SOC+XT SAN Limited Program Overview and Purpose

E10K

**SANbox-8
SANbox-16**
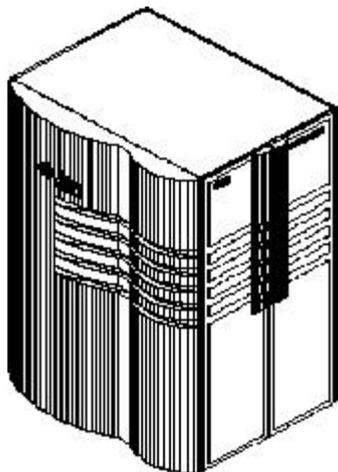
The purpose of the SOC+XT SAN Program is to allow existing SBus servers to attach SAN storage devices in a controlled environment. The focus of this program has been limited to those customers who are **unable** not unwilling to migrate to Solaris 8/beyond. **SunPS approval is required on all SOC+XT configurations**.

The program will have multiple phases as it evolves over time.

Phase 0 started in December 2000.
>   At that time, the program was based on SAN 2.0 "Mamba" features and driver code for the Sun-branded Qlogic switches.
>   Phase 0 allows SOC+ SANs using all storage in SAN 2.0 and SunCluster 2.2.
>   Details on the configuration rules are described on Pages 2-4.

Phase 1 started in November 2001.
>   It is based on SAN 3.0 "Python" features and Sun-branded switch firmware.
>   Phase 1 enhances Phase 0 rules with cascading switches.
>   Details of the configuration rules are described on Pages 5-6.

>   Phase 1 limited some features allowed in Phase 0. The main points:
>   1. Solaris 8 use should be discouraged from the program as much as possible. The X6757A "Ivory" card should be used instead of SOC+.
>   2. A3500FC is not allowed in cascaded switch configs.

Phase 2 started in March 2002.
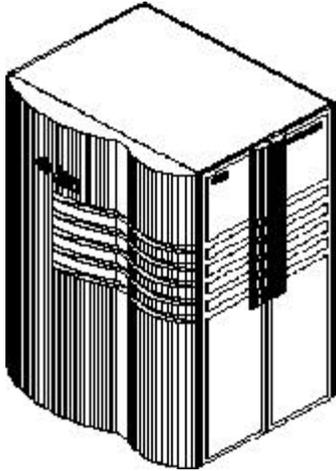>   It is based on SAN 3.1 "Python-XT" features and Sun-branded switch firmware.
>   Phase 2 enhances Phase 1 rules with reverse-TL option for SOC+.
>   Details of reverse-TL are described on Page 7.

SOC+XT SAN Limited Program
Phase 0 Configuration Rules

E10K

SANbox-8
SANbox-16

The following proposed configurations illustrate the limitations of the SOC+XT SAN program. Below are the justifications and rationale behind some of these configuration limitations.

1. Servers are limited to E3X00, E4X00, E5X00, E6X00, and E10Ks running Sol 2.6, 2.7, or 2.8, with SOC+ host bus adapters (X6730A) or I/O boards with integrated SOC+ (FRUs 501-4266, 501-4883, 501-4884) only.
2. Whether using I/O boards with integrated SOC+ or SOC+ HBA, only 1 of 2 ports can be used. The other port cannot be used for any purpose, even if not connected to the SAN. [Ref Bug ID 4525143]
3. Switch functionalities are equal to SAN 2.0 "Mamba".
   The main features include 16-ports and SL zoning.
   Switches include SunSwitch and Brocade Silkworm (QuickLoop required)
4. The only supported JBOD is A5200, max = 3 per SL zone.
5. The HW RAID offerings are A3500FC (1x5 and 2x7 only), T3 and T3+.
   Max = 4 array controllers per SL zone.
6. Storage arrays can NOT be mixed in the same SL zone.
7. All configurations must offer dual physical layer connectivity to ensure no single point of failure.
8. No more than 2 initiators or hosts per SL zone.
9. Cluster software: VCS 1.3, SC 2.2 and SC 3.0 on limited basis only!!!
   See separate SC 2.2 page for details (pg 4).
10. SunPS approval is required. Refer to http://scope.central.

Unsupported components and features for this program include:
A.   Cascading of switches using A3500FC
B.   Non-Solaris operating systems
C.   No tapes, including using bridges
D.   No roadmap growth to Fabric with SOC+ cards

SOC+XT SAN Limited Program
Phase 0 Minimum Firmware

Note: Each pair of switches are zoned into two separate zones. Zone 1 and Zone 2 are separate entities.

Minimum Firmware and Software Revisions (matches SAN 2.0 "Mamba" release).
Note: the latest versions of these software packages and patches are recommended.

Solaris -   2.6, 2.7 or 8 with latest recommended patch clusters from SunSolve

A3500FC FW 3.01.03.60 [ref Sun Early Notifier 20029]
A5200 FW 1.09 [ref Sun Early Notifier 14838]
T3 FW 1.17b, Loop card FW 5.02, EPROM Image 2.10 [ref Sun Early Notifier 14838]
T3+ FW 2.0, Loop card FW 5.02 [ref Sun Early Notifier 14838]
Disk Firmware can be found on Sun Early Notifier 14838

SANbox FW 3.03.51 & SANsurfer GUI 2.07.50 for Phase 0 [ref www.sun.com/storage/san]
     ref SAN 2.0 installation guide 806-6922-10
SANbox FW 3.04.62 & SANsurfer GUI 2.08.30 for Phase 1
     ref SAN 3.0 installation guide 816-0830-10
SANsurfer GUI version 2.07.50 (Sun version) requires JDK 1.2.2.05a
     ref SAN 2.0 installation guide 806-6922-10

STORtools 3.3 for Sol 2.6 & 7
     STORtools 4.0.1 for Sol 8
RM 6.22 (read notices on http://acts.ebay/storage/A3500 if upgrading from 6.11)
Component Manager 2.1 (2.2 recommended)
NWS Agent 2.1 (formerly RAS Agent)

Veritas VxVM 3.0.4 with patches, 3.2 is recommended, VCS 1.3 with patches

SunCluster 2.2 SOC+XT Solaris 2.6
SAN Configuration Rules

The following configuration rules have been approved by SunCluster Marketing, Network Storage Marketing, and Professional Services. These configuration rules follow the standard SOC+ SAN program rules described on Phase Zero (pg 2) except where noted.

1. Servers are limited to E3X00, E4X00, E5X00, E6X00, and E10K (same as SOC+)
   Storage devices are limited to T3WG single brick only.
   Switches are limited to Sun SANbox 8/16 only, no Brocade Silkworm.
   Volume management uses Veritas VxVM only.
2. SunCluster SAN configurations require SunPS active involvement in architecting and implementing these solutions.
3. Sales team is responsible for explaining all limitations of this solution to the customer:
   a.) This solution must upgrade to SunCluster 3.0 within nine months of deployment.
   b.) SunCluster 3.0 requires Solaris 8 running on all nodes.
   c.) SAN components are limited to functionality and features of SAN 2.0"Mamba".
4. Account team is responsible for supplying corporate ES with all configuration information before the Cluster-SAN solution becomes "live".
5. Approval for this Cluster-SAN solution requires local and corporate-level support from ScopeTool (http://scope.central).
6. A RADIANCE case should be opened for each Cluster-SAN solution.
7. This program has only *ONE* upgrade path to SunCluster 3.X and Solaris 8/above.

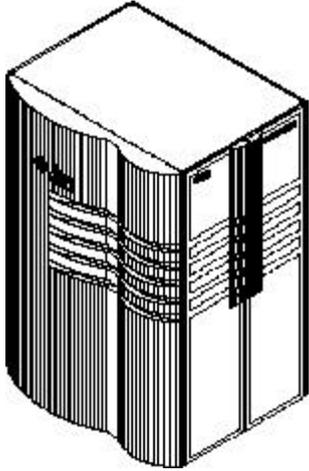The following components and features are not supported under this program:

A. Python SAN features: cascaded switches using A3500FC, fabric.
B. Non-Solaris operating systems.
C. Non-Sun Storage Devices.

Note: SunCluster 3.0 supports the same rules as above, except Solaris 8 only, and VxVM 3.04 or 3.11

SOC+XT SAN Limited Program
Phase 1 Configuration Rules

E10K

SANbox-8
SANbox-16

The following rules describe the SOC+XT SAN Limited Program. Below are the justifications and rationale behind some of these configuration limitations.

