

Antony Steel

Session: a111037 - Data Centre migrations (101)

Thanks to:
Various customers who allowed testing



`#include <std_disclaimer.h>`

These notes have been prepared by an Australian, so beware of unusual spelling and pronunciation. All comments regarding futures are probably nothing more than the imagination of the speaker and are IBM Confidential till after GA.



- An introduction to migrating workloads between Data Centres - exploring some of the options through looking at some recent customer examples

Data Centre (DC) migrations....

- Is a common problem (migrating workloads from one DC to another)?
 - Out grow DC; DC closes; contracts reviewed; Customer moves
- Procedure
 - Analyse customer requirements
 - Current infrastructure, RTO, RPO, availability and redundancy configuration and requirements
 - Outage windows and change control; Access/security requirements
 - How many customer teams? Coordination / communication issues
 - Audit environment
 - OS; patches; Management & monitoring; Applications & Application dependencies
 - Network / Storage design – don't want surprises (for example storage units in wrong racks)
 - Review options <=> match with requirements
 - Timings; back-out plans; OS / firmware / application dependencies
 - Develop plan / coordination
 - What can go wrong? What will go wrong? Back-out / contingency plans
- Perform
- Post migration review / lessons learnt

Data Centre (DC) migrations....a common problem?

- Migrating workloads from one DC to another
 - Out grow existing DC
 - DC closes
 - Review of contracts
 - Customer moves
 - Consolidating DC / hardware upgrade
 - Move p5-p7 to p8
 - Move “Out of support” AIX levels to in support
 - Upgrade of hardware / move to private cloud
 - Migration of mixed Power to POWER8 for traditional workloads / HANA
 - Migration of mixed x86 linux to Linux on POWER
 - New linux workloads on POWER
 -

Procedure - Analyse customer requirements

- Review the current infrastructure
 - RTO, RPO
 - availability and redundancy configuration
 - Availability and redundancy requirements during the migration
 - Sufficient CPUs / licensed cores
 - Tools – FLRT; LPM verification; iostat
- What constraints around timings
 - Outage windows – fit into plan
 - change control – prepare to avoid delays
 - Access/security requirements
- Managing the process
 - How many customer teams?
 - Fit your migrations in with other parts of the organisation (you are unlikely to be the only migration)
 - Coordination and communication issues
- Testing
 - Reproducible tests
 - I/O performance for storage (# mirror copies, Different storage systems / sites)
 - Simulated load for LPM

- Conduct an audit of the current environment
 - What Operation Systems and patch levels
 - What virtualisation in place
 - What management & monitoring
 - What applications & dependencies
 - Current network design
 - Current storage layout
- Audit demonstrate that current environment meets customers standards?
- Audit expose any problems for migration?
 - We found both storage and network inconsistencies during the audit and were able to highlight some OS exposures that could be patched prior to migration

Procedure – Review options

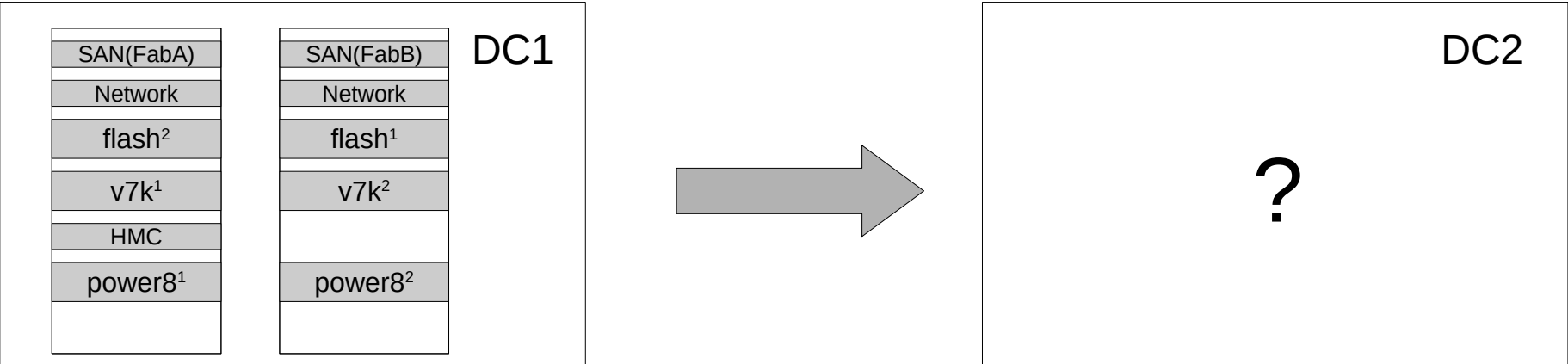
- Review possible LPAR migration options
 - What limitations do the customer requirements impose
 - Outage? Availability? Timing windows? Network / storage access?
- Review possible Storage migration options
 - What limitations do the customer requirements impose
 - Outage? Availability? Timing windows? Bandwidth and data volumes
 - What are the options generally for DR?
 - Storage mirroring (sync, async, hyperswap)
 - OS mirroring (LVM, GLVM)
 - Application mirroring (Log shipping)
- For each option review:
 - Timings
 - back-out plans
 - OS / firmware / application dependencies

- Develop plan
- Coordination / communication across teams
- What can go wrong? What will go wrong? Back-out / Rollback / contingency plans

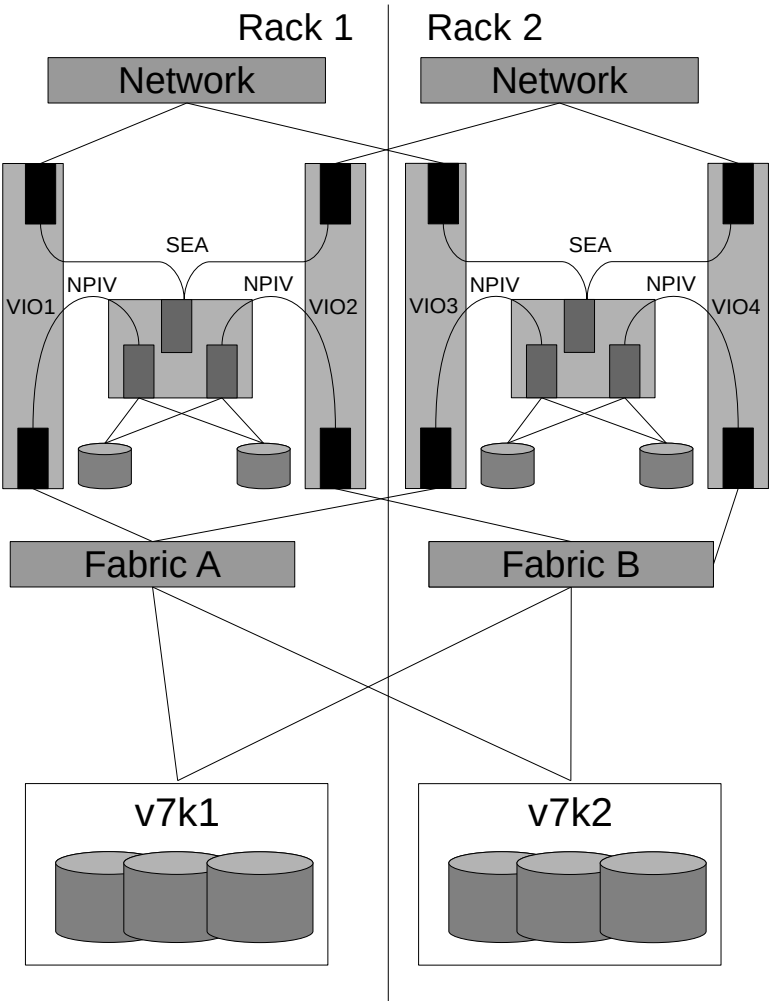
- Perform migration(s)
- Post migration review
- lessons learnt

