

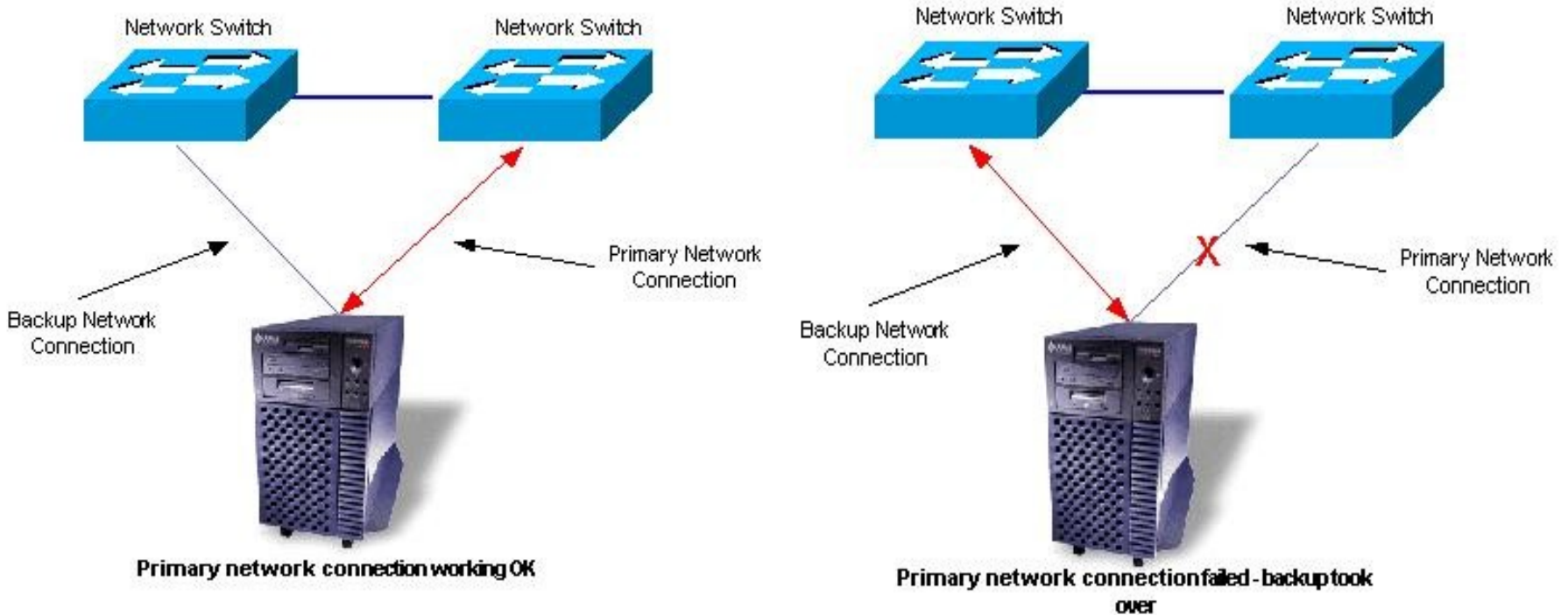
UNIX Days 2007

Solaris 10 - Networking and storage HA solutions

Gdańsk, 30 maja 2007 r.

- Network HA – IP Multipathing.
- Storage HA - I/O multipathing.

IP Multipathing (IPMP) overview.



- reliability,
- availability,
- network performance,

with two or more network interfaces.

- IPMP group,
- IPMP addressing,
- IPMP interface configurations,
- failure detection and failover.

- one or more physical interfaces,
- the same NIC media type,
- two types of addresses: data addresses and test addresses.

- active interface and standby interface,

```
# ifconfig ce0 10.0.1.1 netmask + broadcast + group ipmp0 up  
# ifconfig ce1 group ipmp0 standby up
```

- in.mpathd daemon can detect NIC "failure" and "repair",
- link-based failure detection (IFF_RUNNING flag on NIC),
- probe-based failure detection (ICMP Echo requests).

```
cat /etc/default/mpathd
#
#pragma ident    "@(#)mpathd.dfl 1.2    00/07/17 SMI"
#
# Time taken by mpathd to detect a NIC failure in ms. The minimum time
# that can be specified is 100 ms.
#
FAILURE_DETECTION_TIME=10000
#
# Failback is enabled by default. To disable failback turn off this option
#
FAILBACK=yes
#
# By default only interfaces configured as part of multipathing groups
# are tracked. Turn off this option to track all network interfaces
# on the system
#
TRACK_INTERFACES_ONLY_WITH_GROUPS=yes
```