1. Similarities to the existing SOC+ Program: (previous pages)
   same servers, operating systems, HBA (SOC+ only), no tape support

2. Major differences between SOC+ and SOC+XT:
   a) cascading of switches using short or long-wave GBICs allowed
   b) only T3 and A5200 (no 9GB drives) storage arrays permitted
   c) upgrading existing SOC+ accounts require new Approvals through ScopeTool (http://scope.central)
   d) min firmware levels and config rules follow SAN 3.0 "Python" release

3. SL-mode (QuickLoop for Brocade) across an ISL restrictions
   a) maximum active initiators per addressable LUN = 2
   b) maximum number of linear switch hops between SOC+ and LUN = 1
   c) max number of A5200s in any zone still = 3

4. Major enhancements with SOC+XT (Phase 1):
   a) One interswitch link/hop permitted (single cascade) [Phase 1]
   b) ISL's may use short-wave or long-wave GBICs
   c) servers and storage may reside in the same SL zone
      even with 1 ISL
   d) nameserver zones may exist on the same switch with non-SOC+ devices

5. SunPS approval is required. Refer to http://scope.central

SOC+XT SAN Limited Program
Phase 1 Minimum Firmware

Minimum Firmware and Software Revisions matches SAN 3.0 "Python" release
Note: the latest versions of these software packages and patches are recommended.

Solaris -   2.6, 2.7, or 2.8 with latest recommended patch clusters from SunSolve [Early Not 14838]

A5200 FW 1.09 [ref Sun Early Notifier 14838]
T3 Controller FW 1.17b, Loop card FW 5.02, EPROM Image 2.10 [ref Sun Early Notifier 14838]
T3+ FW 2.0, Loop card FW 5.02 [ref Sun Early Notifier 14838]

SANbox FW 3.04.62 (Sun 3.0 "Python" version)
    SANsurfer GUI version 2.08.30 (Sun version) requires JDK 1.2.2.05a
Brocade Silkworm 2400/2800 firmware 2.4.1f (Sun multi-platform SAN 3.0 "Python" version)

STORtools 3.3 for Sol 2.6 & 7
    STORtools 4.0.1 for Sol 8
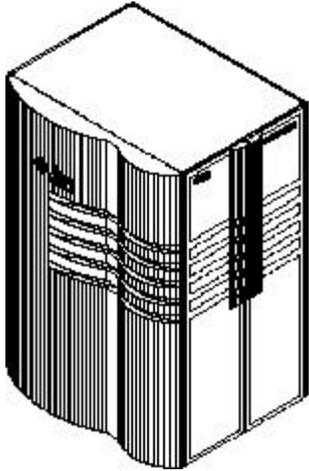Component Manager 2.2
NWS Agent 2.1

Veritas VxVM 3.1.1 with patches, 3.2 is recommended

SOC+XT SAN Limited Program
Phase 2 Configuration Rules

E10K

SANbox-8
SANbox-16

The following rules describe the SOC+XT SAN Limited Program. Below are the justifications and rationale behind some of these configuration limitations.

1. Similarities to the existing SOC+ Program: (previous pages)
   same servers, operating systems, HBA (SOC+ only), no tape support.

2. Differences to the existing SOC+ Program: (previous pages)
   reverse-TL or Fabric-Assist mode supported with T3/T3+ only.
   A5200s not supported in TL-mode, SL-mode only.
   A3500FCs not supported in cascaded configurations.

3. What is reverse-TL mode and why is it useful?
   Reverse-TL allows the SOC+ HBA to appear as a fabric initiator in the SAN.
   This allows a SOC+ zone to co-exist with another fabric HBA, like Leadville.
   The benefit is simplified zoning based on HBA limitations.

4. What is the minimum firmware required?
   Same storage array firmware as in Phase 1 (pg 6).
   Sun SANbox firmware *must* come from SunPS

5. Does this apply to Brocade?
   Yes. Brocade refers to this feature as Fabric-Assist mode.

6. Use of reverse-TL or Fabric-Assist requires a new submission to ScopeTool.
   [refer to http://scope.central]

SOC+XT SAN Limited Program Configuration Descriptions

For simplicity, only the 8-port switch configurations are drawn. However, these designs could be doubled to use the 16-port switch as well. These examples illustrate the rules.

Single server with 8-port switches and zoning configurations (Phase 0-2) :

Config 1 pg9 uses A5200

Config 2 pg10 uses A3500FC

Config 3 pg11 uses T3WG

Config 4 pg12 uses T3ES

Dual servers with 8-port switches and zoning configurations (Phase 0-2):

Config 5 pg13 uses A5200

Config 6 pg14 uses A3500FC

Config 7 pg15 uses T3WG

Config 8 pg16 uses T3ES

Config 9 pg 17 describes Ping-Pong effect using SAN 2.0 "Mamba" release.

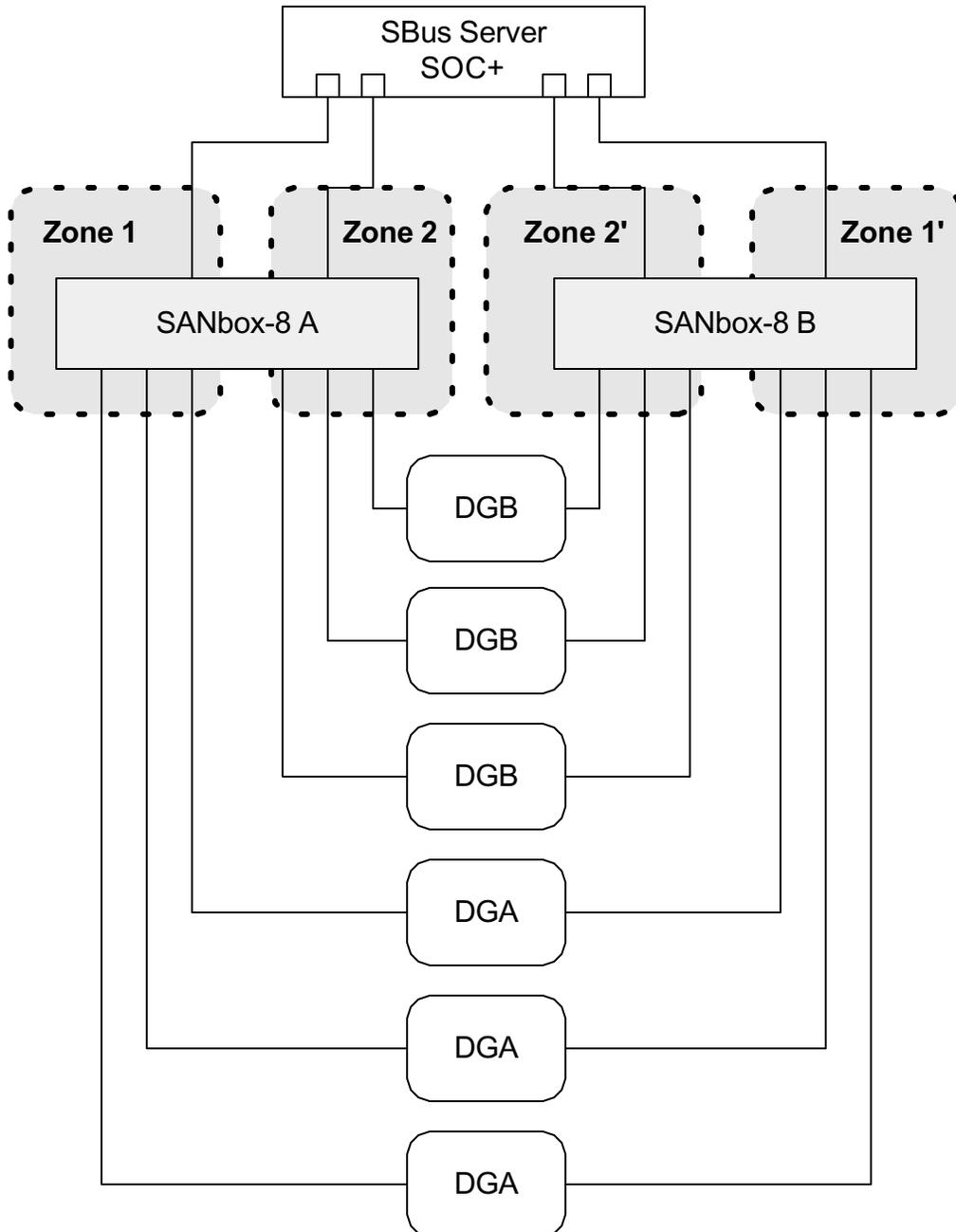Config 10 pg18 Chris Woods' explanation on T3 Ping-Pong

Dual servers with a pair of ISLs (Phase 1 and 2)

