NASHVILLE

Zerto CON

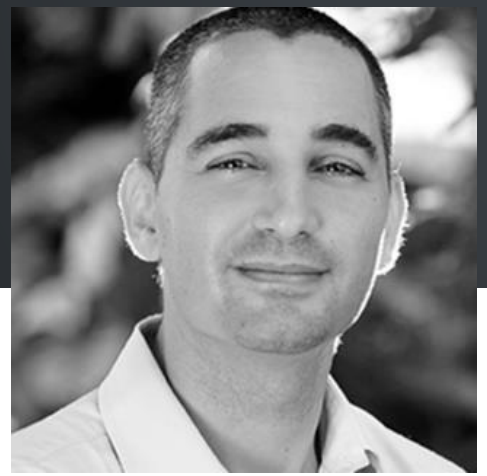
2019

Journey to the Center of the Journal

Oded Kedem
David Aldrich

NASHVILLE Zerto CON 2019

Presenters



Oded Kedem

Chief Architect

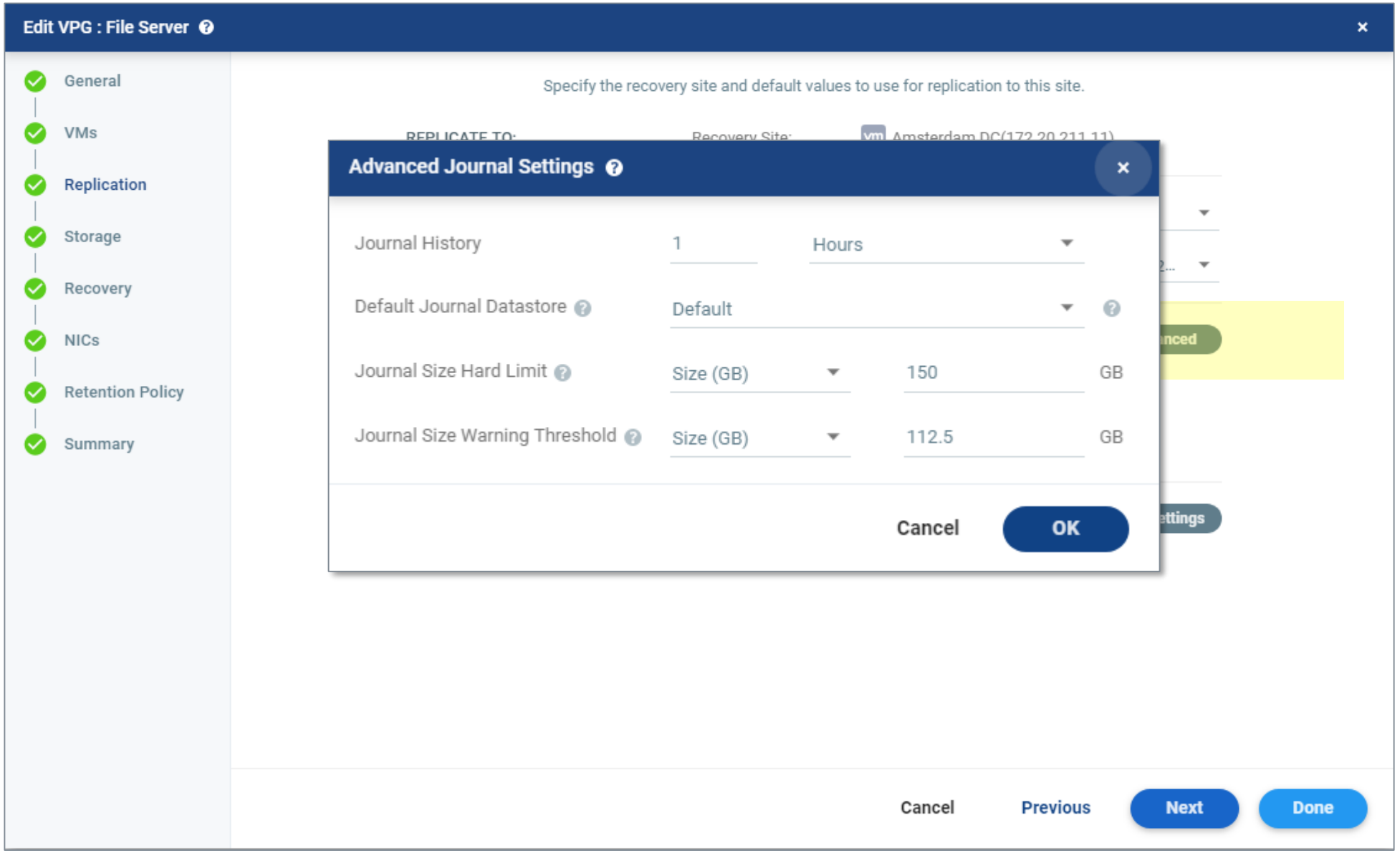
David Aldrich

Sr. Manager, Systems Engineering

NASHVILLE Zerto CON 2019

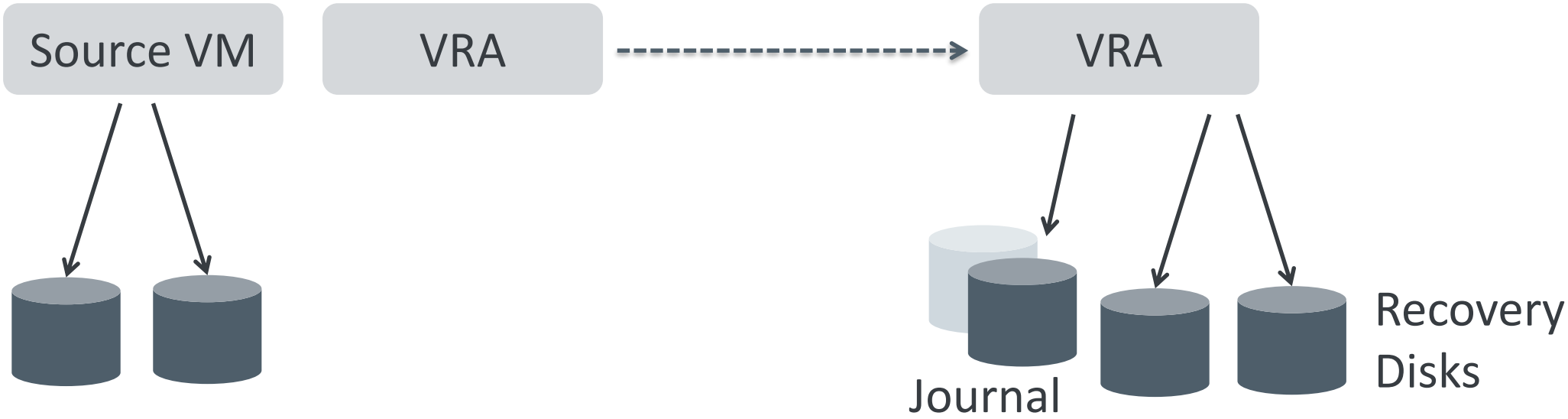
Agenda

- Journal Entities and Data Flow
- FIFO Journal Structure
- Failover
- Synchronization without snapshots
- Public Cloud (Azure & AWS)



Replication Entities

- Recovery volume per source virtual disk
- “Journal” per replicated VM
  - Comprised of a dynamic number of disks
  - Contains multiple independent FIFO queues