■ active-active interface configuration:

```
# ifconfig ce0 host-1-ce0.srv netmask + broadcast + group hanfs  
up addif host-1-ce0-probe.srv -failover deprecated netmask +  
broadcast + up
```

```
# ifconfig ce1 host-1-ce1.srv netmask + broadcast + group hanfs  
up addif host-1-ce1-probe.srv -failover deprecated netmask +  
broadcast + up
```

- Use probe-based failure detection,
- DNS names for test addresses,
- /etc/hosts with test and data addresses,
- use `if_mpadm` for detach/failover interface

- All members of a group must be on the same subnet and should, if possible, be connected to physically separate switches.
- A physical interface must be a member of only one group.
- Physical interfaces on different subnets must be in different IPMP groups.
- Multiple IPMP groups may be configured on the same subnet.
- It is permissible to have a group with a single physical interface as its member.
- Each physical interface must have its own unique MAC address.
- Physical interfaces in the same IPMP group must be of the same type.
- All interfaces in the same group must have the same STREAMS modules configured in the same order.

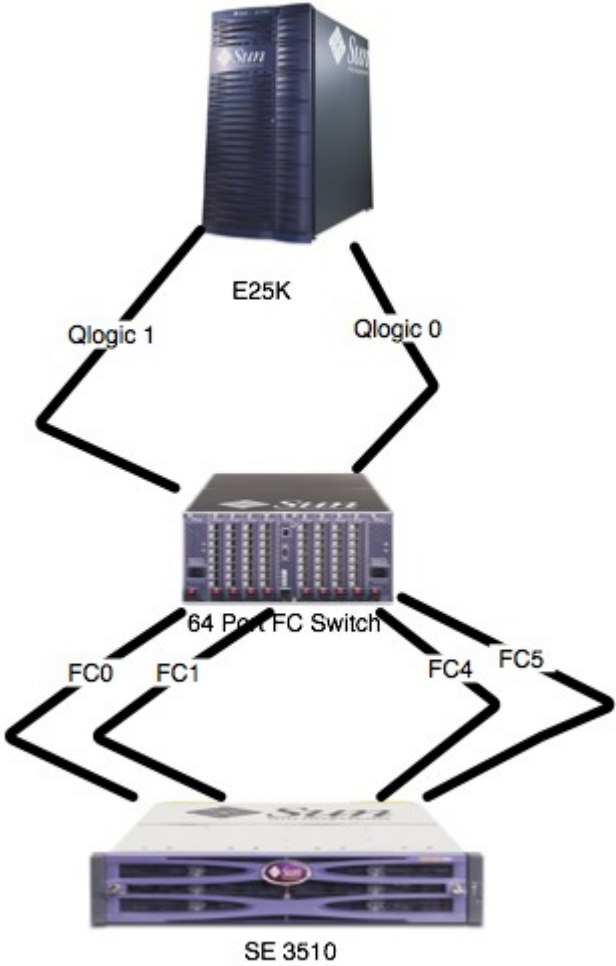
<http://docs.sun.com/app/docs/doc/816-4554/6maoq027c?q=IPMP>