Config 11 pg19 uses Sun SANbox switches

Config 12 pg 20 use Brocade Silkworm switches

Config 1: Single Host A5200 SAN (Phase 0-2)

SBus Server SOC+

Zone 1 **Zone 2** **Zone 2'** **Zone 1'**

SANbox-8 A **SANbox-8 B**
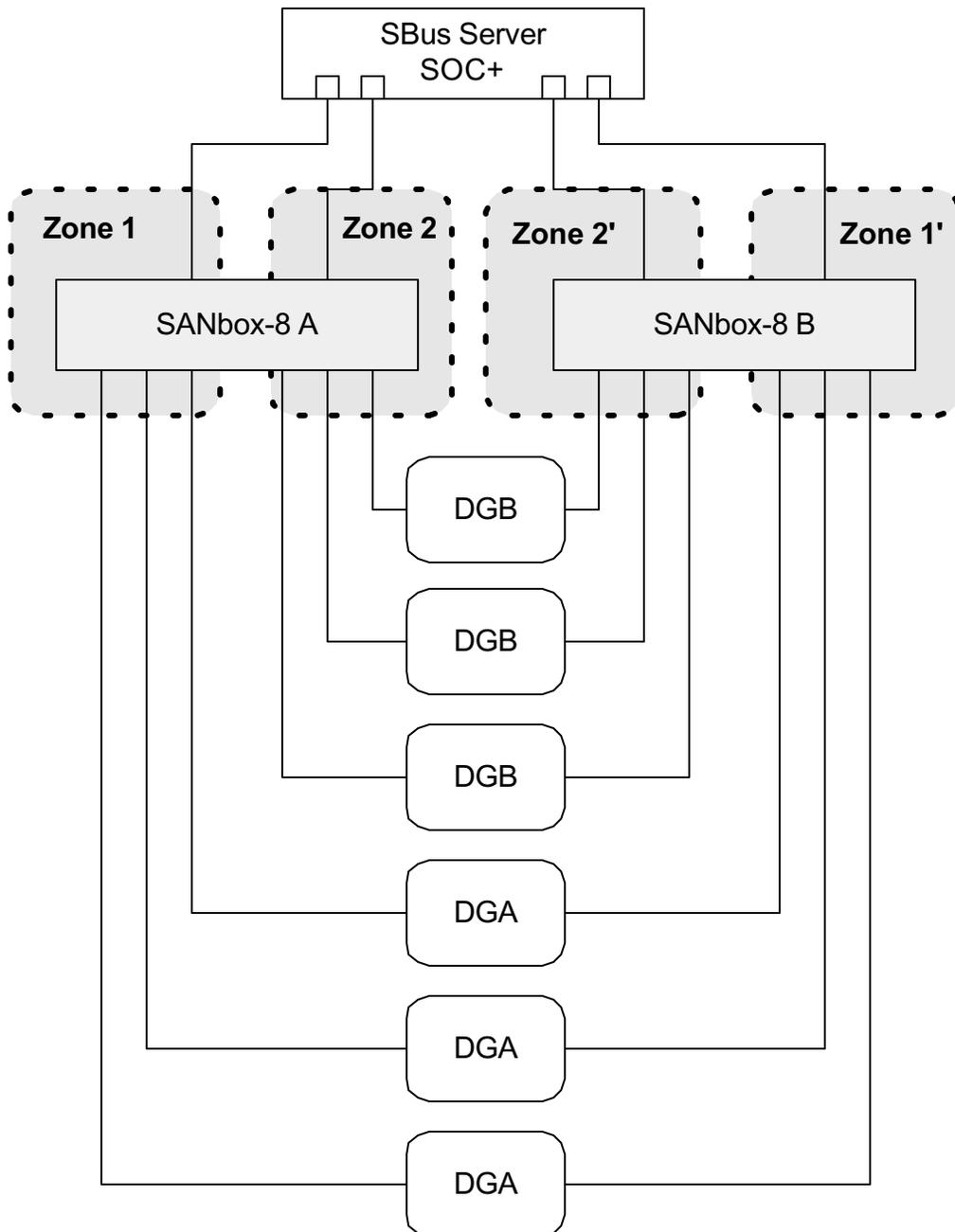
DGB

DGB

DGB

DGA

DGA

DGA

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each A5200 is configured as a single loop, split loop is not permissible.
- Three A5200's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- DGA may be mirrored to DGB using VxVM across switches.
-
- Maximum available storage:
- RAID 0+1, no hot spares
- 132 spindles
- 1201GB with 18GB disks
- 2402GB with 36GB disks
- Footprint is 1 rack cabinet
-
- Note: Two IB boards per array must be used for redundancy.
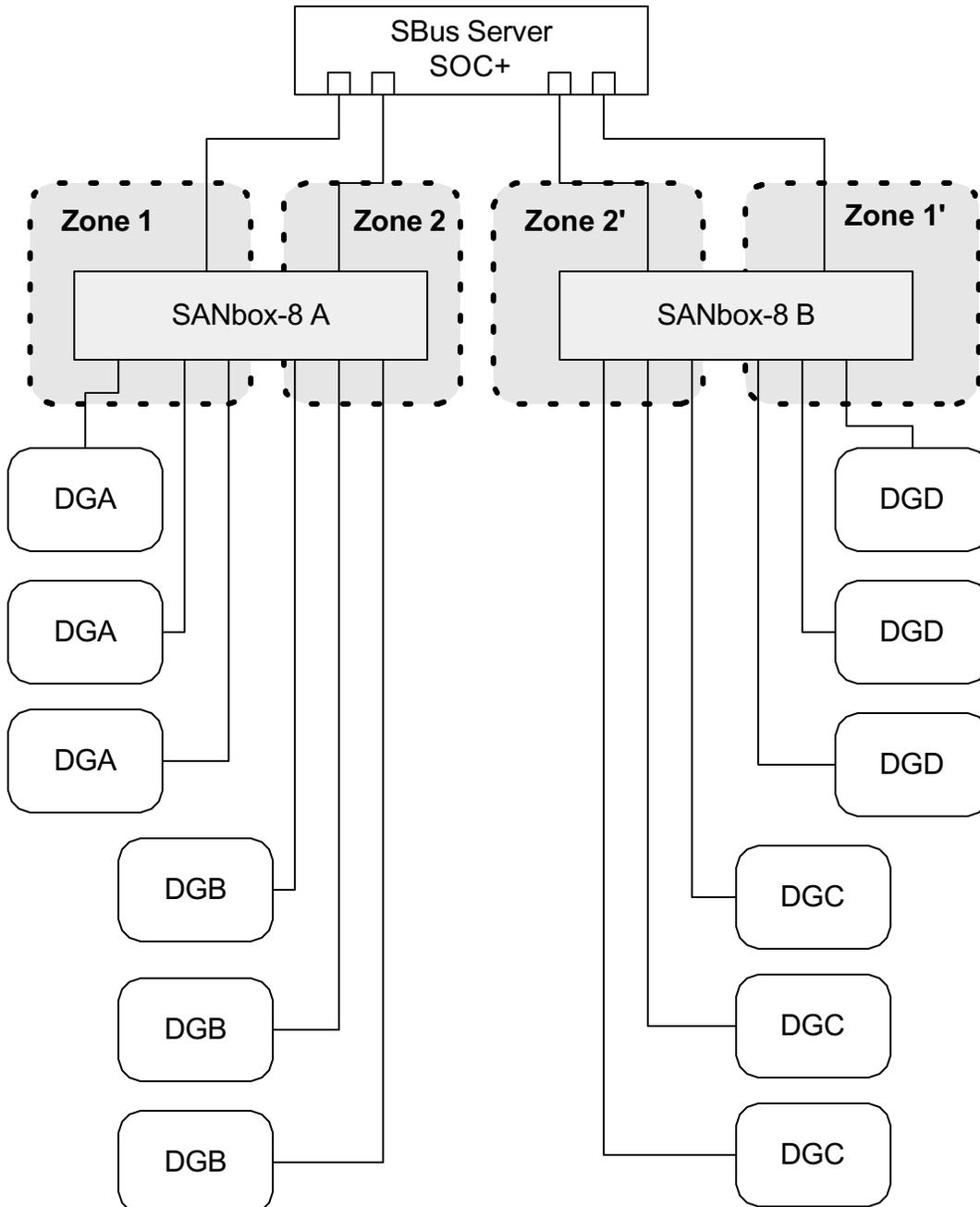
Config 2: Single Host A3500FC (1x5) SAN (Phase 0-2)

