Calculating IO loading for array devices
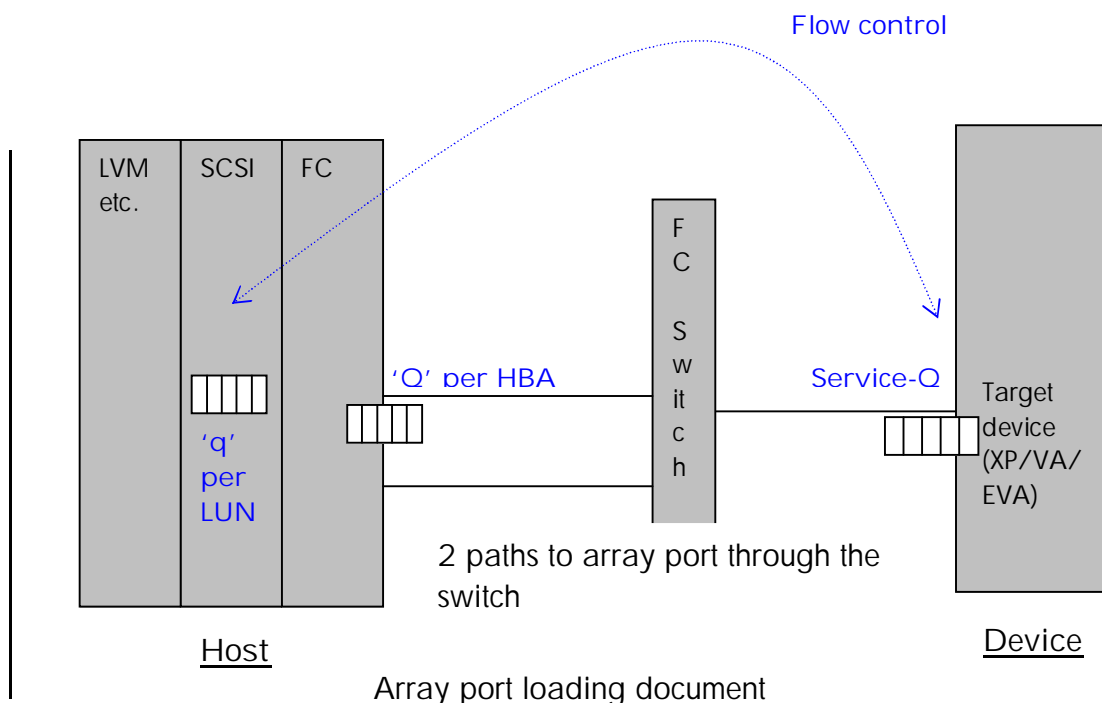
An HP Whitepaper

Objective:

The goal of this document is to provide simplified methods for assessing IO loading on array devices in a lab or customer SAN environment. The formulations presented should also help guide connectivity limitations on configured LUNs, switches, switch ports and host ports.

Scope:

The document presents fundamentals of IO loading with VA, XP 256, XP 512, XP128/1024 & EVA examples. Special features like LUN security and host groups are also presented through examples. The model should be extendable to other devices in the SAN as well. Loading is presented from the target device port loading perspective only. Total SAN loading and analysis are not addressed by this document. The concepts and variables presented here are general SCSI terminologies. The specific driver/OS variables will need to be ascertained by the reader, e.g., in Solaris the driver 'q_depth per LUN' is depicted by the variable 'sd_max_throttle'.

1. IO loading fundamentals

Flow control

LVM etc. | SCSI | FC

'Q' per HBA

F C S w it c h

Service-Q

'q' per LUN

Target device (XP/VA/ EVA)

2 paths to array port through the switch

Host

Device

Array port loading document

The key factor in determining loading in any mass storage SAN setup is to understand the factors used. The main factors used in any mass storage setup are at the SCSI layer level both in the initiator (host) & the target (device). These factors provide an IO flow control mechanism between the initiator and the target. The host flow control variable is called 'q-depth per LUN' & the device flow control variable is 'service Q-depth'. The above diagram depicts the placement of these variables in the IO architecture.

Five main variables are used to assess the loading factor on a Device port, namely,

- H = Number of Host connected to that array device port.
- P = Number of Host paths connected to that array device port.
- q = Q-depth per LUN on the hosts (for the host port), i.e. meaning maximum I/Os outstanding per LUN from the host at any given time. This is a SCSI driver parameter. For HPUX, it is 8.
- L = Number of LUNs configured on that array device port being seen by a particular host path.
- Service-Q = maximum Q-depth per Array device port which signifies the maximum number of I/Os outstanding (can be handled) on that port at a given time. , i.e. as defined below Service-Q value = 2048 for EVA.
- Q = Q-depth per HBA port i.e., maximum I/Os an HBA can have outstanding at an instant. This parameter should only be needed if 'q' (q-depth per LUN) is not available.

