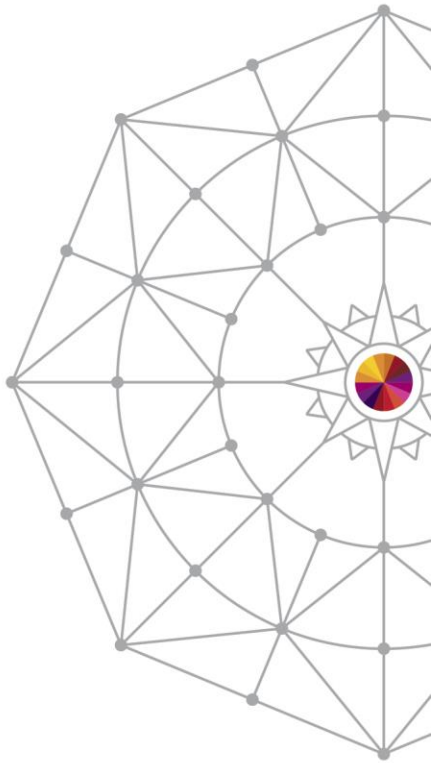
Disk Library for Mainframe User Experience in a High Availability Environment

Jim Keller - USAA
Brian Kithcart – EMC
Keith Murray - EMC

August 4, 2014
Session # 15770



#SHAREorg

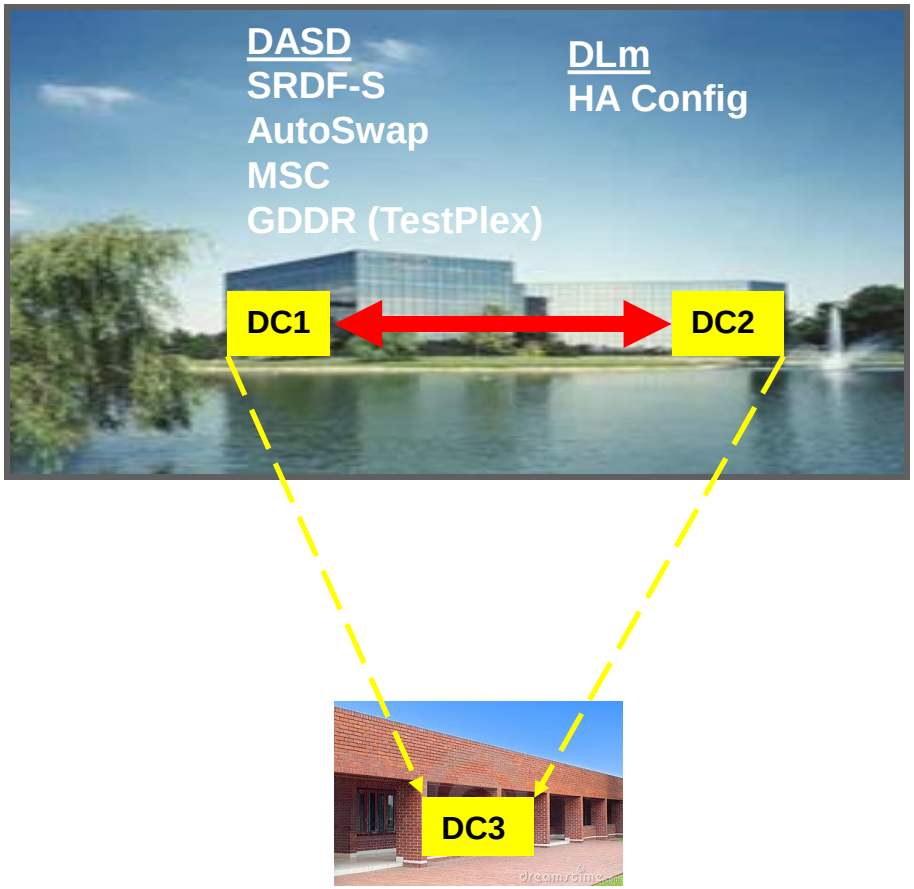
SHARE in Pittsburgh 2014

The Company

- Founded 1922
- Mission:
  - USAA's mission is to facilitate the financial security of its members, associates, and their families through provision of a full range of highly competitive financial products and services; in so doing, USAA seeks to be the provider of choice for the military community.
- Product Offerings
  - Insurance, Banking, Investing Real Estate, Retirement Planning, Health Insurance, Shopping & Discounts

USAA – 3 Site Replication

Campus Environment - HA



DR Site

USAA – 2008 Mainframe Tape Goals

- Migrate toward a “Tapeless” Infrastructure while Improving Service Levels
  - Reduce/Eliminate Tape Libraries, Tape Drives and Tape Media
  - Reduce Physical Tape Library Footprint (TCO/ROI Benefits)
- **Create a HA Tape Environment**
- Improve Processor Utilization by Eliminating Tape/HSM Delays
  - Reduce overall HSM CPU Consumption
- Reduce Cost of Tape H/W and S/W Maintenance, Environmentals
- Leverage IP Replication with DLm to Reduce bandwidth

Characterization of USAA mainframe tape data

Overall “Tape” Profile of Largest Plex (multiple plexes)

- At least 500 TB’s is under HSM control
  - 300 TB’s BU
  - 200 TB’s ML2
- 300 TB’s associated with backups, including image copies and archive logs
- 600 TB’s of other (Reports, Syncsort, Batch processing, Long Term Retention, etc…)

Tape data is critical!

Physical tape replaced in 2009

- DC1
  - 5 Tape libraries with 4 VTL's and 70 3590 type drives
  - 20 3490E type stand alone physical drives
  - 4 3480 type stand alone physical drives
- DC3
  - 2 Tape libraries with 1 VTL and 32 3590 type drives
  - 16 3490E type stand alone physical drives
  - 4 3480 type stand alone physical drives
  - Ficon channel extenders were being used to write to remote tape libraries, eliminating physical offsite vaulting

Initial DLm4080 Experience

- Key objectives for DLm 4080 (replace physical tape)
  - High availability in DC1/DC2 , D/R to DC3
  - Substantial reduction in footprint, maintenance cost, moving to DLm
  - High scalability and redundancy
  - No FICON channel extenders

- Key metrics
  - Nightly batch window reduced by 40%
  - 99.9% of all tape mounts fulfilled in less than 1 second
  - Over 4 to 1 compression
  - ~ 1786 TB's in each DLm4080 DC1 & DC2, ~ 798 TB's in 3rd site DLm4080
  - Each DLm4080 capable of 600 MB/sec

If life was so good, why move to the DLm8100?

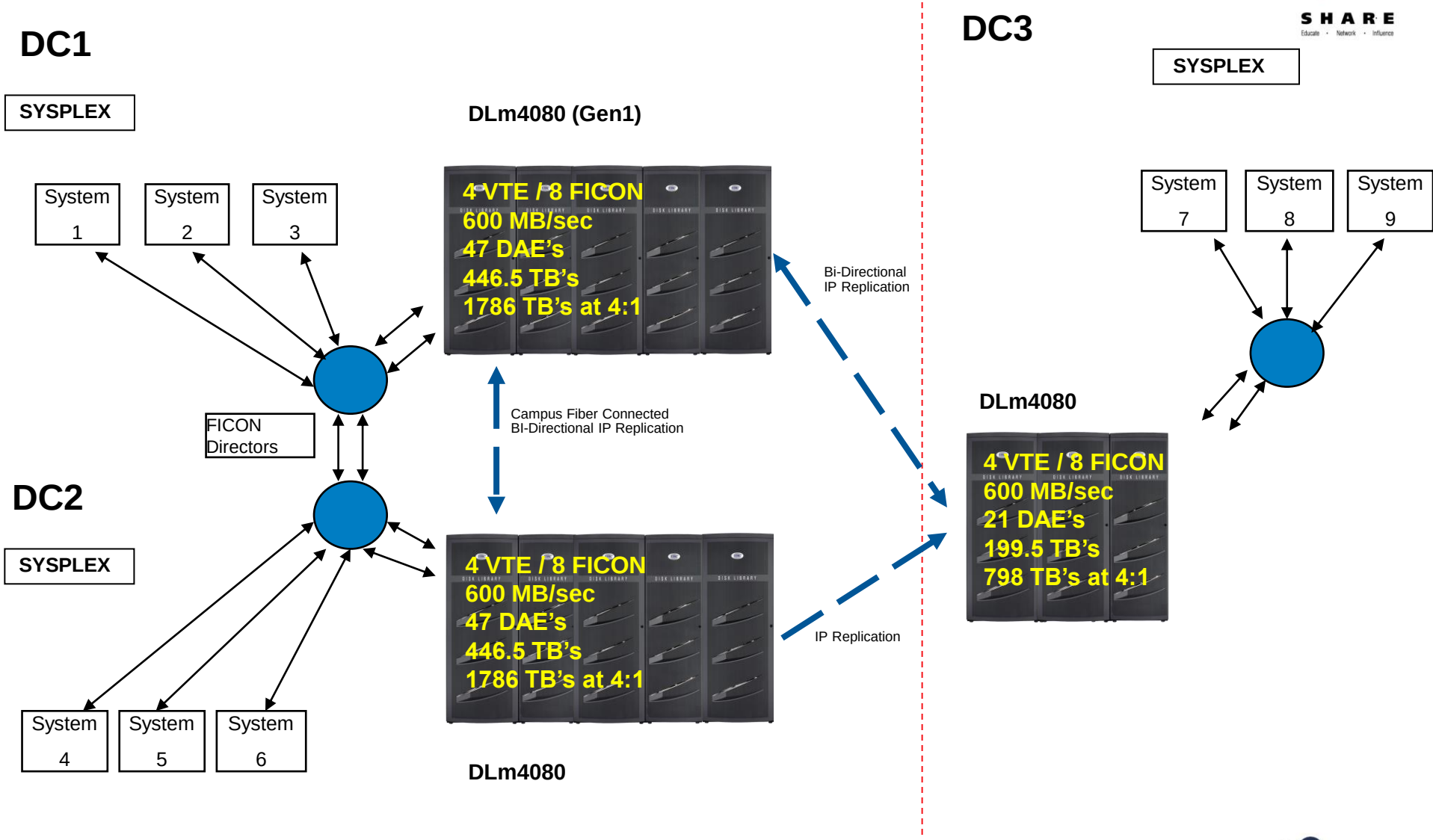
Benefits of DLm8100 Migration

- Significantly more data ingest realized with 8Gb Ficon
  - New DLm's can scale to 4.8 GB/sec Frontend FICON
    - 8x DLm4080's 600 MB/sec
  - Up to 8 VTE's per DLm
- A single DLm8100 can scale to 2560 TB's
  - 14.592 PB's at 5.7:1 compression – lot's of room for growth!