SBus Server SOC+

Zone 1 **Zone 2** **Zone 2'** **Zone 1'**

SANbox-8 A **SANbox-8 B**

DGB

DGB

DGB

DGA

DGA

DGA

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each A3500FC is configured as RAID 5 with 1 hot spare per tray.
- Three A3500FC's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- Mirroring of Array Groups may be accomplished using VxVM.
-
- Maximum storage available:
- RAID 5, one hot spare per tray
- 360 spindles
- 4804GB with 18GB disks
- 9608GB with 36GB disks
- Footprint is 6 rack cabinets
-
- Note: The A3500FC is configured as 1x5.
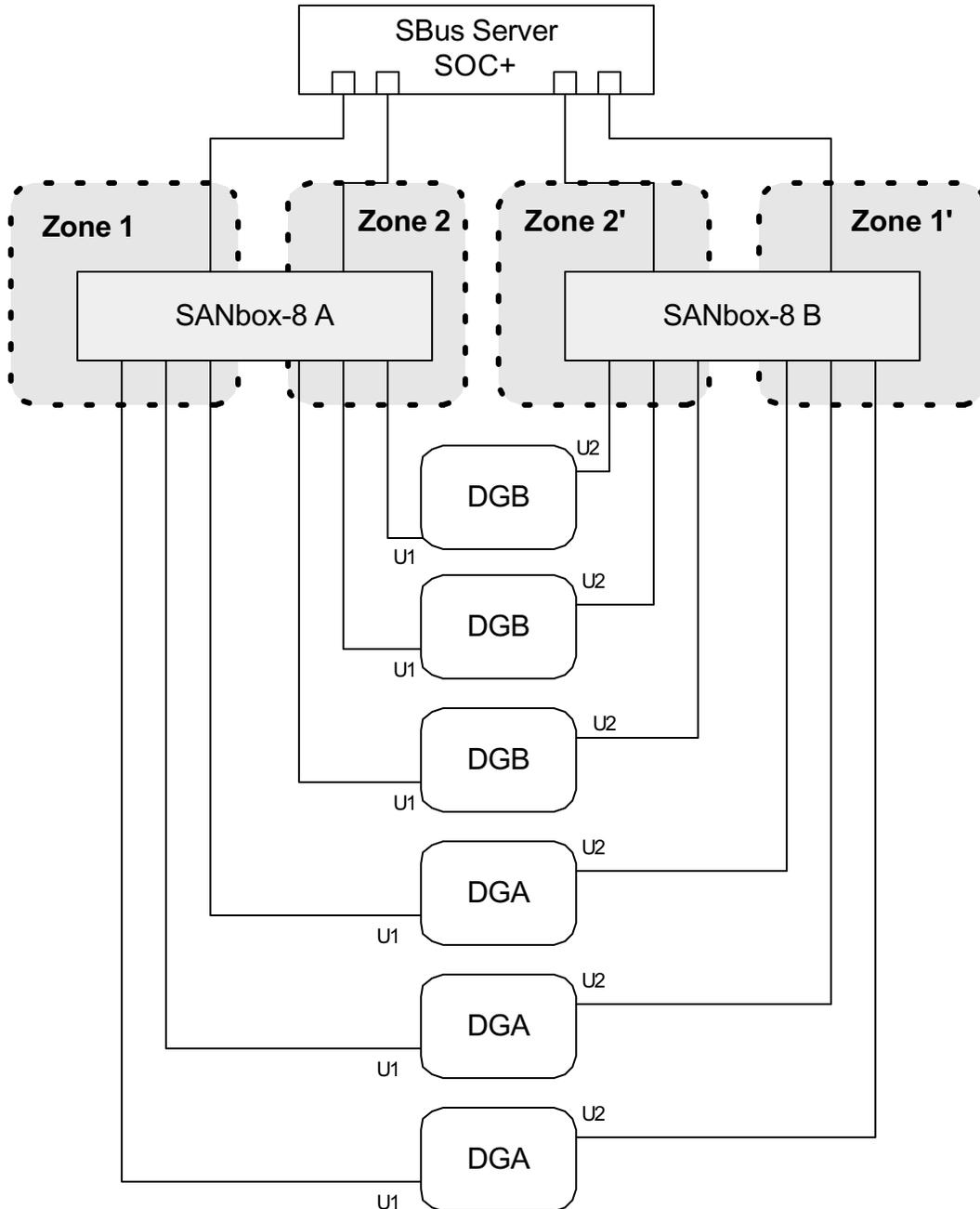- The A3500FC must be configured in Multi-initiator controller mode.

Config 3: Single Host T3WG SAN (Phase 0-2)

SBus Server SOC+

Zone 1 — Zone 2 — Zone 2' — Zone 1'

SANbox-8 A — **SANbox-8 B**

DGA, DGA, DGA, DGB, DGB, DGB

DGD, DGD, DGD, DGC, DGC, DGC

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each T3WG is configured as RAID 5 with 1 hot spare.
- Three T3WG's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- Array Group C = DGC
- Array Group D = DGD
- DGA may be mirrored to DGD using VxVM
-
- Maximum available storage:
- RAID 5, one hot spare
- 108 spindles
- 764GB with 18GB disks
- 1528GB with 36GB disks
- 3066GB with 73GB disks
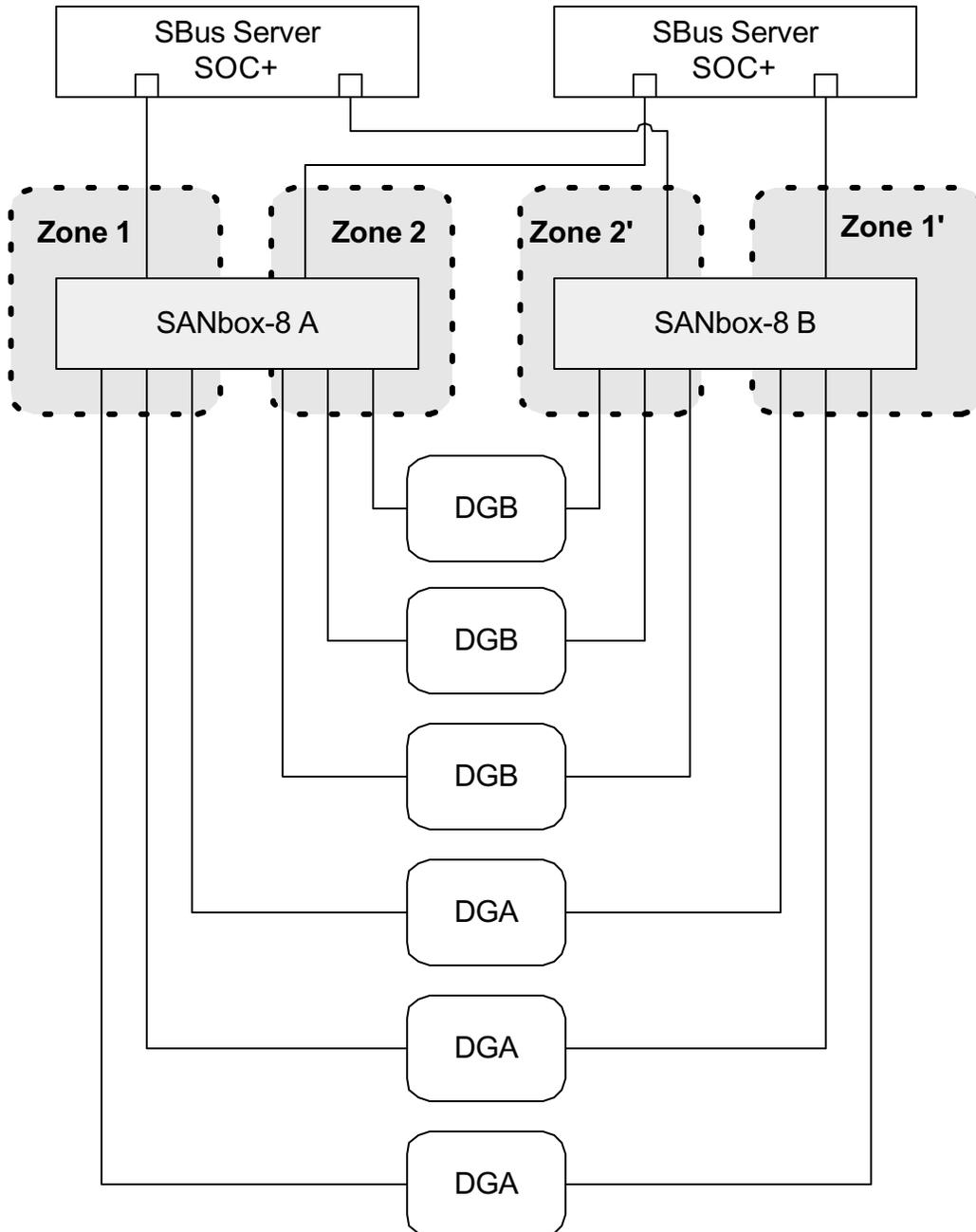- Footprint is 3 rack cabinets