Example

- Customer requirement
 - Move operations from existing DC to new DC in same city
 - 2 racks of power and storage
 - Fully virtualised; mixed patch level (OS / f/w)
 - LVM mirroring between racks
 - Issues
 - Needed to apply some patches (HMC; VIO; AIX; firmware)
 - Some mirrors inconsistent
 - No outage highly desired; work on production only during restricted hours on Sunday morning
 - Backup no redundancy and ran nightly (therefore physical move during the day)



Customer starting position



Options to migrate OS instances

- Hardware
 - Need to maintain customer's availability requirement (2 copies of storage; 1 standby server)
 - Need inter-site link to handle data replication within window mixed with system activity, while not impacting the system activity
- Options to migrate OS
 - LPM
 - Pro: No outage; easy to roll back
 - Con: HMC; Network; Storage visible; Network latency
 - Simplified Remote Restart
 - Pro: faster if memory access high; less requirements; easy to roll back
 - Con: Short outage
 - Create new instance at new site, test; move data LUNs for migration
 - Pro: Less network/storage requirements; Can perform OS updates in move
 - Con: Longer outage
 - PowerHA
 - There was no PowerHA in the customers environment

Used this option for one customer, nimadm migration for OS LUNs rsync for data LUNs Change freeze for OS, regular sync to keep last sync to a minimum for cutover.

- Take the opportunity to perform AIX migration / install code or updates?
 - If migrating the OS: use NIM *nimadm*
 - If installing code etc, and OS level etc correct – AIX Live Kernel update
- *nimadm**
 - The process does the following:
 - Create a copy of the rootvg on a spare client LUN (similar to alt disk install *alt_disk_copy*)
 - Migrate the newly created copy of rootvg to the new version of AIX, installing additional filesets as required – While the system is still running on the current version of AIX, ie no disruption
 - At convenient time schedule a reboot of the client – choosing the new rootvg as the target.
 - The advantages are:
 - Minimum disruption on the client (just a reboot) – no outage for the migration
 - As the process of the migration is run from NIM, it primarily uses NIM resources (and some network bandwidth)
 - Easy back-out and debugging. Can roll back to the original copy of the rootvg and investigate the updated rootvg to resolve problem – In a fully virtualised environment can even assign the rootvg to another LPAR!

* Used for customer in Singapore AIX 5.3 on old h/w to POWER8

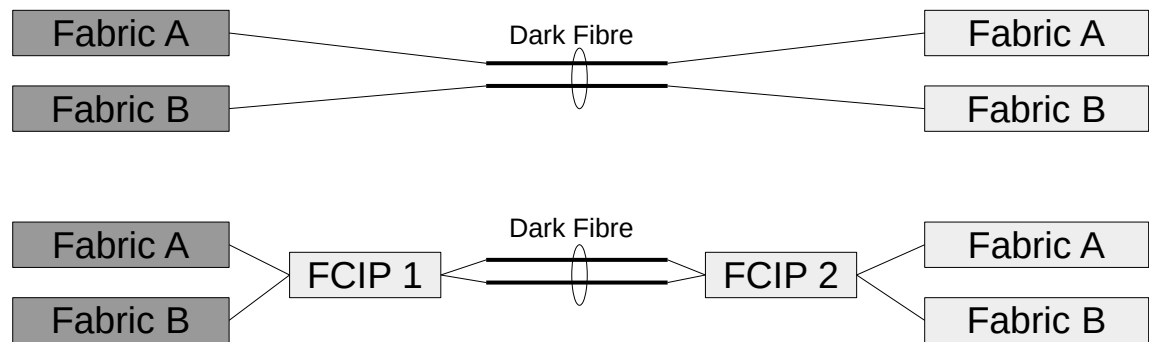
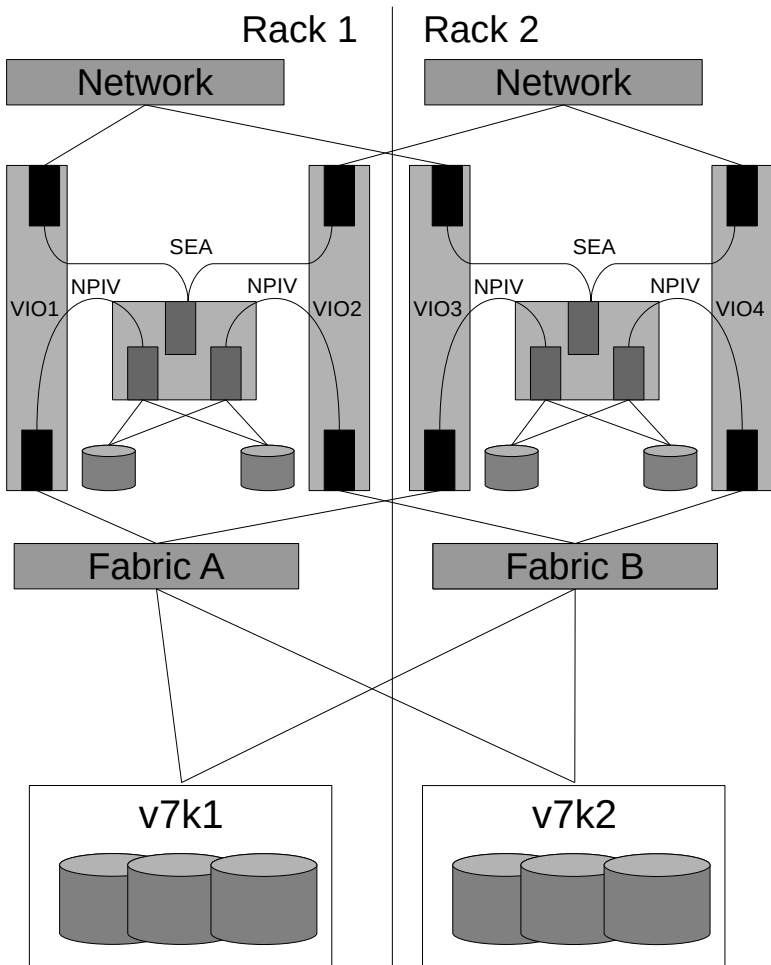
Migrate storage

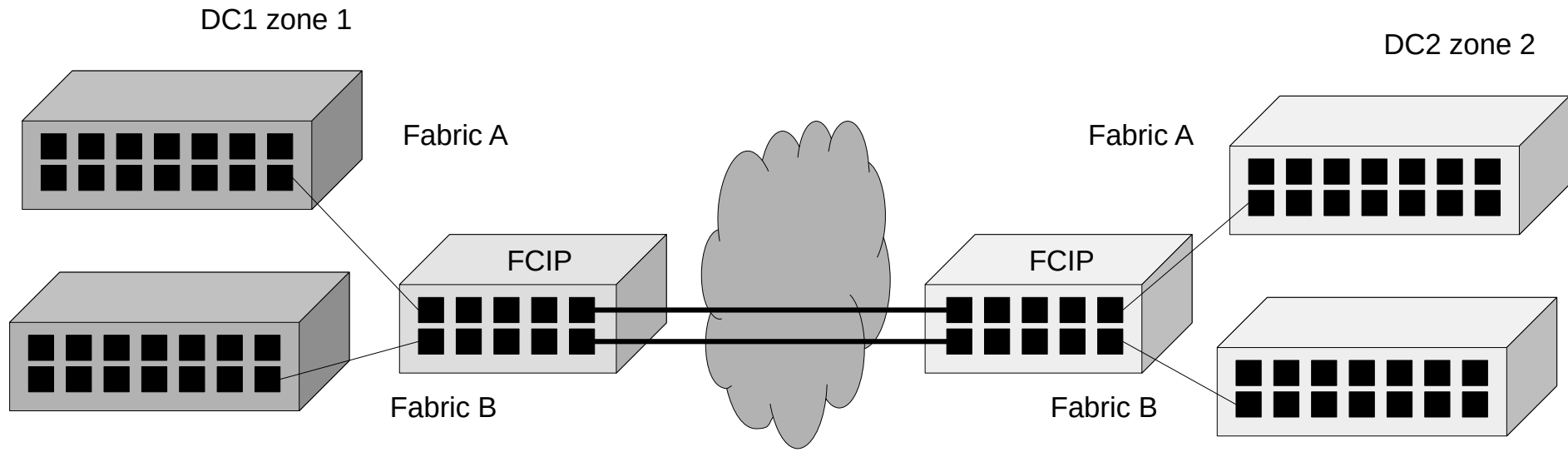
- Options to maintain at least 2 copies of storage
 - Use Storage subsystems mirroring/replication
 - Pro: consistent across all platforms; offload data copying so less impact on CPU
 - Con: Expensive; license requirements
 - Use LVM
 - Pro: cheap and easy (already use LVM mirroring within site)
 - Con: mirroring is fine for AIX/Linux..
 - Use GLVM
 - Pro: only requires IP link
 - Con: AIX specific
 - Log shipping / replication
 - Pro: Less bandwidth requirement
 - Con: Application specific

