Security, Integrity and Choices for NVMe over Fabrics

Nishant Lodha
Marvell
Agenda

- NVMe-oF®, the choices and the confusion
- Use Cases by Fabric
- Securing NVMe-oF
- Key Takeaways
NVMe-oF
Scaling our NVMe Requires a (Real) Network

- Many options, plenty of confusion
- Fibre Channel is the transport for the vast majority of today’s all flash arrays
  FC-NVMe Standardized in Mid-2017
- RoCEv2, iWARP and InfiniBand are RDMA-based but not compatible with each other
  NVMe-oF RDMA Standardized in 2016
- FCoE fabric is an option
- NVMe/TCP – is here! Standardized in NOV2018
RDMA Use Cases by Application

- HyperConverged Infrastructure
- Disaggregated Storage – NVMe-oF
- Disaggregated Storage – iSER
- Disaggregated Storage – SMB
- VM Migration
- RDMA Accelerated CEPHS
- NFS over RDMA
- Low Latency VMs
- NIC with RDMA
- iWARP
- RoCE
- RoCEv2
- RDMA Accelerated CEPHS
- CEPHS
- RDMA
- NFS over RDMA
- Low Latency VMs
NVMe-oF™ RDMA – potential challenges

**Infrastructure and Skillset change?**

- **Not Automatic**
- **Not Precise**
- **Not for everyone**

**Congestion**

- **Keeping the network ‘lossless’**

**RDMA/OEFD expertise**

**Skillset Requirements**

- **RNIC Upgrade Required**
- **RDMA Camps**

**Creates Islands**

**Backward Compatibility**

---

2019 Storage Developer Conference. © Marvell. All Rights Reserved.
Relationship Status: Microsoft and RoCE

After endless support calls with customers struggling with the configuration complexity of RoCE, we have updated our RDMA network recommendations:
docs.microsoft.com/en-us/windows-...

See the Microsoft Blog – comparing the RDMA types
NVMe Transport Performance Comparisons

iSCSI adds 82% more latency, Delivers fewer IOPS

NVMe-oF Latency Comparisons
4KB Random Reads Single Thread and IO Depth

NVMe-oF IOPS Comparisons
32KB Random Reads 8 Threads and 32 IO Depth

Latency (us)

IO Operations Per Second (IOPS)

Local NVMe Software iSCSI NVMe-oF RoCE Local NVMe iSCSI NVMe-oF 25GbE
FC-NVMe!

“NVMe” Over Fibre Channel

Low Latency, High Throughput

Transport NVMe Natively over Fibre Channel

Low Latency

Reliable, Secure, Available

Leverage Existing Investments in Fibre Channel

FC-NVMe T11 Committee

Ecosystem Ready
FCP vs. FC-NVMe

FCP vs. FC-NVMe: 4KB RD & 4 Jobs / DP to 1 LUN/NS per port

FC-P vs. FC-NVMe: 4KB RD & 4 Jobs / DP to 1 LUN/NS per port

FCP

FC-NVMe

FC-NVMe Scales in performance
## Use Cases by Fabric

No one size fits all!

<table>
<thead>
<tr>
<th>DAS, HPC, AI/ML</th>
<th>Enterprise Applications</th>
<th>All Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVMe/RDMA (Ethernet)</td>
<td>FC-NVMe (Fibre Channel)</td>
<td>NVMe/TCP (Ethernet)</td>
</tr>
</tbody>
</table>

- **DAS, HPC, AI/ML**
  - NVMe/RDMA (Ethernet)
  - Logos are indicative of workload characteristics only.

- **Enterprise Applications**
  - Leverage existing infrastructure. Reliability is key.

- **All Applications**
  - Simplicity is key. Balance of performance and cost.

Performance at the cost of complexity

2019 Storage Developer Conference. © Marvell. All Rights Reserved.
NVMe/TCP
NVMe-oF: NVMe/TCP

- What: Defines a TCP Transport Binding layer for NVMe-oF
- Promoted by Facebook, Google, Intel, Marvell etc.
- Not RDMA-based, Standardized on 15NOV18
- Why:
  - Enables adoption of NVMe-oF into existing datacenter IP network environments that are not RDMA-enabled
NVMe-oF Driver Stack

Target

Initiator

2019 Storage Developer Conference. © Marvell. All Rights Reserved.
Offloading NVMe/TCP

Host

NVMe Host (nvme)

Send Queue

Completion Queue

Read/Write Data Buffer

IO stack

TCP/IP Stack

TCP Header

TCP Payload

TCP Header

TCP Payload

L2 Driver (qede)

