Washington Systems Center - Storage

**Accelerate with IBM Storage:
Cisco / IBM c-type SAN Analytics**

Art Scrimo

Product Manager Cisco

January 2020

Accelerate with IBM Storage Webinars

The Free IBM Storage Technical Webinar Series Continues in 2020…

Washington Systems Center – Storage experts cover a variety of technical topics.

Audience: Clients who have or are considering acquiring IBM Storage solutions. Business Partners and IBMers are also welcome.

To automatically receive announcements of upcoming Accelerate with IBM Storage webinars, Clients, Business Partners and IBMers are welcome to send an email request to accelerate-join@hursley.ibm.com.

Located on the Accelerate with IBM Storage Site: https://www.ibm.com/support/pages/node/1125513

Also, check out the WSC YouTube Channel here:
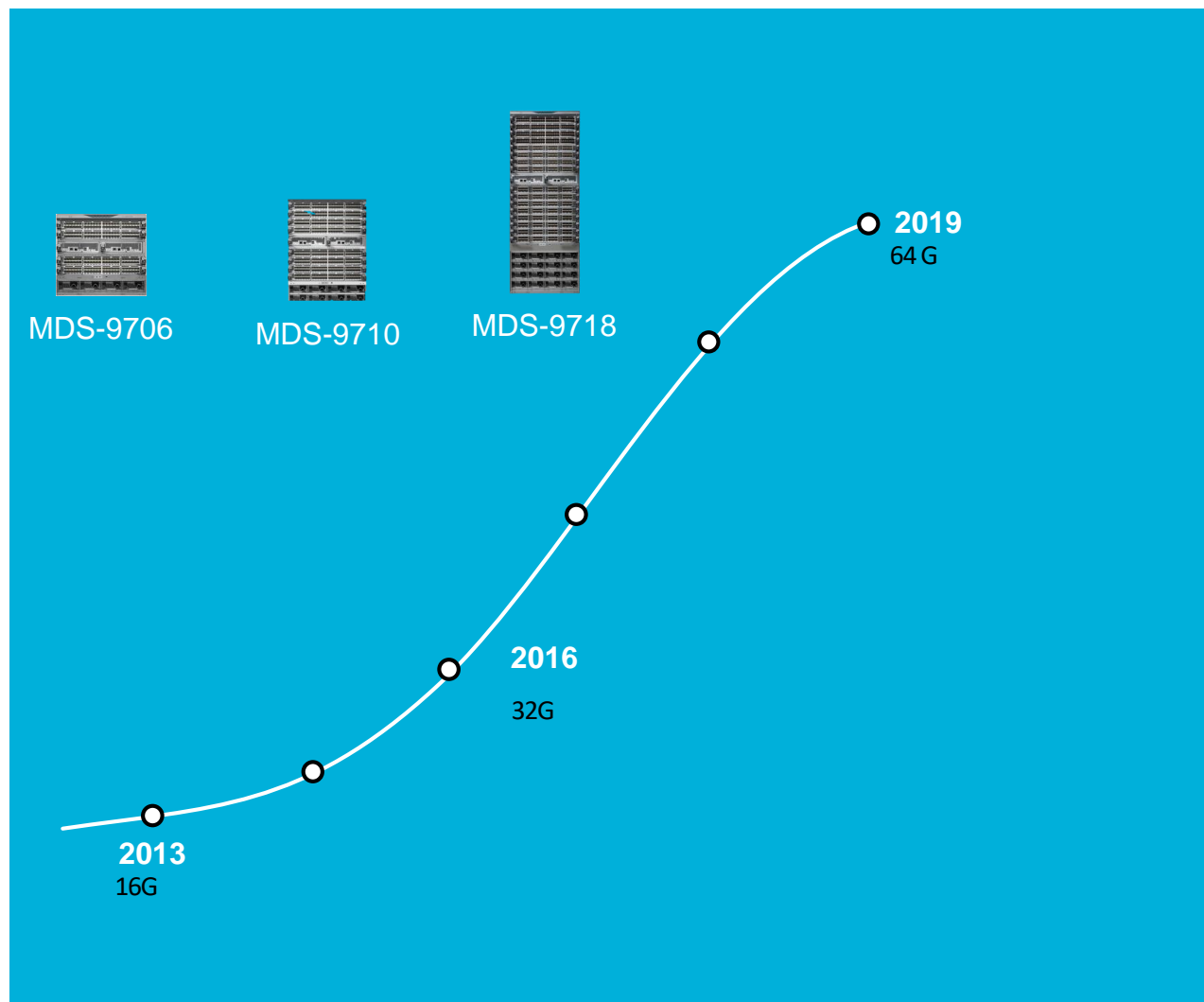https://www.youtube.com/channel/UCNuks0go01_ZrVVF1jgOD6Q

2020 Upcoming Webinars:

January 14 – Brocade / IBM b-type SAN Mondernization
     Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=ed703072fdd739b7e1384ed84869aed8d

January 21 - Cisco / IBM c-type SAN Analytics
     Register Here: https://ibm.webex.com/ibm/onstage/g.php?MTID=eabc0e050b2ff8bdcea4e012c30dfbf71

Built to last

MDS-9706
MDS-9710
MDS-9718

2019
64 G

2016
32G

2013
16G

NEW

SPEED

Investment protection for NVMe/FC and all flash arrays
- 64G ready director
- No forklift upgrade
- Built for the most demanding storage environments

NEW

VISIBILITY

Gain actionable insights
- Industries first NVMe/FC analytics
- FC-SCSI and NVMe/FC support
- Built for customer choice and flexibility

NEW

AUTOMATION

Reduce operational complexity
- Extending devops support for IT automation: ANSIBLE
- Reduce OPEX; simple integration
- Built for advanced SAN automation

Built to last

The good...

- All flash arrays are great for performance, smaller footprint, less power, higher density
- NVMe is great for performance and will have an impact on future deployments
  - Pods and application specific deployments to start

and the bad...

- Are your hosts ready?
  - Typically we see slow drain caused by slower host attached devices
- Not many NVMe supported devices today
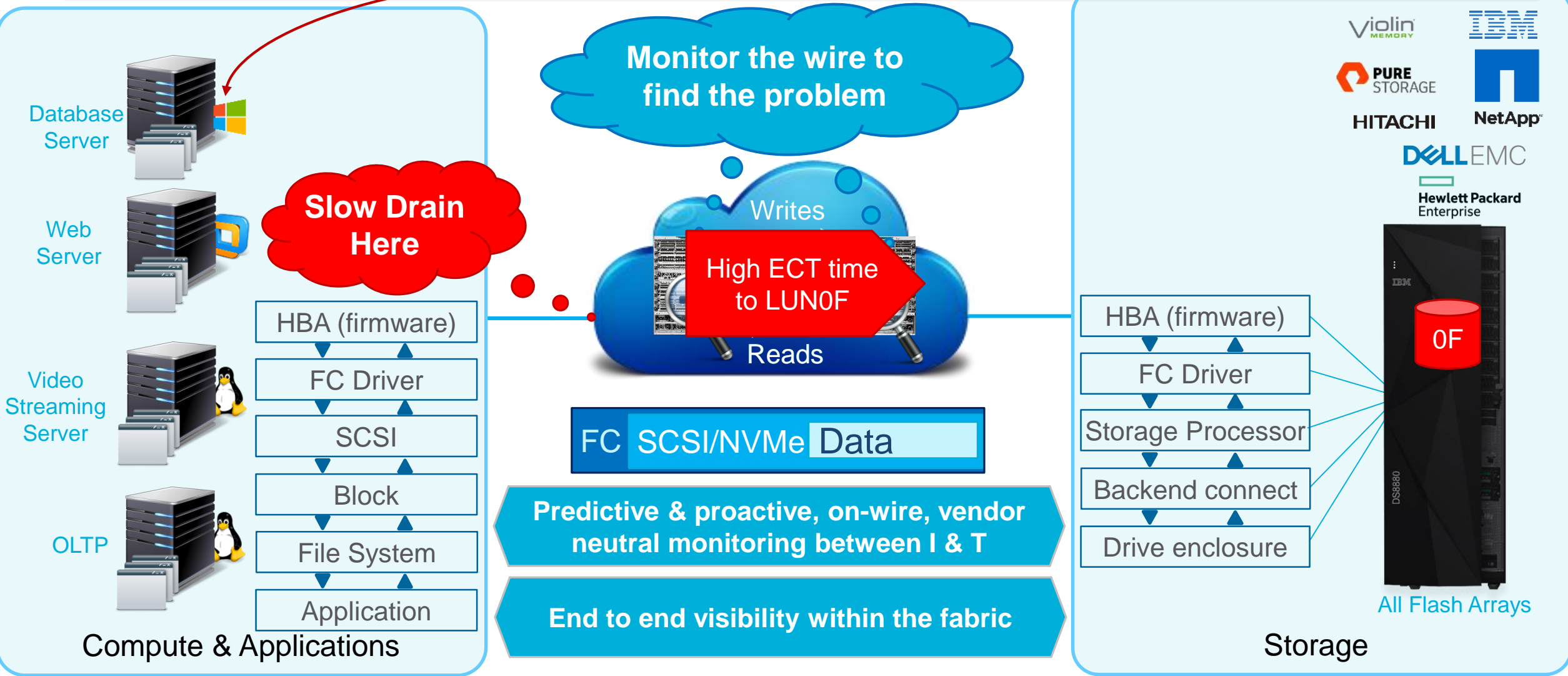 - Can you plan around the future deployments today?