Config 4: Single Host T3ES SAN (Phase 0-2)

SBus Server
SOC+

Zone 1

Zone 2

Zone 2'

Zone 1'

SANbox-8 A

SANbox-8 B

DGB — U2 / U1

DGB — U2 / U1

DGB — U2 / U1

DGA — U2 / U1

DGA — U2 / U1

DGA — U2 / U1

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each T3ES is configured as RAID 5 with 2 hot spares.
- Three T3ES's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- DGA may be mirrored to DGB using VxVM
- LUN failovers will be handled by the T3ES through the back FC connections.
-
- Maximum available storage:
- RAID 5, two hot spares
- 108 spindles
- 1528GB with 18GB disks
- 3056GB with 36GB disks
- 6132GB with 73GB disks
- Footprint is 1.5 rack cabinets

Config 5: Dual Hosts A5200 SAN (Phase 0-2)

SBus Server SOC+

SBus Server SOC+

Zone 1   Zone 2   Zone 2'   Zone 1'

SANbox-8 A

SANbox-8 B

DGB

DGB

DGB

DGA

DGA

DGA

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each A5200 is configured as a single loop, split loop is not permissible.
- Three A5200's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- DGA may be mirrored to DGB using VxVM across switches
- All storage is accessible in the event of a server failover
-
- Maximum available storage:
- RAID 0+1, no hot spares
- 132 spindles
- 1201GB with 18GB disks
- 2402GB with 36GB disks
- Footprint is 1 rack cabinet
-
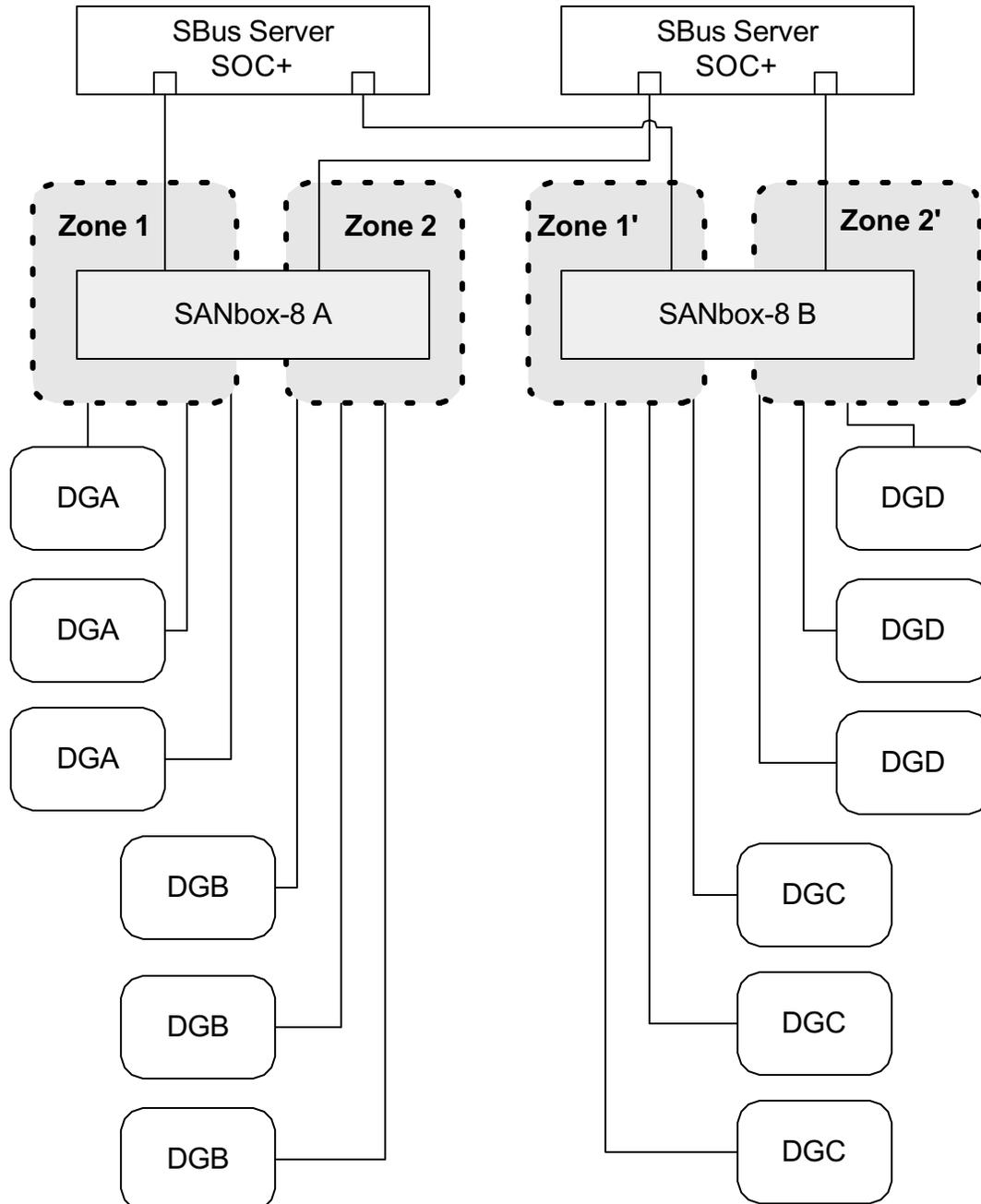- Note: Two IB boards per array must be used for redundancy.

Config 6: Dual Hosts A3500FC (1x5) SAN (Phase 0-2)
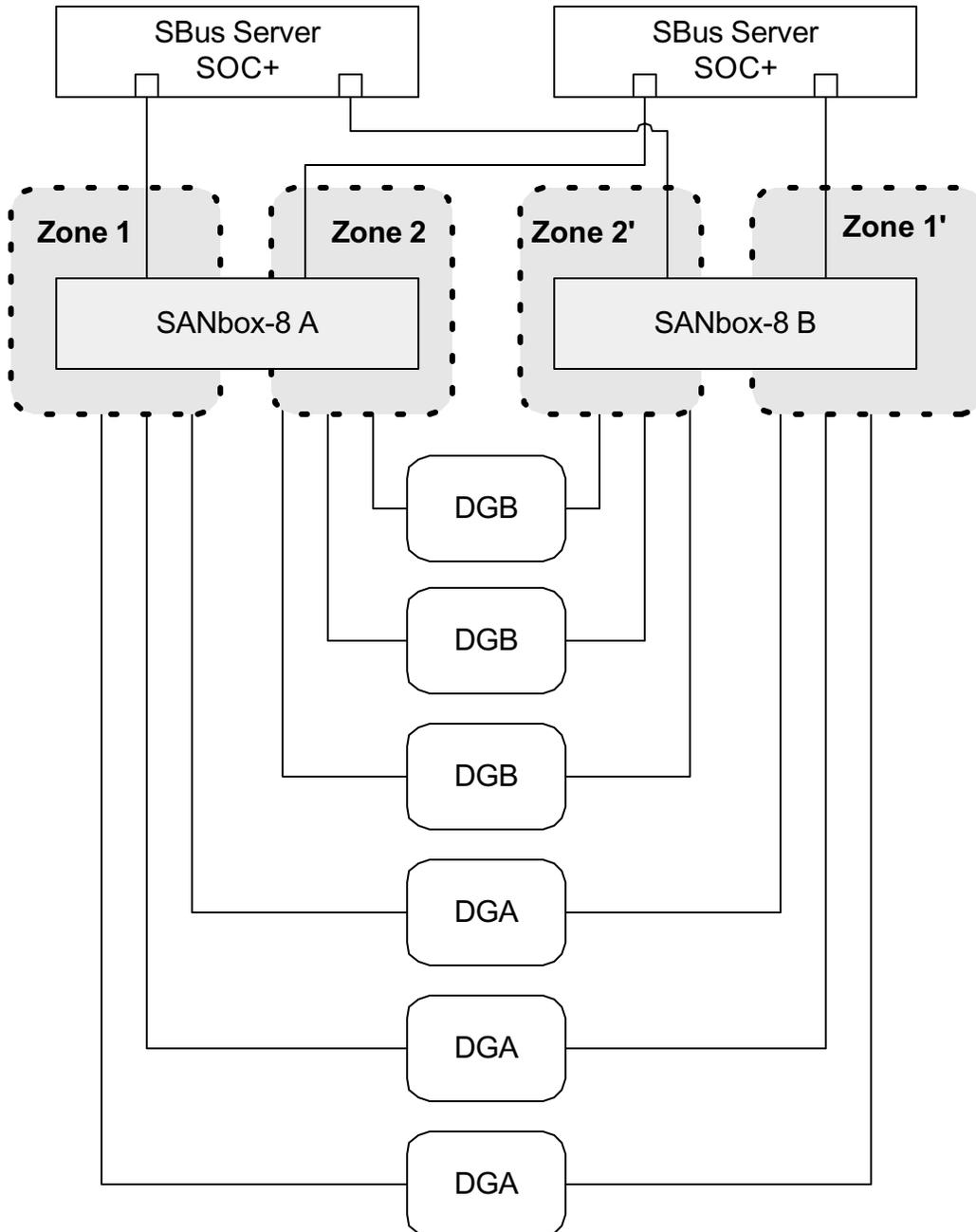
- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each A3500FC is configured as RAID 5 with 1 hot spare per tray.
- Three A3500FC's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- Mirroring of Array Groups may be accomplished using VxVM
- All storage maybe accessible in the event of a server failover
-
- Maximum available storage:
-    RAID 5, one hot spare per tray
-    360 spindles
-    4804GB with 18GB disks
-    9608GB with 36GB disks
-    Footprint is 6 rack cabinets
-
- Note: The A3500FC is configured as 1x5
- The A3500FC must be configured in active-active controller mode for cluster.

