

## Architecting Power Systems for I/O

By Jaqui Lynch

Whenever I architect Power Systems infrastructure for a client, I do two things. The first is to ensure the backplane on the server (where offered) is split. This gives the client more flexibility should they want to run two VIO servers or LPARs booting from an internal disk. The cost is minimal and it is much easier to do it upfront than to try and back-port it. But additionally, I pay attention to where the adapters get installed in the server and I/O drawers. Architecting the system correctly upfront saves a significant amount of time, money and effort. It also ensures that you don't run into issues where adapters require certain slots that other adapters may also require.

Whenever you purchase a new system you should always download two manuals: The Technical Overview Redbook and the PCI Adapter Placement Guide. These two manuals provide a wealth of information on the server and on the I/O connections. Using the E850C as an example, I will look at designing for I/O below. Remember, that for every high-performance port that is active (anything 8GB and above), you will need to add extra memory to the system. If using NPIV there is additional memory needed for every VFC (virtual fibre client). Details on this are provided in the VIO server presentation in the references.

### ***CEC (The Server Itself)***

The CEC in the E850C consists of two, three or four processor modules. The number of available slots depends on the number of processor modules installed. Peak bandwidth supported by x16 slots is 16GBps in simplex mode and 32GBps in duplex mode. The x16 slots provide half that bandwidth.

<i>Processor Modules</i>	<i>PCIe X16 slots</i>	<i>PCIe x8 slots</i>
2	4	3
3	6	3
4	8	3

The slot numbering is not intuitively obvious. In documentation IBM sometimes refers to the slot numbers and slot addresses interchangeably. The table below provides both for translation purposes:

<i>Module</i>	<i>Slot No</i>	<i>Slot Address</i>	<i>Type</i>
0	9	C10	x16
0	10	C11	x8
0	11	C12	x16
1	5	C6	x8
1	6	C7	x8
1	7	C8	x16
1	8	C9	x16
2	3	C3	x16
2	4	C4	x16
3	1	C1	x16
3	2	C2	x16

At least one PCIe ethernet adapter must be installed in the CEC. The slot identified for this is C11 (aka slot 10). Slots C6 and C7 share bandwidth and are also shared with the USB controller. For every adapter, there is a slot priority list provided in the adapter placement guide. The best performance is obtained for an adapter by ensuring it is in the highest priority adapter slot. Redundancy across processor modules can be obtained by putting adapters in the highest priority slot in two separate modules.

As an example: In the E850C the EN0A 16GB 2 port fibre card lists the following slot priority for a four-module server: 9, 7, 4, 2, 11, 8, 3, 1, 5, 6. So, with two cards, the best performance and redundancy would be obtained by putting one card in slot 9 (C10) and the second in slot 7 (C8). I normally map out all the cards based on which are the highest priority. This depends on workload—for some workloads the network cards are more important.

Slot assignment is also affected by external I/O drawers and the specific cards selected to support those drawers. As an example, the EJ14 RAID card that is used to connect to the external disk drawer is only supported in slots 9, 7, 4, 2 in the CEC. They are supported in different slots if you are using a PCIe3 adapter drawer. Gathering this information ahead of time and mapping it out in a spreadsheet can ensure a much smoother installation and implementation.

### ***PCIe3 I/O Drawer***

The EMX0 PCIe3 adapter I/O drawer is typically connected using two EJ08 cable adapters. These can go in any slot in the CEC, however the priority order is 9, 7, 3, 1, 11, 8, 4, 2. I would put these in slots 9 and 7 to provide the best performance for the adapters in the I/O drawer. If you are adding an EXP24S disk drawer you may need to use different slots for this drawer as the disk drawer has more limited options. The EMX0 drawer consists of the drawer itself plus one or two fanout modules. Each fanout module provides six PCIe3 slots as follows:

Fanout	Slot	Slot Address	Type
1	1	P1-C1	x16
1	2	P1-C2	x8
1	3	P1-C3	x8
1	4	P1-C4	x16
1	5	P1-C5	x8
1	6	P1-C6	x8
2	1	P2-C1	x16
2	2	P2-C2	x8
2	3	P2-C3	x8
2	4	P2-C4	x16
2	5	P2-C5	x8
2	6	P2-C6	x8

The left side of the drawer (from the rear) is P1 and the right side is P2. Adapters should be split across the fanouts for redundancy. For consistency, I like to split them to the same slot on each side. For example, if I'm using P1-C1 for one of my fibre pairs then I will use P2-C1 for the other. The PCIe adapter placement guide provides the list of slot priorities for the EMX0 I/O drawer.