Conversion from DLm4080 to DLm8100

- 1.6 PB and 530,000 tape volsers
  - Accomplished in 6 weeks!
- Pre-conversion used DLm capacity was 313 TB
  - ~4.1:1 compression
- Post conversion used DLm capacity is 200 TB
  - ~5.7:1 compression
- 36% compression rate increase over the DLm4080
- Nightly batch window reduced by 30%
- I/O intensive processing (Full Volume Dumps) reduced by 40%

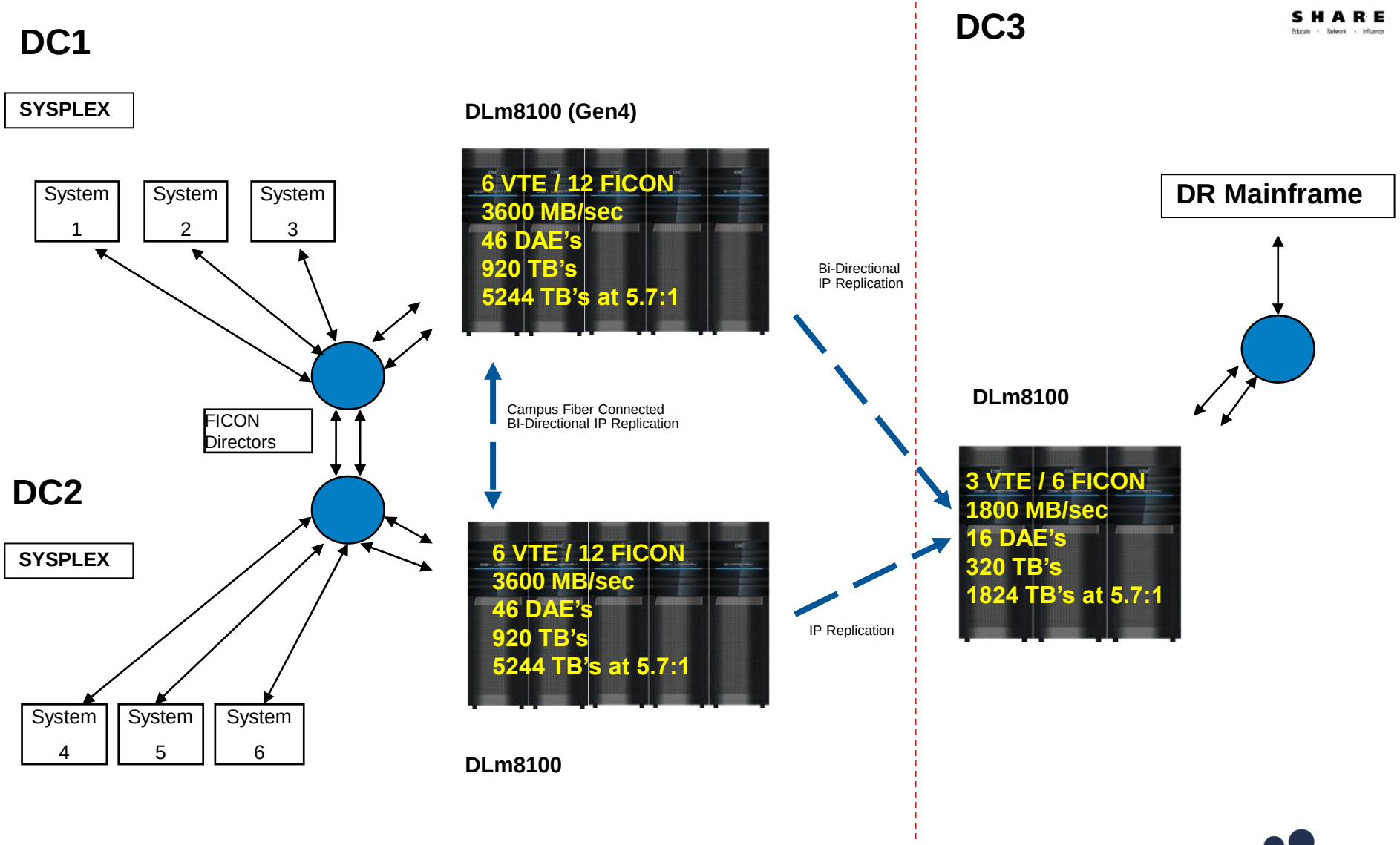
High Availability Tape & DASD - 2008



Initial DLm 4080 Implementation

- Site 1 & 2 are ~ 1 mile apart
- Site 3 is ~ 300 miles

High Availability Tape & DASD - 2014



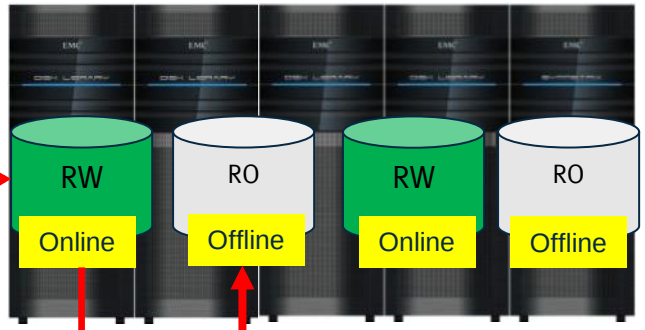
Today

- Site 1 & 2 are ~ 1 mile apart
- Site 3 is ~ 300 miles

HA SMS Configuration – STG1 & STG2

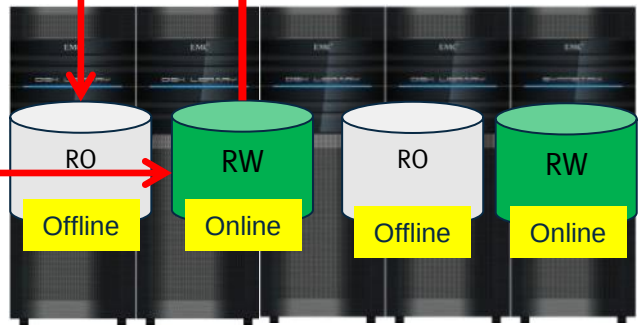
DC1

MTL1/STG1
Writes from host
to RW volser's



MTL1
STG1

MTL2
STG2



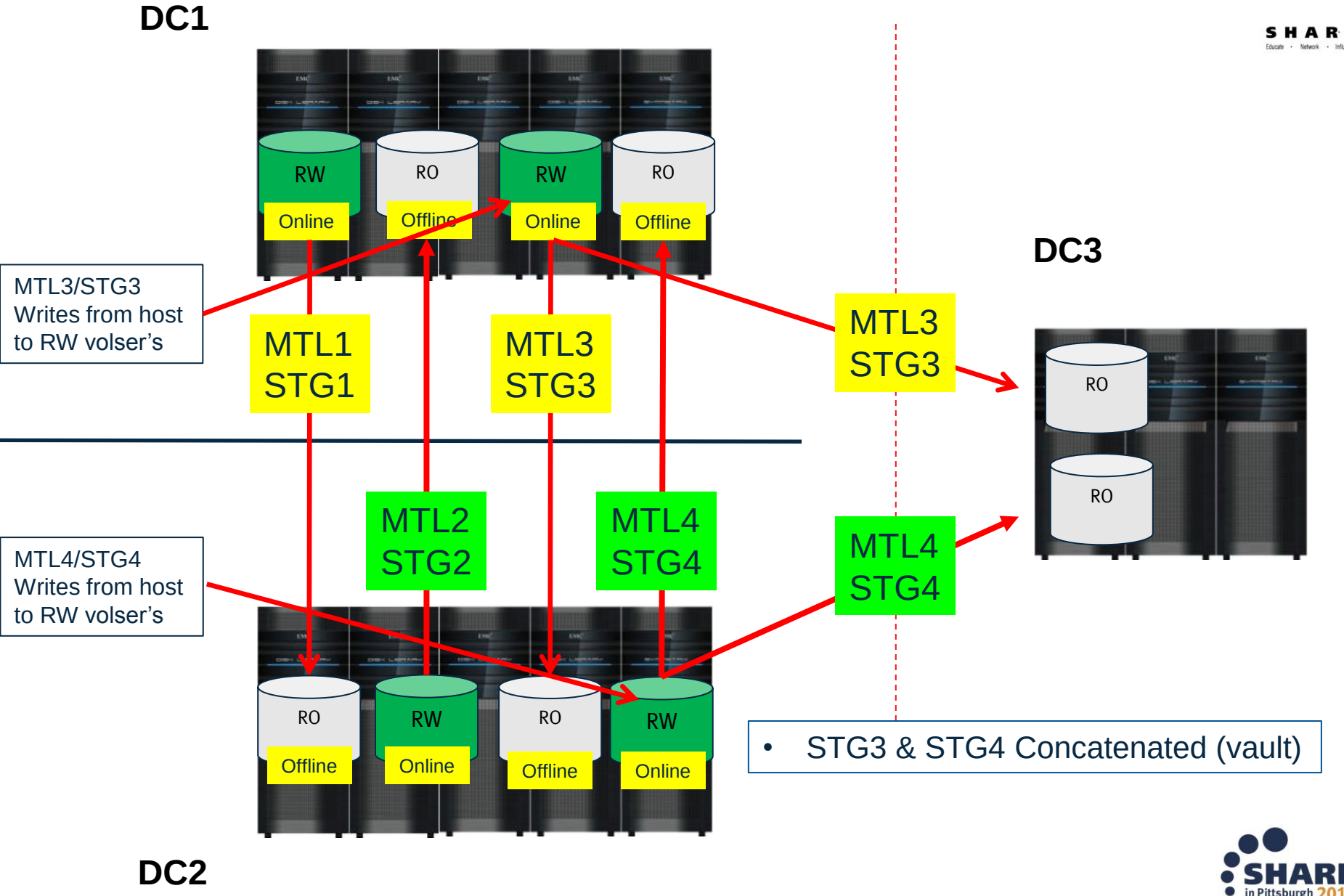
MTL2/STG2
Writes from host
to RW volser's

DC2

DC3

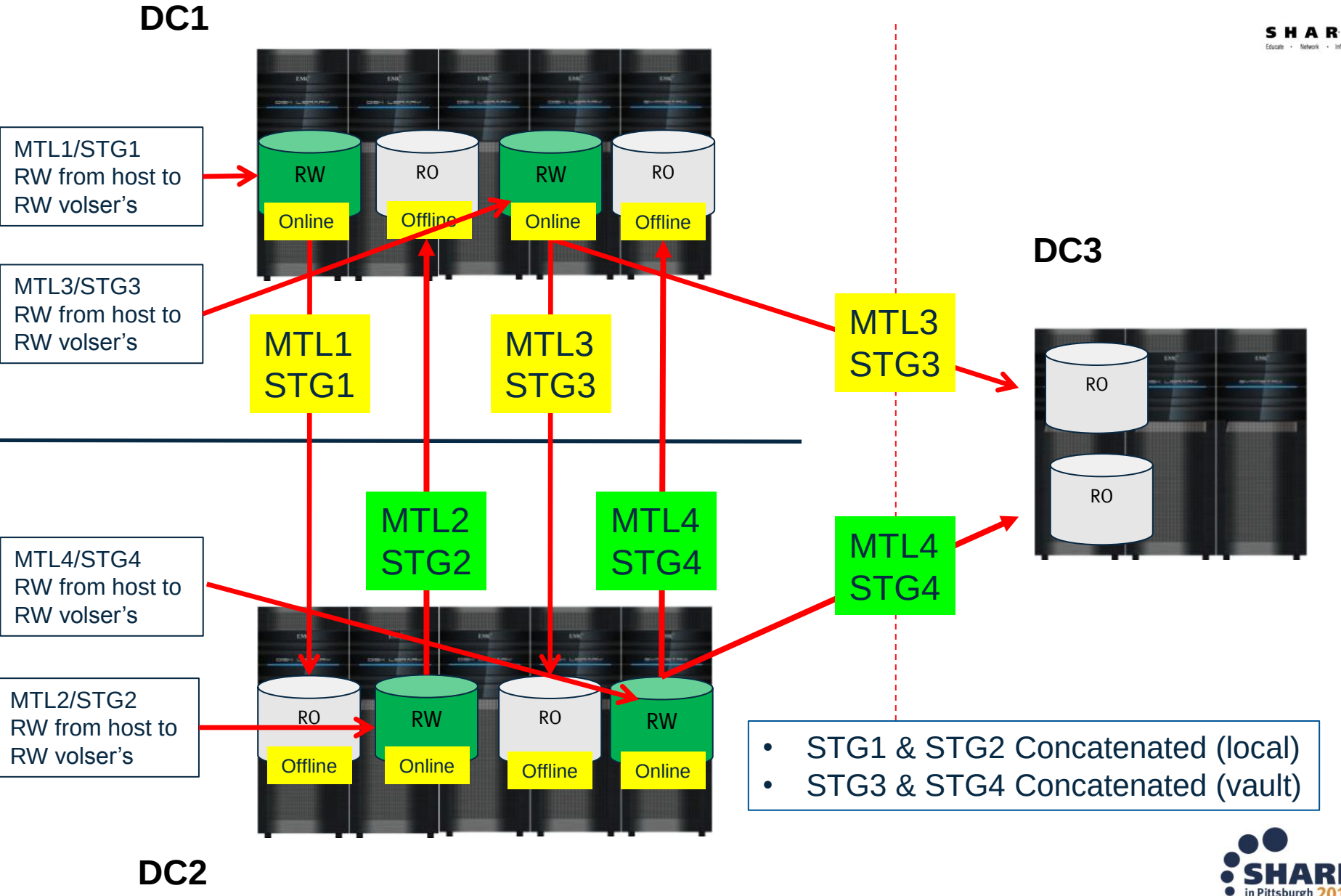
- STG1 & STG2 Concatenated (local)