Config 7: Two Hosts T3WG SAN (Phase 0-2)

SBus Server SOC+

SBus Server SOC+

Zone 1

Zone 2

Zone 1'

Zone 2'

SANbox-8 A

SANbox-8 B

DGA

DGA

DGA

DGB

DGB

DGB

DGD

DGD

DGD

DGC

DGC

DGC

- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each T3WG is configured as RAID 5 with 1 hot spare.
- Three T3WG's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- Array Group C = DGC
- Array Group D = DGD
- DGA may be mirrored to DGC using VxVM
- DGB may be mirrored to DGD using VxVM
- All storage is not accessible in the event one server fails.
-
- Maximum available storage:
- RAID 5, one hot spare
- 108 spindles
- 764GB with 18GB disks
- 1528GB with 36GB disks
- 3066GB with 73GB disks
- Footprint is 3 rack cabinets
-
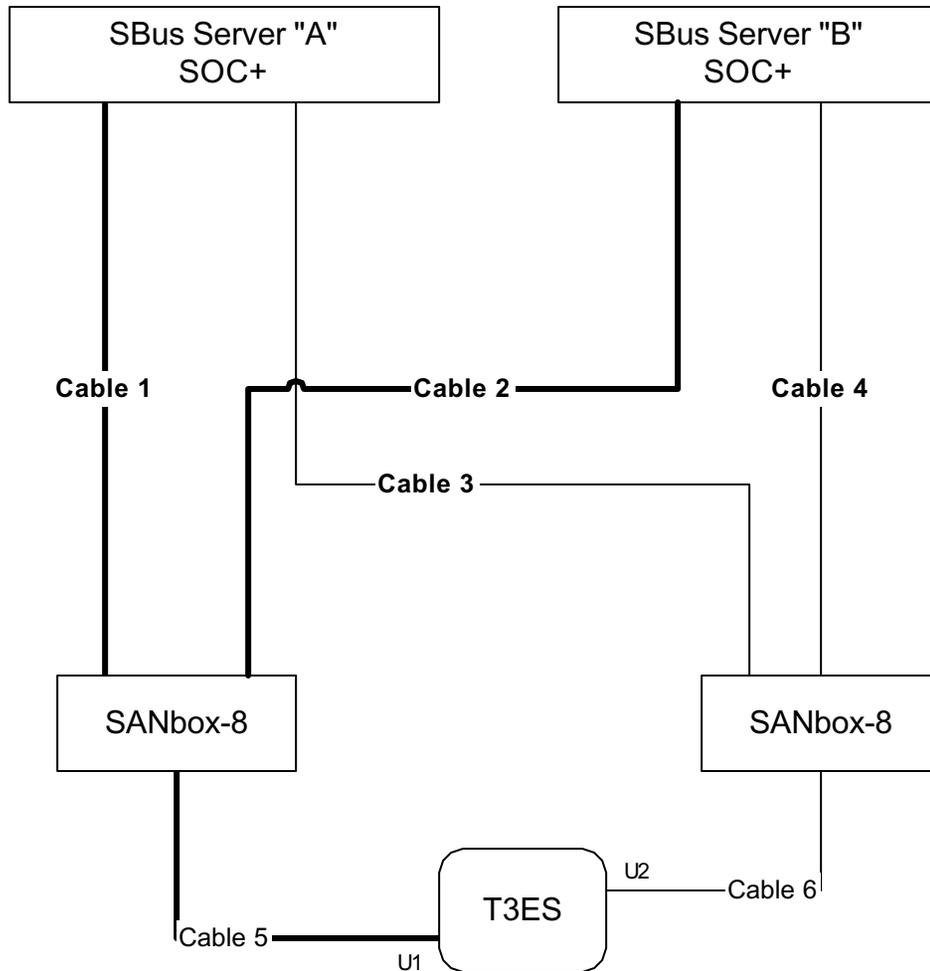- Clustered config would be different.

