STORAGE DEVELOPER CONFERENCE

SDC 23

BY Developers FOR Developers

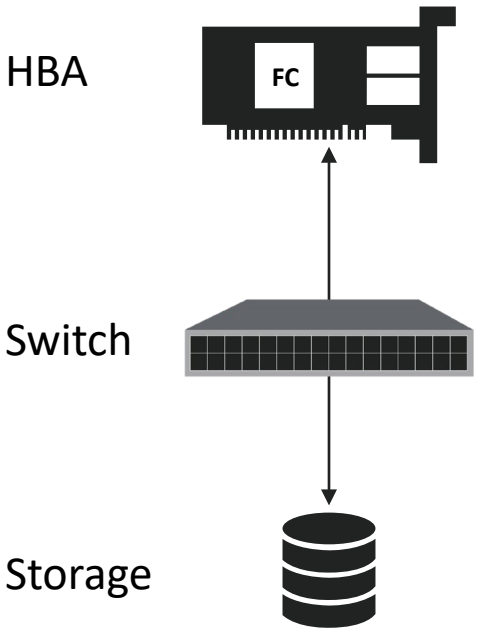
Trucks Keep Right!

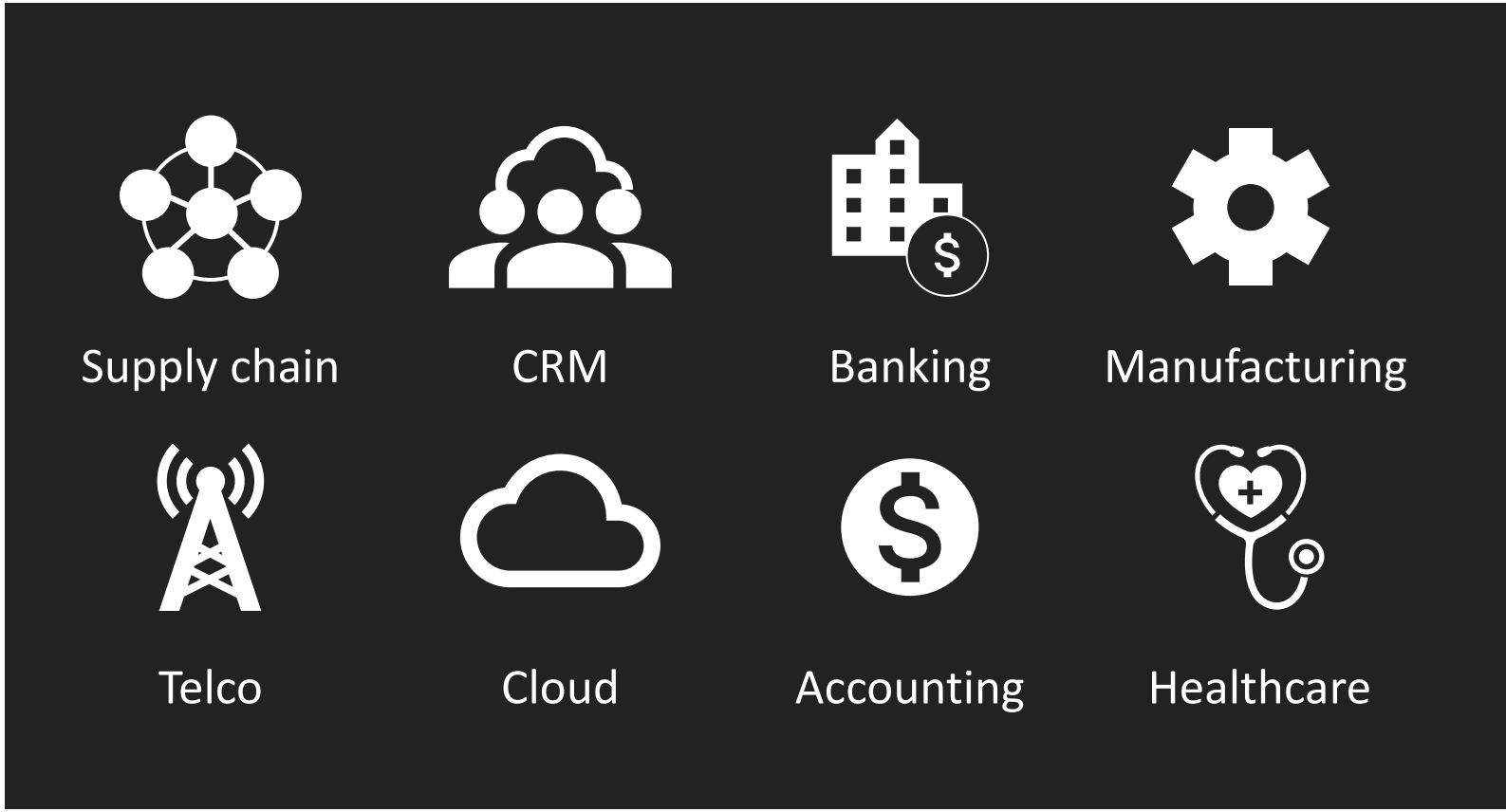
Maximizing Efficiency with Fibre Channel Virtual Lanes