- Option 1: Lift and shift - Move 1 rack at a time
 - Con: As we shift each rack, we do not have redundancy
- Option 2: Storage replication – 2 way Storwize Remote Copy
 - Con: The cutover from source storage on site 1 to target storage on site 2 is disruptive
 - IBM Geographically Dispersed Resiliency (GDR) could be used to reduce the failover time
 - IBM PowerHA-XD could be used to reduce the failover time
 - Need to use either native IP replication or stretch the fabric using FCIP routers or dark fibre
- Option 3: Storage replication – 2 way Storwize HyperSwap
 - Con: The cutover from source storage on site 1 to target storage on site 2 is disruptive
 - IBM PowerHA-XD could be used to reduce the failover time
 - Need to use either native IP replication or stretch the fabric using FCIP routers or dark fibre
- Option 4: Server replication – 3 way Logical Volume Mirroring
 - Compatible with Live Partition Mobility = online migration
 - Options to present LUNs from new DC to OS images (Stretch SAN; FCIP Routers; iSCSI; GLVM)
- Option 6: Introduce iSCSI as an option for mirroring
 - Con: Major change in operation of site, keeping 2 copies of data at all time difficult to manage

Extending the fabric

- Few options need to be considered:
 - Leverage iSCSI host to storage protocol instead of the FCP protocol currently used
 - Leverage FCIP routers to interconnect FCP across both data centres without merging SAN fabrics (FCIP routers might only be used to simply prevent the SAN fabrics on each site from merging... they do not have to be used to encapsulate FCP frames into TCPIP frames)
 - Decision: 2 loan switches to maintain redundancy





Summary

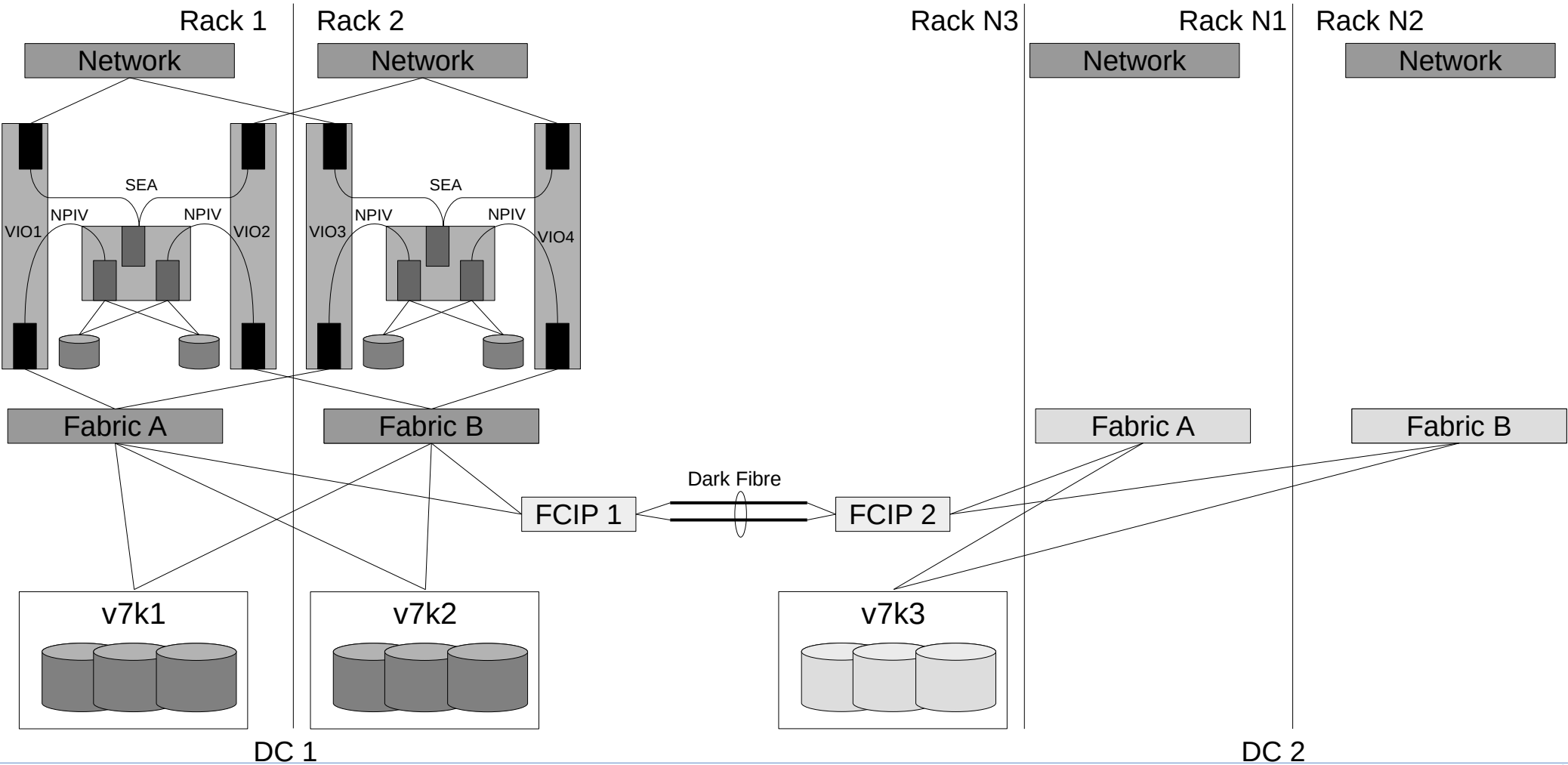
	iSCSI	FCIP Router	Cross site interconnect (dark fibre w/wo WDM)		
			SAN switch	Host Connection	Storwize Remote Copy
Compatibility with LVM	Yes	Yes	Yes	Yes	No
Disruptive	No	No	No	No	Yes
Redundancy	Yes	Yes	Maybe ¹	Maybe ²	Maybe ³
Performance	*	****	***	**	*****
Cost	\$\$	\$	\$\$\$\$	\$\$\$\$\$	\$\$\$
LW SFPs	No	No	Yes	Yes	Yes
Migration steps	+++++	++++	++	+	+++
Stability/Resiliency	**	****	*	**	***

1. If you add a SAN switch per site during the relocation, than SAN fabrics A & B can be stretched, otherwise it would mean interconnecting both SAN fabrics.
2. If you have many dark fibres or active DWDM with many channel then all host connections on site 1 can be connected to the SAN switch B on site 2 and reversely all host connections on site 2 can be connected to the SAN switch A on site 1.
3. Recommendation is at least 2 ports per Storwize node canister