The current state of storage I/O problems

Application issues identified

Not my problem

Not my problem

Not my problem

Compute & Applications

- Database Server
- Web Server
- Video Streaming Server
- OLTP

| HBA (firmware) |
| FC Driver |
| SCSI |
| Block |
| File System |
| Application |

SAN

Writes

Reads

- Too many components involved
- Everybody is limited by their own view
- Virtualization adds complexity
- Hybrid-shared environments
  - Bare-metals & virtualized servers
  - Spinning disks & All flash arrays
  - Multiple speed (2/4/8/16/32Gbps)

Storage

| HBA (firmware) |
| FC Driver |
| Storage Processor |
| Backend connect |
| Drive enclosure |

Violin MEMORY

IBM

PURE STORAGE

NetApp

HITACHI

DELL EMC

Hewlett Packard Enterprise

DS8880

Accelerate with IBM Storage

The current state of storage I/O problems

Application issues identified

Database Server

Web Server

Video Streaming Server

OLTP

Slow Drain Here

- HBA (firmware)
- FC Driver
- SCSI
- Block
- File System
- Application

Compute & Applications

Monitor the wire to find the problem

Writes

High ECT time to LUN0F

Reads

FC | SCSI/NVMe | Data

Predictive & proactive, on-wire, vendor neutral monitoring between I & T

End to end visibility within the fabric

Violin MEMORY

IBM

PURE STORAGE

NetApp

HITACHI

DELLEMC

Hewlett Packard Enterprise

0F

- HBA (firmware)
- FC Driver
- Storage Processor
- Backend connect
- Drive enclosure

All Flash Arrays

Storage

Accelerate with IBM Storage

IBM 32GB Portfolio with SAN Telemetry Capabilities

Cisco MDS 9700
(IBM SAN768C-6,
SAN384C-6, SAN912C-6)
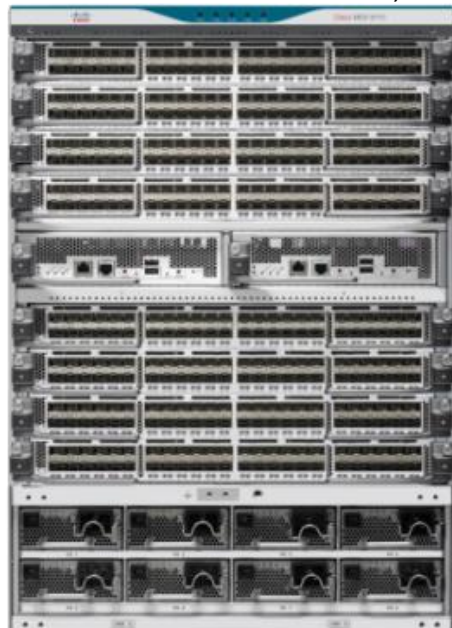
MDS 9700 32GB Line Card

MDS 9132T 32GB 1U Switch
(IBM SAN32C-6)

MDS 9148T 32GB 1U Switch
(IBM SAN48C-6)

MDS 9396T 32GB 2U Switch
(IBM SAN96C-6)
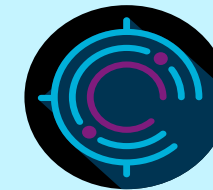
- Hardware Integrated Telemetry

- End-to-end metrics

* These features are available for all Cisco 32Gbps products, not available on the 16Gbps products

Cisco Telemetry - Architecture

Data Collection

- Inbuilt data collection capability into 32G FC port-ASIC
- No impact to data traffic
- Collects only FC & SCSI/NVMe headers, not data

Data Processing

- On-board Network Processing Unit (NPU) on 32G FC module
- Receives headers of frames from port-ASIC
- Extracts metadata from headers
- Stores meta-data in multiple views

Data Analytics and Visualization

- DCNM receives meta-data processed by NPU
- Stores data for long duration
- Analyzes & provides end-to-end visibility, trending, variations, etc
- Correlation with host & target data

Internals of 32G FC module/switch

32G FC Port ASIC

FC & SCSI/NVMe headers

NPU

Metadata to supervisor

32G FC Port ASIC

DCNM or 3rd party Apps

Accelerate with IBM Storage
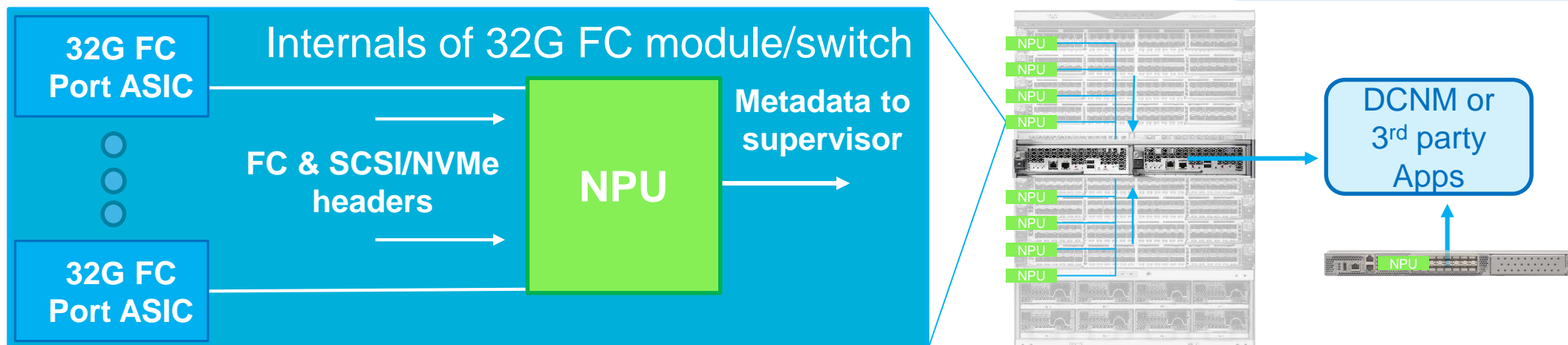
60+ metrics per flow