Girish Basrur and Nishant Lodha, Marvell

Fibre Channel

Reliable, secure, **block storage connectivity** for business-critical systems



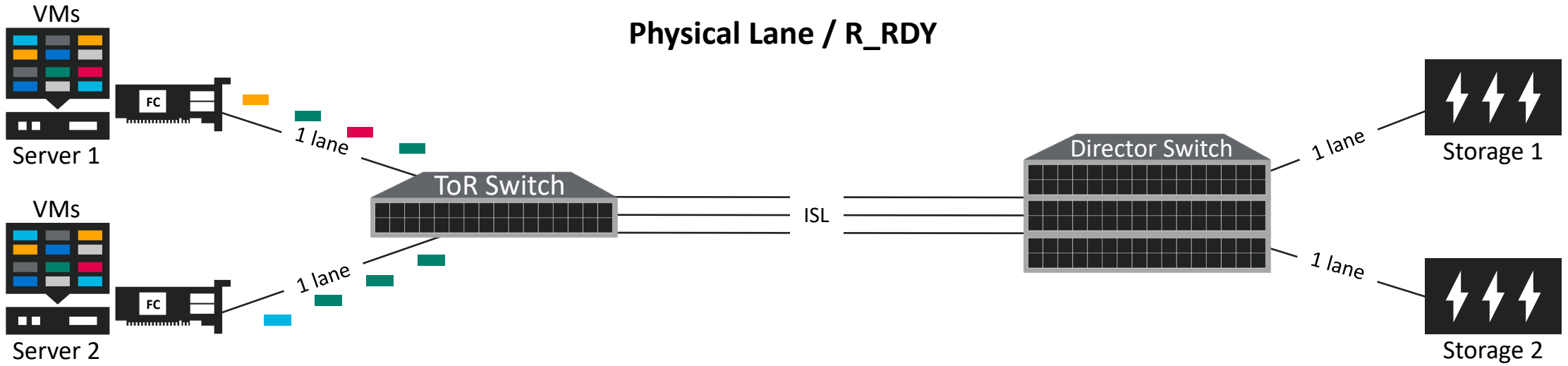


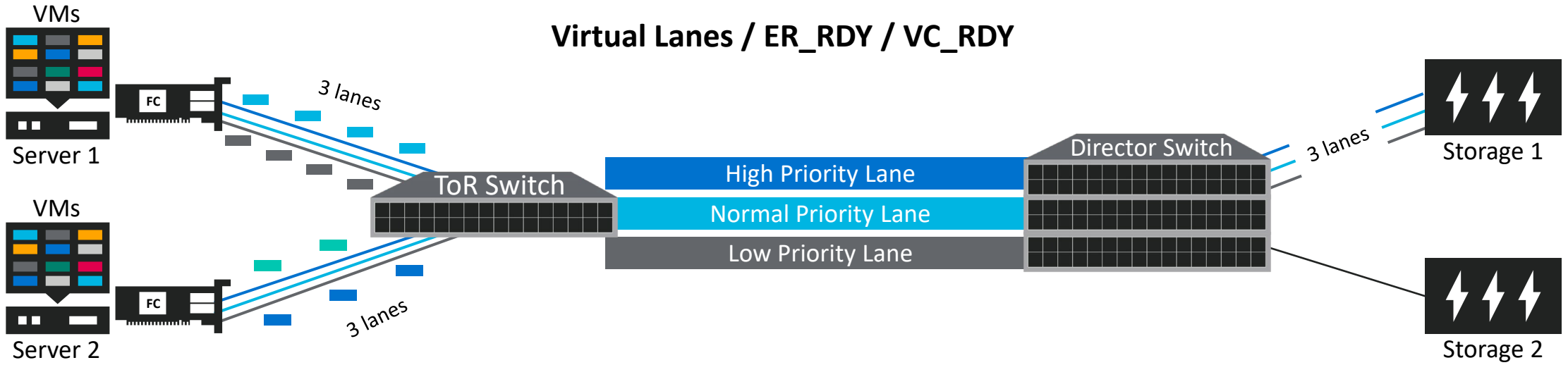
Work on a "Credit"-based mechanism

Virtual Lanes

The technology and its use cases

Credit based flow control in Fibre Channel



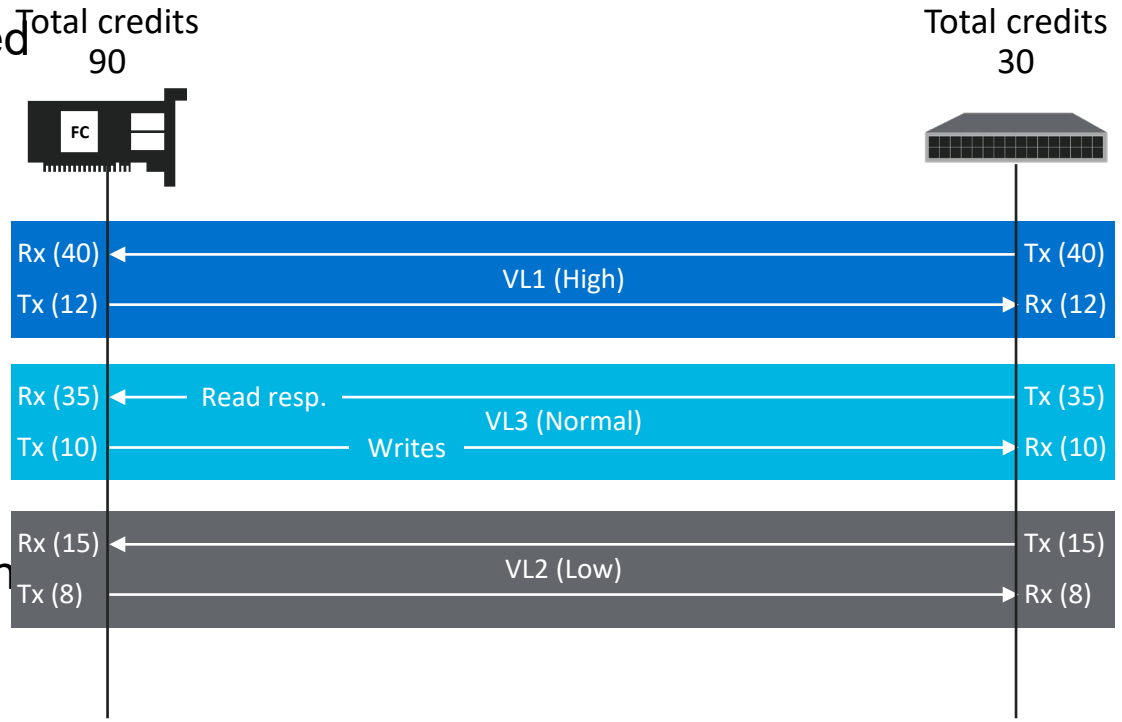


FC Virtual Lanes

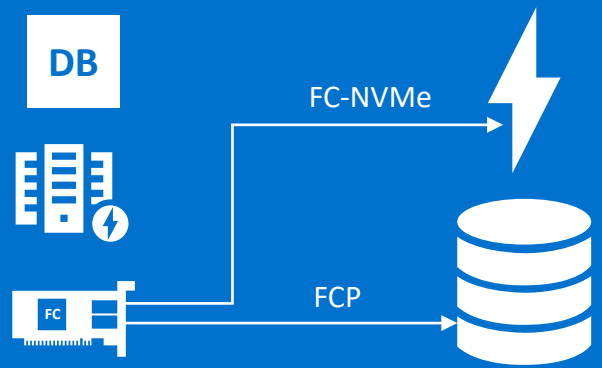
- HBA to Switch, typically a one lane highway
- Slow traffic impedes everyone else
- Very difficult to root cause, impacts SLAs
- Marvell Virtual Lane Technology (standards based
  - Three lane highway
  - Slow traffic in slow lane, fewer credits
  - No head of line blocking
  - Dedicated lane from initiator to target
- Many use cases around segregation, QoS, traffic priorities

How do Virtual Lanes work?

- **Configure**: Physical FC Link (and credits) divided into multiple independent lanes
  - Cisco MDS Virtual Links (VLs)
  - Brocade FC Virtual Channels (VCs)
- **Negotiate:** HBA ⇔ Switch ⇔ Storage @Login
  - Do you VL?
  - How many?
  - Rx BB_CRs per VL?
  - Tx BB_CRs per VL?
  - Priority Range values per VL?
- **Operate:** Frames are tagged (CS_CTL for Cisco, Primitives for Brocade) and transmitted on specific VLs