Considerations

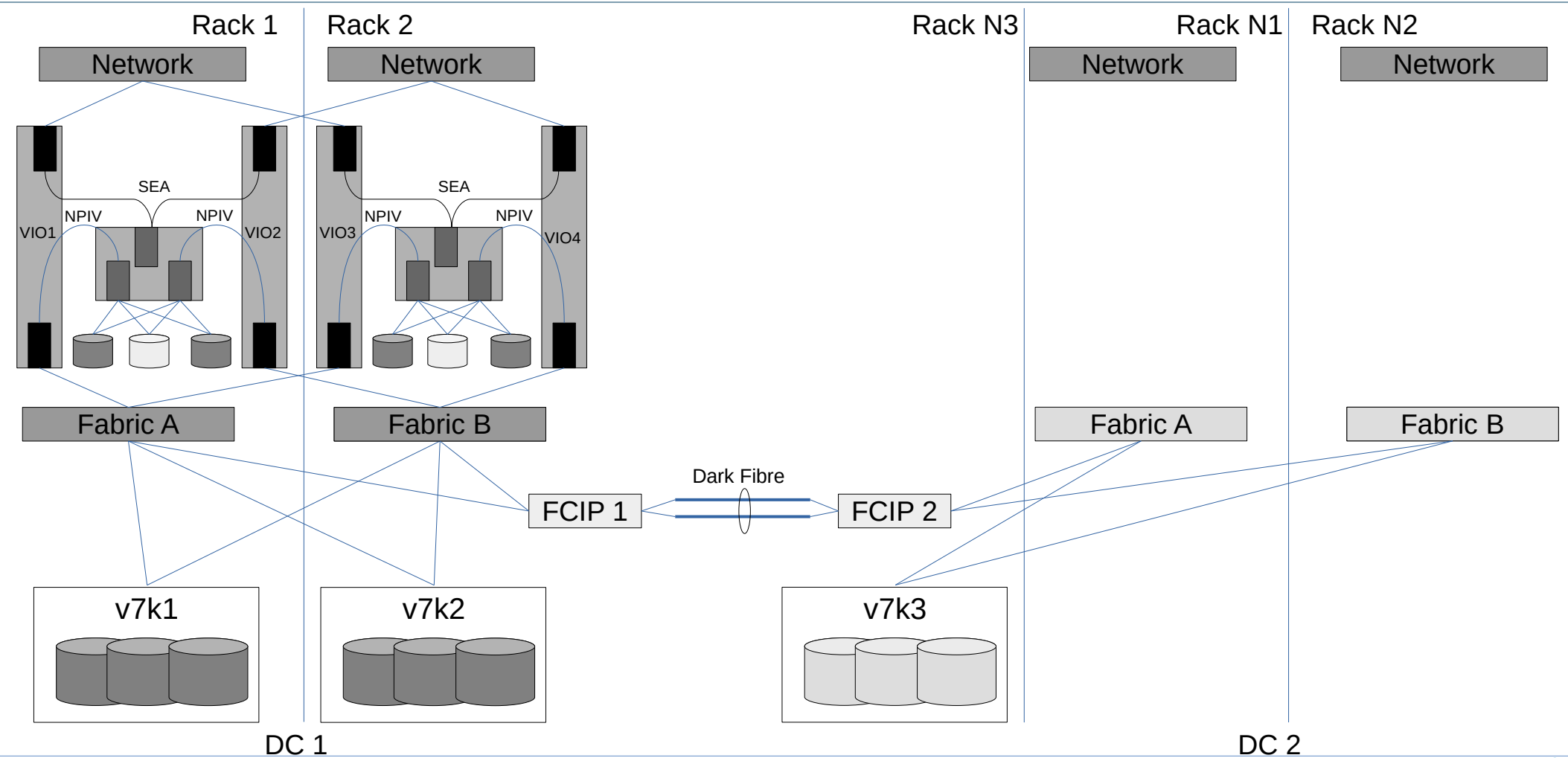
- Why use FCIP routers
 - The key advantage is because of the integrated routing function that isolates 2 SAN fabrics while still interconnect them. That is connecting but not merging, so any topology change on one side does not propagate to the other side (for example a registered state change notification (RSCN) storm does not propagate outside the EX or VEX (routed) port.
 - It was easy to harvest both the SAN switch configuration and the Storage / LUN configuration to script a consistent build of the remote SAN switches and storage.
- Bandwidth requirements
 - Audit of environment
 - Audit of LVM and Filesystem layout provided the total amount of data to be moved
 - Performance audit of the virtual machines gave an indication of the amount of data updated
 - This gave us some flexibility, as depending on cable type, link speed and distance, we have a choice of SFPs and may require multiple links across multiple protocols.
 - Depending on your fibre cable type, the link speed and the distance, you might need to use LongWave SFPs as well as WDM equipment. For this customer, using LW SFPs was a must, and depending on how many dark fibre links we were provided and our expected throughput, we may need to share one or multiple links across multiple protocols.

Extending the fabric



- Confirm the configuration of the loan switches (copy and edit from the clients switches)
- Confirm the creation of all the LUNs on the loan storage (scripted).