- These variables relate to each other through a simple equation:

    Service–Q => H * P * q * L   <u>Equation 1</u>

    OR

    Service-Q => H * P * Q   <u>Equation 2</u>

    <u>Please note: Equation 1 or 2 can be used alternatively in the equations provided thus far, depending on the parameters available at the customer site.</u>

For heterogeneous hosts connected to the same port on the device,

$$\text{Service-Q} \Rightarrow \text{Host OS}_1(H * P * q * L) + \text{Host OS}_2(H * P * q * L) + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + \text{Host OS}_n(H * P * q * L)$$

Using an example with VA – Consider an HPUX K-class connected as follows to a VA: Two host FC cards are connected to a switch. The switch then is connected to one port on the VA. Thusly there are two paths connected to one VA port, hence ''P' is 2. 'q' value for HPUX is 8. Let's assume that the number of LUNs, 'L' assigned to the VA port is 10. The Service-Q value we get is

$$\text{Service-Q} \Rightarrow (1 * 2 * 8 * 10 = 160)$$

** for VA this is a service-Q loading of 21.33% , i.e., 160/750 * 100

Recommendations:

1. If a certain level of loading is being achieved, either decrease the number of host ports connected to the device port or number of LUNs assigned to the device port.
2. If a new setup is being configured all caution must be taken to stay below the published Service-Q value of the array device.

(Please note: host default q-value should be changed under advisement from the Host vendor and as a last resort.)

ASIDE: This configuration should work OK without problems. However, if the Q-value is higher than 750, VA will not be able to handle IOs in access of 750 and will respond back with an IO check condition of Q-full. The Q-full then may or may not instigate the host to throttle the I/Os depending on the OS being used, i.e., currently all OS versions of HPUX do not do control flow/IO throttling.

1.1 MAX SERVICE Q-DEPTH VALUES for devices

XP128/1024	1024
XP512/48	1024
XP256	1024
VA7100/7400/7410	750
EVA	2048

1.2. Hazard limitations:

While testing Hazard, MAXDEV variable, determines the maximum LUNs, hazard launches IOs against at one time. Hazard will start 4 raw and 4 filesystem IOs per LUN for c-8 & c16. Thus the maximum load emanating from hazard would be calculated as "MAXDEV X 8". This would be a limiting factor for exercising the load on an array port.

e.g., in the example above, our Service-Q was 160, if we used MXDEV= 15, the Hazard stress would be a maximum of 120, which would not fully exercise the array port capability. A MAXDEV of 20 and above would exercise the port optimally.
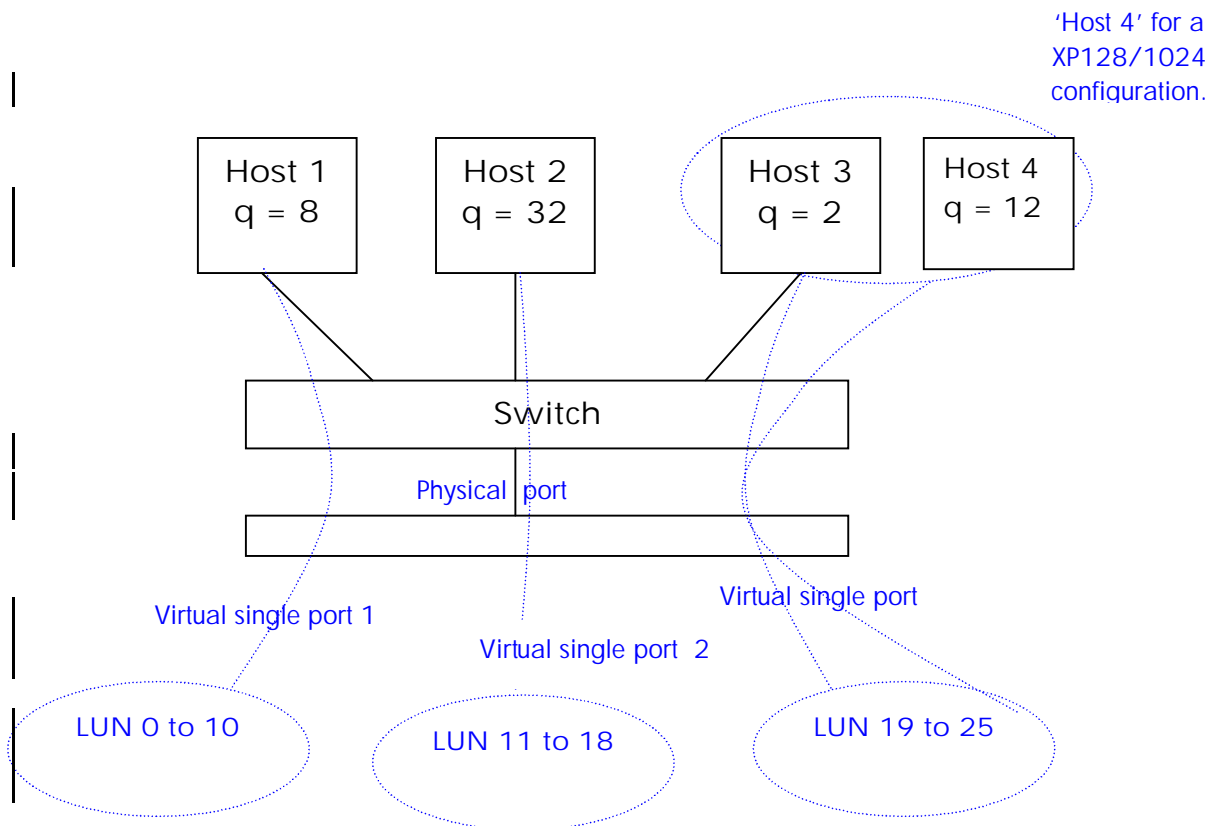
1.3. SPECIAL CASES:
How to calculate loading on an array port when using LUN security, host groups in XP128/1024, Heterogeneous connectivity & multi-clusters?