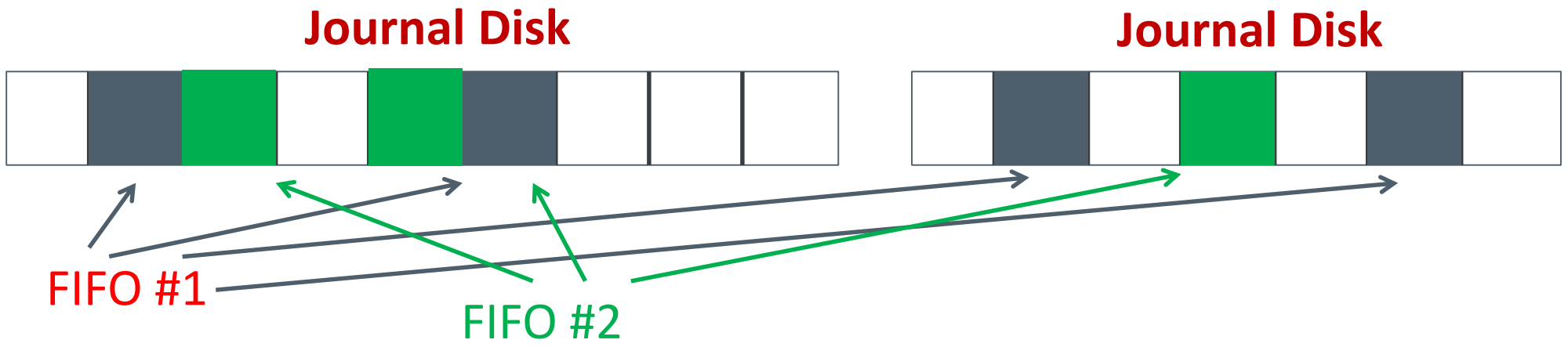


Disk Boxes Extend VRA Disk Count

- Hypervisors limit the number of virtual disks connected to a single VM
  - Virtual disks required for recovery volumes and journals
- Disk box (shadow VRA) has an empty OS
  - Consumes no CPU and almost no memory

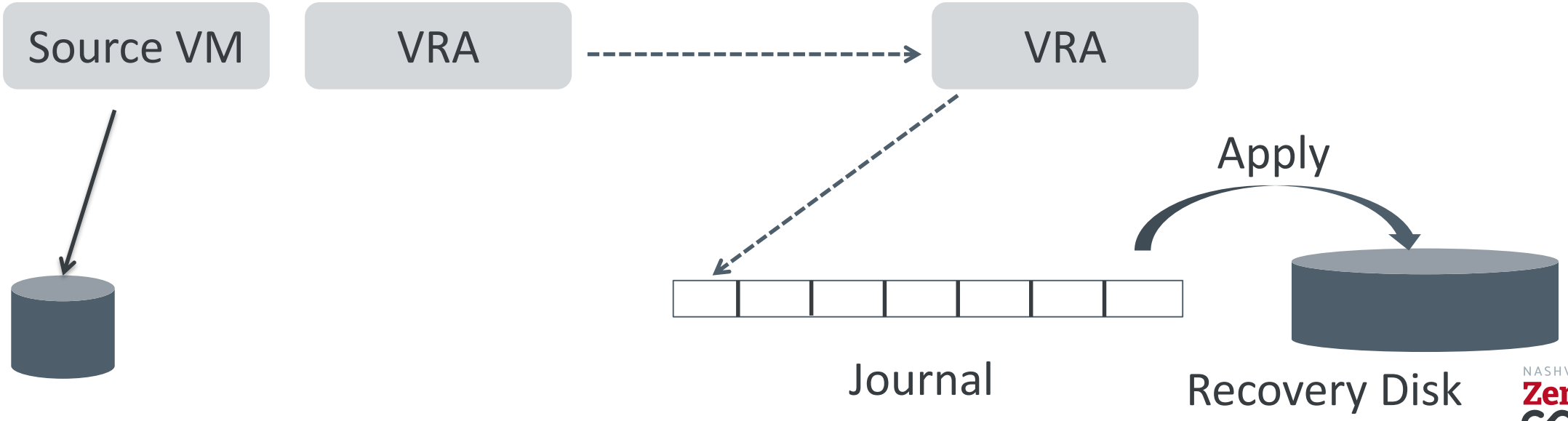
The Journal

- One or more disks connected to the VRA
- Dynamic number of logic FIFO queues
  - The FIFO is a circular queue
- Lists of 16MB Grains
- Disk count and disk sizes are changed dynamically