<http://sunsolve.sun.com/search/document.do?assetkey=1-9-70062-1>

<http://sunsolve.sun.com/search/document.do?assetkey=1-9-86869-1>

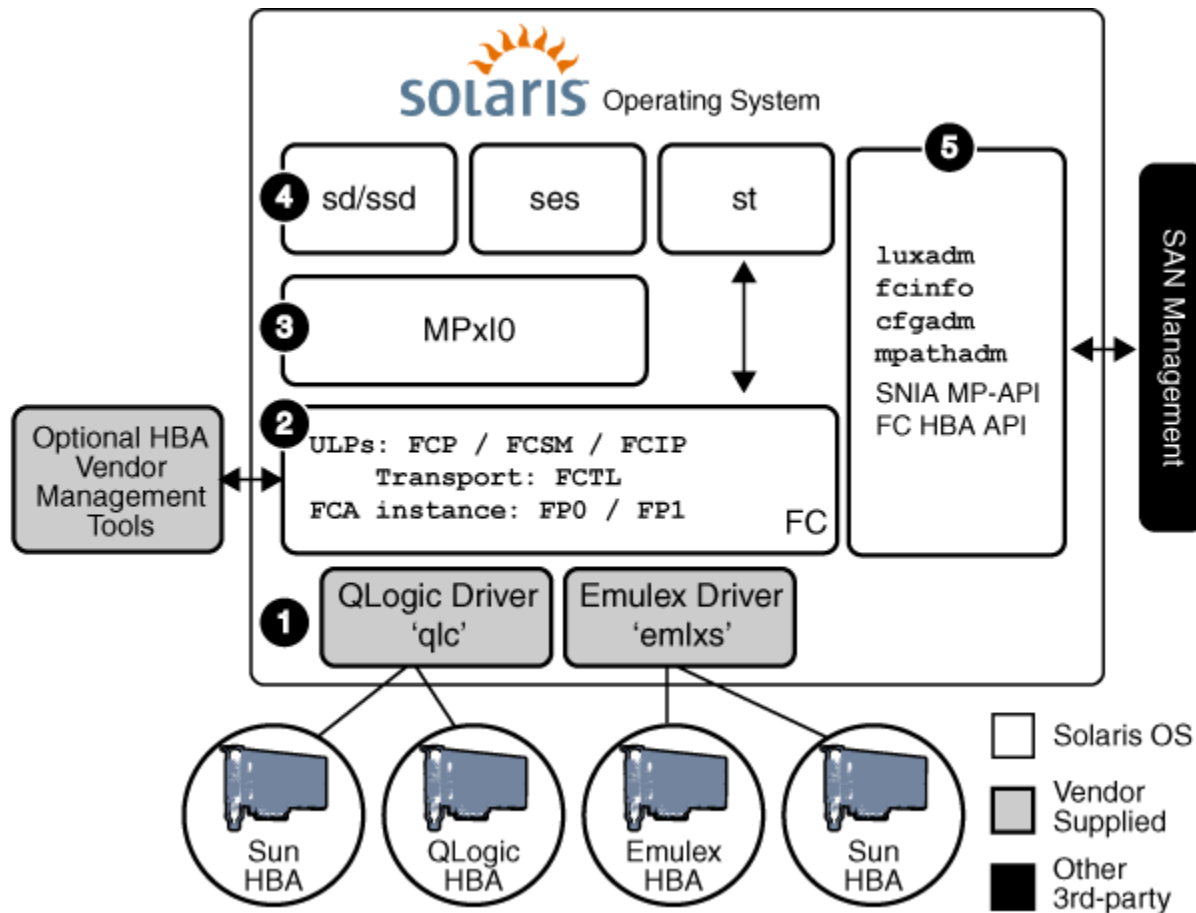
<http://sunsolve.sun.com/search/document.do?assetkey=1-25-41553-1>

SAN example.



- multiple redundant paths to a storage system,
- single device instances,
- symmetrical/asymmetrical device support,
- load balancing (round-robin algorithm),
- failover.

MPxIO overview.



- By default on x86 platform,
 - /kernel/drv/fp.conf, /kernel/drv/scsi_vhci.conf

- By manual configuration on SPARC platforms,
 - /kernel/drv/fp.conf, /kernel/drv/scsi_vhci.conf
 - /usr/sbin/stmsboot

- `mpathadm show mpath-support libmpscsi_vhci.so`
- `luxadm inquiry /dev/rdisk/cXtYdZs2`
- `/kernel/drv/scsi_vhci.conf`

```
# For enabling MPxIO support for 3rd party symmetric device need an
# entry similar to following in this file. Just replace the "SUN      SENA"
# part with the Vendor ID/Product ID for the device, exactly as reported by
# Inquiry cmd.
#
# device-type-scsi-options-list =
# "SUN      SENA", "symmetric-option";
#
# symmetric-option = 0x1000000;
```

- /kernel/drv/fp.conf (on x86)
mpxio-disable="yes";
- stmsboot -d (on sparc)
- stmsboot -e (on sparc)

Reboot required!

■ /kernel/drv/fp.conf

```
mpxio-disable="yes";
```

```
name="fp" parent="/pci@6,2000/SUNW,q1lc@2" port=0 mpxio-  
  disable="no";
```

```
name="fp" parent="/pci@13,2000/pci@2/SUNW,q1lc@5" port=0  
  mpxio-disable="no";
```

Reboot required!