Write CMD

MDS MDS

XFER_RDY

Exchange Completion Time

Data Access Latency

STATUS

Read CMD

MDS MDS

Exchange Completion Time

Data Access Latency

STATUS
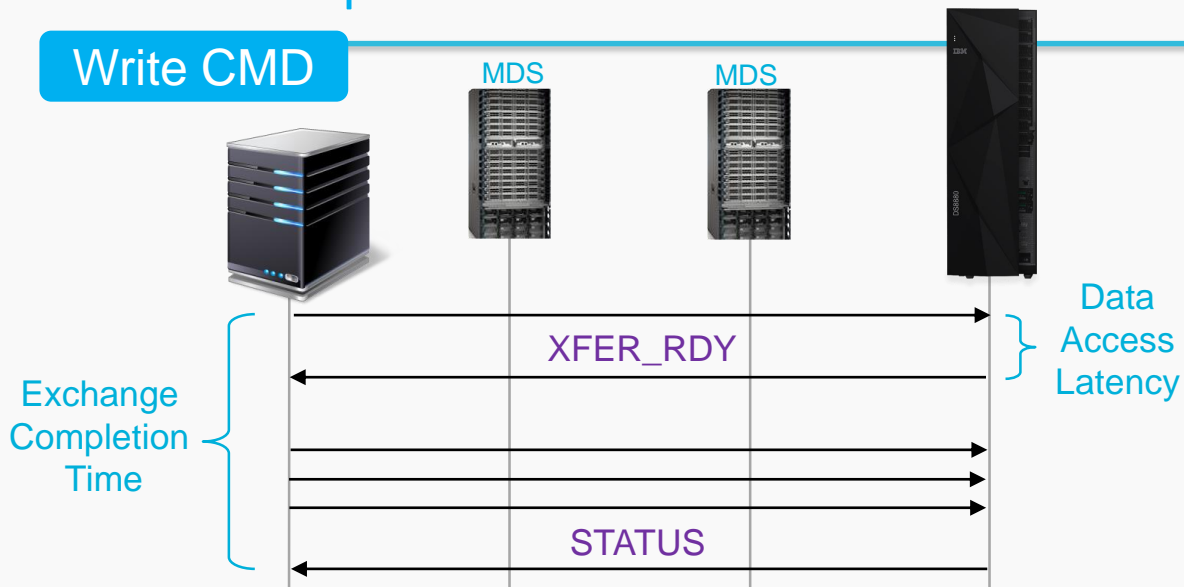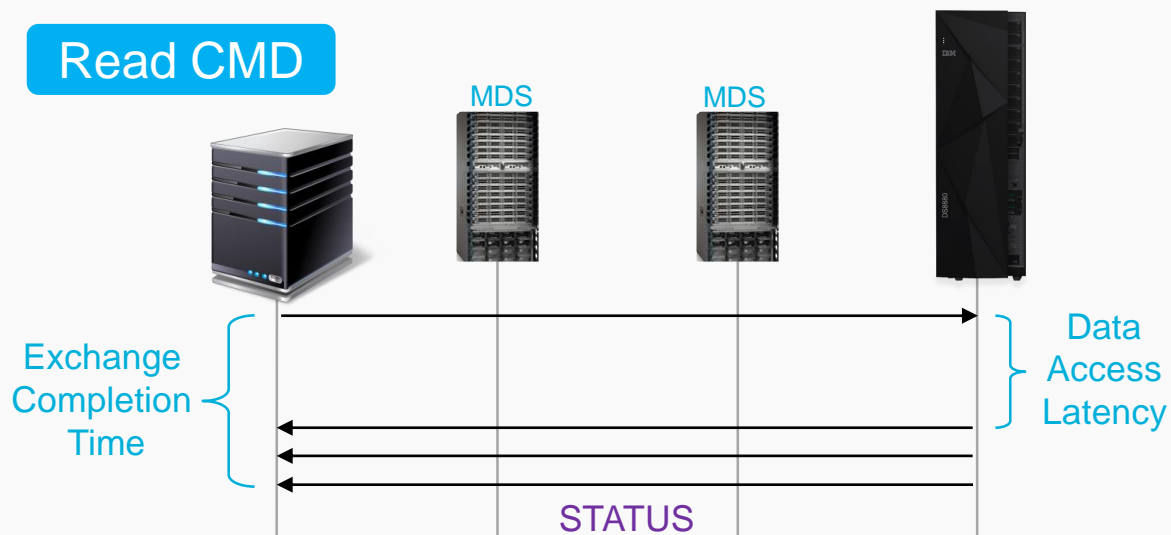
Exchange Completition Time (ECT)

Data Access Latency (DAL)

Outstanding IO (Queue Depth)

IOPS, Detailed flow level counters

Per frame timeout drop frames

Failed FC exchanges, IO retransmissions

SCSI/NVMe Error conditions (Aborts, Rejects, ect)

IO block size and other flow level stats

All metrics at granularity of SID – DID – LUN

32GB Line Card Deployment models

Location of 32G FC module

Storage Edge
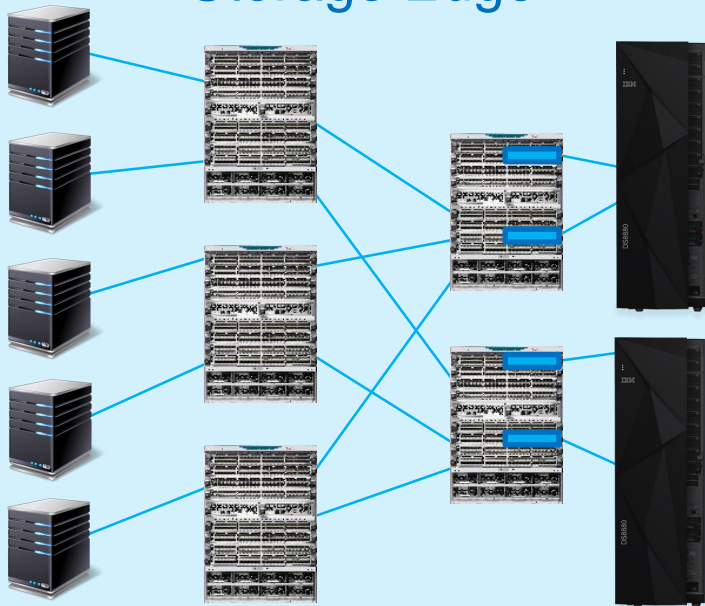
- Closest to storage
- 32G FC module in data path for all flows
- Best for existing MDS 9700 deployment

Core Deployment
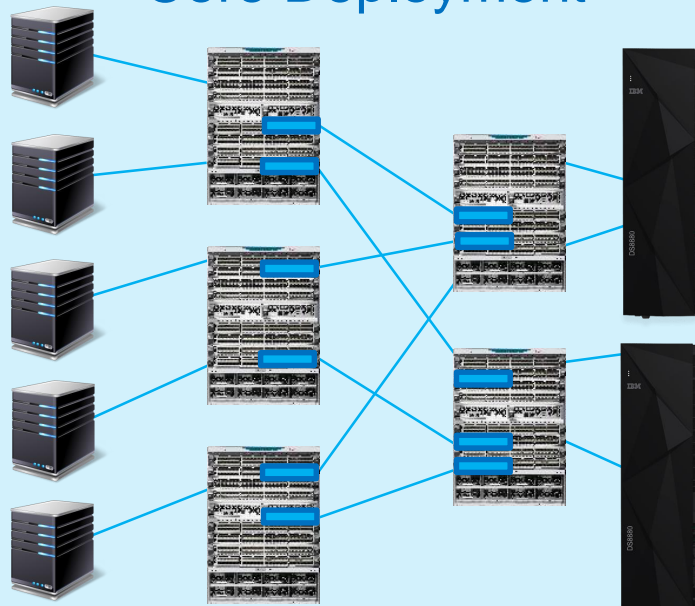
- High capacity core with 32G FC speed
- 32G FC module in data path for all flows
- Best for existing MDS 9700 deployment

Full Fabric
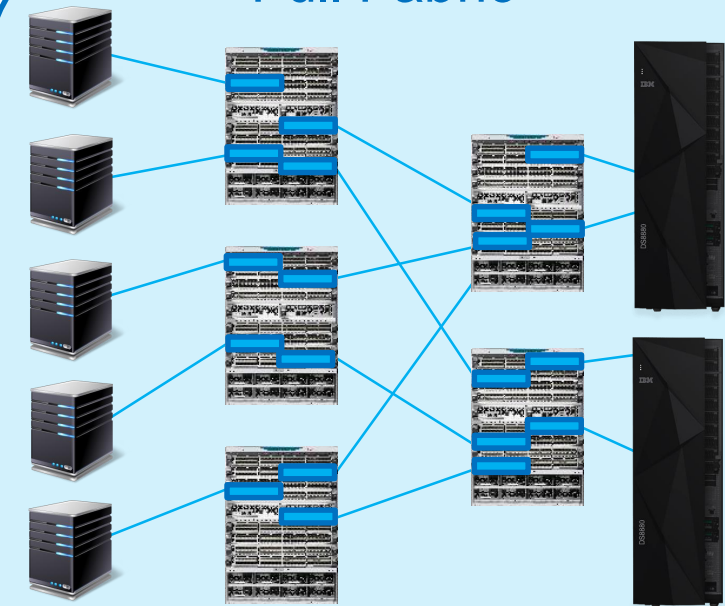
- Best model for end-to-end visibility and correlation
- 32G FC module in data path for all flows
- Best for upgrading from MDS 9500