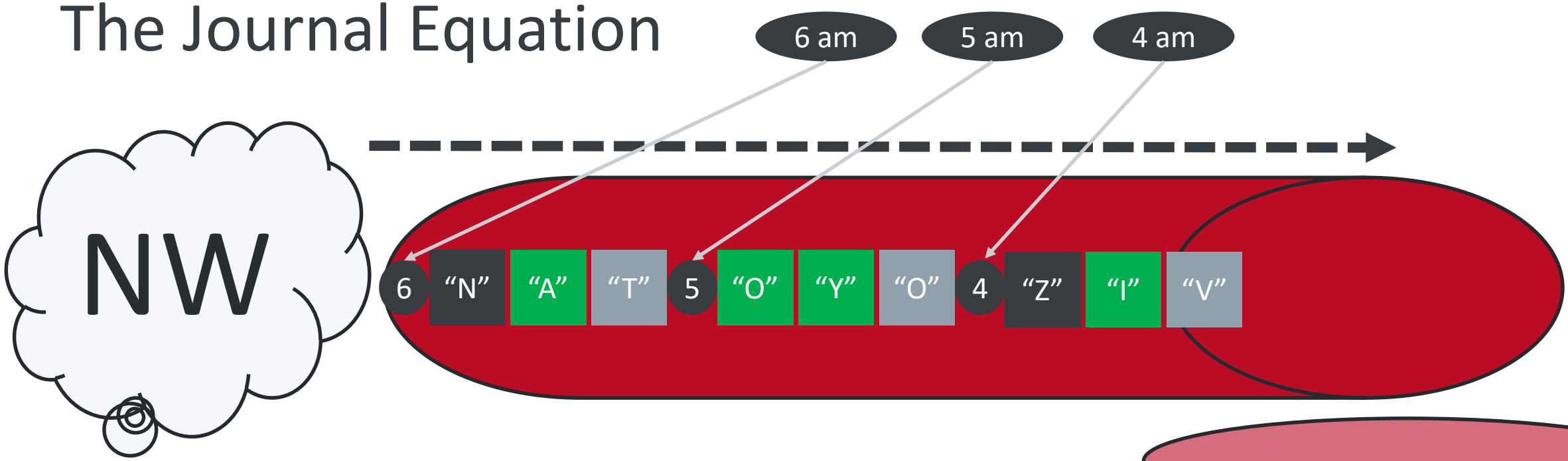


Data Flow

- Source VRA sends data & meta data
- Target VRA journals data
- Data is applied to decrease journal size



The Journal Equation

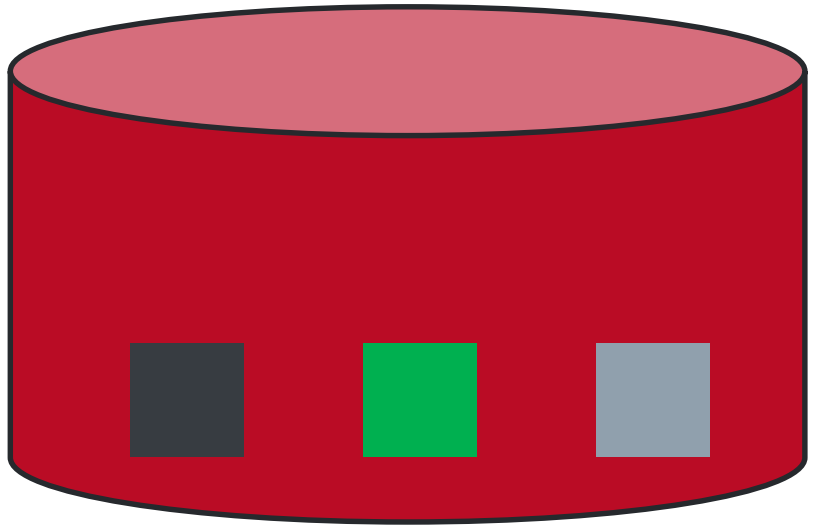


View(CP) = Mirror + Journal(CP)