### **5887 EXP24S Disk I/O Drawer**

If you also have an EXP24S disk I/O drawer this adds a level of complication. If it is a simple mode 1 installation (all disks owned by one LPAR) then it is most likely connected using one or two adapters (some of the adapters must be installed in pairs). The EJ0L and EJ14 adapters are only supported in four slots – 9, 7, 4, 2 of the CEC and only 9 and 7 are available in a two module system. They're also supported in the EMX0 PCIe3 adapter drawer.

When selecting the EXP24s you need to specify disk mode. Disk modes refer to how the disk plane is split up and connected and must be selected at order time. If the disk is mode 2 or 4 then additional adapters are required. The EXP24S has 24 disk bays in groups of 6. In mode 1 all 24 disks are connected to one (or a pair of RAID) adapter/s. In mode 2 they are broken into two sets of 12 and in mode 4 there are four sets of six disks. The drawer can support up to four LPARs if it's in mode 4, but at least four (and possibly eight) adapters are needed along with any required cables. Each increase in mode requires more adapters to support the additional connections. These adapters can be in the CEC or in the EMX0 PCIe drawer.

### **Pre and Post Order**

Even if it's not mandatory I always do a solutions assurance (now called TDA). The intent is to ensure the configuration is valid and will meet the needs of the workloads. Part of that TDA includes a spreadsheet that maps out the location of every adapter card. That spreadsheet can then be used to install the adapters in the predetermined slots prior to loading the operating system. That way everything is set to the best possible options right from the start.

### **Summary**

When trying to plan for I/O on POWER servers it is important to refer to the Technical Overview and Adapter Placement guides to ensure the configuration being built is valid and will meet the needs for the workloads. As an example, if you need to have 6 LPARs boot from internal disk then you will need a split backplane in the CEC and a mode 4 EXP24S disk drawer. This will most likely require one or more EMX0 PCIe3 adapter drawers. It is much easier to get this right when ordering than to try and fix it after the server arrives.

This level of planning isn't onerous but should be done for all new servers in order to get the most out of the new equipment. Although we used the E850C as an example, the adapter guides and overview manuals exist for all the current servers and can be found with a simple Google search. It is well worth investing the time and effort in this level of planning as it allows you to maximize performance and redundancy while also ensuring the solution will work.

### **References**

E850C 8408-e4e Adapter Placement Guide

[ftp://ftp.software.ibm.com/systems/power/docs/hw/p8/p8eab\\_85x.pdf](ftp://ftp.software.ibm.com/systems/power/docs/hw/p8/p8eab_85x.pdf)

OR

[https://www.ibm.com/support/knowledgecenter/5148-21L/p8eab/p8eab\\_85x\\_slot\\_details.htm](https://www.ibm.com/support/knowledgecenter/5148-21L/p8eab/p8eab_85x_slot_details.htm)

E850C 8408-e4e Technical Overview and Introduction

<http://www.redbooks.ibm.com/abstracts/redp5412.html?Open>

Care and feeding of VIO servers

<https://youtu.be/Z0wShBTjXhM>

<https://www.ibm.com/developerworks/community/wikis/form/anonymous/api/wiki/61ad9cf2-c6a3-4d2c-b779-61ff0266d32a/page/1cb956e8-4160-4bea-a956-e51490c2b920/attachment/4611d5c8-6941-478b-876f-b459205ffd0c/media/CareandFeedingofVIOs-VUG-Aug2117-2up.pdf>