32G FC module must be in frame switching path

Fixed Switched Deployment models

32G FC Connections

16Gbps Fixed Switch Host Edge
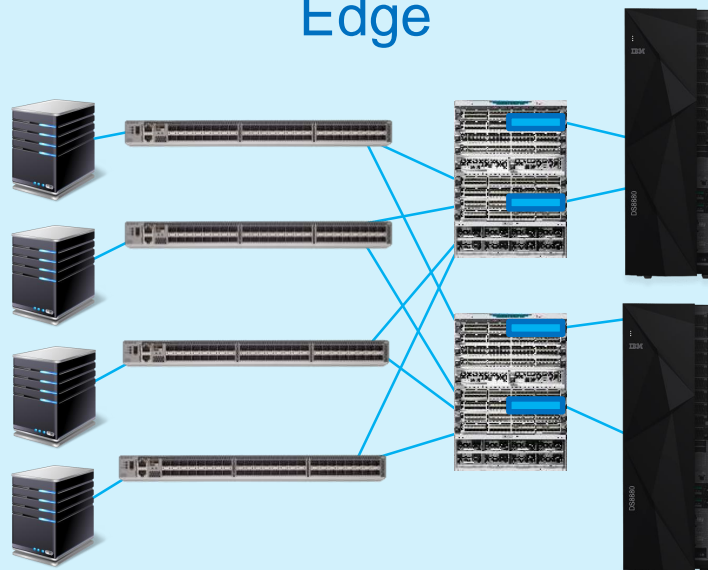
- Closest to storage
- 32G FC module in data path for all flows
- Best for existing MDS deployment

32Gbps Fixed Switch Host Edge
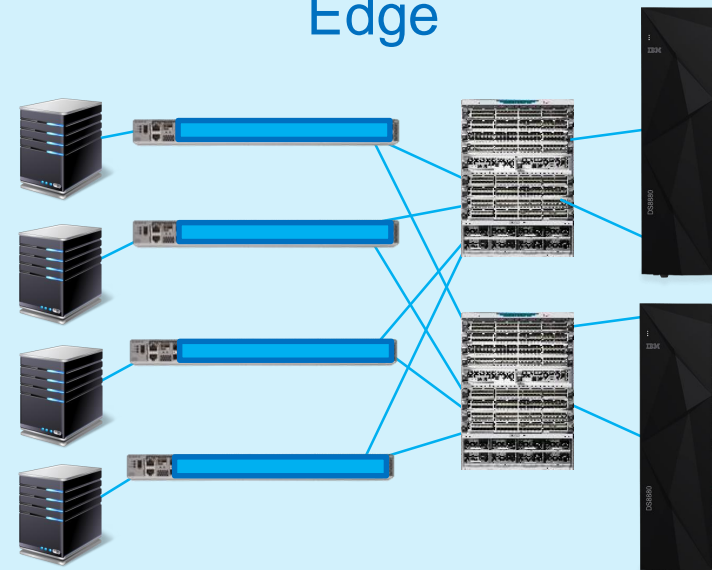
- Closest to the host edge
- 32G FC switches in data path for all flows
- Best for new edge refresh

32G FC product must be in frame switching path

Accelerate with IBM Storage

Metric visualization flexibility

On-Switch CLI

- Flexible SQL-like CLIs to pull raw data from 32G FC module
- Output in key-value format (similar to dictionaries or JSON)
- Unmatched flexibility for advanced user
- Basic trending & correlation Optional standard CLIs
- Output in nice-to-read tabular format, just like any other NX-OS show command output

Remote RESTful access

- Capability to query switch remotely
- Response in JSON format
- Response output depends on the SQL-like input query under the request
- Same as the existing NX-API model
- Works best to track specific metrics
- Extremely flexible

Streaming Telemetry

- MDS streams out the metrics to remote apps regularly
- Data transport and encoding in industry leading open format
- Optimized performance for continuous data (instead of Pull mode) export at high frequency.

DCNM and 3rd Party Apps

SAN Analytics data delivery

Database Server

Web Server

Video Streaming Server

OLTP

Metrics are calculated at microsecond granularity

Writes

SAN

Reads

gRPC

Metrics are exported every 30 seconds

DCNM Server

DS8880
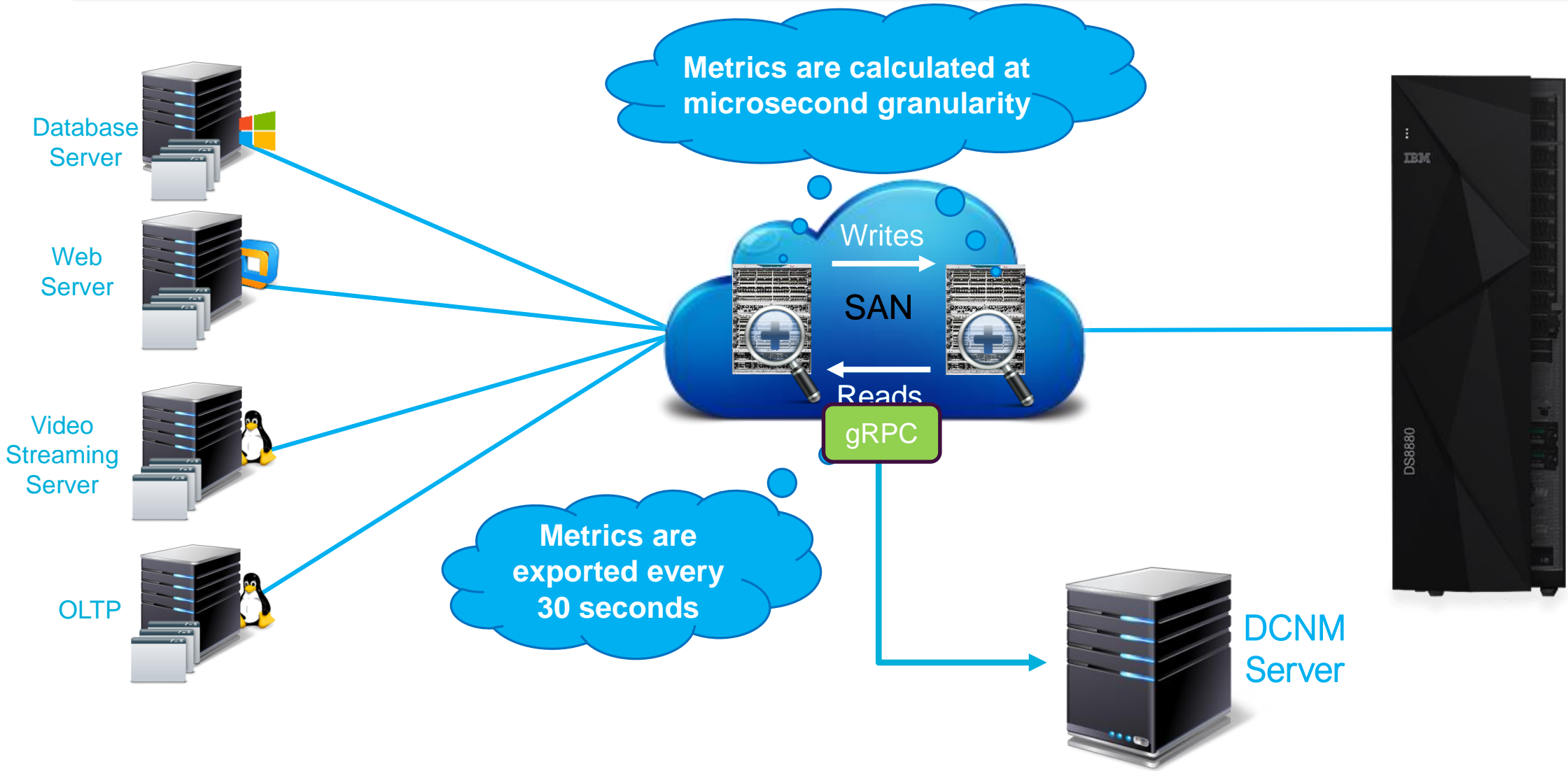
DCNM 11.3(1) Features – SAN Insights – Use cases

Show me which hosts are talking to which storage connections (Different from zoning)

- Select Host Enclosure on the left side
- See all connections to storage from each host
- Check health of each active traffic flow across all fabrics

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Accelerate with IBM Storage

Show me which storage devices are talking to which hosts (Different from zoning)

- Select Storage Enclosure on the left side
- See all connections to Hosts from each Storage Device
- Check health of each active traffic flow across all fabrics

Accelerate with IBM Storage

Show me the poorest performing flows in my environment

- Select Flows up at the top
- Select the Slowest Flow from the left side (top most is worst)
- Check health of each active traffic flow across all fabrics

Find trends and patterns using slider in Monitor/SAN Insights

- User slider to go back in time and find patterns
- One hour increments for every dash
- Watch how things change over time with each click into the past
- Use custom graphing to dive down deeper to see trends

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Accelerate with IBM Storage

Is it a slow SAN or slow storage device?