View(4) = _____ _____ _____

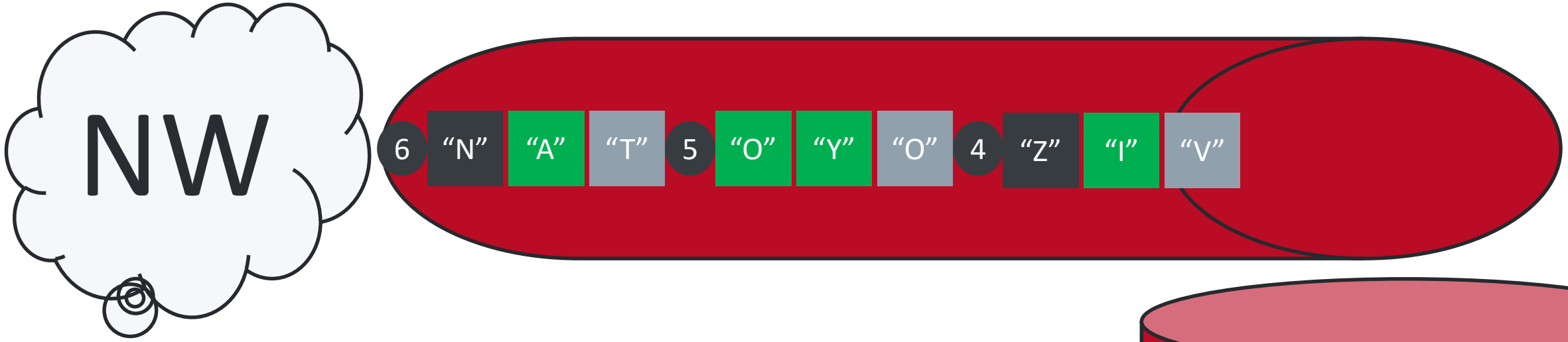
View(5) = _____ _____ _____

View(6) = _____ _____ _____



Advanced: Region Locator (RL)

Do we actually need to overwrite disk to compute views?
To compute view(6), does it matter what was view(5) and view (4)

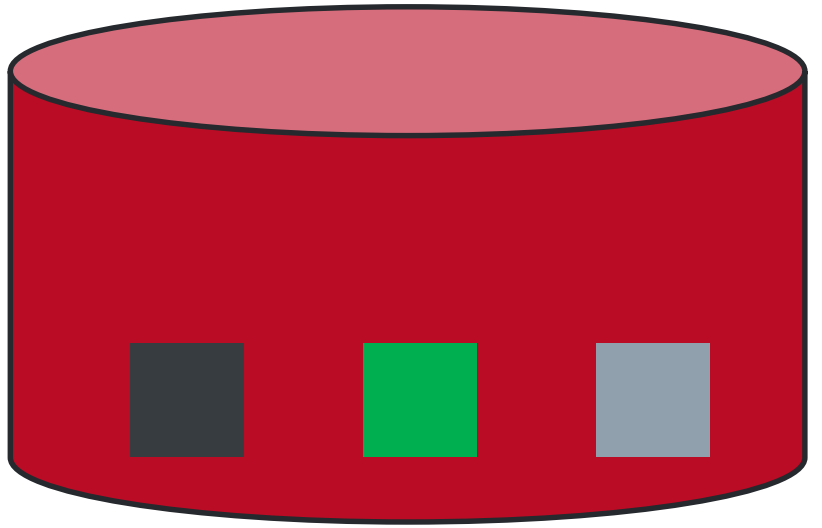


View(CP) = Mirror + Journal(CP)