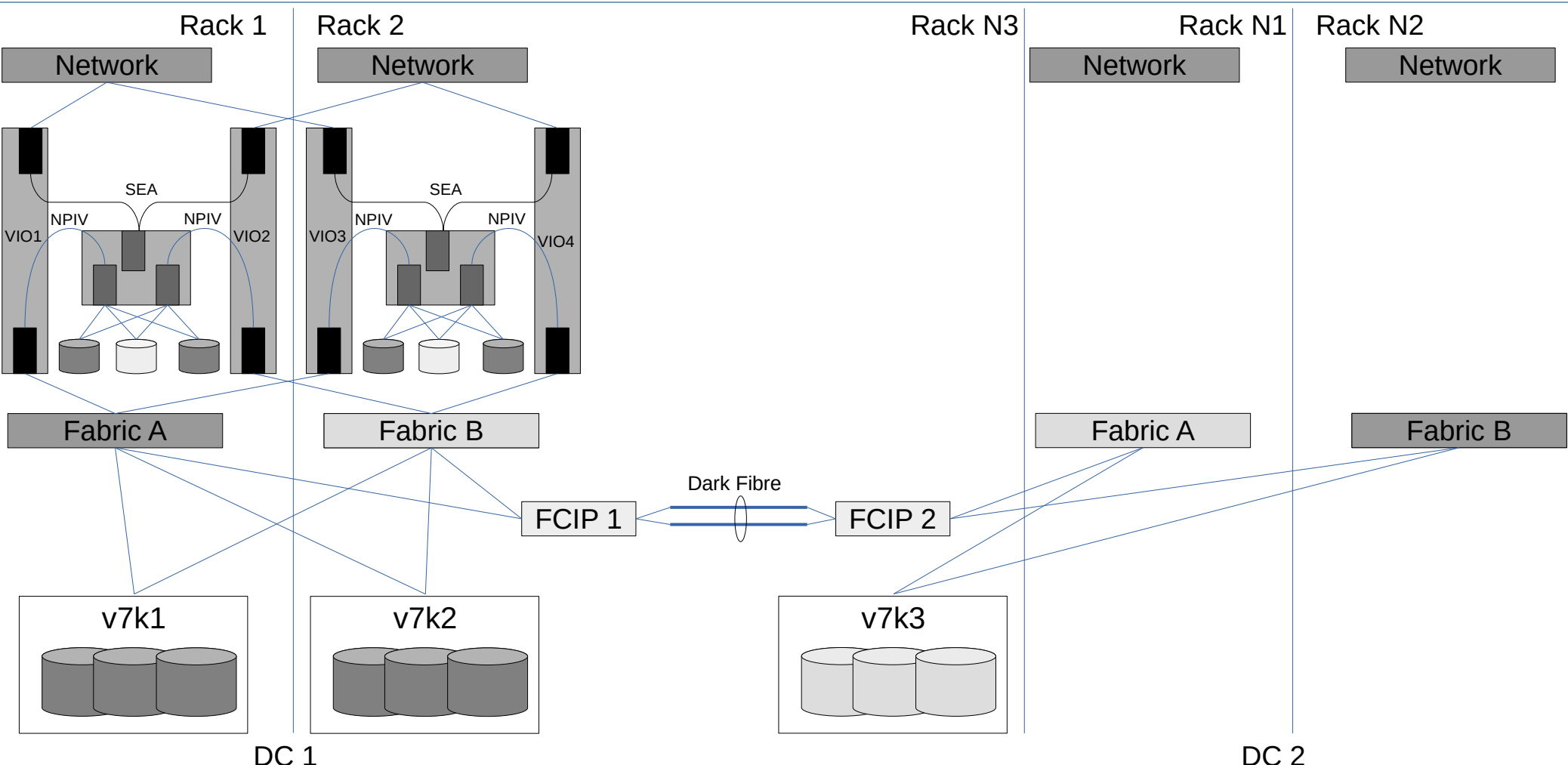
Add third copy to all LUNs



Build and testing

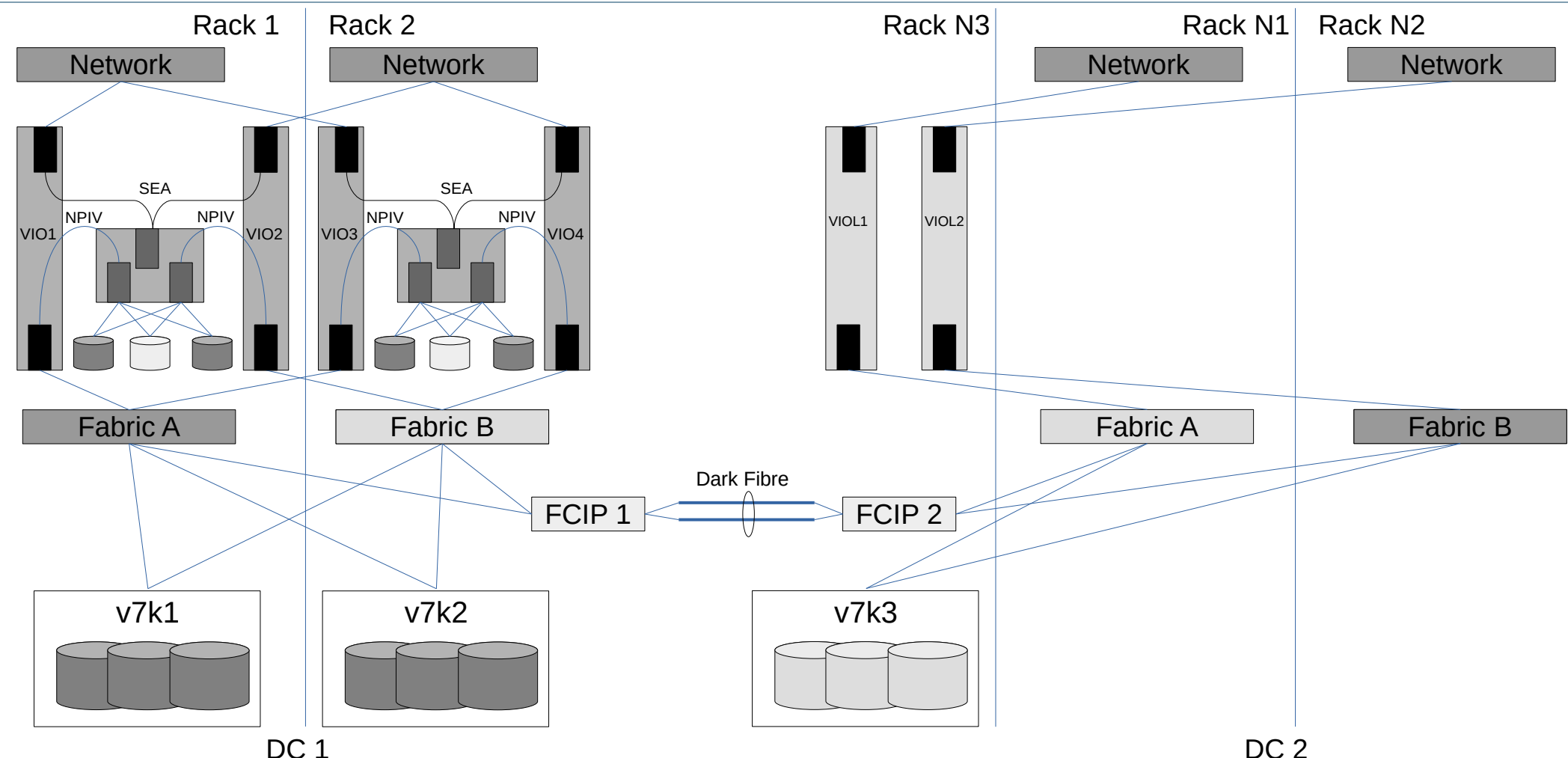
- Add LUNs to test LPAR (testing configuration end to end)
 - Scripts were used on each LPAR to manage the 3rd copy of the mirror and manage each missing copy and re-sync as the underlying storage was removed and then became available during the hardware relocation
- Build I/O baseline performance for test LPAR with 2 way mirror
 - Create hosts and mappings on loan storage for test LPARs
- Add additional LUN(s) to each VG on test LPAR and sync (time to confirm performance)
 - Re-scan the storage, confirm presence and size of new LUNs
 - Run script to add LUNs to appropriate VGs and resync each LV
 - Build table of multiple test runs to establish mean and standard deviation
 - Confirm plans (time to resync data between successive rack moves and fit within backup windows, while reducing impact on day to day operations)
- Test I/O performance for test LPAR with 3 way mirror and compare
- After sign-off add third mirror to all LPARs and script sync during quiet times.

Swap switches (Fabric A stable)