Cisco Data Center Network Manager

Use case : Custom Graphing | Metrics | Search here

REAL TIME GRAPH

50:05:07:60:5e:bf:e9:62 From 2018-09-01 To 2018-09-07

6–Sep–2018 21:46:00
Normalized Write ECT : 0.1049 ms/KB
Normalized Write DAL : 0.0045 ms/KB

— Normalized Write ECT — Normalized Write DAL

DATATABLE [Latest 100 records]

Initiator Enc	Initiator WWN	Target Enc	Target WWN	LUN	Switch IP	Timestamp	Normalized Write ECT (ms/KB)	Normalized Write DAL (ms/KB)
RHEL_17417	21:00:00:0e:1e:18:18:31	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
UCSB_RHEL_17439	20:00:06:25:b5:00:00:2f	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
Win_17416	10:00:00:90:fa:47:6e:51	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0.9323	0.0404
WIN_174233	20:00:78:da:6e:a0:ac:d3	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
UCSB_WIN_17440	20:00:06:25:b5:00:00:3e	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
RHEL_17414	10:00:00:90:fa:50:a1:c7	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
UCSB_RHEL_17439	20:00:06:25:b5:00:00:3f	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
SUSE	21:00:00:0e:1e:18:18:81	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0
UCSB_WIN_17440	20:00:06:25:b5:00:00:2e	IBM_F840	50:05:07:60:5e:bf:e9:62	0000-0000-0000-0000	172.25.174.7	2018-09-7 11:14:23	0	0

- Use Custom Graphing to search by LUN ID, Host or Storage port
- Compare DAL and ECT
- DAL high storage Device causing slowness; ECT High SAN causing slowness

- **Based on this graph all issues are pointing to a SAN related slowness on this storage port**

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Is my Storage Device slow to respond to read or write requests?
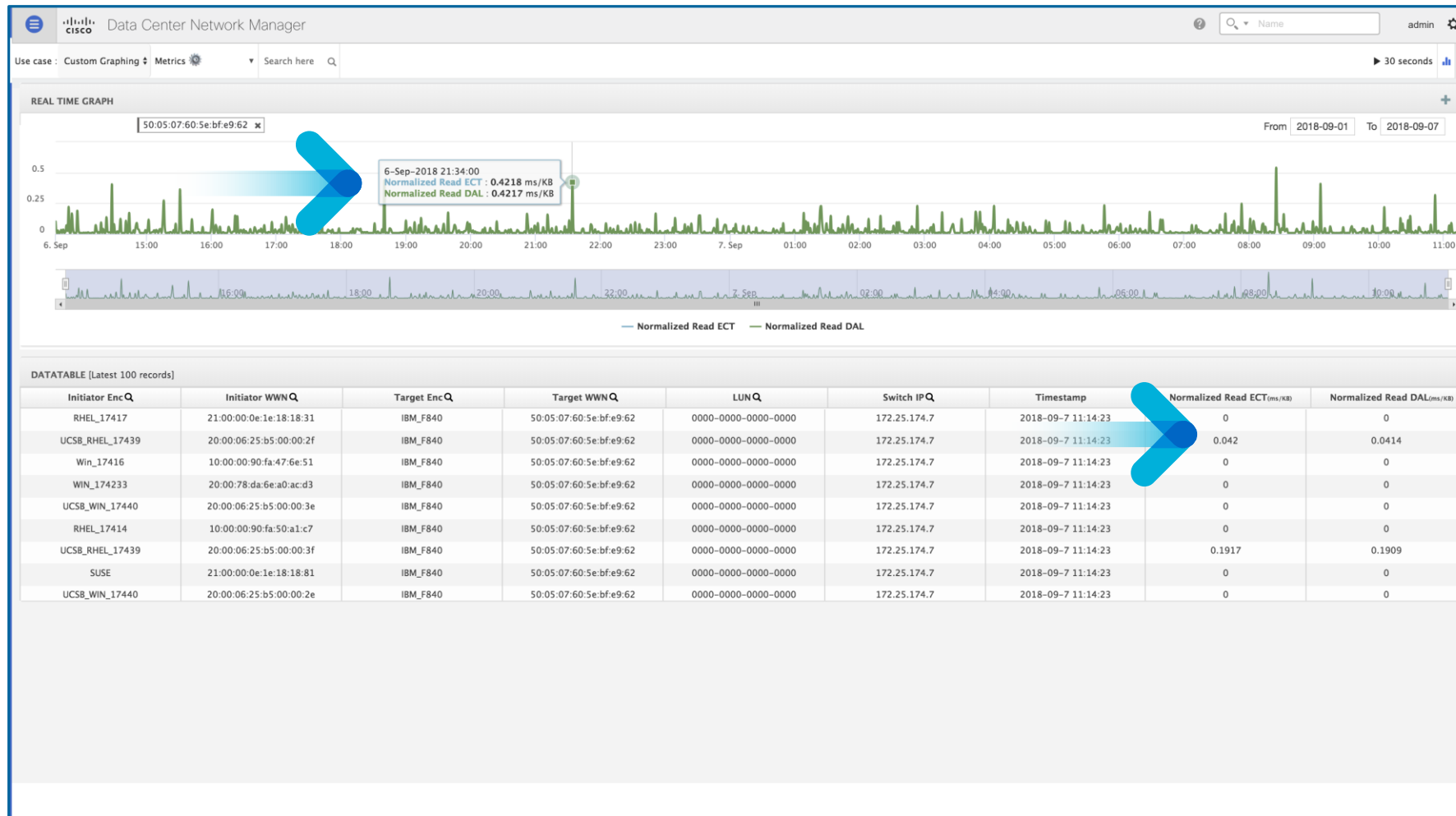
- Use Custom Graphing
- Input WWPN or host name of storage device
- Select Read or Write DAL and ECT
- If they track on a similar trend than the storage device is impacting ECT
- **Based on this graph all issues are pointing to a storage device causing slow response times**

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Accelerate with IBM Storage

Show me how busy my host is

- Use Host Enclosure view
- Find Host in list
- Select Interface in Topology Map
- Check each path under Initiator Target Pairs
- See both Flow based perf metrics and interface based metrics
- Errors, Aborts and other error related metrics provided

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Washington Systems Center - Storage

Show me how busy my storage device is

- Use Storage Enclosure view
- Find Storage Device in list
- Select Interface in Topology Map
- Check each path under Initiator Target Pairs
- See both Flow based perf metrics and interface based metrics
- Errors, Aborts and other error related metrics provided

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Accelerate with IBM Storage

Are there errors impacting my hosts?

- Use Host Enclosure view
- Find Host in list
- Select Interface in Topology Map
- Check each path under Initiator Target Pairs
- Errors, Aborts and other error related metrics provided

Accelerate with IBM Storage

Are there errors impacting my storage device?

- Use Storage Enclosure view
- Find Storage Device in list
- Select Interface in Topology Map
- Check each path under Initiator Target Pairs
- Errors, Aborts and other error related metrics provided