HA SMS Configuration – STG3 & STG4



- STG3 & STG4 Concatenated (vault)

HA SMS Configuration – 4 Storage Groups



HA ACS Routine – Concatenating Storage groups

```
/*-------------------------------------------------------------------*/
/* FOR DLM                                                           */
/*-------------------------------------------------------------------*/
  WHEN (&UNIT = 'DLMLCL')
      DO
        SET &STORGRP = 'STG1, STG2'
        EXIT
        END
/*
  WHEN (&UNIT = 'DLMVLT')
      DO
        SET &STORGRP = ‘STG3,STG4’
        EXIT
        END
```

For “Normal BAU” processing, STG1 & STG2 get concatenated for “Local” work and STG3 & STG4 for “Vault” work.

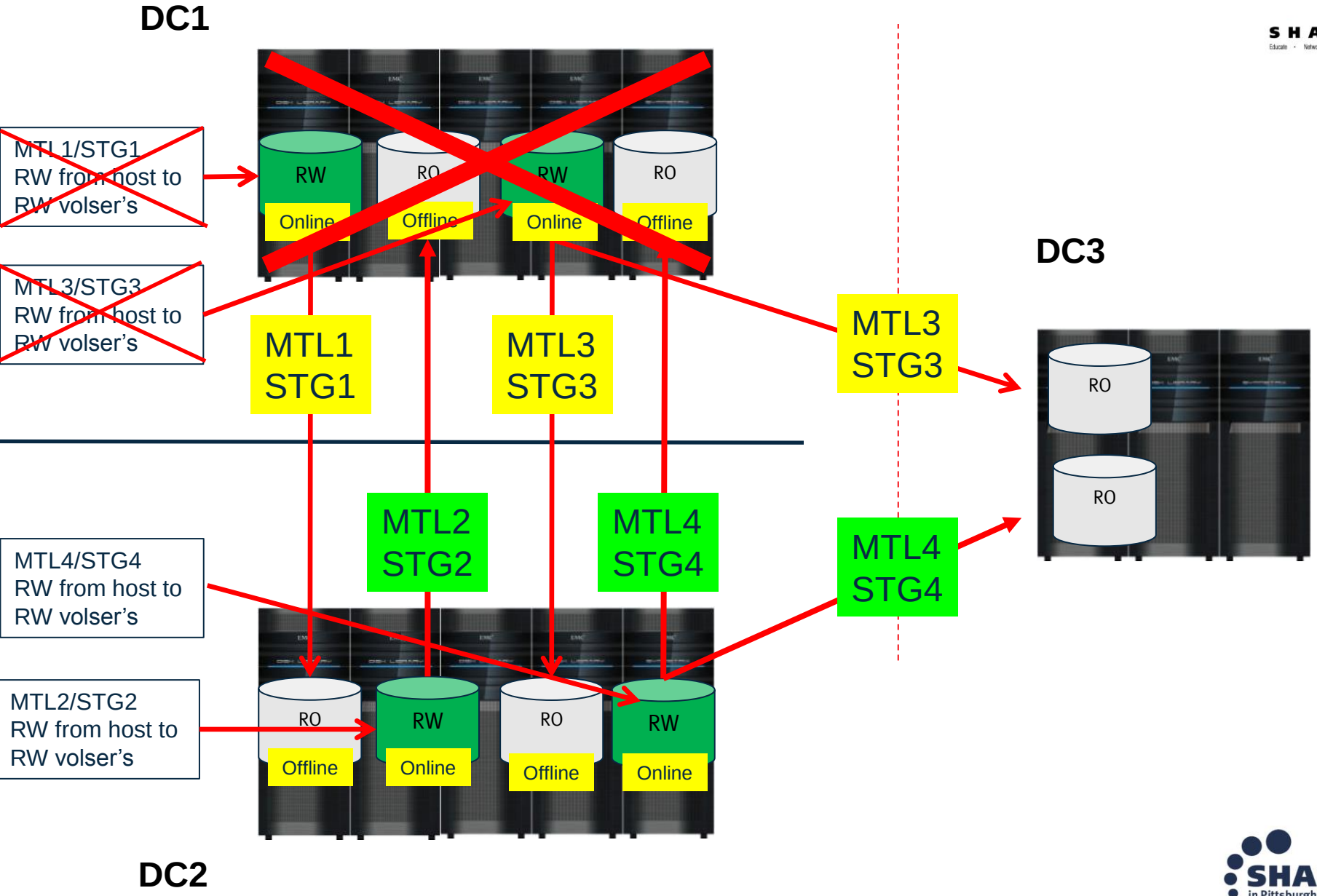
STG1 MTL has Tape drives defined ONLINE to DLm DC1 and OFFLINE in DLm DC2. Likewise, STG2 MTL has tape drives defined ONLINE in DC2 and OFFLINE in DC1.

A similar setup exists for the Vault work for STG3 and STG4.

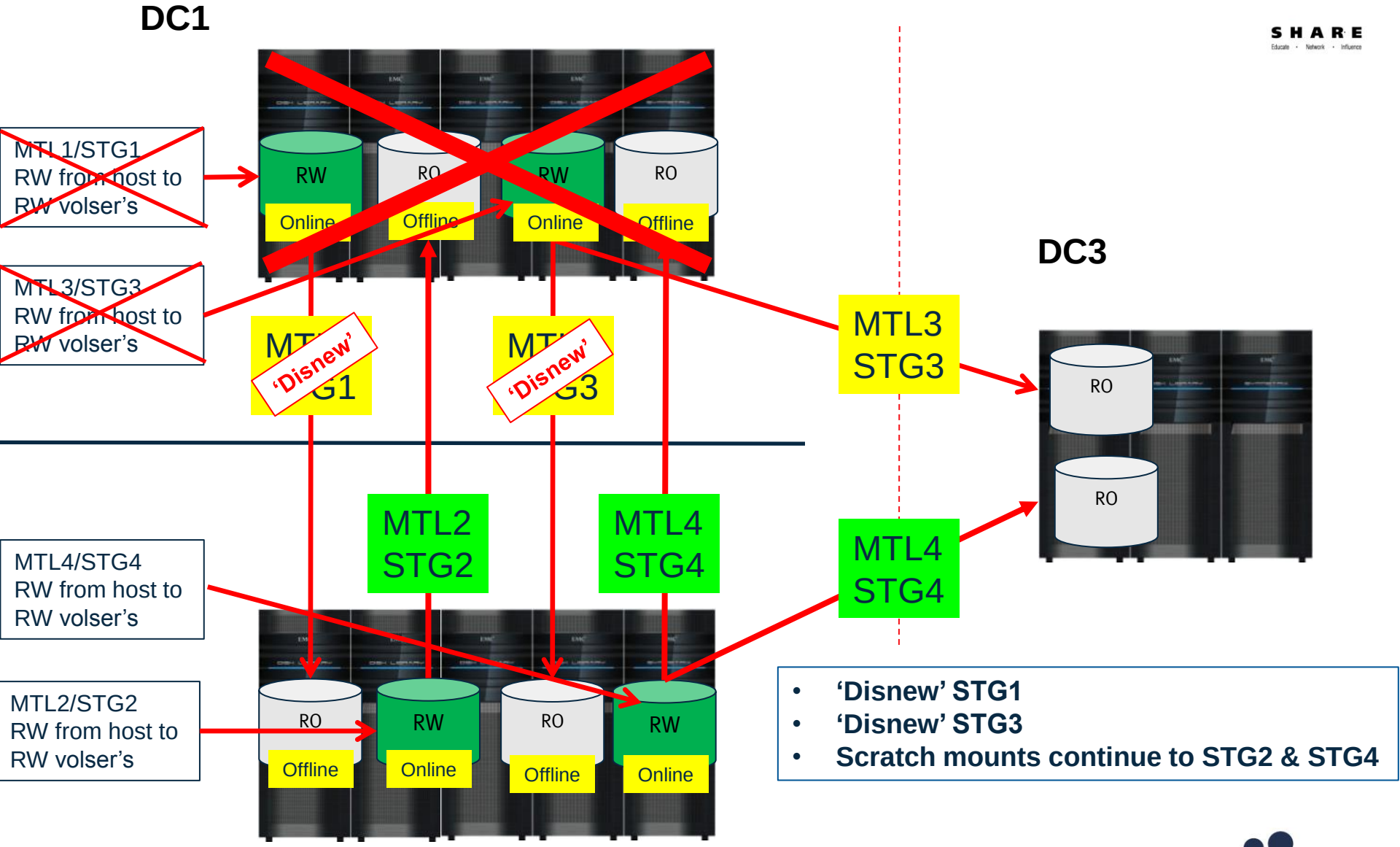
Site 'Unavailable' Scenarios

- Planned or Unplanned Outage (one DLm unavailable)
  - Short Duration (Read only)
    - Implement Local Failover process
  - Long Duration (Read/Write)
    - Load local Failover configuration
    - This will invoke Target DLm8100 as Read/Write
    - Stop Replication
    - Plan for Resynchronization - location may not be known
- Total Local Site Failure
  - What happens at DC3?
    - Failover Replication
    - This will make Target DLm8100 Read/Write
    - Load Disaster Recovery configuration
    - Build out configuration
    - Plan for Resynchronization to alternate remote site