- **Availability**: Marvell 32G and 64GFC HBAs



Use case: Prioritizing Flash traffic
Application I/O spans disk and NVMe storage



Hybrid storage access

Ability to prioritize flash traffic

Situation: DB server transacting with flash for queries and disk for logs

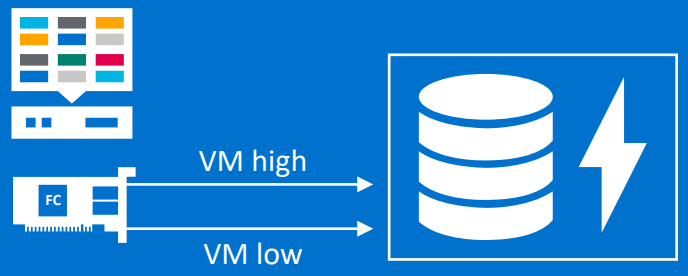
Trigger: HBA driver detects destination media / protocol

Action: Move all NVMe flows to fast VL and HDD flows to slow VL

Result: No head of line blocking of NVMe IOs

Use case: Mapping VM priority

Multiple VMs share the same fabric, map VM priority profile to fabric priorities



Virtualized

Ability to prioritize individual applications

Situation: Virtual Machines sharing the physical storage infrastructure

Trigger: HBA driver receives VM priority along with VM-ID tags

Action: I/Os to high priority VMs mapped to fast and normal VLs

Result: Storage fabric that is aware of administrator set VM priority

Use case: Target congestion

Read and write traffic in progress, storage device is congested

Storage congested

Peer Congestion Notification (FPIN PCN)
(Target in zone is congested)

Situation: Writes to slow target. Target not sinking to media fast enough.