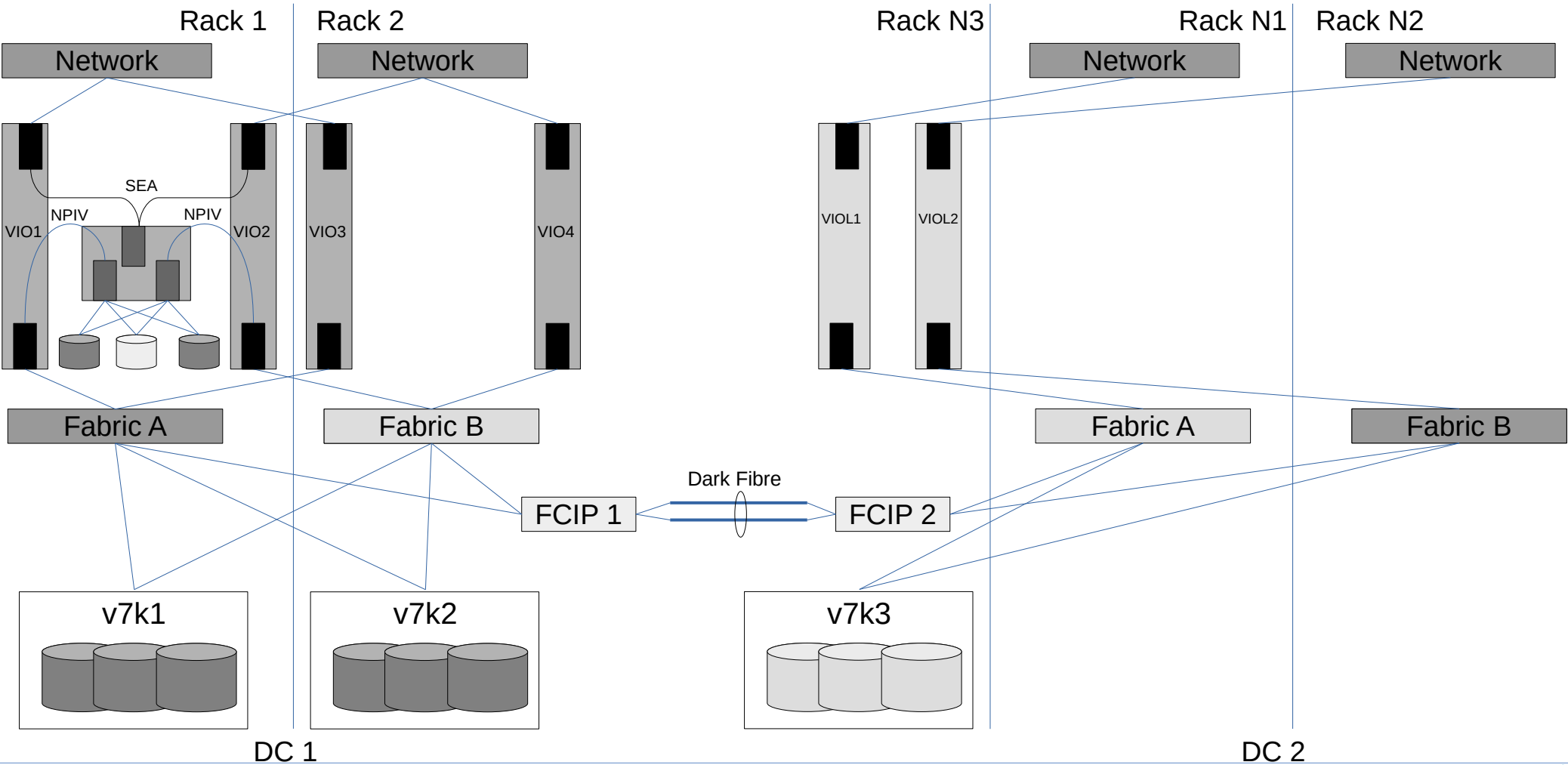
- Confirm all paths, mappings and performance

Adding redundancy for virtual machines

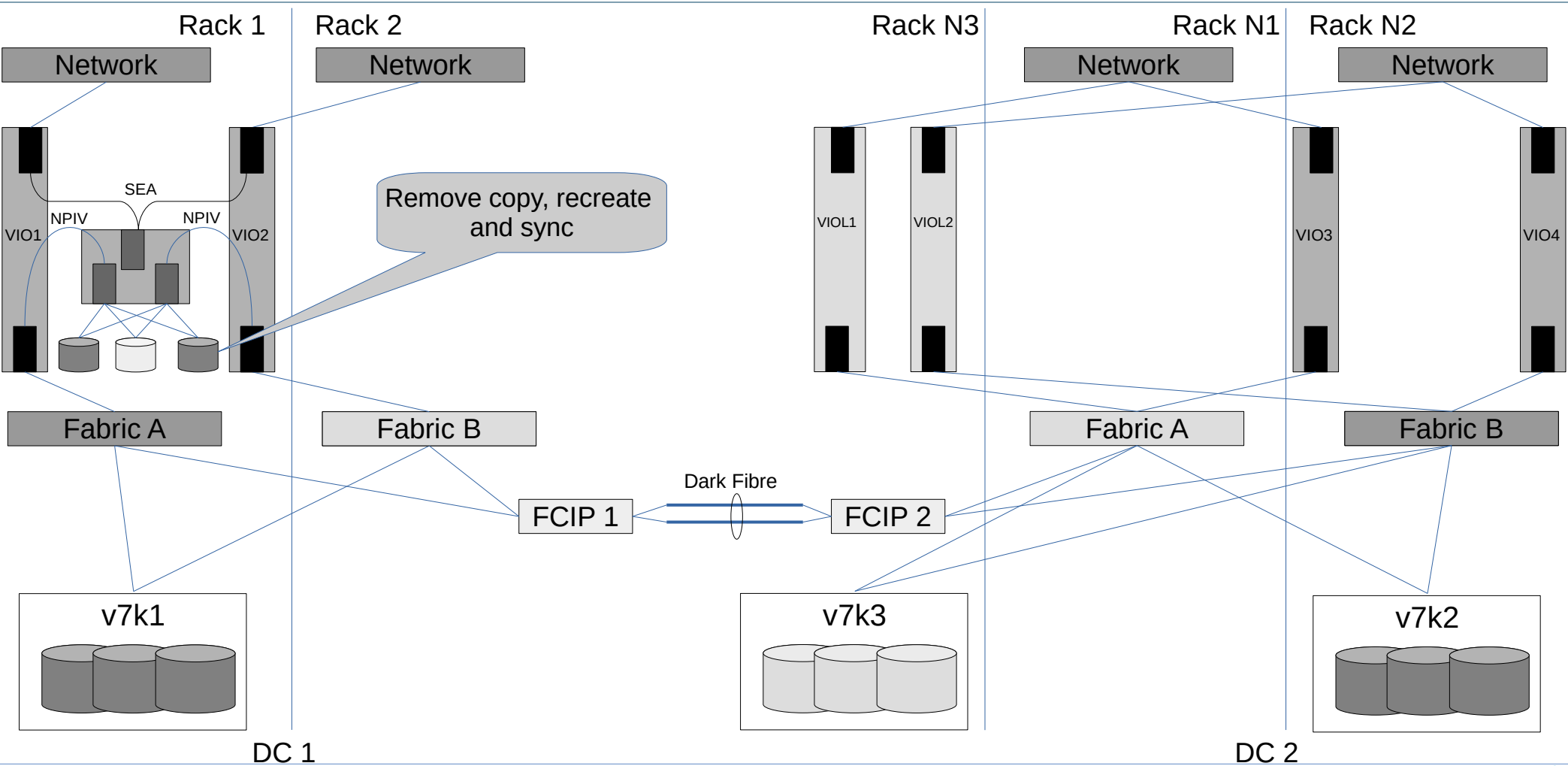


- After establishing a baseline for LPM (measure from Rack 1 to Rack 2 in DC 1 with simulated I/O and memory operations), test LPM from DC1 to DC2 and back again
- Test scripts for removing mirrored LUNs from specific storage units and re-adding / resync'ing. Confirm timings from different Host Servers / sites
- Moving test LPAR between Rack 1, Rack 2 and the loan rack also confirmed zoning / mapping and physical paths.

