiple tests were run with ncpu using varying numbers of processes. The results reported are averages from multiple runs using 128 processes. In all three LPARs, this exceeded the number of threads available to the LPAR and it stressed the CPU accordingly. When comparing the user time from ncpu for 128 processes on each LPAR to the other LPARs, it's clear that the two POWER8 LPARs seem to scale very closely to how the rPerf scales. The POWER7 LPAR did better than expected using 47 percent more user time than the AIX v7 P8 LPAR instead of 65.5 percent more and 34.78 percent more user time than the AIX v6 P8 LPAR instead of 54.65 percent. Similar results were seen using the C program written to drive CPU.

My Conclusions

I should note that these were very limited tests and there are many more tests left for me to run in the benchmarking suites that I am testing. Additionally, they don't test all of the functions of the server and are not a true mixed OLTP workload, which is what most run. However, they do provide some initial data that shows that POWER8 appears to scale as expected, both due to improvements in the memory performance as well as CPU performance. Also the jump from SMT4 to SMT8 provides around a 9 percent boost, which is on a par with what's predicted in the published rPerf. Further, more detailed tests are planned using several test suites for memory, cpu, I/O and network performance. Although preliminary and limited, they provide a window into the performance potential of the new POWER8 scale-out servers.

Overall, the POWER8 experience has been very positive so far. Firmware and HMC updates went smoothly and the server appears to be performing as expected. This provides a level of confidence that you can move from POWER7 to POWER8 while reducing the VPs or cores in LPARs using rPerf comparisons as an approximate scaling factor.