Trigger: Switch sends PCN to FC HBA via Fabric Notifications

Action: Move ALL IOs of this slow flow to a slow VL (isolate it)

Clear: Move flow back to normal/fast VL

Use case in production using Marvell® QLogic® FC HBAs and Cisco MDS switches

Mitigating Congestion with VLs

A deep dive

Fabric Notifications – the foundation to Congestion Mitigation!

Types of Notifications

- Link integrity (Potholes)
  - Awareness: Physical layer errors exceed a threshold
  - Decisive Action: Use an alternative path or adjust for the error rate
- Congestion (NB Traffic)
  - Awareness: Fabric has identified the congestion source (it's you)
  - Decisive Action: Reduce the rate of IO requests or quarantine flows
- Peer congestion (SB Traffic)
  - Awareness: Fabric has identified the congestion source (not you)
  - Decisive Action: Use another path or reduce the transfer rate
- Delivery (can't make it)
  - Awareness: Fabric transmit timeout, cannot deliver
  - Decisive Action: Report IO error immediately (short-circuit protocol timeout)

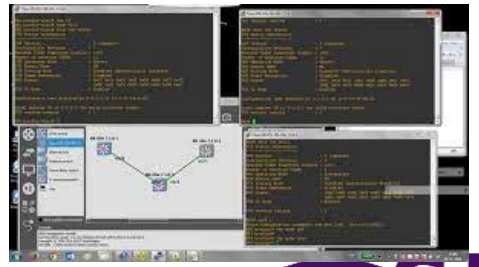
Marvell QLogic
Enhanced 16Gb, 32Gb
or 64Gb HBA