Planned/Unplanned Outage in DC1 – Short Duration

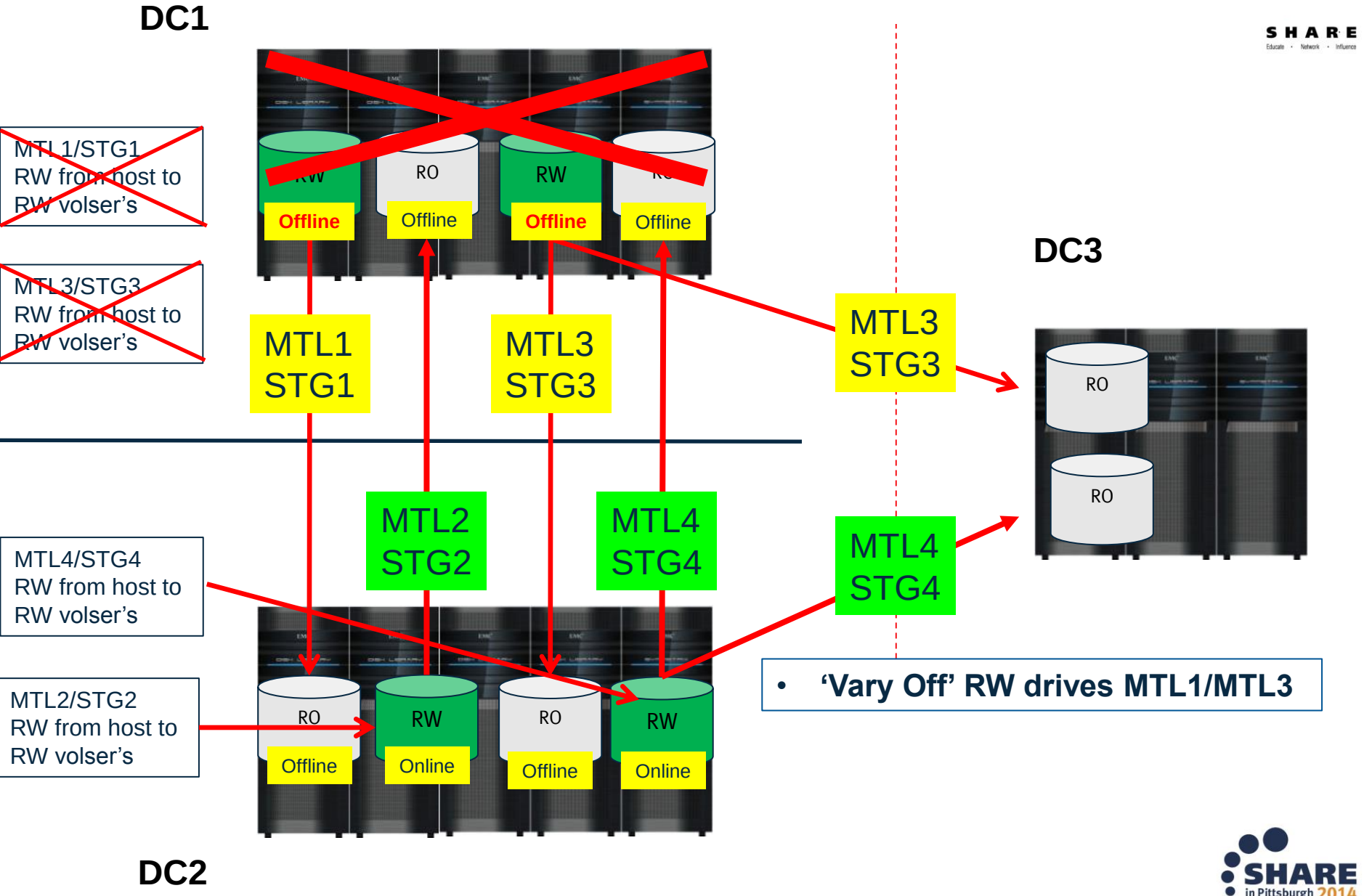


Planned Outage in DC1 – Short Duration

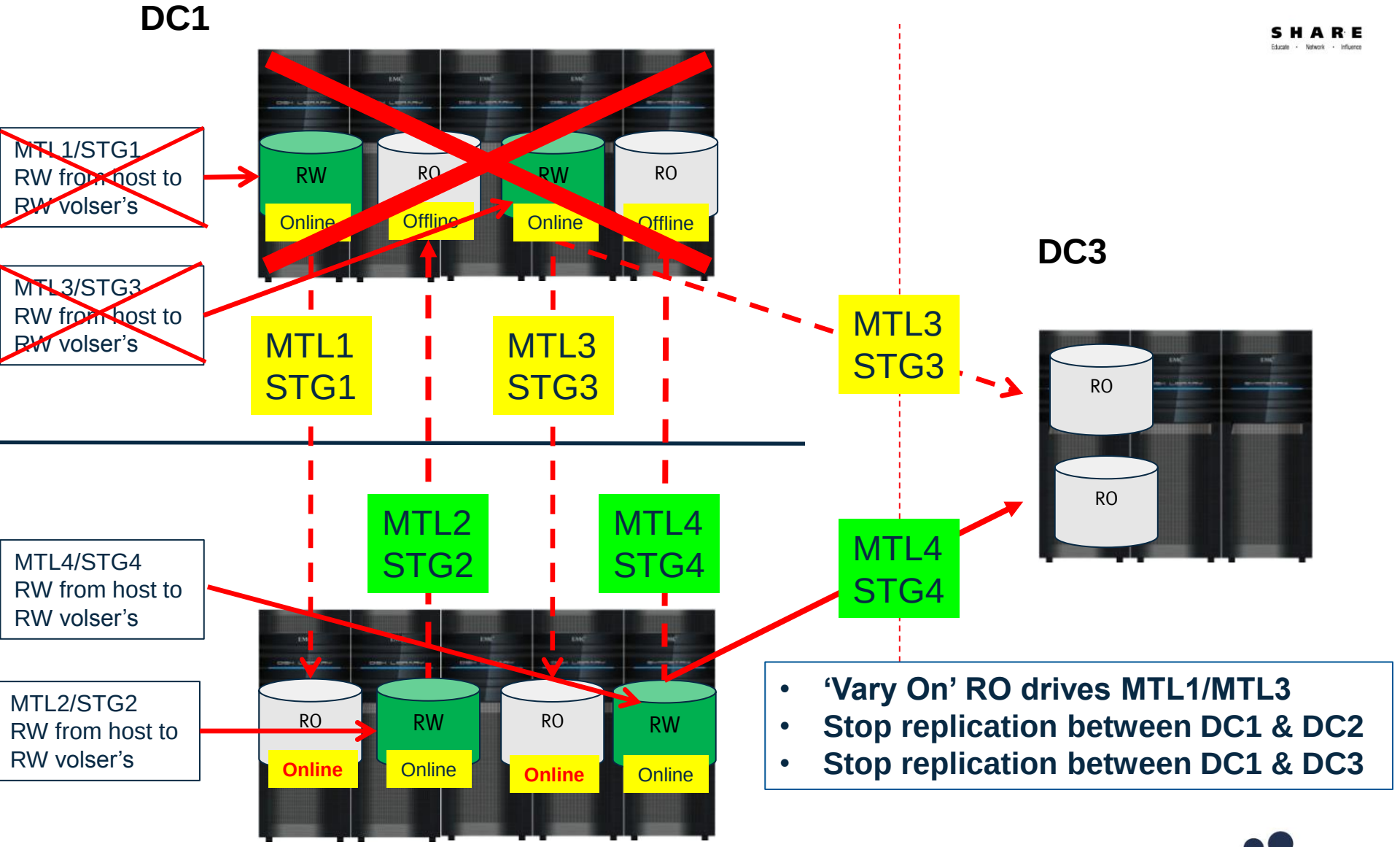


- 'Disnew' STG1
- 'Disnew' STG3
- Scratch mounts continue to STG2 & STG4

Planned Outage in DC1 – Short Duration

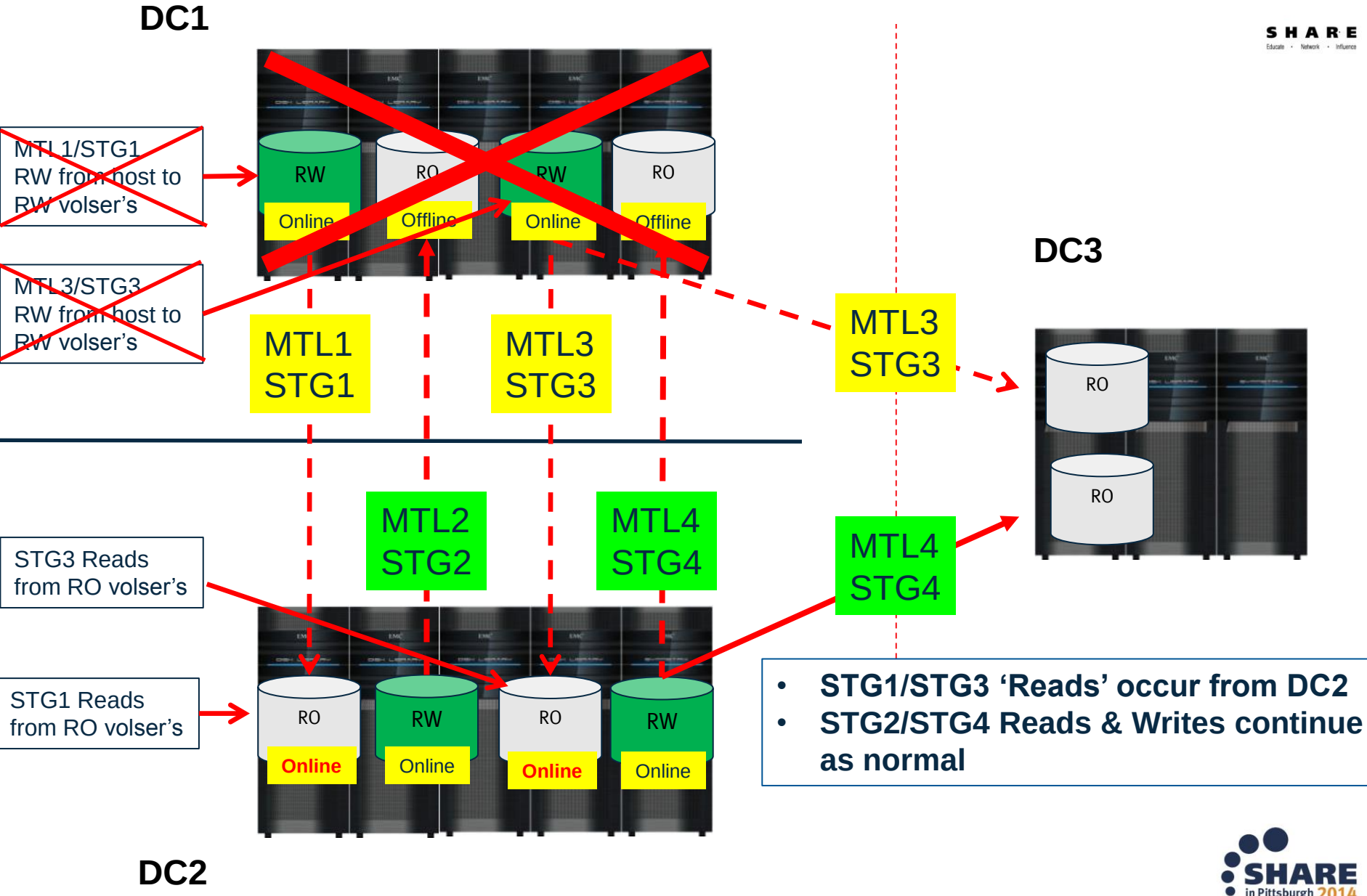


Planned Outage in DC1 – Short Duration



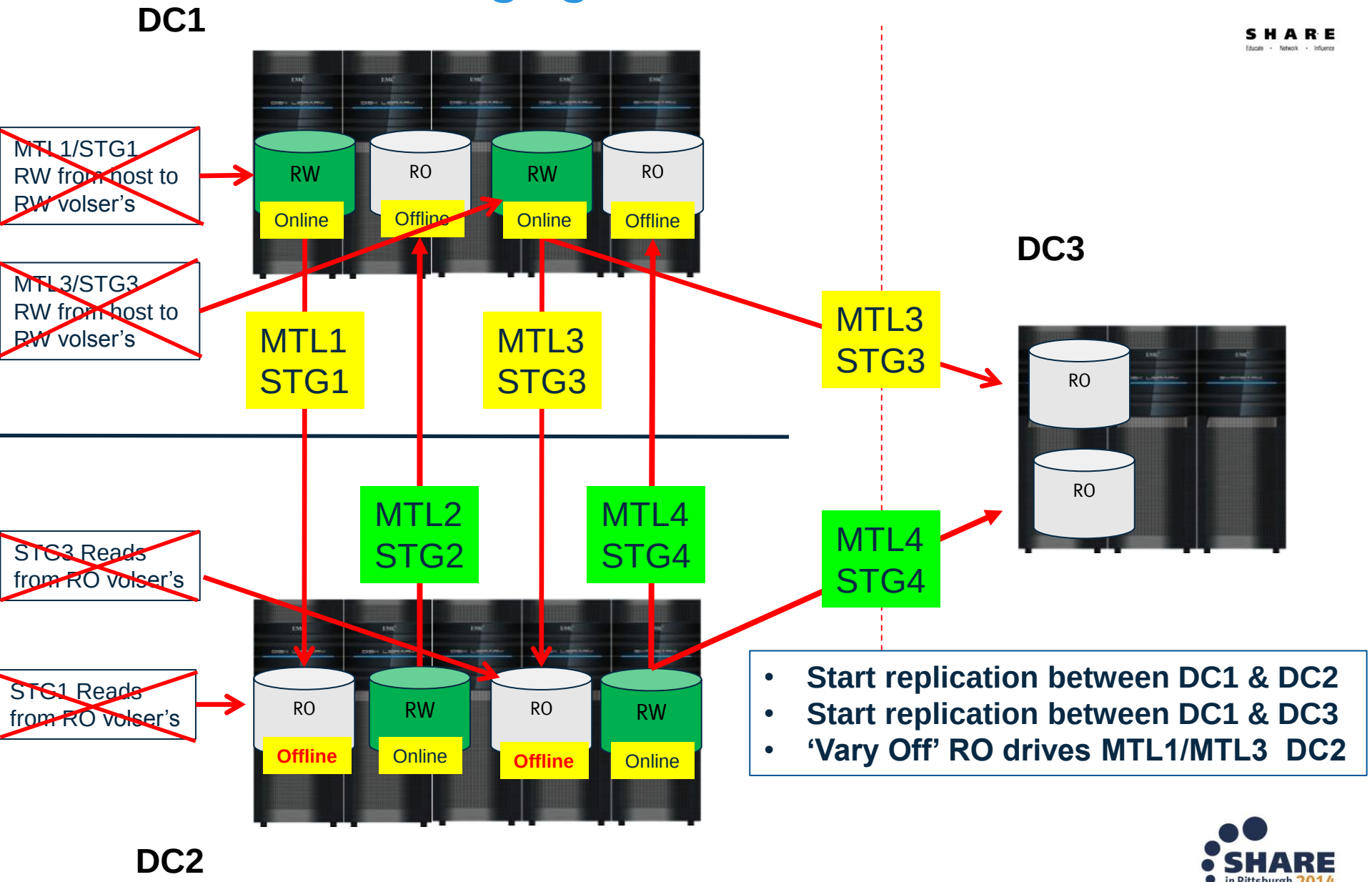
- 'Vary On' RO drives MTL1/MTL3