Config 8: Two Hosts T3ES SAN (Phase 0-2)
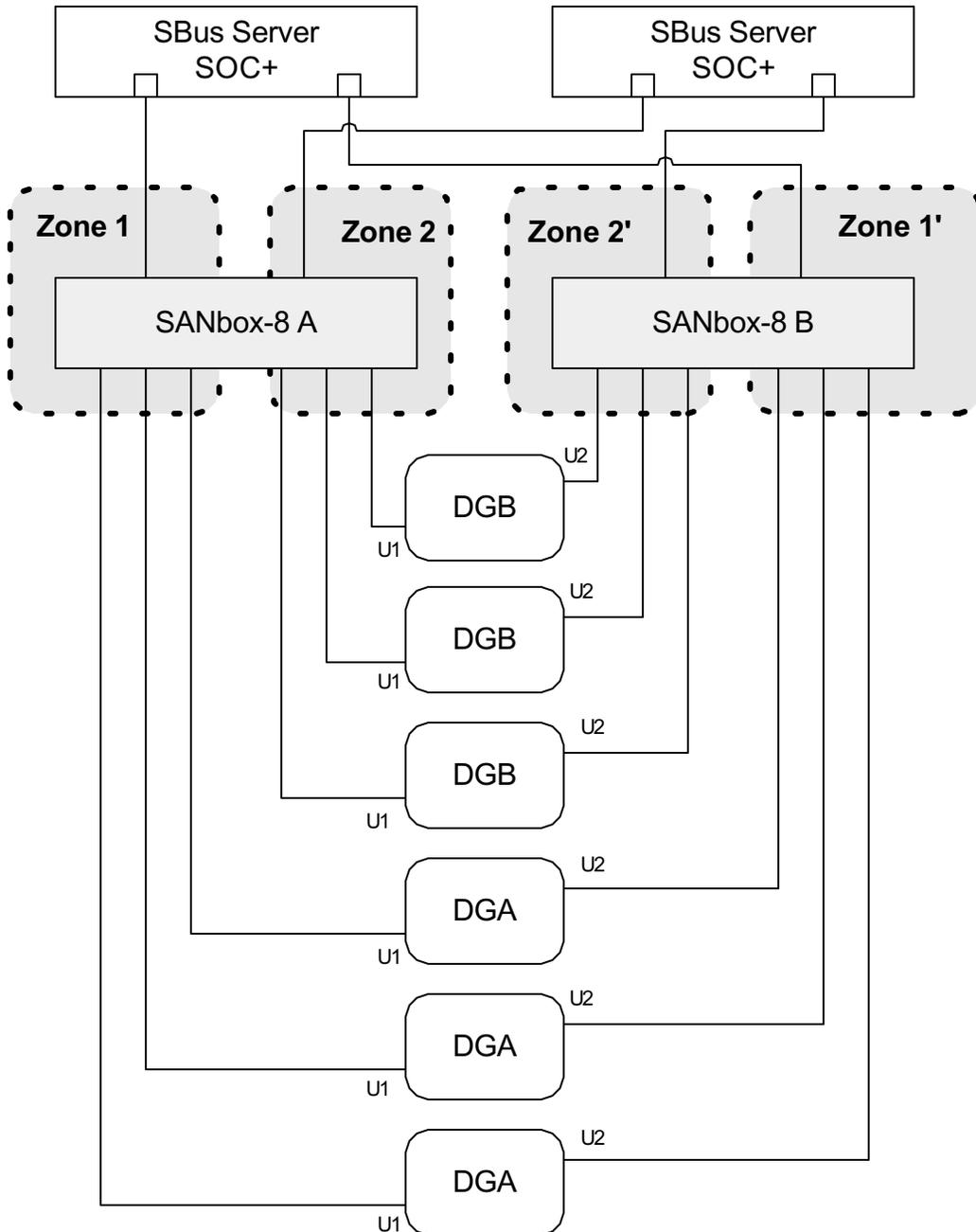
- Each line from the server equals 1 SOC+ card.
- Each switch has 2 zones of 4 ports: 1 initiator and 3 targets.
- Each T3ES is configured as RAID 5 with 2 hot spares.
- Three T3ES's per Array Group
- Array Group A = DGA
- Array Group B = DGB
- Mirroring of data must be done at the host level, and not through Cluster software.
- LUN failovers will be handled by the T3ES through the back FC connections.
-
- Maximum available storage:
- RAID 5, two hot spares
- 108 spindles
- 1528GB with 18GB disks
- 3056GB with 36GB disks
- 6132GB with 73GB disks
- Footprint is 1.5 rack cabinets
-
- Note:
  The T3ES's are connected to the same server. This does not violate the multi-host rule on SC2.2 with Sol 2.6
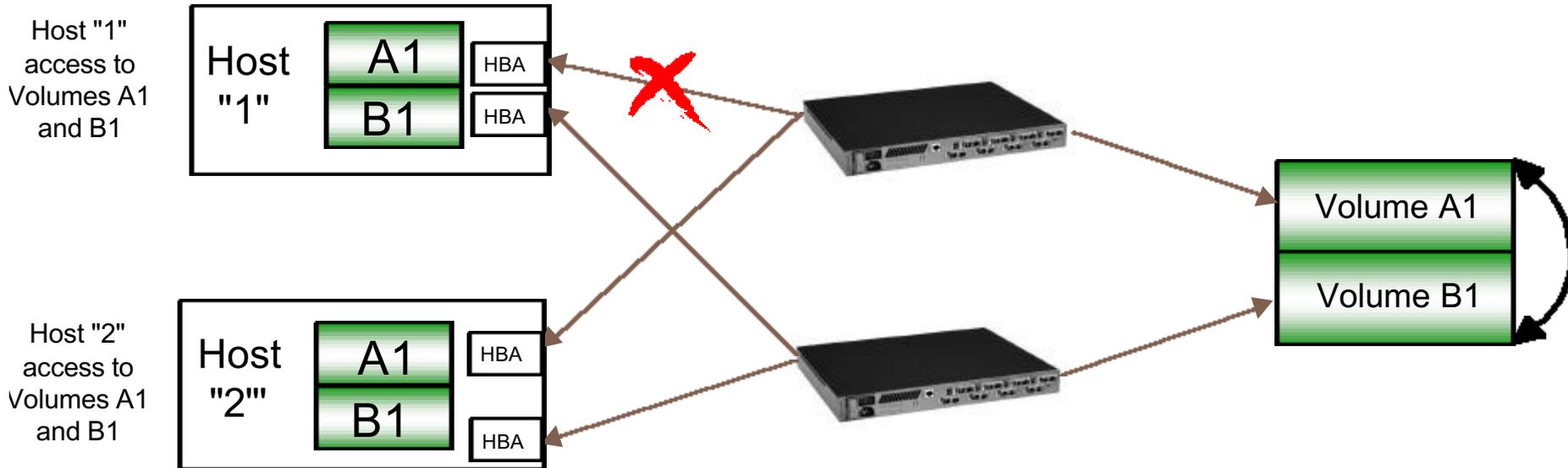
Config 9: Why T3ES cannot be in a SunCluster 2.2 environment with SOC+

SBus Server "A"
SOC+

SBus Server "B"
SOC+

Cable 1

Cable 2

Cable 4

Cable 3

SANbox-8

SANbox-8

U2

Cable 6

T3ES

Cable 5

U1

Primary Data Path for Server A uses Cables 1 and 5
Primary Data Path for Server B uses Cables 2 and 5

- Scenario: Assume concurrent data access: eg. OPS, QFS, SAM-FS, Veritas Cluster FS, RYO.
- Primary datapath for Server A fails upstream from switch (ie. GBIC, cable, HBA)
- VxVM initiates a failover request on Server A to use Cable 3 and 6. (ping)
- T3 receives data transfer request from Server A on cable 6 and initiates an implicit LUN failover.
- However, Server B's upstream path (ie Cable 2) works fine and continues to issue I/O through Cable 2 to 5.
- Receipt of data transfer I/O from Server B on Cable 5 triggers a LUN fail back (pong).
-
- Conceptually, this is at odds with multi-pathing and SANs, which is fixed with SAN 3.0 "Python" release using MPxIO, and PGR/3 firmware addition on T3.
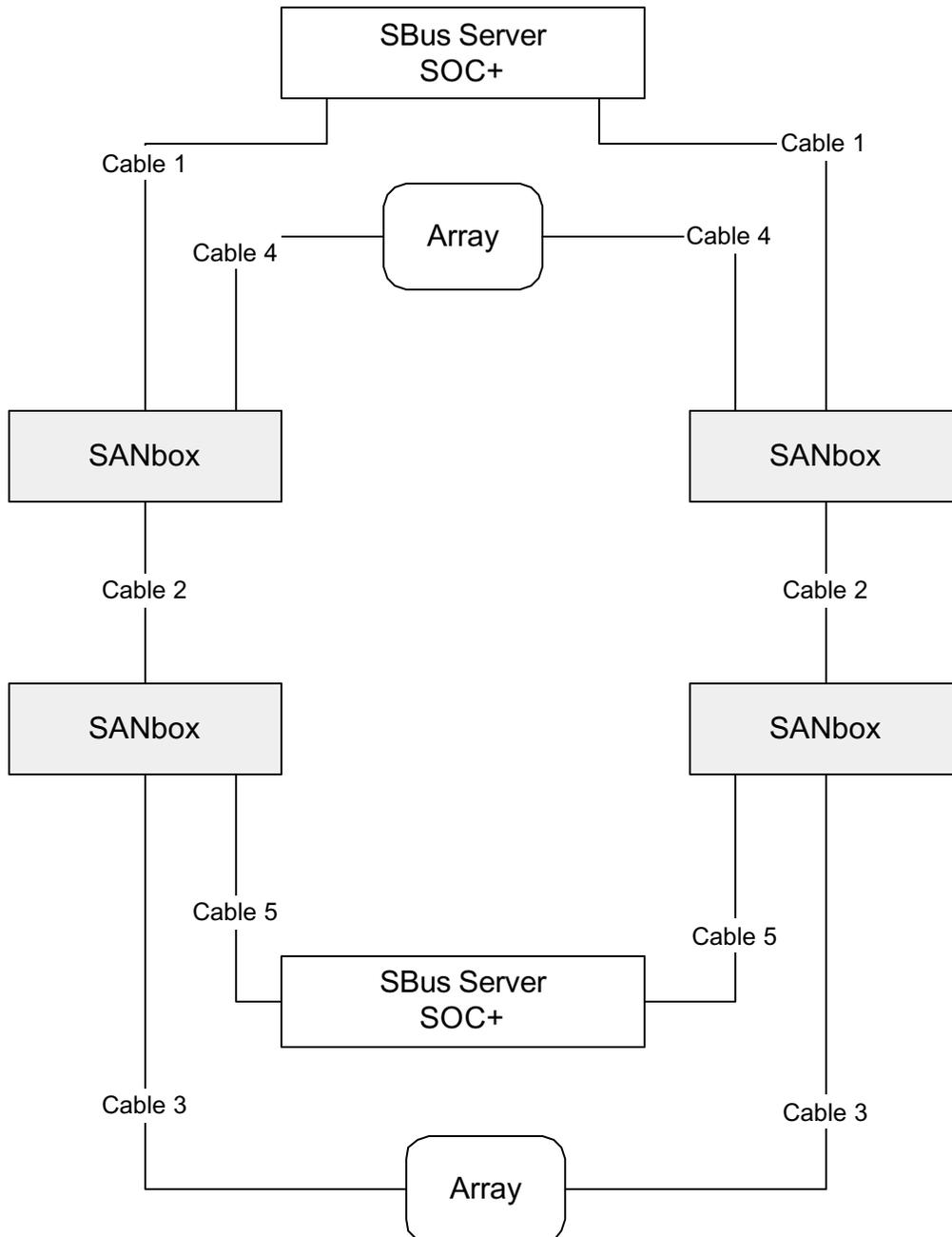
Config 10: T3 Partner Pair Fail-Over in a Multiple Host SAN Environment (Clustered or Otherwise) (Or why I've always hated playing Ping-Pong)