FPIN
Communications

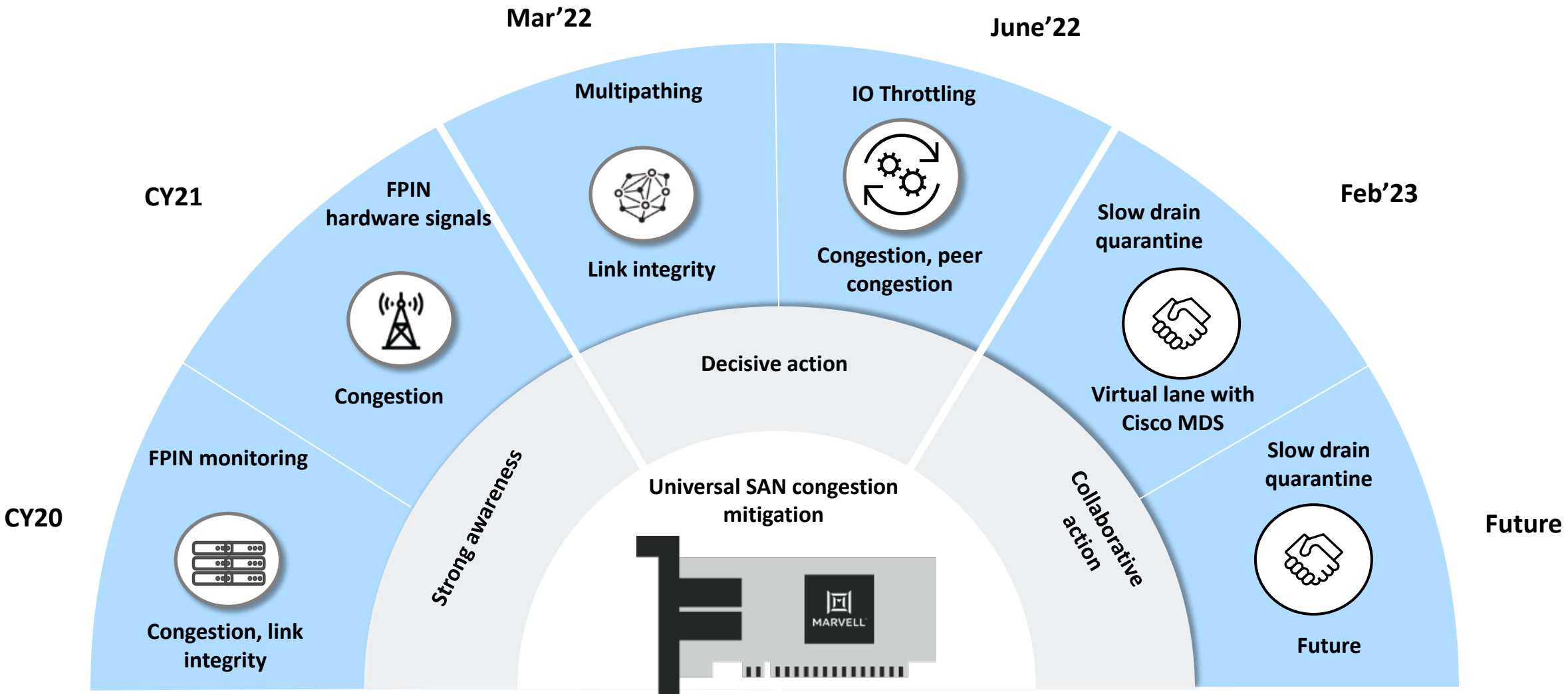
16Gb/32Gb/64Gb FC switch

Brocade	FOS 9 or later
Cisco	NXOS 8.5.1 or later

FC Switch mgmt. console



Congestion Mitigation – The Journey



Marvell QLogic self-driving SAN roadmap

Handling target congestion with Virtual Lanes (VLs)

- Under scenarios where target is congested, VL usage can help in two distinct ways
  - Isolation of traffic to the congested target port on a separate slow lane
  - Auto-throttling of traffic to congested target port
    - Handled with lower credits assigned to slow lanes
- Traffic throttling assists in target recovering from transient congestion conditions
  - Lack of resources
  - Temporary credit stalls

Virtual Lanes – not a silver bullet

- VL usage may not mitigate following scenarios