- Stop replication between DC1 & DC2
- Stop replication between DC1 & DC3

Planned Outage in DC1 – Short Duration



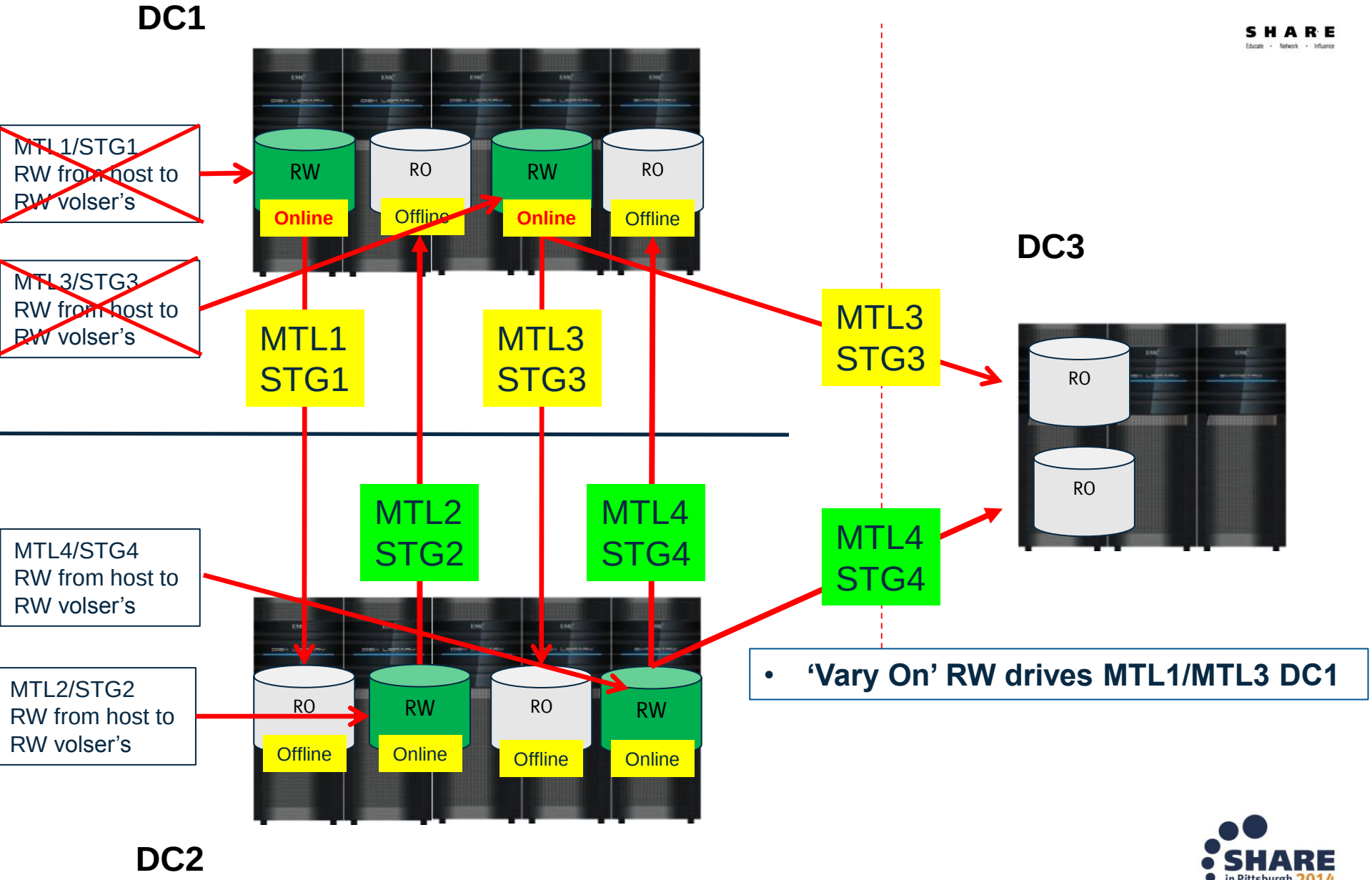
- **STG1/STG3 'Reads' occur from DC2**
- **STG2/STG4 Reads & Writes continue as normal**

Bringing DC1 Back Online



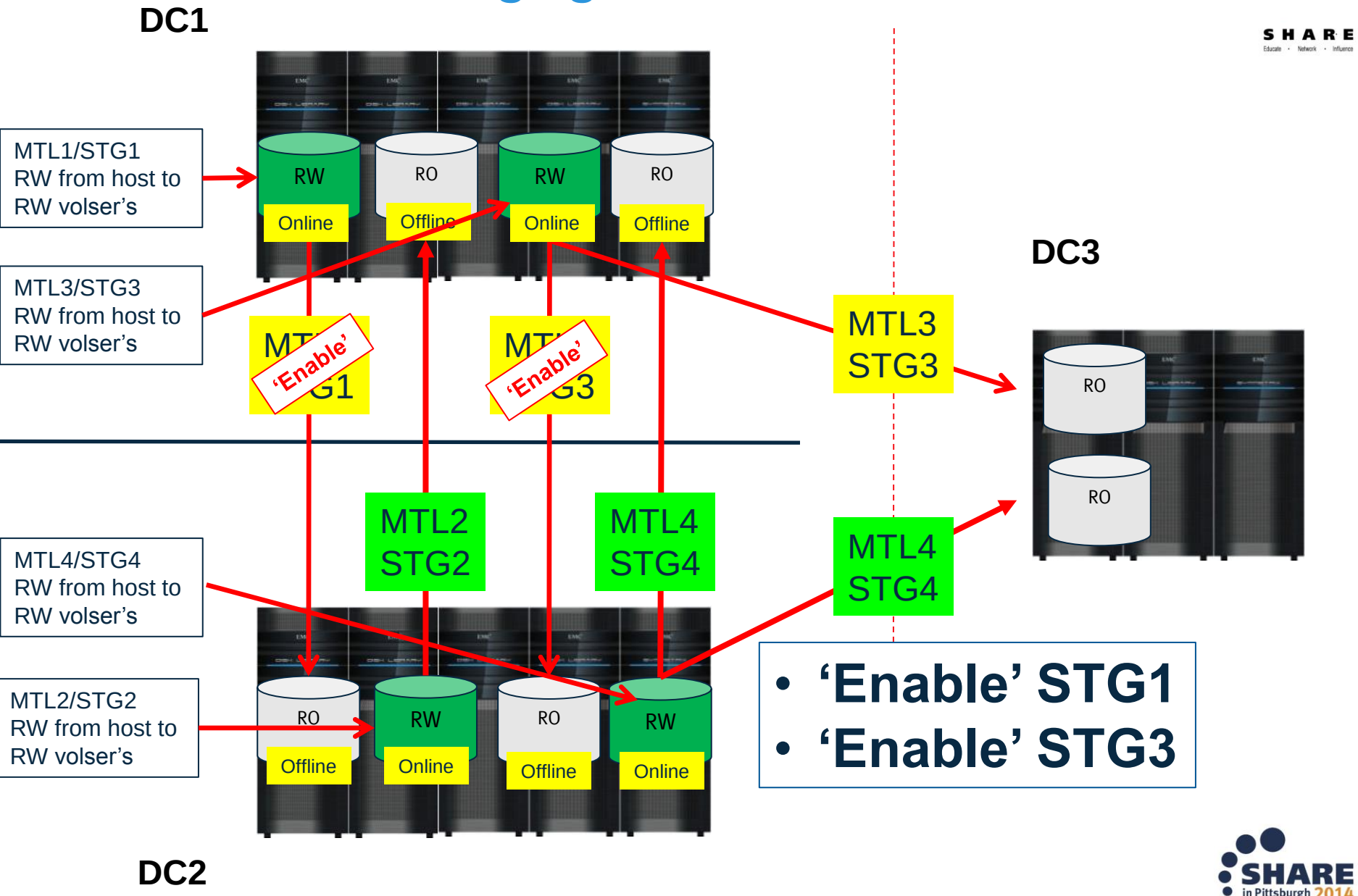
- Start replication between DC1 & DC2
- Start replication between DC1 & DC3
- 'Vary Off' RO drives MTL1/MTL3 DC2