Accelerate with IBM Storage

What is my current I/O Block Size for my host?
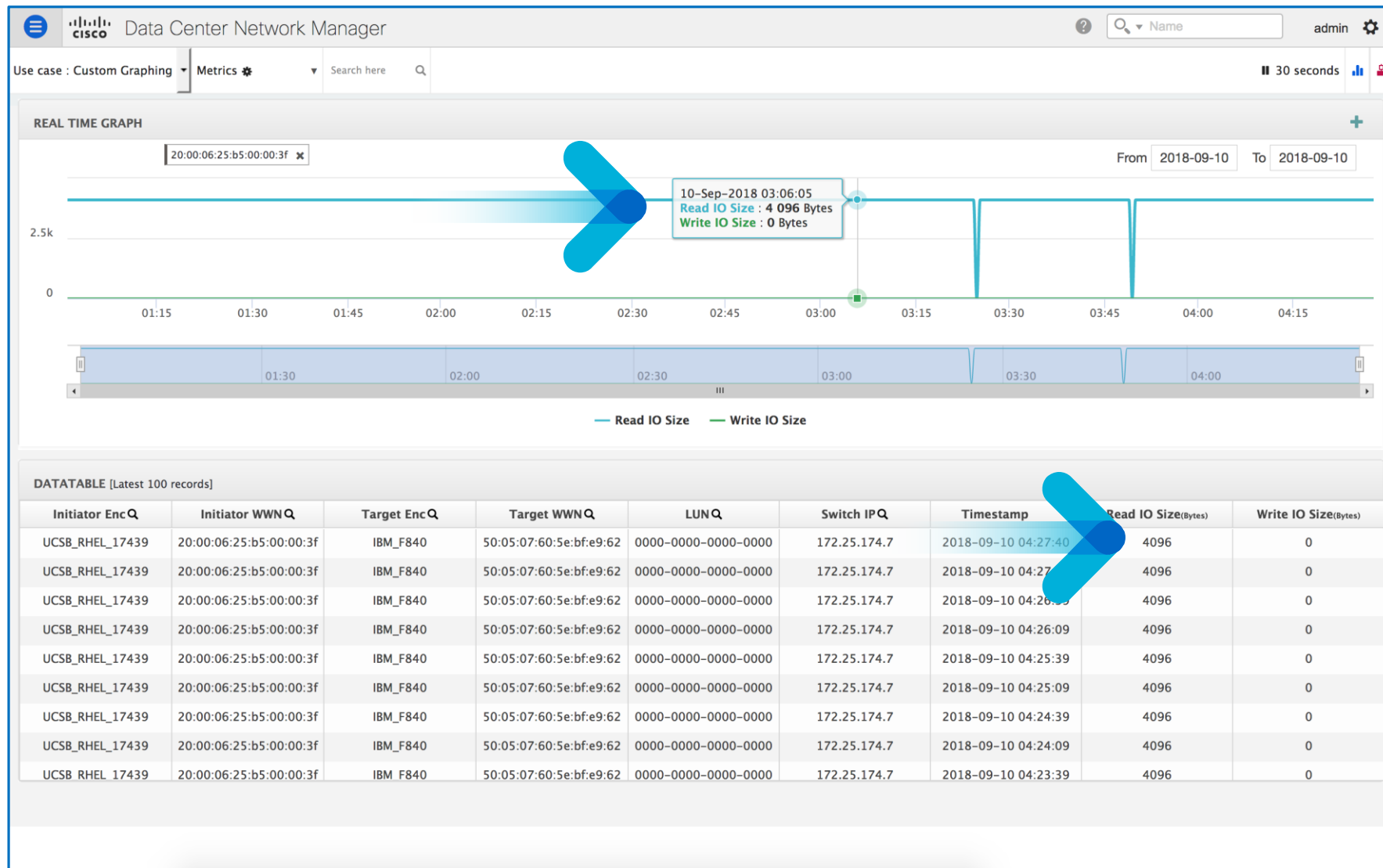
- Use Custom Graphing
- Input WWPN or host name of host
- Select Read or Write IO Size in Metrics
- See block size over time

For more information please check
https://www.youtube.com/watch?v=mIWq7Oko0mI&t=

Accelerate with IBM Storage

What is the Queue Depth for my host?

- Use Monitor/SAN Insights
- Select a Host on the left
- Scroll down and look at Read Active I/O and Write Active I/O
- number of outstanding IO's that are yet to be completed in real-time

Accelerate with IBM Storage

Can I see NVMe Analytics data?

- Use Monitor/SAN Insights
- Select NVMe Radio button
- See all NVMe Flows

Accelerate with IBM Storage

Show me what is using the most throughput?
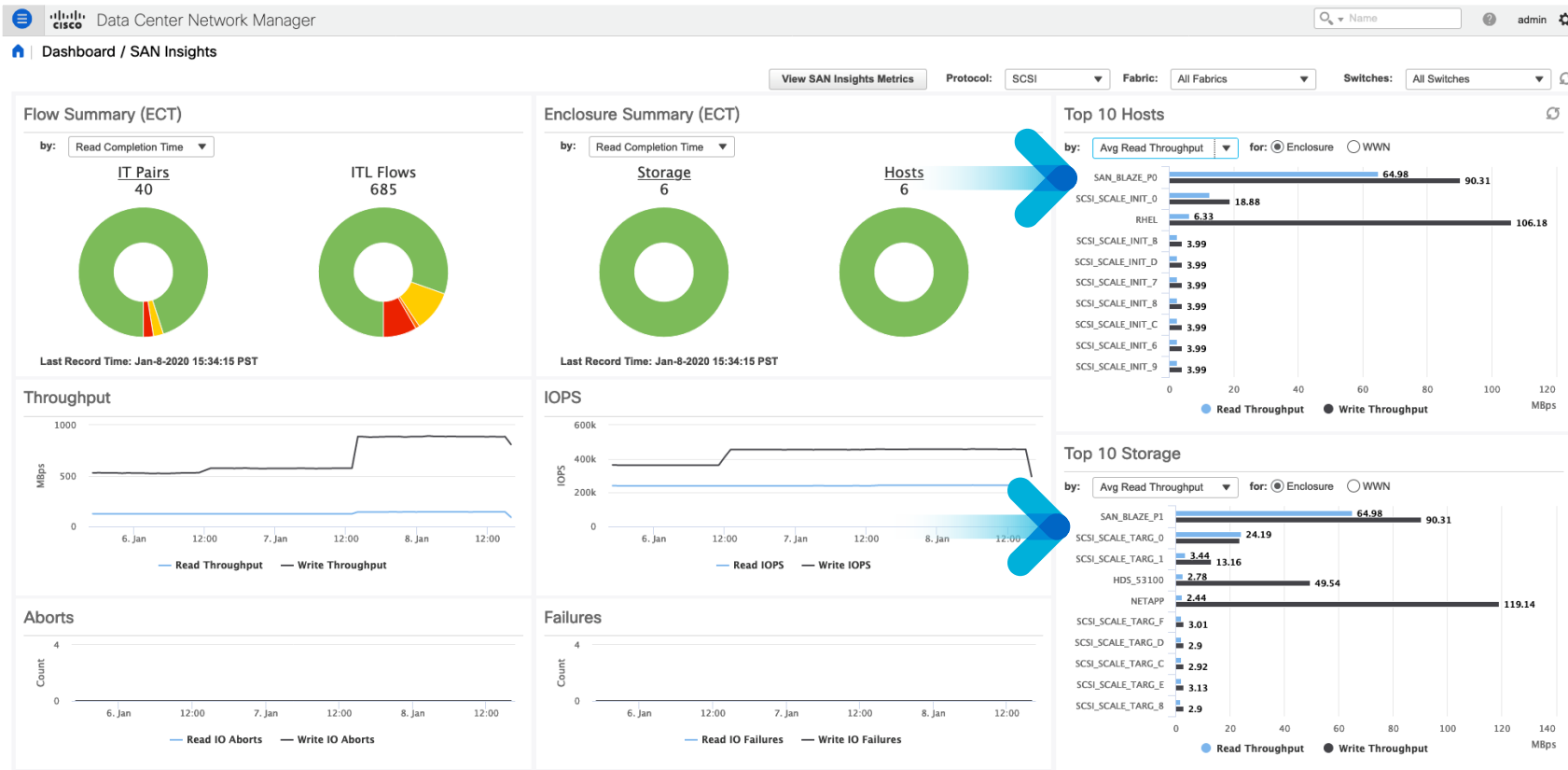
- Use Dashboard/SAN Insights
- Select Read or Write Throughput for Hosts
- Select Read or Write Throughput for Storage