Remove workload from 2nd rack

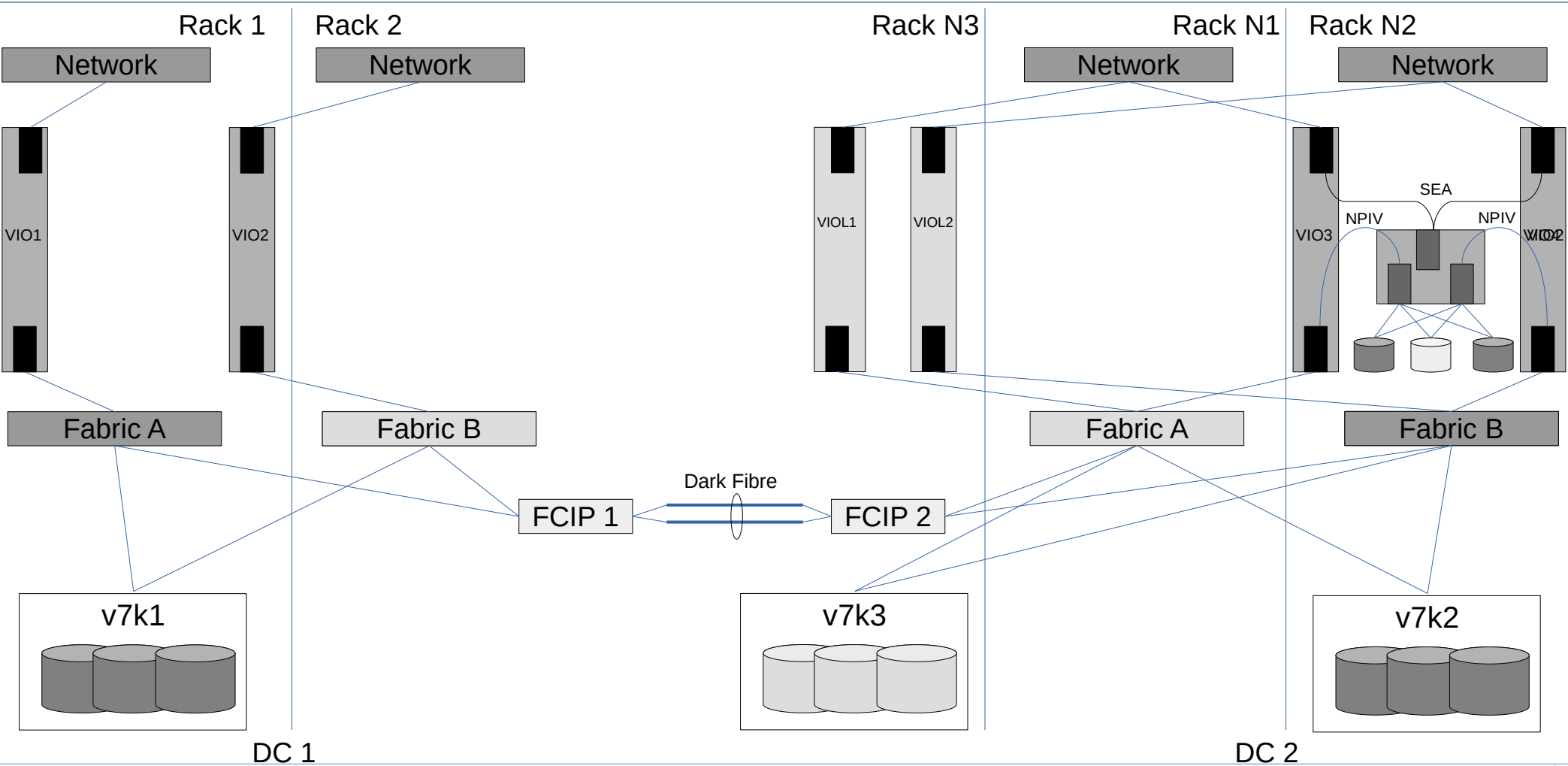


Move Rack 2

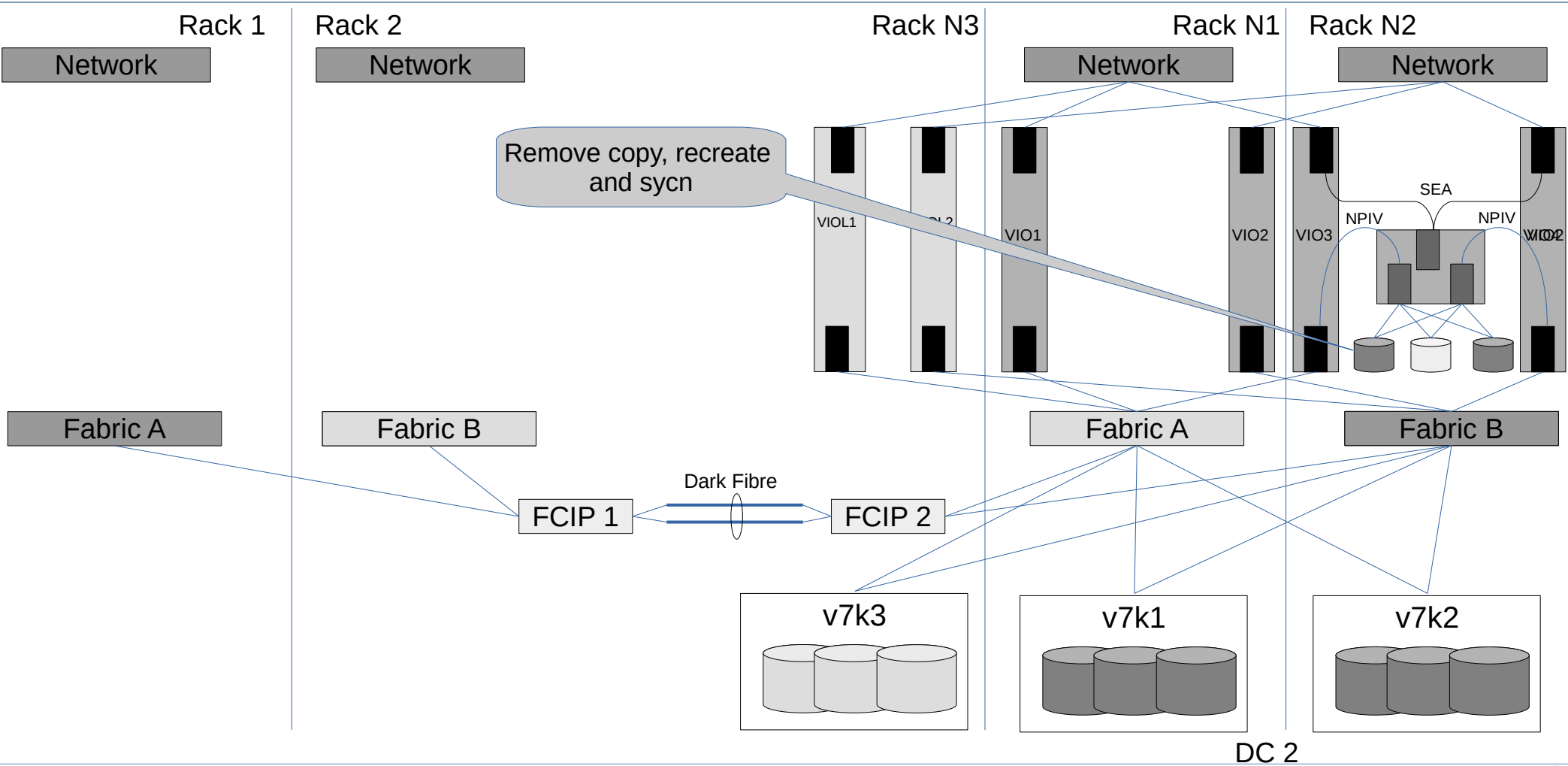


- Confirm all paths
- Run hardware diagnostics
- Check all error logs
- Confirm test LPAR LPM to Rack 2 in DC2