Marvell NIC
L2 Firmware

Target

NVMe Target (nvmet)

Send Queue

Read/Write Data Buffer

Completion Queue

NVMe Target Transport (qedn)

Marvell NIC
NVMeTCP Target Firmware

TCP Offload
Accelerating NVMe/TCP

4K Read IO - 1 pending latency [usec]

4K Write IO - 1 pending latency [usec]
Significant CPU Savings with NVMe/TCP Offload
Security FC-NVMMe
# Drivers for FC-NVMe Security

## Security and Privacy Sensitive Verticals

<table>
<thead>
<tr>
<th>Vertical</th>
<th>Healthcare</th>
<th>Financial</th>
<th>Government</th>
<th>Defense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy/Regulatory</td>
<td>HIPAA</td>
<td>GDPR</td>
<td>ISO27001</td>
<td></td>
</tr>
<tr>
<td>New Deployment Use Cases</td>
<td>DR</td>
<td>Secondary Cloud Storage</td>
<td>Hybrid Cloud</td>
<td>Multi-tenant Data Centers</td>
</tr>
</tbody>
</table>
Cost of a data breach and Recent events

None of these breaches have been directly attributed to Fibre Channel

Source: IBM Security
2019 Storage Developer Conference. © Marvell. All Rights Reserved.
Isn’t FC Secure Already?

**Trusted Storage Interconnect for Decades**

<table>
<thead>
<tr>
<th>Physical Security</th>
<th>• Data Centers are physically secured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segregation</td>
<td>• Fibre Channel SANs are segregated networks</td>
</tr>
<tr>
<td>Partitioning</td>
<td>• FC Zoning ensures fabric partitioning</td>
</tr>
<tr>
<td>Masking</td>
<td>• LUN masking restricts access to specific LUNs</td>
</tr>
<tr>
<td>Management</td>
<td>• Out-of-Band Management (IP) is secure, OS Controls</td>
</tr>
</tbody>
</table>
Yes, But…

- New Data Center Architectures bring new threats
  - Distributed data centers - Remote replication and DR backups may be accessed by different users over Fabrics that span several sites
  - Multi Tenant data centers – Need to segregate and protect data traversing the same wire
- Increasing scale of FC SANs
  - Networks can be misconfigured
  - Fabric configuration databases are shared, have WKAs
- Existing mechanisms may not be enough
  - Switches are the sole entity that grant/deny access
    - Authorization based
    - “Segmentation” tools being used to implement “Security”
    - Soft zoning, LUN Masking
Potential DC Storage Security Threats

Sniffing Storage Traffic
Storage Masquerading
Data Corruption
Session Hijacking

Mitigated by Fibre Channel SAN Security
FC-SP-2: What and Why?

Why? : Need to transition SANs from Authorization and segmentation based FC security to authentication and encryption based security!

What? FC-SP-2 is a ANSI/INCITS standard (2012) that defines protocols to –

- Authenticate Fibre Channel entities
- Setup session encryption keys
- Negotiate parameters to ensure per frame integrity and confidentiality
- Define and distribute security policies over FC

Designed to protect against several classes of threats
<table>
<thead>
<tr>
<th>Authentication Infrastructure</th>
<th>Authentication</th>
<th>Security Associations</th>
<th>Crypto Integrity Confidentiality</th>
<th>Authorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secret, certificate, password and pre-shared key based architecture</td>
<td>Protocol to assure identify of communication entities, negotiation of security requirement and protocol</td>
<td>Protocol to establish shared key between communicating entities, Based on IKEv2 (RFC4595)</td>
<td>Frame by frame encryption, replay protection, origin authentication, ESP_Header or CT_Authentication</td>
<td>Fabric policies that control which entities can connect with each other, management access to the fabric</td>
</tr>
</tbody>
</table>
**FC-SP-2 ESP_header**

- **ESP_header** (optional) is a layer 2 security protocol that provides
  - Origin authentication
  - Integrity
  - Anti-replay protection
  - Confidentially

- Encapsulating Security Payload (ESP) is defined in RFC 4303
- FC-FS-3 defines optional headers for Fibre Channel, FC-SP defines how to use ESP in Fibre Channel
- Similar protections exist for CT_Authentication
Silicon Root of Trust

Protecting the Integrity of Fibre Channel Firmware

Signing Fibre Channel firmware

Verifies firmware signature using Silicon Root of Trust

Process at Factory/Manufacturing
Key Takeaway
Making the right “fabric” choice!

Not “just” about “fabrics” performance

Culture and Install Base

Use Cases and Security
That’s it!