Host Oversubscription in SAN

Host-edge switch needs more
time to transmit the frames

Response at line rate

64K read request

16G

16/32G

AFA- 1

Host-1

8G

NFA-2

8G

64K read request

Response NOT at line rate
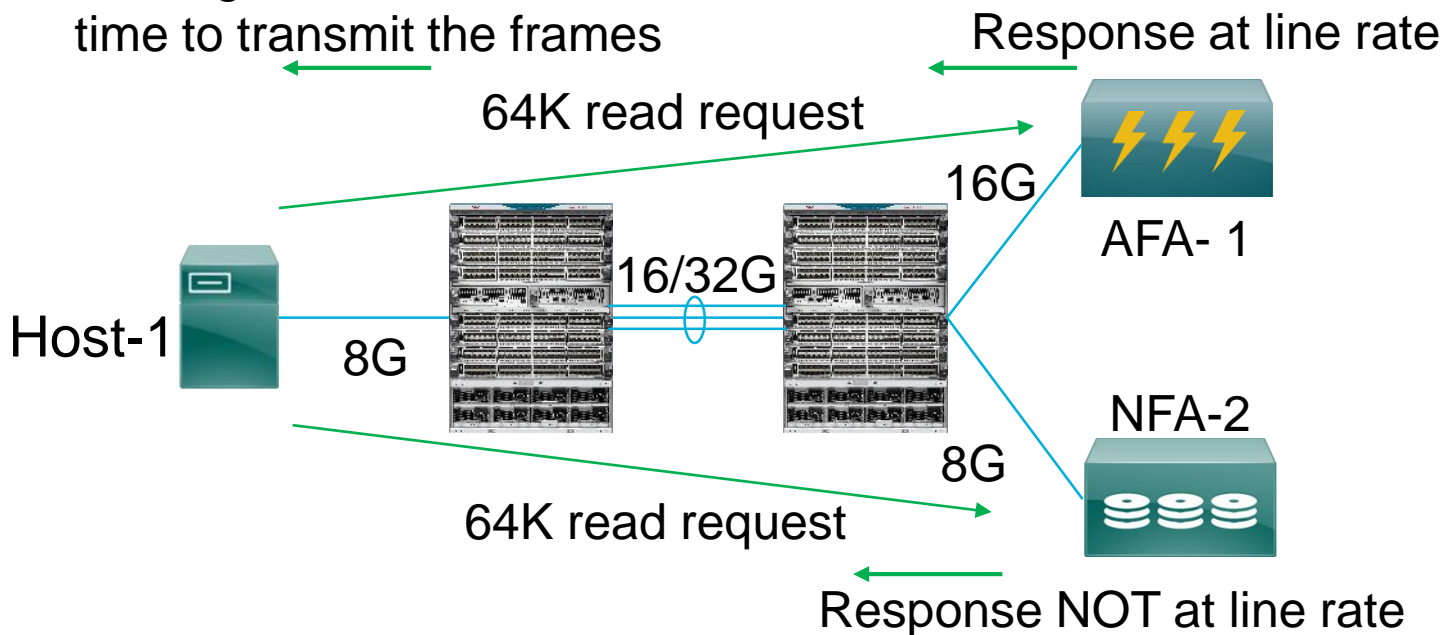
Time to transmit 32 full size FC frames

4G

8G

16G

40µs   80µs   160µs
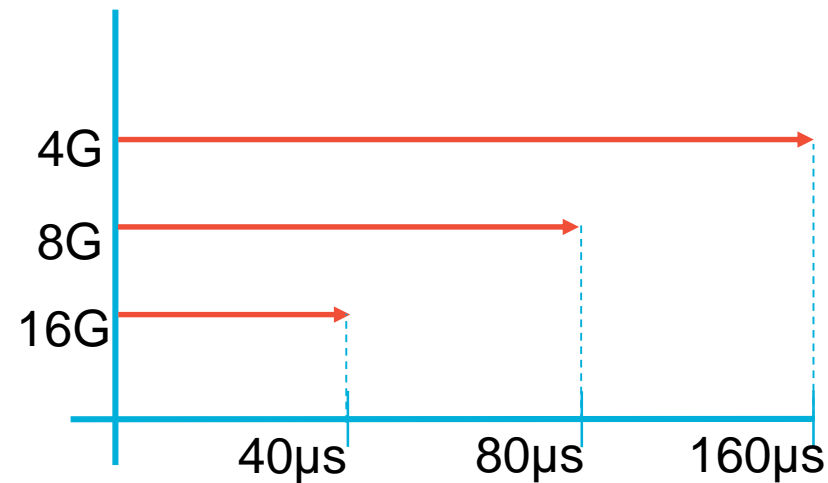
- AFA are extremely fast

- Frames are transmitted at line rate (of the directly connected interface)

- Backpressure is created if downstream links have less bandwidth

- Faster responses from AFA may not leave enough BW on host-edge for responses from NFA

Case study: Large European bank
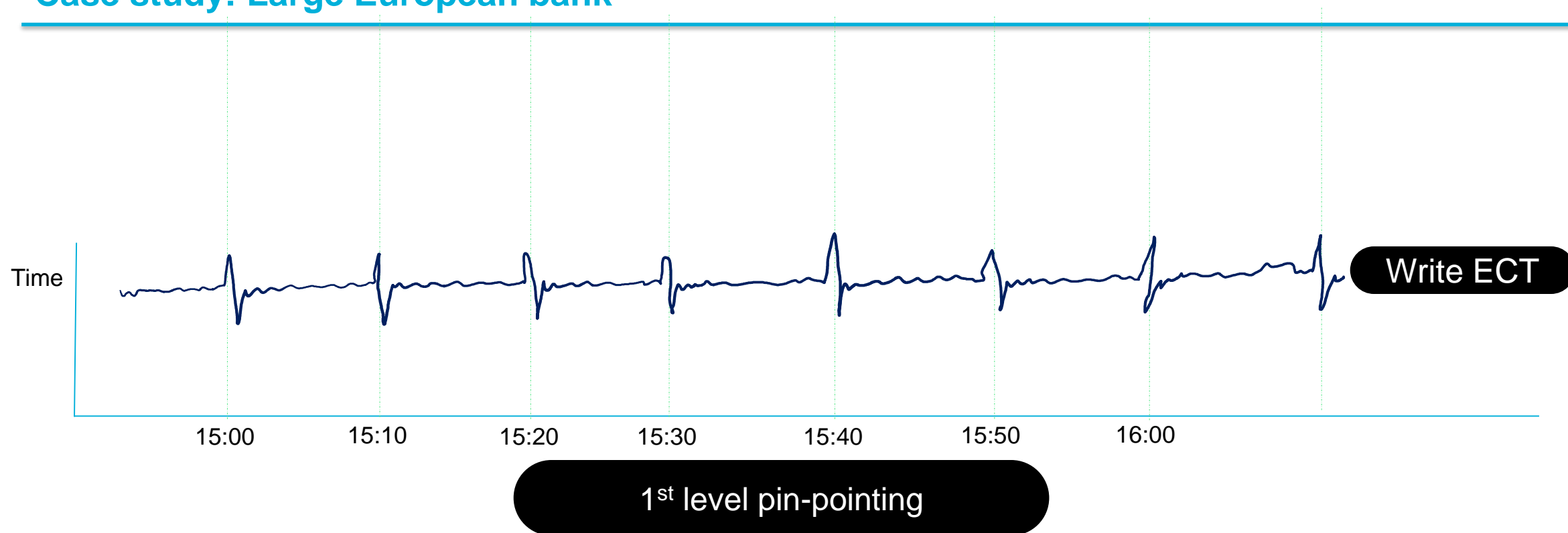
- Write ECT spike followed by dip
  - May be the cause of application performance issues
- Frequency – every 10 minutes

Case study: Large European bank
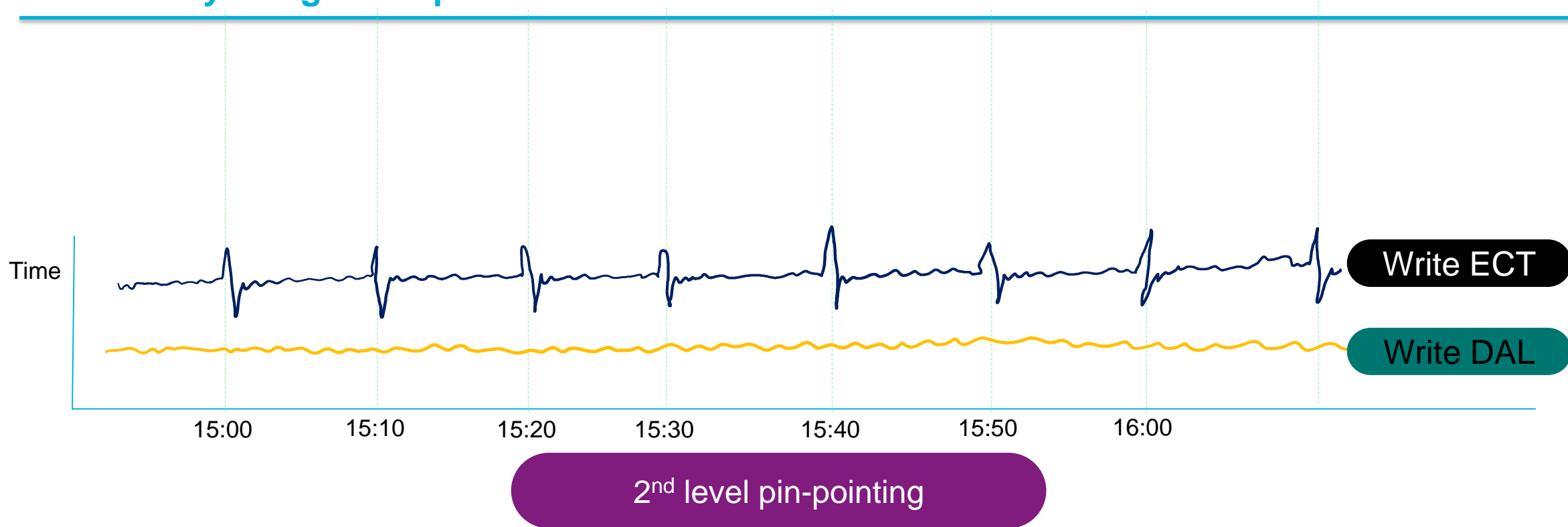
Time

Write ECT

Write DAL

15:00   15:10   15:20   15:30   15:40   15:50   16:00

2nd level pin-pointing

- Write ECT spike followed by dip. Frequency – every 10 minutes
- DAL is stable (no change)
  - Not a storage array issue