Host "1" access to Volumes A1 and B1

Host "1"
A1
HBA
B1
HBA

Volume A1

Volume B1

Host "2" access to Volumes A1 and B1

Host "2"
A1
HBA
B1
HBA

- Scenario: (Assumes concurrent data access: e.g. OPS, QFS, SAM-FS, Veritas Cluster FS or RYO.)
- Host "1" sustains a failure "upstream" from the T3 pair (At the cross) HBA, Gbic, Cable or Switch Port.
- VxVM initiates a fail over alternate path to allow continued access to Volume A1 (upper). (Ping)
- T3 receives data transfer request for Volume A1 on Host 1 alternate path and initiates an implicit lun fail over.
- However, Host 2's "upstream" path to Volume A1 is good. It continues to issue I/O on the primary path.
-     Receipt of data transfer I/O on primary (upper) path triggers a lun fail back (Pong)
- This is at odds with entire concept of multi-pathing and SANs.
-     Requires implementing a Path Virtualization "black box" solution or re-architecting the T3 fail over.
-   Many luns (Lun slicing) exacerbates this issue.

Config 11: SOC+XT SAN using Sun SANbox switches (Phase 1-2)

SBus Server
SOC+

Cable 1

Cable 1

Array

Cable 4

Cable 4

SANbox

SANbox

Cable 2

Cable 2

SANbox

SANbox

Cable 5

Cable 5

SBus Server
SOC+

Cable 3

Cable 3

Array

- The SANbox ports for cables 1, 3, 4 & 5 are set to "SL-port". Therefore:
  SOC+(FCAL) -- SL to SANbox
  SANbox SL -- Array (FCAL)

-

- Cable 2 connects the two SANbox switches together using short-wave or long-wave FC cable, up to 10km. The SANbox ports are set to "T-port".
  Cable 2 is optional.

-

- The storage array may be one of the following types:
  A5200 (no 9GB drives allowed)
  T3ES and T3+ES have no concurrent data access in SunCluster configs.
  T3WG and T3+WG may be used but only a single connection to the switch is allowed.

-

- Both SBus servers and both Arrays may be in the same SL zone.

-

- All devices are running at SAN 3.0 Python revision levels.

-

- Reverse-TL requires SunPS Approval.

Config 12: SOC+XT SAN using Brocade Silkworm (Phase 1-2)
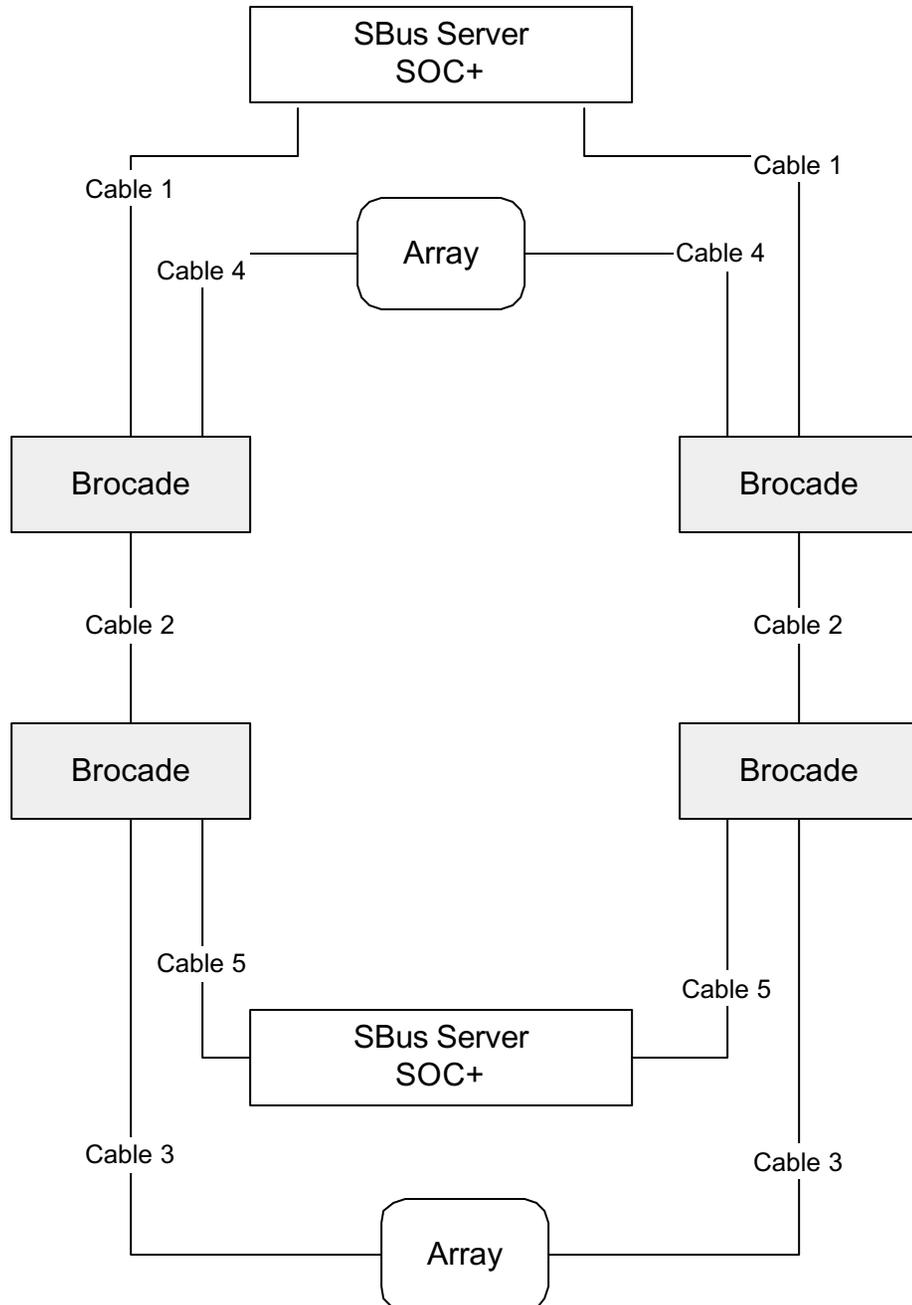
- The Brocade Silkworm ports for cables 1, 3, 4 & 5 are QuickLoop enabled. Therefore:
  SOC+(FCAL) -- QuickLoop Brocade
  Brocade QuickLoop -- Array (FCAL)

- Cable 2 connects the two Brocade switches together using short-wave or long-wave FC cable, up to 10km. The Brocade ports are automatically set to "E-port".
  Cable 2 is optional.

- The storage array may be one of the following:
  A5200 (no 9GB drives allowed)
  T3ES and T3+ES have no concurrent data access in SunCluster configs.
  T3WG and T3+WG may be used but only a single connector to the switch is allowed.

- Both SBus servers and both Arrays may be in the same QL zone.
- Brocade automatically logs in the array. QuickLoop enables the SOC+ to see the QL nodes.

- All devices are running at SAN 3.0 Python revision levels.

- Fabric-Assist requires SunPS Approval.

SOC+XT SAN Limited Program Reference Websites

The following websites should be used for references for the SOC+XT Limited Program.

Basic Configuration Rules:
    NWS Mrkt SOC+ homepage: http://webhome.ebay/networkstorage/products/soc/
    SunPS SOC+ homepage: http://sunps.central/Services/storage/tech_resources/

CPRE Support SAN Page: http://cpre-amer.west/nws/products/Switch/
Americas Product Support Services SAN Page: http://acts.ebay/storage/SANs/

SAN Components Homepage (FW, GUI, docs, etc):   http://webhome.ebay/sancomponents

Documentation:
    main: http://edist.central
                   http://inforserver.central/data/sshandbook/Systems/system-list.html
    switch: http://cpre-amer.west/nws/products/Switch/documentation.html

Diagnostics Software:
    RASagent: http://webhome.central/RASAgent
    STORtools: http://webhome.central/stortools/

Patch Matrices:
    A5200/T3:   Early Notifier 14838
    A3500:   Early Notifier 20029
    SunCluster: Early Notifier 24617
    E10K   http://cpre-amer.west/esg/hsg/starfire/patches.html