Bringing DC1 Back Online



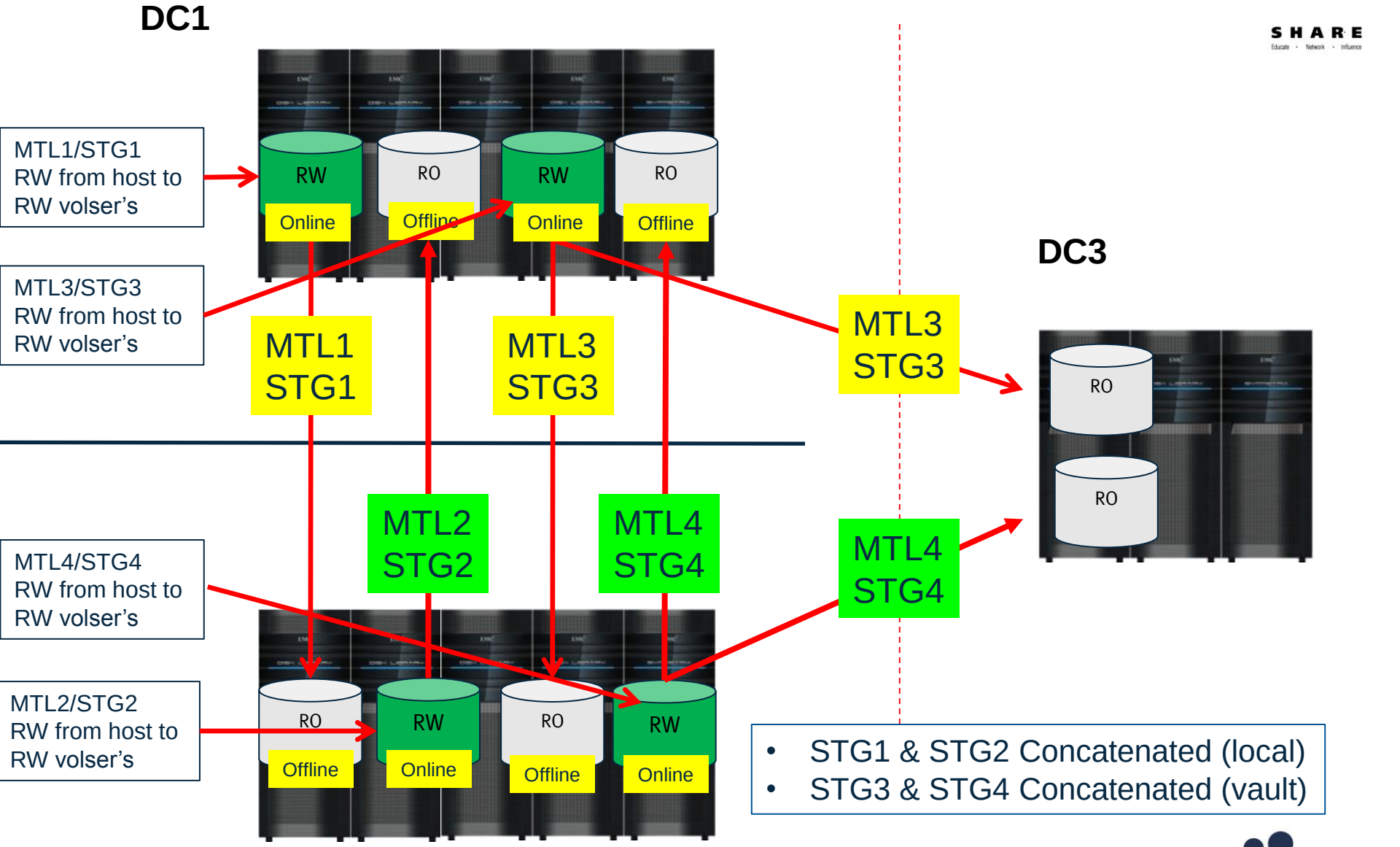
- 'Vary On' RW drives MTL1/MTL3 DC1

Bringing DC1 Back Online



- 'Enable' STG1
- 'Enable' STG3

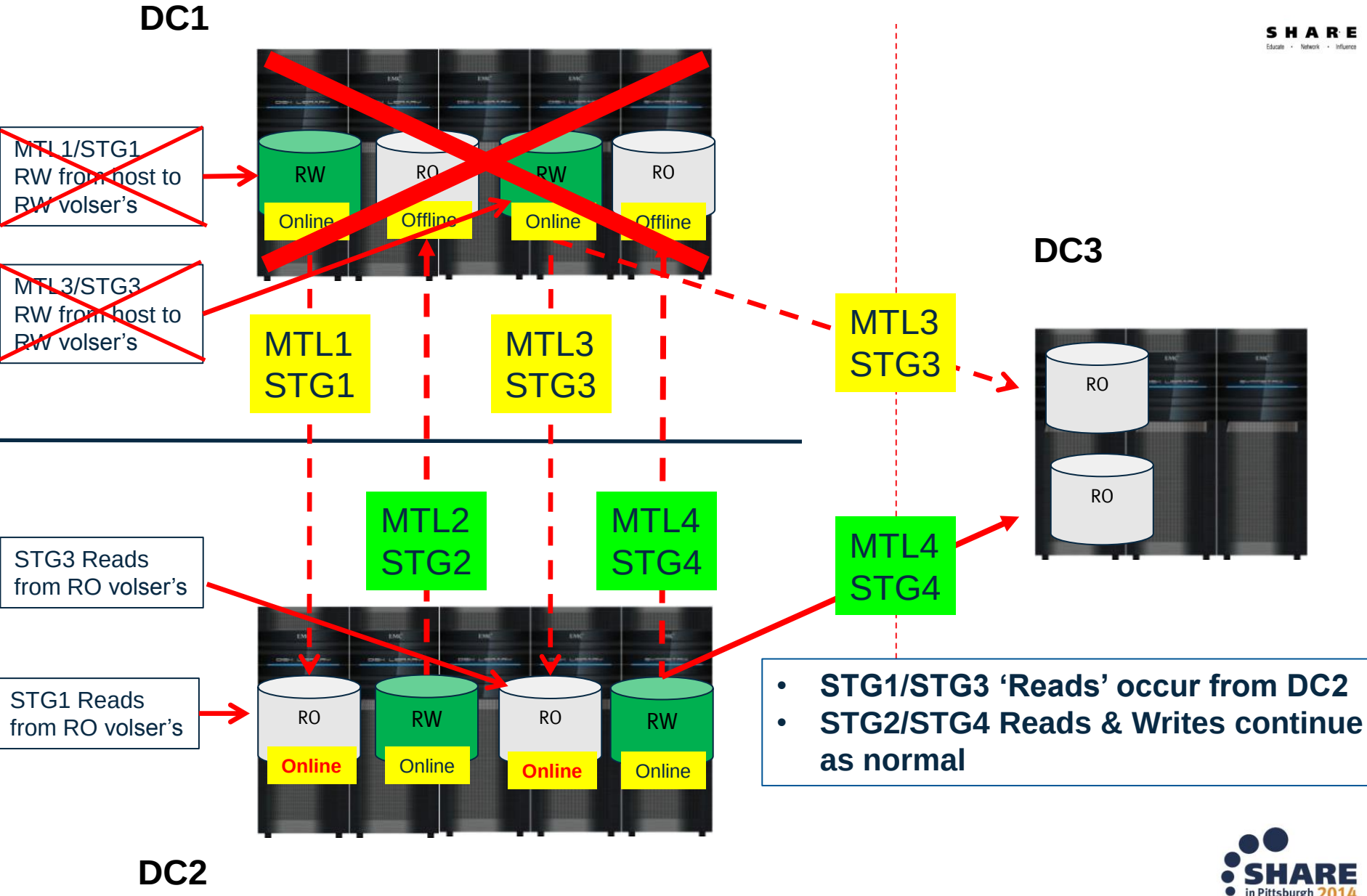
DC1 – Back Online



- STG1 & STG2 Concatenated (local)
- STG3 & STG4 Concatenated (vault)

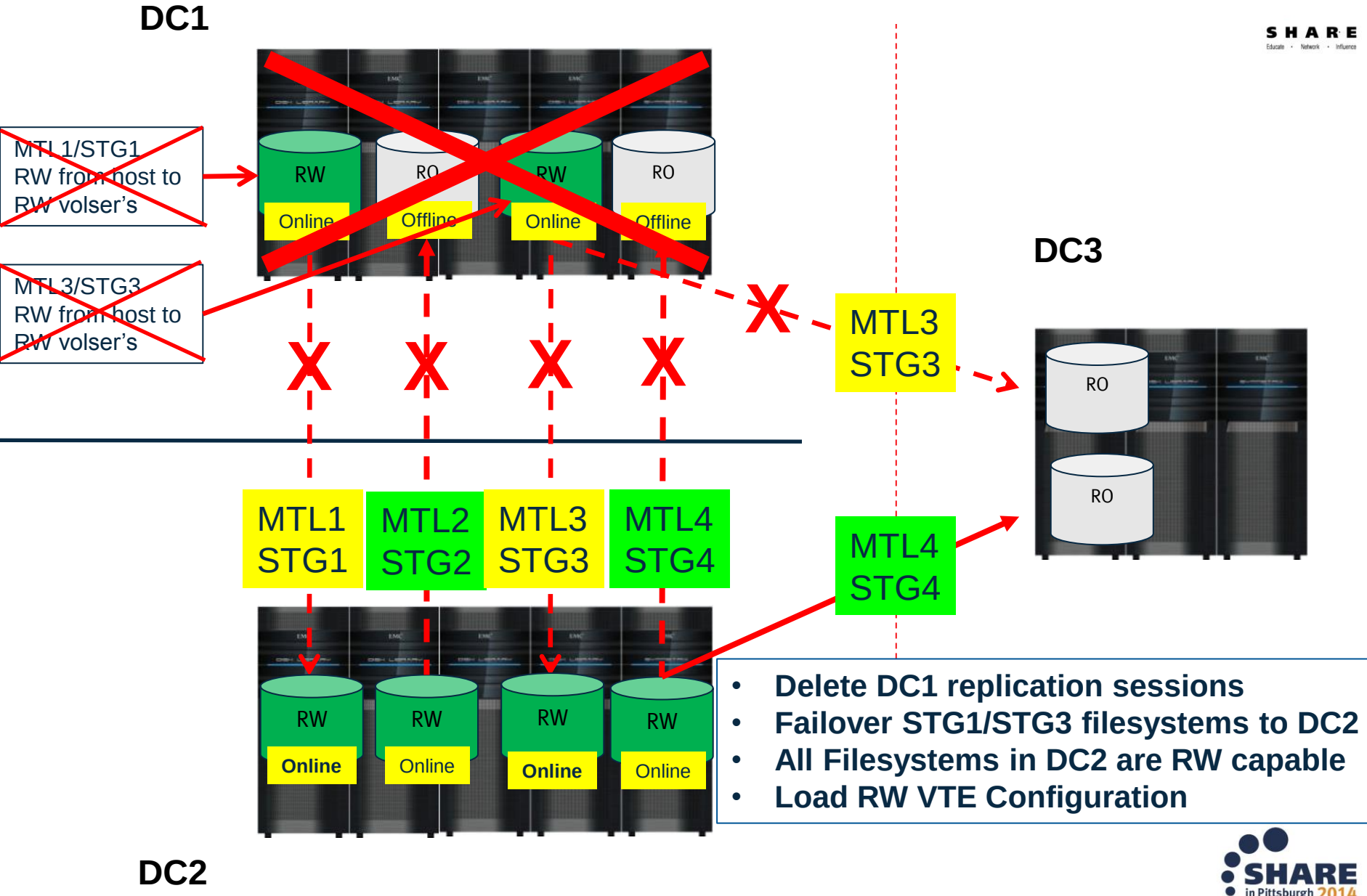
What if it is now going to be Long Duration?