View(4) = _____ _____ _____

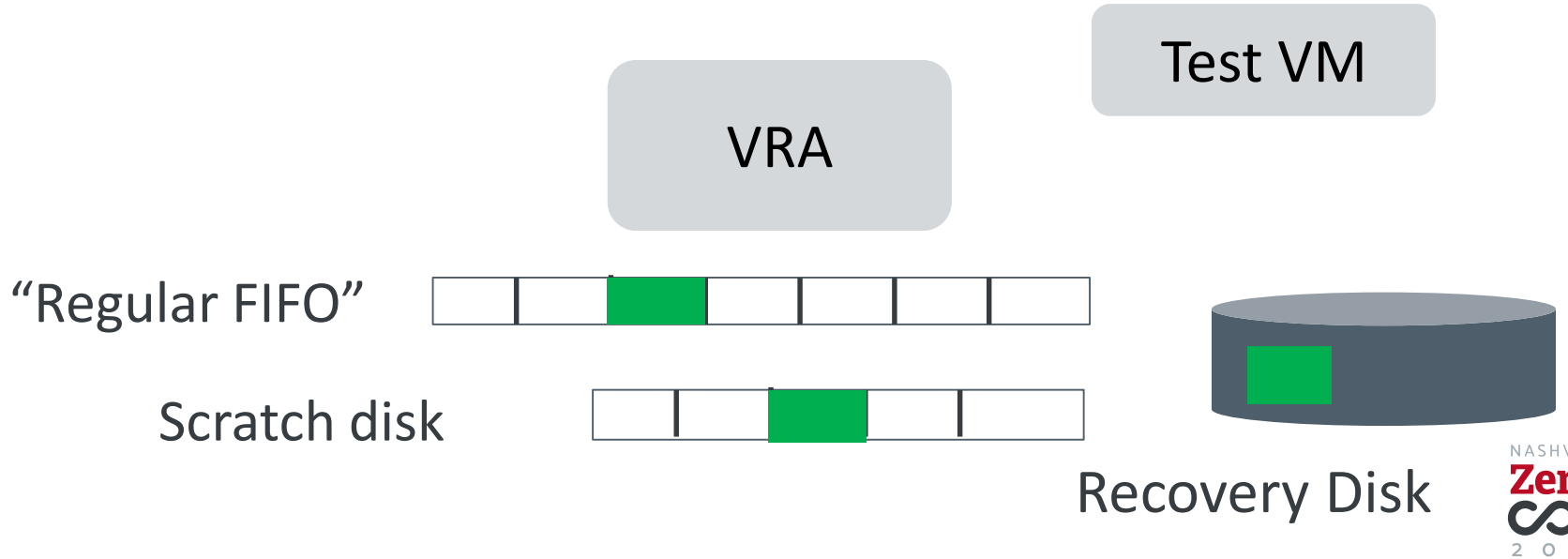
View(5) = _____ _____ _____

View(6) = _____ _____ _____



Failover Test

- Region locator maps the blocks
- Recovery disks → test/target VM
- Journaling continues as long as checkpoint isn't applied
- New writes saved in scratch disk
  - Scratch disk is disposed when test completes



Live Failover (and Move)

Always perform test before failover

Quick Quiz: What's the difference between Failover-Test and Test-Before-Failover?

Commit

- Promote: Efficient & destructive apply
- Reverse replication

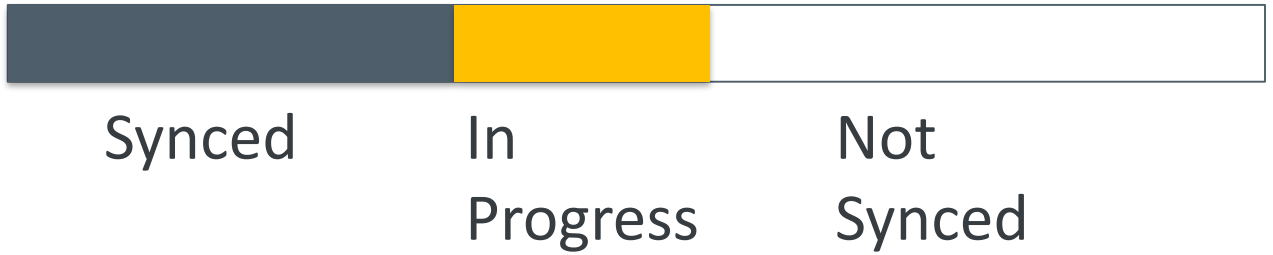
Synchronization without Snapshots

Synchronization Types

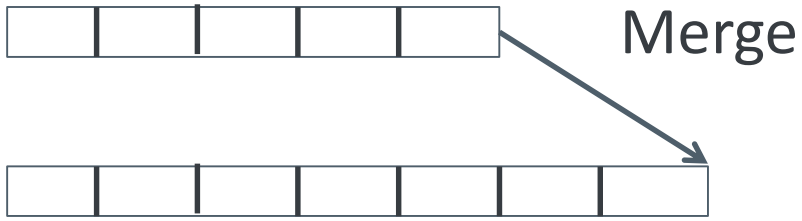
- Initial sync
- Delta Sync
- Bitmap sync

- Differences between sync types
  - Which blocks to sync?
  - Do I have a delta to compare to?