Move virtual machines

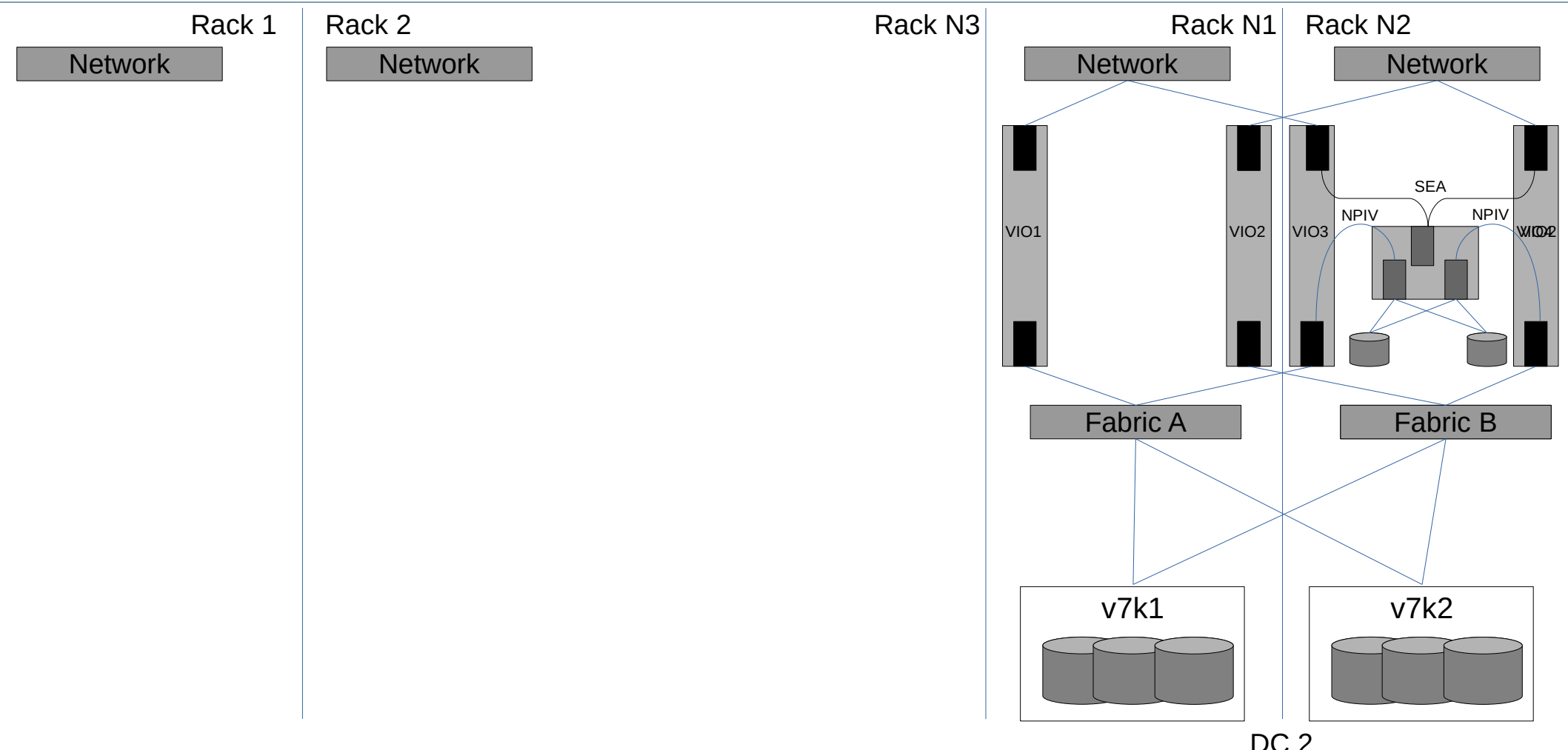


Move rack 1



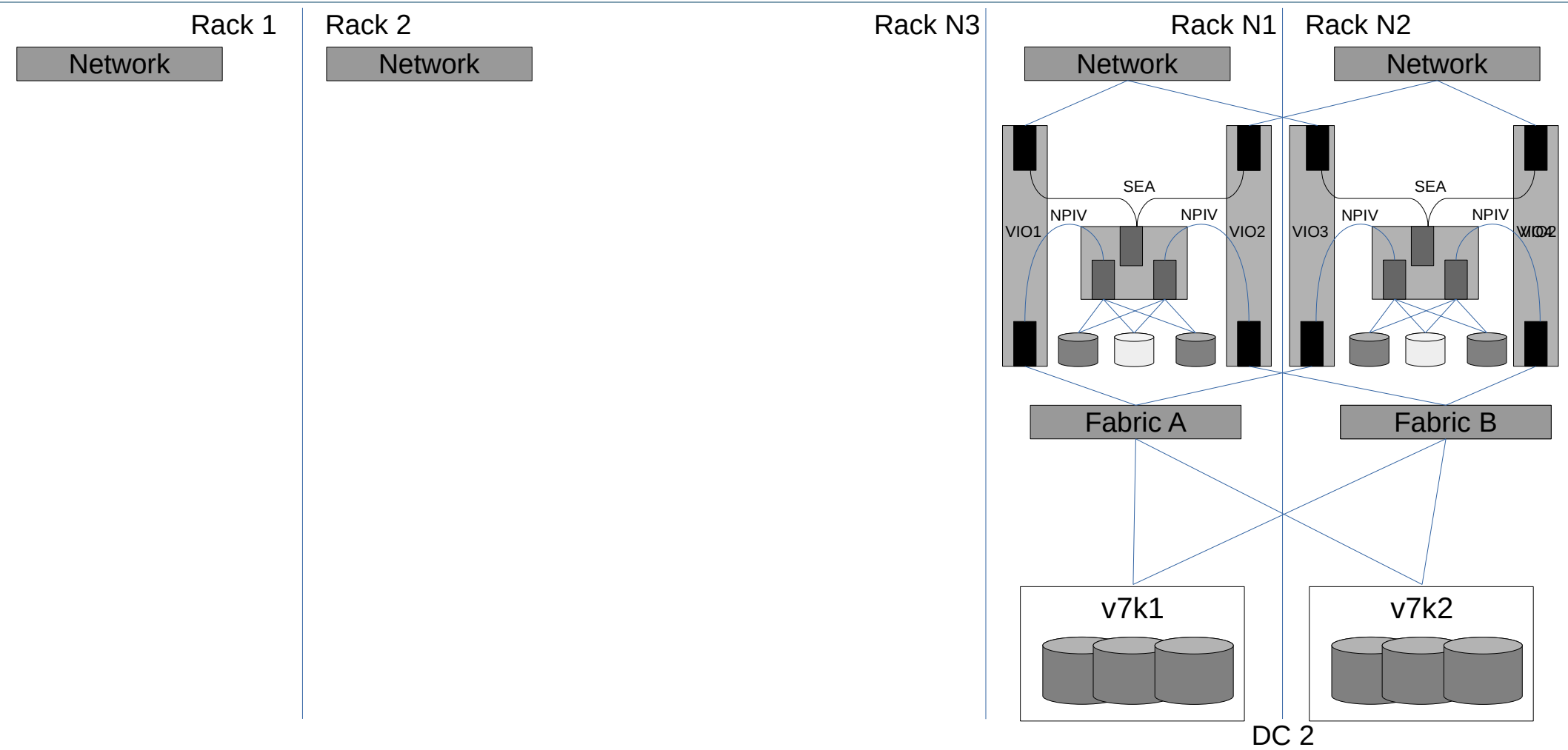
- Confirm all paths / LUNs available
- Confirm all Logical volumes syncd and available
- Confirm baseline performance and LPM between 2 new racks

Swap switches (Fabric B stable) and remove loan equipment



- Confirm baseline I/O performance on test LPAR with 2 way mirror matches pre-migration performance
- Confirm LPM performance between the two new racks
- Remove all customer configuration from the loan equipment
 - Scrub all the disks on the Loan storage (learned lessons about performance of *urandom* and *yes*)
 - Remove all configurations from the loan switches and factory reset
 - Remove all configurations from the FCIP routers and factory reset

Balance workload in DC2



- Customer happy (Power was the only architecture that was migrated without an outage – or impact on performance)
- Lessons learned
 - While technical team is comfortable with the technology and plans, not all of the customers teams are that familiar with the features we took advantage off – communicate!
 - Control communications – While it is important for everyone to be kept up to date, don't swamp with too much information
 - You can never test to much
 - You can never plan too much (and have a good imagination when dreaming up potential issues – access as DC is still a worksite?)

- Please feel free to contact me for more specific details about anything mentioned above.

Session: a111037

Data Centre migrations (101)

¿ Questions ?

Thanks!

Your feedback about this session is very important to us.

Please remember to submit a survey

For further information....

Contact:

Antony (Red) Steel

antony.steel@belisama.com.sg

+65 9789 6663

