

New POWER8 E850 rocks!

By Jaqui Lynch

On May 11, 2015 IBM made a set of announcements expanding the enterprise server line of POWER8 servers by adding options to the E870 and E880 and also by adding a new server, the E850 (84-8-E8E). The E850 combines enterprise qualities into a 4 socket 4U server that is incredibly powerful. Additionally it provides for COD (capacity on demand) which makes this very attractive to customers who want flexible performance in a compact but reliable package.

Specifications

The E850 comes in 3 different core options. Each socket is populated by one DCM (dual chip module). The first core option uses 3.02ghz cores with 12 cores per socket for a maximum total of 48 cores. The second uses 3.35ghz cores with 10 cores per socket for a maximum of 40 cores. The third option offers 3.72ghz cores with 8 cores per socket for a maximum of 32 cores. A minimum of 2 sockets must be filled and at least half the cores must be activated. Using PowerVM these cores can be virtualized at 20 micro-partitions per core and can take advantage of all the services offered by PowerVM as well as the other operating systems (currently AIX and Linux).

With respect to memory there are 8 CDIMMs per socket for a total of 32 CDIMMs if all 4 sockets are populated. Today, the maximum capacity is 2TB per server although there is an SOD to go to 4TB. Minimum memory to be installed is 128GB. At least half the installed memory has to be activated and a minimum of 8 CDIMMs have to be populated. As with all servers maximum memory bandwidth (192 GBps per socket) is only attained if all the CDIMMs for that socket are filled.

There are two x16 PCIe Gen3 slots per socket with a maximum of 8. Additionally there are 3 x8 PCIe Gen3 slots but one of these contains a 2 port 10Gb LAN card. This means that with 2 sockets full there are 7 PCIe slots, with 3 sockets there are 9 slots and with 4 sockets there are 11 slots. I/O bandwidth also differs depending on the number of sockets filled:

2 sockets	192 GBps
3 sockets	256 GBps
4 sockets	320 GBps

To further extend slot options it is now possible to add PCIe Gen3 I/O drawers to the E850. One full I/O drawer can be added per populated socket so between 2 and 4 I/O drawers can be added. Each drawer adds a total of 10 PCIe Gen3 slots for a combined maximum of 51 slots if 4 drawers are added to a 4 socket system. These drawers are for PCIe cards only (no disk) so if additional disks are needed then either an EXP24S SAS drawer needs to be added or a fibre HBA to connect

to the SAN. Up to 64 x EXP24S drawers, each holding up to 24 SFF disks or SSDs, can be added. This combination makes the E850 incredibly flexible.

With respect to disk there are 8 hotswap SFF SAS disk drive bays (2.5") that can contain disks or SSDs. Additionally there are 4 x 1.8" SSD bays. There is also a bay for a DVD drive but there is no place for a tape drive. Easy Tier can be used to migrate between hard drives and SSDs to improve performance.

There are three options for the storage backplane. There is a dual SAS controller backplane with a 7.3GB write cache, a dual SAS controller backplane with no write cache and a split disk backplane with no write cache. I normally configure the split disk backplane as this allows you to boot two VIO servers from internal disks or to split the disks between two LPARs. When the split backplane is used each controller handles 4 SFF disk bays and 2 of the SSD bays. The DVD is assigned to one of the controllers only.

The server requires 200-240 volt AC power and takes 4U in a rack. It must be mounted with the rack rails that it is shipped with. The default warranty is Monday to Friday, 9 hours per day for one year. Warranty uplifts are available to 24 x 7 plus extensions in time. In terms of power the E850 has a maximum power source loading of 3.57 kVA and a maximum power consumption of 3500 watts. The maximum thermal output is 11,940 BTU/hour. The server weighs a maximum of 152lbs so it is not a one person job to rack the server.

Other exciting features

In addition to the specifications listed above, the E850 offers capacity on demand for both cores and memory, AME (Active Memory Expansion) and Active Memory Mirroring for the Hypervisor. It includes all the POWER8 chip RAS (reliability, availability and serviceability) capabilities along with memory protection, hot plug SAS bays, hot plug PCIe slots, redundant and hot plug power supplies and cooling fans, hot plug Time of Day battery and a highly resilient architecture for power regulators.

Operating System Support

The E850 can run as a standalone system or be partitioned using either IVM (Integrated Virtualization Manager) or an HMC (Hardware Management Console). The preferred method for most customers would be the HMC, which is required if you want to use redundant VIO servers or most of the advanced features of PowerVM. If an HMC is being used then it needs to be at least a 7042-CR5 or 7042-C08 and at least 8GB of memory is recommended although it can run (slowly) with 4GB. The E850 will require the latest HMC code which is 8.8.3.0 or later – that code can be downloaded from IBM Fixcentral.

The E850 also requires the latest levels of PowerVM and AIX. For PowerVM the supported version is 2.2.3.51 or later. For AIX v6.1 and v7.1 the minimum supported level depends on whether all I/O is virtualized or not. The required

levels are listed in section 2.14.2 of the E850 Technical Overview and Introduction redbook. Multiple levels and flavors of Linux are also supported (see section 2.14.3 of the redbook). IBM i is not supported at this time.