Sync flow

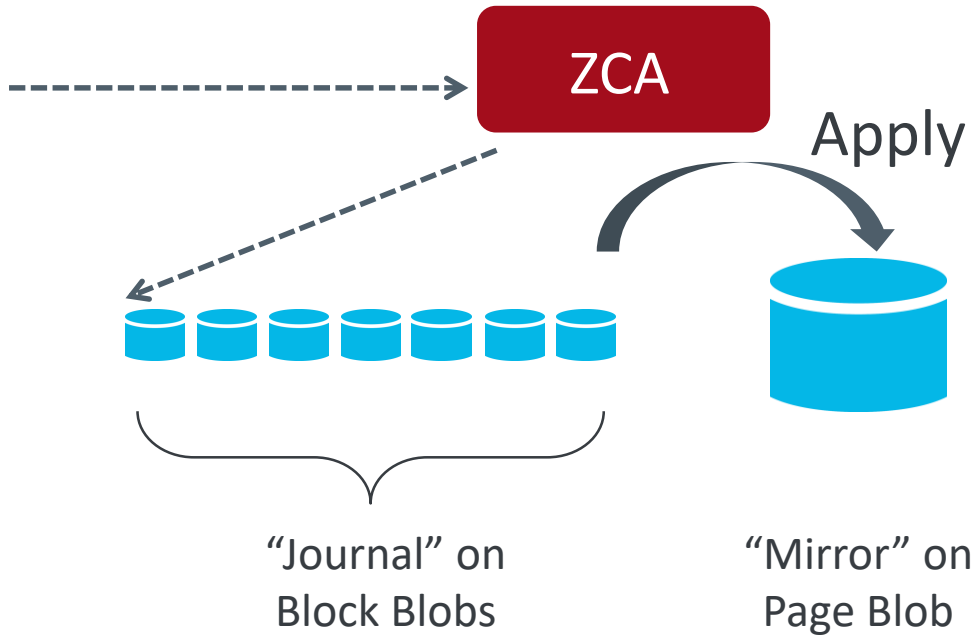


- Partition the disk to
  - Already synced
  - Currently in sync
  - Will sync in the future
- Work on a specific segment
  - “New IOS after sync IOs”
  - Merge between segments
- 2 FIFOs: for new IO and for sync blocks



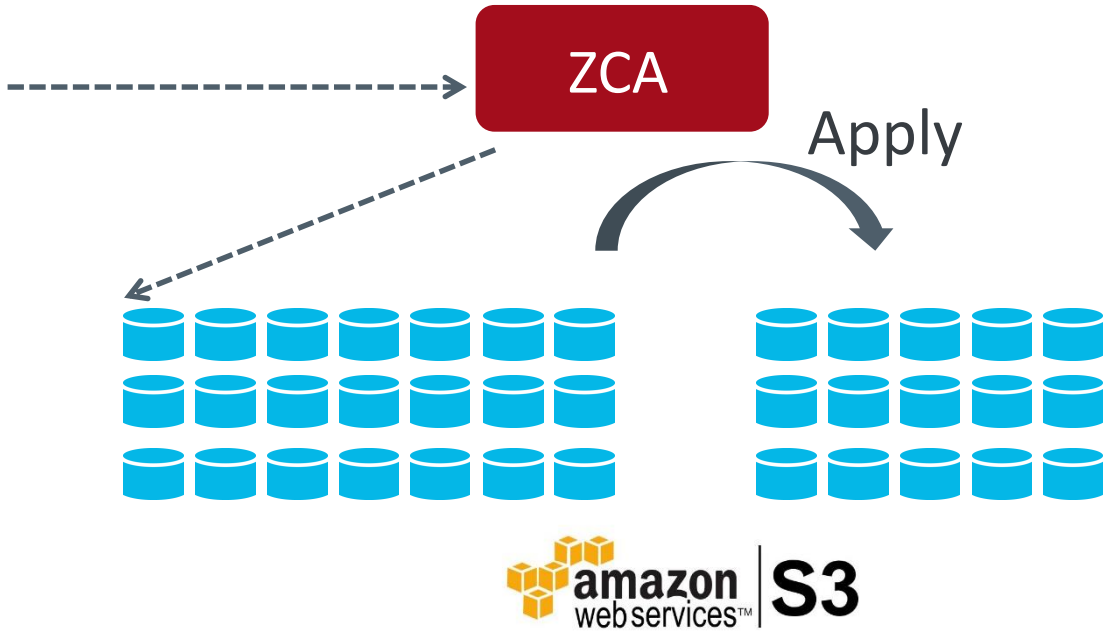
Journaling in Microsoft Azure

- Block Blob Storage provides infinite & Elastic Journal
- Page Blob – Random access storage for recovery