Case study: Large European bank
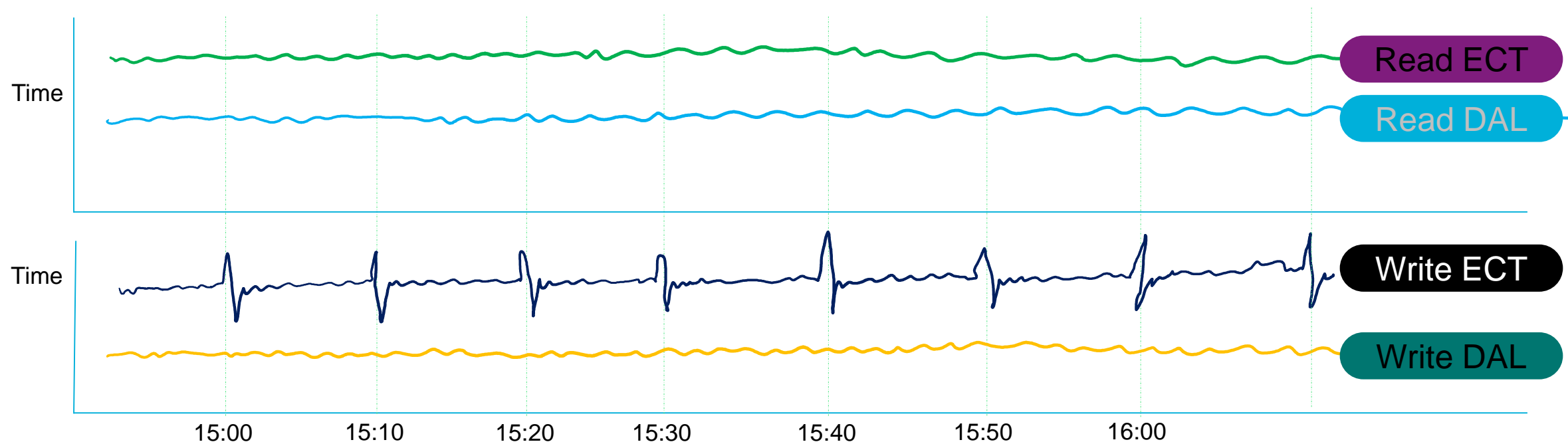
Read ECT

Read DAL

Write ECT

Write DAL

Time

Time

15:00 15:10 15:20 15:30 15:40 15:50 16:00

3rd level pin-pointing

Write ECT spike followed by dip. Frequency – every 10 minutes

DAL is stable (no change). Not a storage array issue

No changes in Read ECT and DAL. No fabric congestion observed.

No indication of fabric delay. Indication of delay within host.

Accelerate with IBM Storage

Case study: Large European bank

Read ECT

Read DAL

Write ECT

Write DAL

15:00   15:10   15:20   15:30   15:40   15:50   16:00

3rd level pin-pointing

Write ECT spike followed by dip. Frequency – every 10 minutes

DAL is stable (no change). Not a storage array issue

No changes in Read ECT and DAL. Not a fabric issue

Delay within host ➔ Resulted in detection of an unpatched Oracle app on host

Accelerate with IBM Storage

DCNM 11.3(1) Features – SAN Insights - Licensing

DCNM SAN Insights and SAN Telemetry Streaming Licensing

SAN Insights – This is a switch-based license that enables the switch to provide both on-board and outbound access to the telemetry data provided by the 32GB module with the added functions of DCNM visualization. This is provided by the enablement of the SAN_ANALYTICS_PKG on each switch.

SAN_ANALYTICS_PKG is a switch-based license (tied to the serial number of the switch). A single license allows you to enable ITL metric collection on all the ports on the licensed switch.
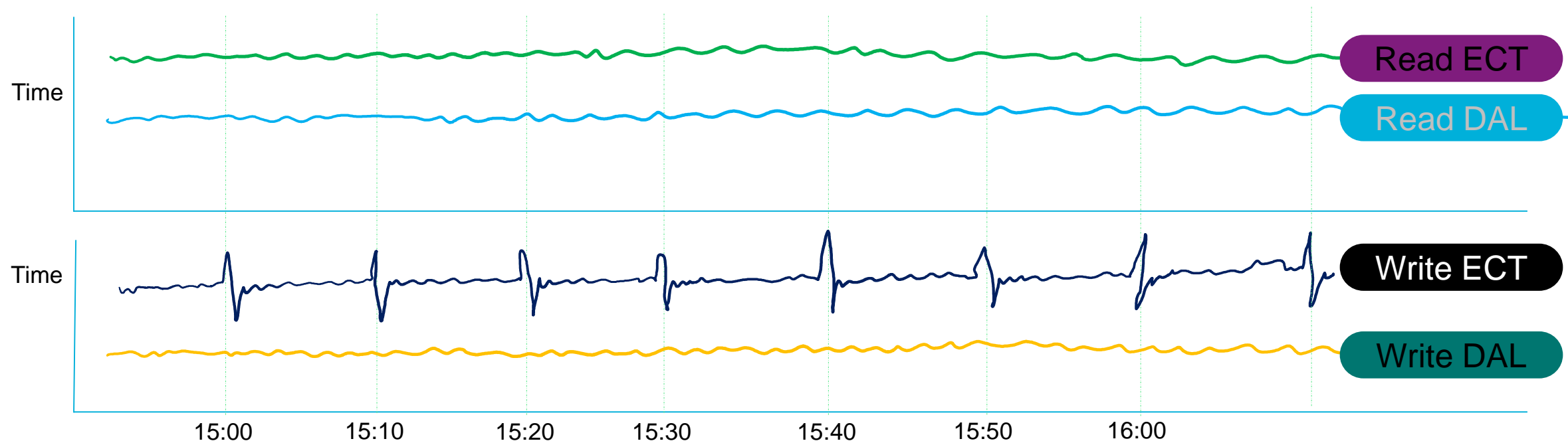
SAN_ANALYTICS_PKG is a time-based license which is offered in 3 years and 5 years term. At the end of the term of the license the switch will provide a message 30, 60 and 90 days before the expiration date. DCNM will also show the expiration of the license in the license views. You need to buy a similar license at the expiry of the term to continue using the functionality.

Since only the ITL flow metric collection is licensed but export is not, SAN_ANALYTICS_PKG is enough to export metrics to DCNM.

DCNM SAN Insights and SAN Telemetry Streaming Licensing

Switch export mechanism (can deliver non-SAN Analytics metrics)	3rd Party solutions VI, Splunk, customer built data platform	DCNM with SAN Insights (data can also be sent to VI, Splunk or other platforms)
		DCNM Advanced License **R**
	SAN Analytics License (Analytics Data) **R**	SAN Analytics License (Analytics Data) **R**
Free Streaming Telemetry	Streaming Telemetry	Streaming Telemetry

MDS 32Gbps portfolio
MDS9132T, MDS9148T, MDS9396T, MDS9700 (with 32Gbps line cards)

R= Required License

DCNM 11.3(1) Features – SAN Insights - Summary

Session summary

- Hardware and software-based solution

- Deploy anywhere in the fabric as long as 32Gbps line card or 32Gbps switch is in the path

- Uses industry standard formakes for compression and transport

- Supports CLI, pull and push methods for Analytics data delivery

- 60+ metrics

- DCNM is the destination for SAN Analytics data and is contain within a feature called SAN Insights

- SAN Analytics is licensed per switch (telemetry is free)

- Supports 20K flows per DCNM server

- Must be at NX-OS 8.3.1 and DCNM 11.2.1 minimum

Thank You!