Convert Current Short Duration Config to Long Duration Config



- **STG1/STG3 'Reads' occur from DC2**
- **STG2/STG4 Reads & Writes continue as normal**

Convert DC1 STG1 & STG3 Filesystems to RW



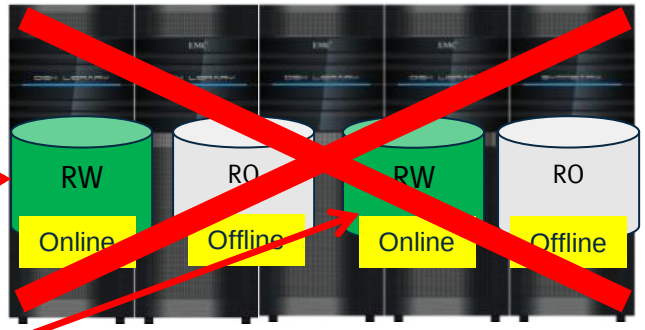
- Delete DC1 replication sessions
- Failover STG1/STG3 filesystems to DC2
- All Filesystems in DC2 are RW capable
- Load RW VTE Configuration

SMS Enable STG1 & STG3 for Host Writes

DC1

MTL 1/STG1 RW from host to RW volser's

MTL 3/STG3 RW from host to RW volser's



MTL3 STG3

DC3

Make STG3 RW

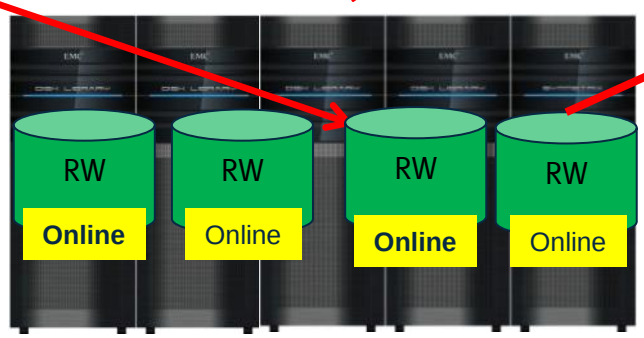
Make STG1 RW

'Enable'

'Enable'

MTL2 STG2

MTL4 STG4



MTL4 STG4

- 'Enable' STG1
- 'Enable' STG3

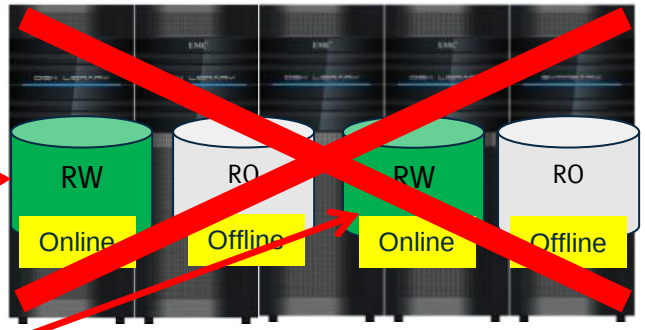
DC2

SMS Enable STG1 & STG3 for Host Writes

DC1

MTL 1/STG1
RW from host to
RW volser's

MTL 3/STG3
RW from host to
RW volser's



DC3

MTL3
STG3

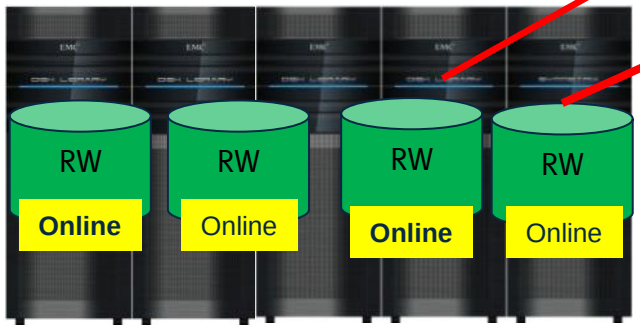
MTL4
STG4

MTL1
STG1

MTL2
STG2

MTL3
STG3

MTL4
STG4



DC2

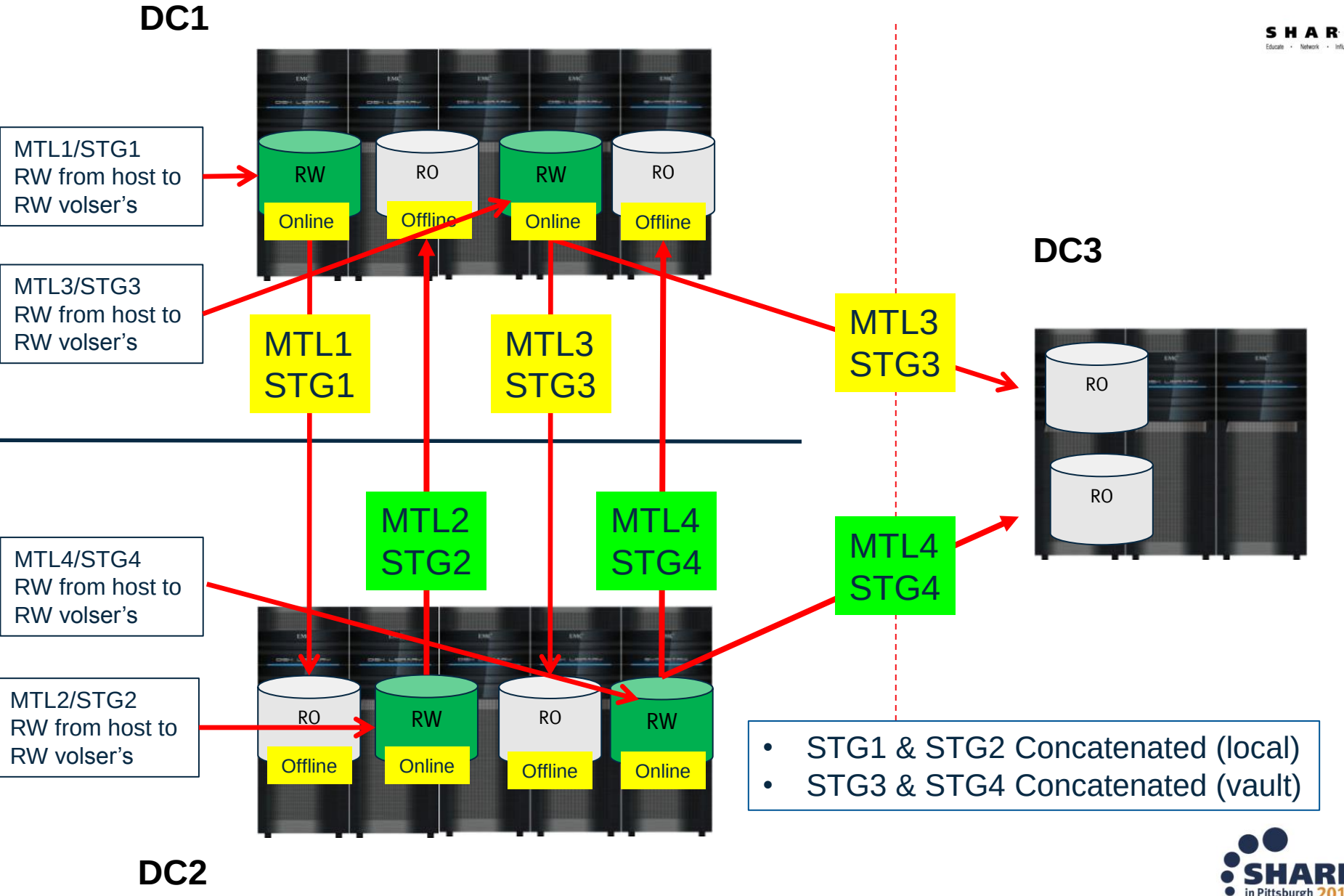
- **Establish replication for STG3 filesystems**

Unplanned vs. Planned Outage Differences

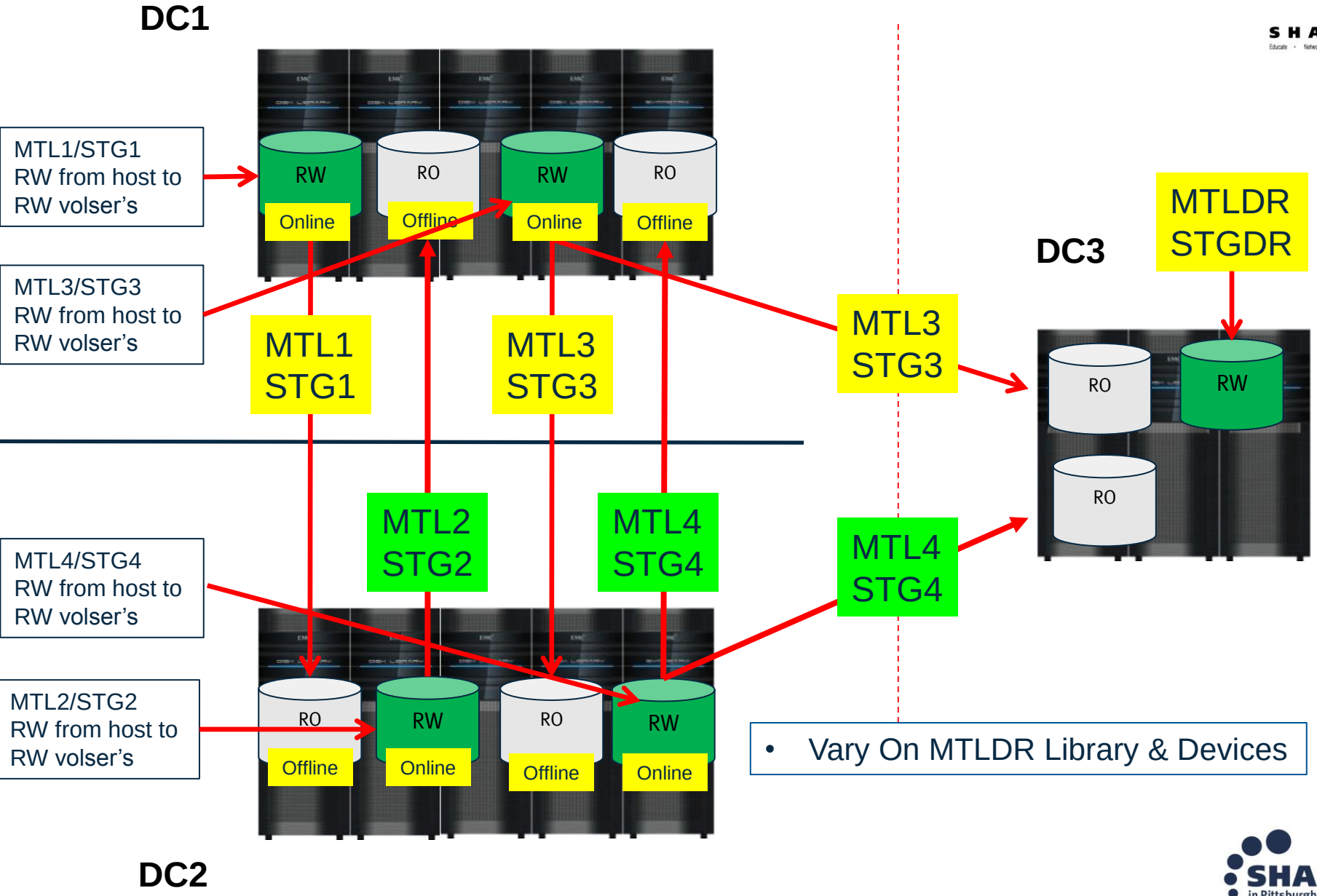
- Replicated Data in flight may have been lost
  - Tools available to identify
- Stop replication on surviving DLm's (source site may not have access)
- Plan for Resynchronization to lost DLm