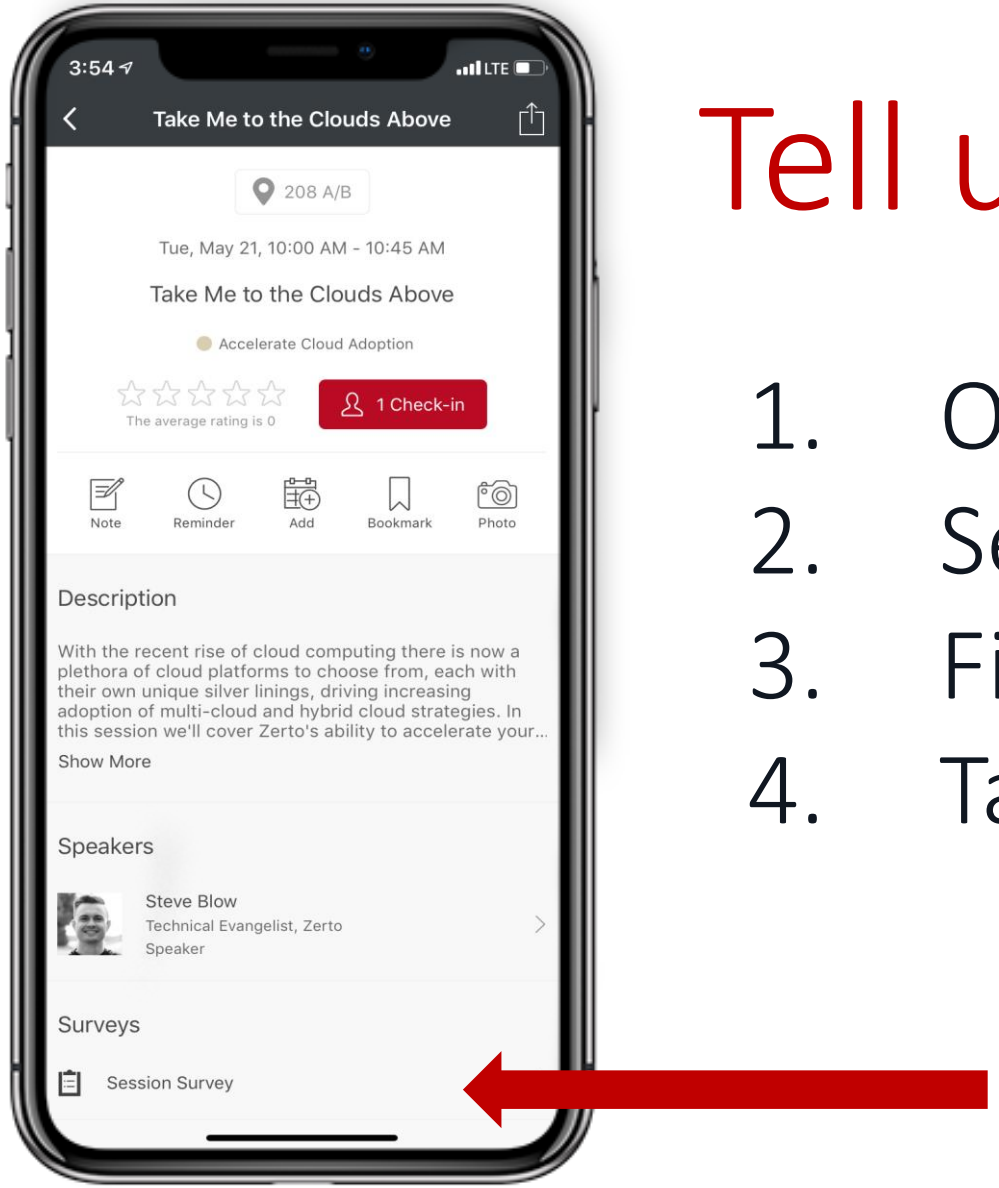


Journaling in AWS

- All data resides on Amazon S3
  - Infinite & Elastic
  - Cost effective
- S3->EBS conversion on recovery



Questions?



Tell us what you think

1. Open the ZertoCON app
2. Select the schedule
3. Find this session
4. Take the brief session survey

Thank You

Zertø