Capacity on Demand (COD)

The E850 also provides for COD. The server can be ordered with some cores and memory inactive and they can be activated later as necessary. This reduces the initial costs as activations and the associated software licenses are only paid for when they are needed. COD comes in several different offerings – CUoD (Capacity upgrade on demand), elastic COD, Utility COD and trial COD. Additionally it is possible to turn dark cores and memory into Linux IFLs (Integrated facility for Linux). Most COD functions, with the exception of CUoD, require an HMC.

CUoD

CUoD allows for the permanent activation of inactive cores and memory. Additional cores and memory are purchased upfront, however, they are not activated until they are needed. Activations are available in 1 core and/or 1GB increments. With CUoD it is also necessary to purchase software licenses for the additional cores when they are activated as these are permanent activations.

For elastic COD, utility COD and trial COD, core activations include incremental licensing for some IBM software and operating systems (see 2.6.5 of the red book). This means that additional licenses would not need to be purchased when the activations are done for cores. These three all require an HMC.

Elastic COD allows for the temporary activation and deactivation of processor cores and memory. These activations are for a period of 24 hours from the time the activation is enabled. Increments are in 1 core and/or 1GB and charges are based on usage reporting which is collected monthly. There is no limit on the number of times processors and memory can be activated or deactivated, however there is a number of activations where it would be better to permanently activate the memory or cores for financial reasons.

Utility COD allows for the temporary activation of cores into the shared processor pool. These cores can only be used for LPARs using micropartitioning. Usage is charged on a per minute basis only when 100% of the permanently activated processors is exceeded. Utility COD requires PowerVM Enterprise Edition to be active on the server. As with elastic COD usage is monitored and reported back to IBM and bills are then generated based on reported usage.

Trial COD allows for a 30 day activation of up to eight cores and/or 64GB of memory. This allows you to deal with emergency situations or to test whether activating cores and memory will solve a problem. There is an additional exception request which allows you to activate all cores or all memory (or both) for 30 days – this can only be done once in the lifetime of the server.

Linux IFLs

IBM has made a very strong commitment to Linux and it shows clearly in the implementation of IFLs. If there are inactive cores and memory on the E850 it is possible to order an IFL package. This provides 4 core activations, 32GB of memory activations, and 4 PowerVM for Linux entitlements. These are at a significantly reduced price to the normal activation and entitlement costs. These are permanent activations and can count towards the minimum number of activations required for the server.

Performance

Not only does the E850 have great technical specifications and reliability, it also exhibits phenomenal performance and scalability. The E850 ranges from 304.5 rPerf for 2 sockets and 16 cores at 3.7ghz to 746 rPerf for 4 sockets and 48 cores at 3.02ghz (SMT8). For SMT4 this range is 284.51 (16 cores at 3.7ghz) to 697.87 (48 cores at 4.08ghz). Comparatively a Power7+ 770 D model with 16 x 3.8ghz cores is rated at 219.16 rPerf and with 48 x 4.2ghz cores is rated at 612.70 rPerf.

A key point to note is that the performance within the E850 is consistent as it is a one node system, whereas with the 770 the memory bandwidth drops significantly once you have to cross to the other nodes to use memory. As an example the theoretical memory bandwidth on the E850 is 192GBps per socket with a potential maximum of 768 GBps if all 4 sockets are populated. The 770 provides for up to 68.22GBps per socket for memory (272.88GBps per 4 socket node), however the interconnects run at 158.02GBps so memory performance is reduced significantly if cross-node access is necessary.

The maximum theoretical bandwidth for I/O for the E850 is also dependent on the number of sockets installed with 192GBps for 2 sockets, 256GBps for 3 sockets and 320GBps if all 4 sockets are populated. For the 770 it is around 78.848GBps for a single 4 socket enclosure with a theoretical maximum for I/O bandwidth of 315.392GBps with all 4 nodes in use.

These performance and reliability characteristics show why applications that are latency or memory sensitive will be great performers on the E850. The E850, with its one node design will provide consistent and fast memory and I/O access within a 4U package. The memory performance on the E850 is significantly better for 4 sockets on the E850 than it is on the 770 and the I/O performance in the E850 for 4 sockets is the equivalent to the I/O performance on the 770 for all 16 sockets across 4 nodes.

Summary

The E850 is a very attractive server, combining small packaging with great performance and pricing. It fits nicely into the POWER8 lineup, completing the

range. Anyone who would normally consider the 770 server should take a close look at the E850 as it may well address all the items of interest including scalability, performance and reliability.

References

E850 Overview

<http://www-03.ibm.com/systems/power/hardware/e850/>

E850 Specifications

<http://www-03.ibm.com/systems/power/hardware/e850/specs.html>

E850 Performance Data

<http://www-03.ibm.com/systems/power/hardware/e850/perfdata.html>

E850 Datasheet

http://www.ibm.com/common/ssi/cgi-bin/ssialias?subtype=SP&infotype=PM&appname=STGE_PO_PO_USEN&htmlfid=POD03108USEN

E850 Technical Overview and Introduction

<http://www.redbooks.ibm.com/redpieces/pdfs/redp5222.pdf>

5 things to know about the new E850

https://www.ibm.com/developerworks/community/blogs/5things/entry/5_things_to_know_about_the_ibm_power_system_e850_server?lang=en

IBM Fixcentral

<http://www.ibm.com/support/fixcentral>