  - Target is oversubscribed and is in constant state of congestion
    - Inherent I/O throttling in slow lane may not suffice
- VL can cause lane switch back and forth
  - Target congestion is mitigated as long as traffic is on slow lane
  - Switch concludes congestion condition is resolved
  - Host switches back to normal lane, potentially re-triggering congestion condition
  - Exacerbated when multiple initiators implement VL

Best of both worlds…

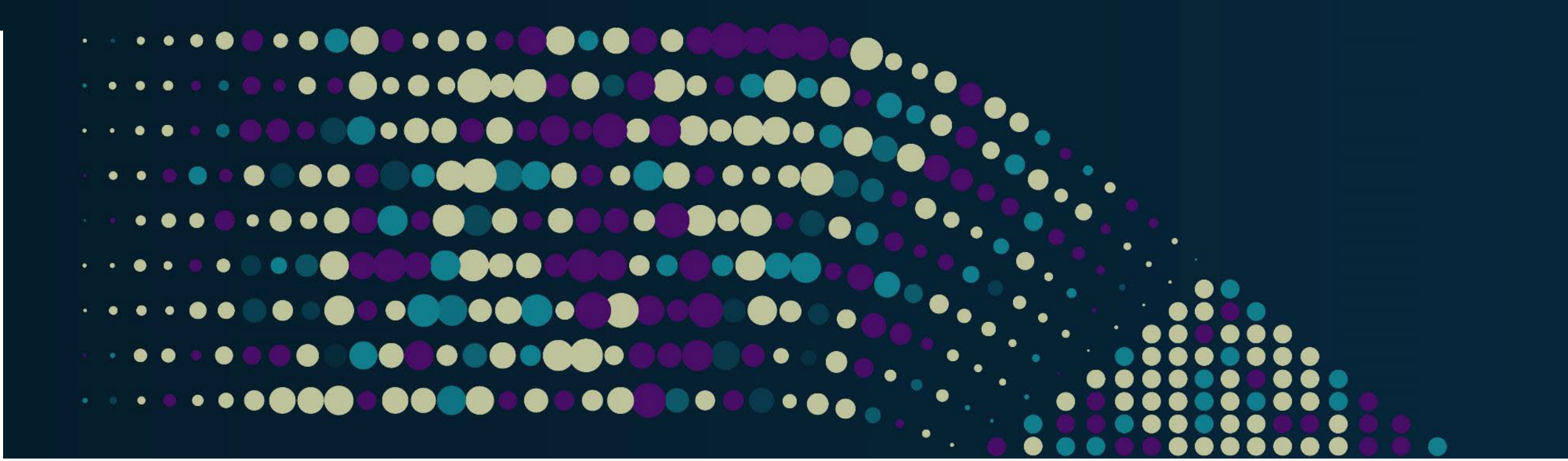
- Hosts use VL as first line of defence in congestion mitigation
  - Will isolate traffic to congested port, leaving other flows undisturbed
- Complement VL with other mitigation strategies
  - Host based I/O throttling
    - Allows for more intelligent throttling and controlled switch back
  - Use of “smart” multipath load balancing policies