Rule of thumb: In these configurations, think of the 'array port' as a combination of 'virtual single port' connections to the host ports. Calculate the loading on each 'virtual single port' connection using q * L, add it for each host then equate it to the service Q value on the array port. i.e.

For all LUN secured host ports connected to the array port calculate
[q(host 1) * L(host1) + q(host 2) * L(host2) + ...+ q(hostN) * L(hostN)]
<= Q(service Q value of the array port)

'Host 4' for a
XP128/1024
configuration.

| Host 1 q = 8 | Host 2 q = 32 | Host 3 q = 2 | Host 4 q = 12 |

Switch

Physical port

Virtual single port 1

Virtual single port 2

Virtual single port

LUN 0 to 10

LUN 11 to 18

LUN 19 to 25

The mentioned example can be applied to LUN security setups in XP256, XP512 and VA and the host group paradigm in XP128/1024.

Considering the figure to be a LUN security setup on XP512 with only Host1, Host2 ad Host3. Loading due to host 1 can be calculated as 11(L) *

8(q) = 88, host2 = 8(L) * 32(q) = 256, and host 3 = 7(L) * 2(q) = 14. Thusly, for an XP512 loading is,

88 + 256 + 14 <= 1024 (optimal service Q on XP512).

In case of a XP128/1024, we add Host 4 to the same host group as Host3 so that both the hosts see the same number of LUNs. The loading due to Host3 & Host 4 now becomes 7(L) * 2(q) + 7(L) * 12(q). Thusly, for a XP128/1024 the loading is

88 + 256 + (14 + 84) $_{host\ group}$ <= 1024 (optimal service Q on XP128/1024).

As can be seen, XP128/1024 provide special LUN security features that allow multiple host groups per port, with each host group addressing a separate set of LUNs.

While calculating the loading, special attention should be paid to array port service Q value, which is a maximum of 1024. (Please note: this value is for the whole port and not for each host group.)

Another simpler way to calculate the loading for XP128/1024 would be to exercise the above rule of thumb per host group on an array port and then calculate loading per virtual single port followed by summation to max. array port can handle, i.e., 1024.

So the above equation now becomes,

$Host\_group_1$(PER HOST SUM OF(N * P * q * L)) + $Host\_group_2$(PER HOST SUM OF (N * P * q * L))+ …………. +$Host\_group_k$(PER HOST SUM OF(N * P * q * L)) <= 1024.

EXAMPLE 2:

If we have 3 host groups on 1 XP128/1024 port such that
$host\_group_1$ has 50 LUNs and 1 host (WWN1)
$host\_group_2$ has 20 LUNs and 5 hosts (WWN2 to WWN6)
$host\_group_3$ has 30 LUNs and 2 hosts (WWN4,WWN5)

Thus, the calculation in this case becomes,

50 * 1 * 8 ($host\_group_{1)}$ +

20 * 5 * 8 (host_group$_{2)}$ +
30 * 2 * 8 (host_group$_{3)}$ >= 1024.

So, the result is an overloaded array port i.e.,

1680 >= 1024, being over by 656.

The easiest way to reduce the load is to reduce the number of LUNs in host_group$_2$ and host_group$_3$.

Please note:
1. The term WWN above, refers to the host port WWN i.e. single path to the host port.
2. In this example, all hosts have same q value of 8.
3. Using Example1 and 2, one can use mixed equations to ascertain loading in heterogeneous, homogeneous & multi-clustered configurations.