Current DR Testing Process

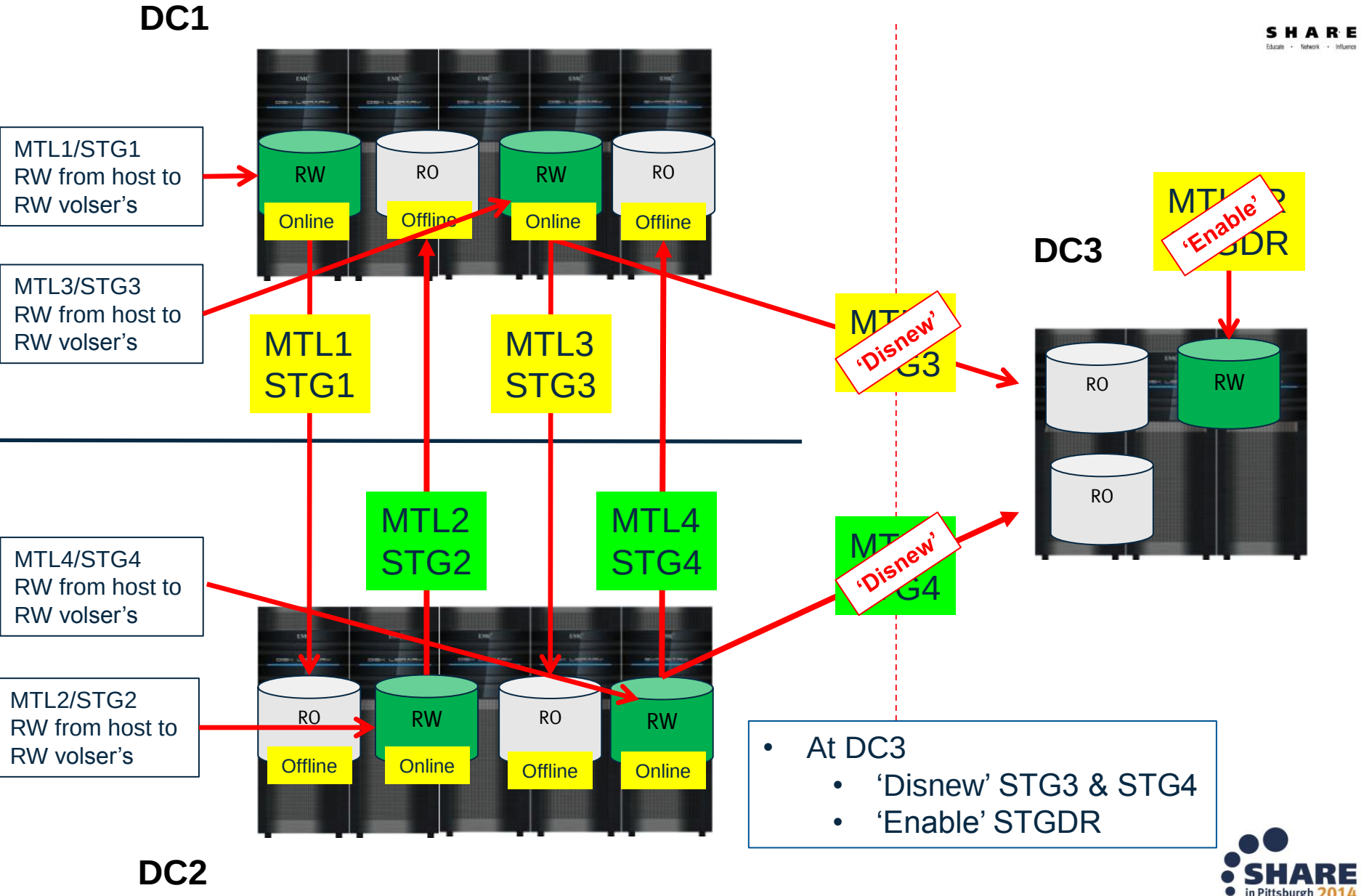
HA SMS Configuration – BAU



HA SMS Configuration – DR Test



HA SMS Configuration – BAU



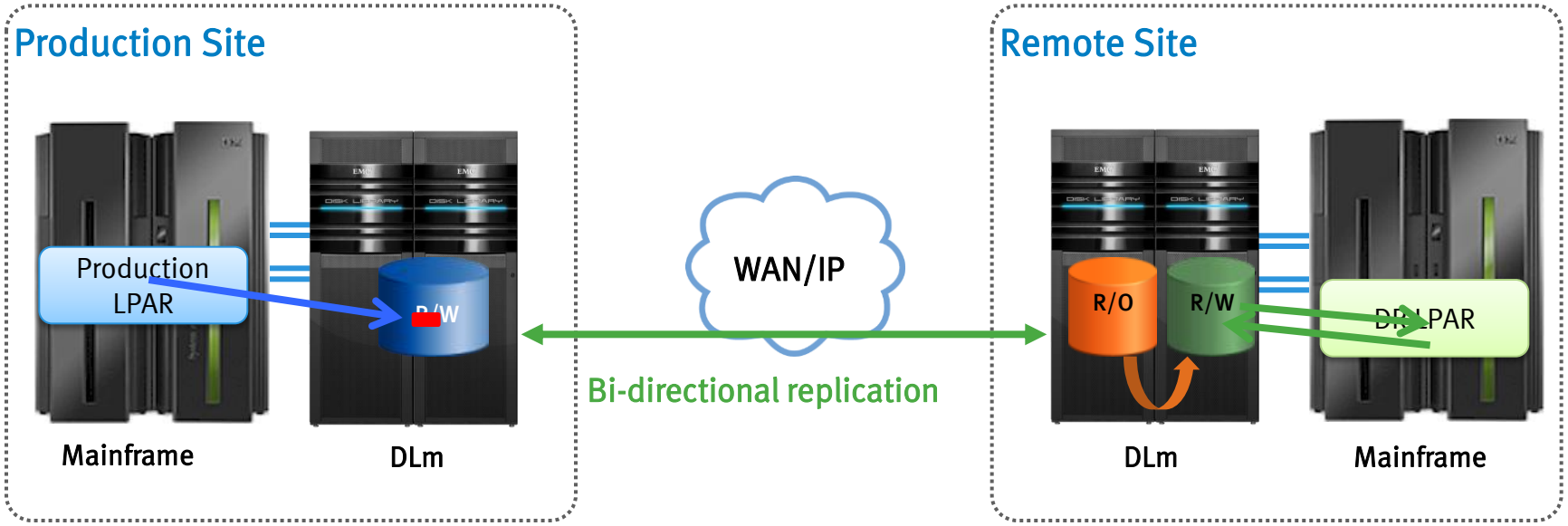
- At DC3
  - 'Disnew' STG3 & STG4
  - 'Enable' STGDR

ACS Routine – DR Test

```
/*-------------------------------------------------------------------*/
/* FOR DLM                                                           */
/*-------------------------------------------------------------------*/
  WHEN (&UNIT = 'DLMLCL')
      DO
        SET &STORGRP = 'STG1','STG2',’STGDR’
        EXIT
        END
  WHEN (&UNIT = 'DLMVLT')
      DO
        SET &STORGRP = 'STG3','STG4','STGDR'
        EXIT
        END
```

The STGDR storage group is setup as unique MTL on the DR DLm. This MTL has filesystems mounted as Read/Write and is used for tape output during a DR test. STG1,STG2,STG3,STG4 storage groups are setup as DIS NEW on the DR side so testing can take place with a minimal amount of manipulation from the production setup.

Testing the Disaster Recovery Environment



Read-only mounts

- Disk arrays allow instant "read-only" copies
- Confirm that tapes can be mounted and all required data can be accessed
- No incremental storage capacity required

Snapshots

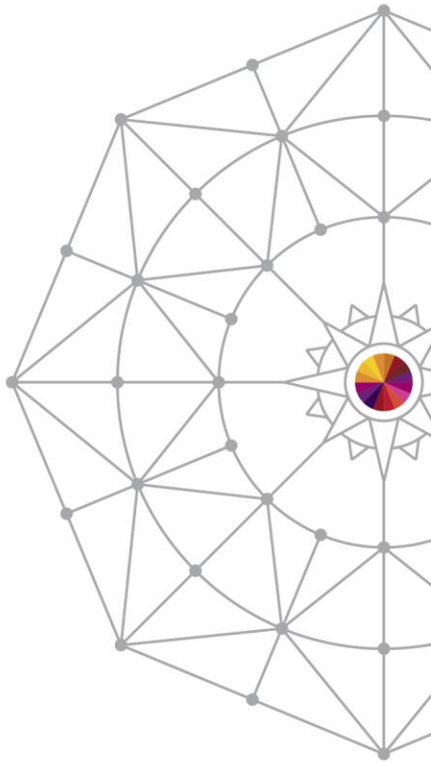
- Disk arrays allow creation of "read-write" snapshot
- Confirm operation at the disaster recovery site
- Some incremental storage capacity required

Remote replication is uninterrupted during testing

Possible Future Projects Enabled by DLm8100

- DASD & Tape Data Consistency (same congroup)
- Flexible Growth for New Client Offerings
- Regulatory Requirements
  - WORM
  - Encryption at Rest

Questions?

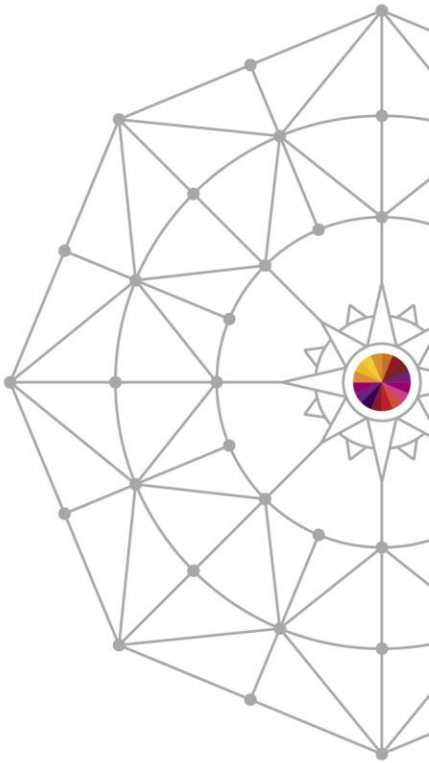


#SHAREorg

Disk Library for Mainframe User Experience in a High Availability Environment

Please do not forget to fill out course eval!

August 4, 2014
Session # 15770



#SHAREorg

SHARE in Pittsburgh 2014