Summary and Next Steps

Bring your use cases!

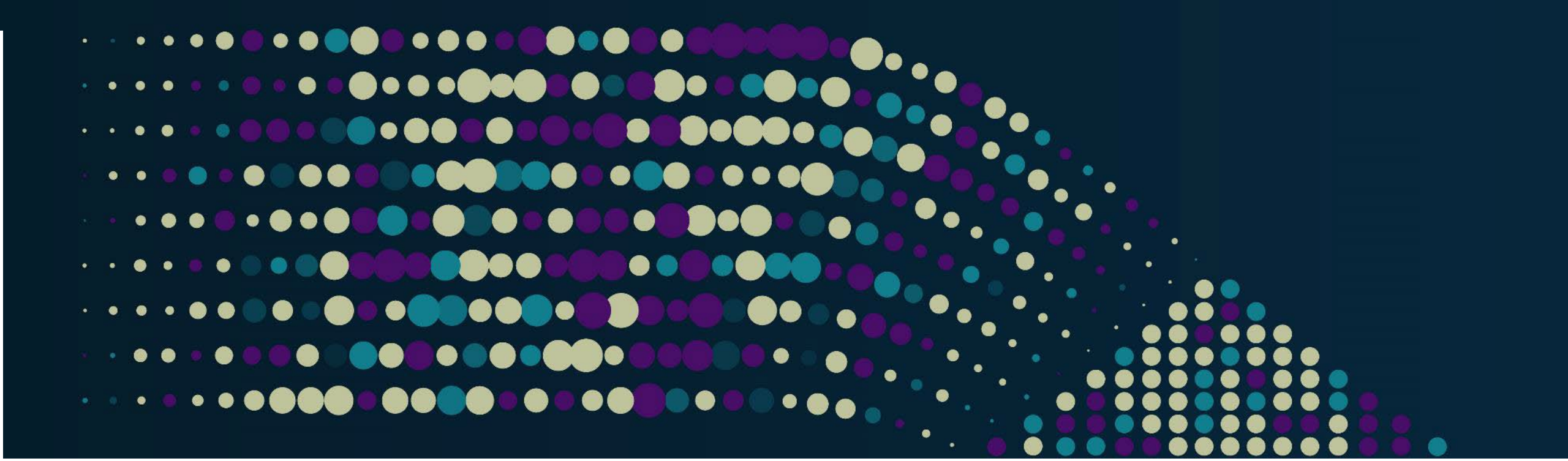
Virtual Lanes avoid delays

- **Congestion** issues and lack of a workload aware storage fabric introduces delays
- **Virtual lanes** enable physical layer priority enforcement for Fibre Channel traffic
- **Futures**: New use cases, dynamic credit allocation, hypervisor and array integration



Please take a moment to rate this session.

Your feedback is important to us.



Q&A

