

# Re: Tradeoff between Single Enclosure VC Domain and Multi Enclosure VC Domain

After my last article on the [Tradeoff Factors in choosing a SED and Multi-Enclosure Domain](#), I have received some good technical reasoning points. We have discussed these things internally and I tried to answer them as much as possible. I thought it would be better if I write an article about it.

So here is the verbatim of what we have gone through.

- 1) What is the difference in the east / west traffic bandwidth ?
- 2) Scaling out – what aspect is better in single enclosure domains ?
- 3) What aspect of domains reduces the availability of running blades / critical workloads ?
- 4) When you refer to domain maintenance that makes the domain unavailable, what specifically are you referring to ?

My answer to these questions are in line as below:

1. There would be much difference in the East-West Traffic because they would of course traverse the core. It would look as repulsive traffic going to the core and coming back. So we would always prefer having multiple enclosures stacked together so that we can have the benefit of east west traffic traversing the stack link (10Gb Copper or 10Gb SFP+ with LR-LR OM2/OM3 Cable). This is the main bake off factor in between Cisco FEX -> Interconnect -> Nexus 5K and HP FlexFabric Multi-Enclosure stacking architecture. Cisco sells their UCS with Fabric Interconnect as Single Point of Management. But to me that does not scale and it is actually a SPOF. In Cisco they are over subscribed, not even on Ethernet, but also on SAN IO. They are the killer of SAN IO. They are much more Network centric and always try to sell more network ports (at least 24 10Gb ports in a stack scenario, where in HP need 20 10Gb ports).

Another factor in Single Enclosure Domain is the number of independent uplinks required per enclosure. Now 10Gb uplink ports does not come cheap. It is always at premium. In case of multi-enclosure stack domain we just need half of the uplinks to the Distribution/Core. So we would always get inclined towards the Multi-Enclosure domain keeping this factor in mind.

So in a nutshell we can easily rule out the SED (Single Enclosure Domain) if we are looking out scale out architecture and routing System Management traffic (vMotion, DRS, FT) within the enclosures.

2. Scale out in a SED does not look good at any time. It would add more Mgmt work when you scale out. We can manage domains (upto 250 domains and 1000 enclosures) through VCEM (Virtual Connect Enterprise Manager) but it comes with a premium cost. So looking at the cost benefit it does not make sense to get VCEM for a small/moderate number of Enclosures/Workload.

So in a nutshell it is better to go for the Multi-Enclosure domain if we are really looking out for the moderate amount of scale out opportunity.

3. SED increases the risk of domain outage. In case of a domain failure we are at risk if we have business critical workloads. So, as I said in my blog, if we have real business critical workloads then these must be located on clusters (Host or Hypervisors) that span at least 2 VC domains. To keep it simple, we can create small SED and migrate our workload to the different domain in case of maintenance, provided we are not looking for a scale our architecture to a great extent and bear the cost of required number of ports at Distribution/Core.

4. VC Domain Maintenance is a useful way to perform updates on a particular VC Domain. (This is in reference to the Domain locked by VCEM)

Some of the useful domain-level operations enabled during VC Domain Maintenance include:

- Upgrading firmware
- Backing up VC Domain configuration
- Administering local user accounts
- Setting LDAP directory settings
- Changing VC Domain configuration
- Domain name
- Static IP address
- Setting SSH
- Setting SSL Certificate
- Resetting Virtual Connect Module (soft reset)
- Monitoring network ports

- Configuring networks
- Configuring storage

Some of the useful network-level operations enabled during VC Domain Maintenance include:

- Monitoring network ports
- Changing network configurations

Some of the useful storage level operations enabled during VC Domain Maintenance include changing storage configuration.

So from an Infrastructure Operations perspective we prefer Single Enclosure VC Domains, which does not require much planning, does not hold ample amount of complexity in frame it.

I had some good discussion on this with HP's CloudSystem Architect Ken Henault (Twitter -> @bladeguy). He made some good points too. I would like to show case what is he upto while talking about the Failure Domain and Mgmt. Domain.

Selecting single or multi-enclosure domains is like any other architectural decision. Factors like performance, cost, manageability and reliability must be considered. While these are highly available systems, designed for no downtime at the enclosure level, accidents can happen. In a properly design multi-enclosure stack like I described, an entire enclosure can go down, but the remaining enclosures will still have connectivity, and continue to operate.

If an entire enclosure were to suffer an outage, bringing the servers within that enclosure back online would take higher precedence over managing the servers that were still on-line. When the servers are restored, presumably the Virtual Connect Manager will also be restored.

The more important consideration in this type of interconnected system is the size of a potential failure domain compared to the management domain. In the example listed here, and in most foreseeable c-Class failure scenarios the failure domain is 8-16 servers. Compared to a management domain of 64 servers with Multi-Enclosure stacking, or 16,000 servers with Virtual Connect Enterprise Manager. The failure domain is significantly smaller than the management domain.

Compare that to blade server designs that focus management and connectivity at the top of the rack, and you see the failure domain and the management domain are the same size. While eight servers down is bad enough, the failure domain in a top of rack design can be up to forty times larger, 320 servers.

Large management domains can increase productivity and flexibility in these interconnected systems. Large failure domains can be disastrous.