■ /kernel/drv/scsi_vhci.conf

```
load-balance="round-robin";
```

```
load-balance="none";
```

```
auto-failback="enable";
```

```
auto-failback="disable";
```

Reboot required!

■ /kernel/drv/fp.conf

```
# pwn-lun-blacklist="target-port-wwn,lun-list"
```

```
pwn-lun-blacklist="5006016908065109,0,10";
```

Reboot required!

■ stmsboot -L

non-STMS device name

STMS device name

/dev/rdisk/c2t40d3

/dev/rdisk/c5t600C0FF000000000098FD57F9DA83C00d0

/dev/rdisk/c2t40d2

/dev/rdisk/c5t600C0FF000000000098FD55DBA4EA000d0

/dev/rdisk/c2t40d1

/dev/rdisk/c5t600C0FF000000000098FD516E4403200d0

/dev/rdisk/c2t40d0

/dev/rdisk/c5t600C0FF000000000098FD514A1D9AF00d0

/dev/rdisk/c2t41d2

/dev/rdisk/c5t600C0FF000000000098FD54DDFB18300d0

/dev/rdisk/c2t41d1

/dev/rdisk/c5t600C0FF000000000098FD511BA5C8000d0

/dev/rdisk/c2t41d0

/dev/rdisk/c5t600C0FF000000000098FD52BABFF1D00d0

/dev/rdisk/c3t42d2

/dev/rdisk/c5t600C0FF000000000098FD54DDFB18300d0

■ mpathadm list initiator-port

```
Initiator Port:  iqn.1986-03.com.sun:01:0003baf6a39d.45deef44,4000002a00ff
Initiator Port:  210100e08ba21f2f
Initiator Port:  210000e08b821f2f
```

■ mpathadm show initiator-port 210100e08ba21f2f

```
Initiator Port:  210100e08ba21f2f
Transport Type:  Fibre Channel
OS Device File:  /devices/pci@1d,700000/SUNW,qlc@2,1/fp@0,0
```

■ mpathadm list lu

```
/dev/rdisk/c5t600C0FF00000000098FD57F9DA83C00d0s2
```

```
    Total Path Count: 2
```

```
    Operational Path Count: 2
```

```
/dev/rdisk/c5t600C0FF00000000098FD55DBA4EA000d0s2
```

```
    Total Path Count: 2
```

```
    Operational Path Count: 2
```

■ mpathadm show lu /dev/rdisk/cXtYd0s2

■ mpathadm modify mpath-support -a on libmpscsi_vhci.so

■ **luxadm display /dev/rdisk/c5t600C0FF00000000098FD52BABFF1D00d0s2**

```
/dev/rdisk/c5t600C0FF00000000098FD52BABFF1D00d0s2
```

```
/devices/scsi_vhci/ssd@g600c0ff00000000098fd52babff1d00:c,raw
```

```
Controller          /devices/pci@1d,700000/SUNW,qlc@2/fp@0,0
```

```
Device Address      216000c0ff998fd5,0
```

```
Host controller port WWN 210000e08b821f2f
```

```
Class              primary
```

```
State              ONLINE
```

```
Controller          /devices/pci@1d,700000/SUNW,qlc@2,1/fp@0,0
```

```
Device Address      226000c0ffa98fd5,0
```

```
Host controller port WWN 210100e08ba21f2f
```

```
Class              primary
```

```
State              ONLINE
```

■ **luxadm display /dev/rdisk/c6t6006016045231B00081FA6901EDDDB11d0s2**

```
/dev/rdisk/c6t6006016045231B00081FA6901EDDDB11d0s2
```

```
/devices/scsi_vhci/disk@g6006016045231b00081fa6901edddb11:c,raw
```

```
Controller          /dev/cfg/c4
Device Address      5006016041e03aa8,8
Host controller port WWN 210000e08b9a5e75
Class             secondary
State             STANDBY
```

```
Controller          /dev/cfg/c5
Device Address      5006016941e03aa8,8
Host controller port WWN 210100e08bba5e75
Class             primary
State             ONLINE
```

- `mpathadm ...`
- `luxadm ...`
- `cfgadm -al -o show_SCSI_LUN`
- `fcinfo ...`
- `stmsboot -L`

- <http://docs.sun.com/source/819-0139/>
- <http://www.sun.com/bigadmin/hubs/storage/>
- http://www.sun.com/bigadmin/.../san_fundamentals.pdf



Ireneusz Szuca <isz@wp-sa.pl>