STORAGE 101:
PAST, PRESENT, FUTURE OF STORAGE.
HARD DRIVE TECHNOLOGIES

- **PATA DRIVES**: Parallel ATA (Parallel Advanced Technology Attachment or **PATA**)
- **SATA DRIVES**: Serial ATA (Serial Advanced Technology Attachment or **SATA**)
- **SAS DRIVES**: Serial Attached SCSI (**SAS**)
- **FATA DRIVES**: Fibre Attached Technology Adapted (**FATA**)
- **SSD**: Solid-State Drive (**SSD**)
- **NVMe**: Non-Volatile Memory Express (**NVMe**)

---

GETTING VIDEO TO THE DESKTOP

• In the 90’s Lockheed Martin was looking to bring video to the desktop.

• What technology did they use?
  -Fibre Channel

• What went wrong?
  -File churn that created lots of extents

• What evolved from the project?
  -Stornext
KEEPING VIDEO TO THE DESKTOP NAS OR SAN

• Starting around the year 2007 post productions
  - studios needed to make a decision NAS or SAN

• Why did we need to use a SAN?
  - Many decided with the SAN as it offered speeds around 400 MB/sec

• Why not use a NAS?
  - Ethernet only offered 1GbE

• What are the disadvantages of a SAN?
  - Difficult to DPX, Fragmentation, many failure points
NVMe vs. SSD

- SSD debuted in 2009
- What’s held SSD back?
  - 6 Gb/SATA (550MB) cable or 12Gb/SAS (1,100MB)
- Why is NVMe so much faster?
  - Connects directly to the PCIe3 X4 bus

- How will we use NVMe in M&E?
  - 2TB delivering 3,000 MB/sec
- Will NVMe be the future?
  - No, it’s the present
THE NUMBERS

• How many drives to fill a 100GbE pipe?
  • SATA HDDs – 250
  • SATA SSDs – 24
  • SAS SSDs – 10
  • NVMe – 4
WHAT IF SOMEONE WENT EXTREME?

• 36 NVMe Drives in 1 RU
• 10 Millions IOPS
• 576 TB
The technology was first defined by the IEEE 802.3ba-2010 standard. 40GbE and 100GbE made their debut.

What can we do with this technology:
- 100GbE from NAS to 10GbE Switch
- Support Image Sequences (DPX, TIFF, OpenEXR) and demanding bitrate(s)
THANK YOU

STEVE GRAPPONE
SYSTEMS ARCHITECT, DIGITAL GLUE
steve.grappone@digitalglue.com
Mellanox Ethernet Switches Power All Platforms

- **3,000+** unique customers worldwide
- **Over 1M** ports deployed
- **OEM’d** by global server OEMs
- **Open** Network Operating Systems

![Diagram showing various industries and technologies supported by Mellanox switches](image-url)
Mellanox Accelerates Businesses’ Success!

Reducing CPU Usage Drastically with Advanced Network Offloads
IP-Broadcasting Applications and Many More on 25/100GbE!

100K subscribers from a Single Caching Server!
Netflix Open Connect Appliance with Mellanox ConnectX-4

Mellanox 40GbE-based Network to Support Clustered File Systems
Live Broadcast Video Editing

Mellanox 10/40GbE-based Storage Network
Offline Broadcast Editing

Only End-to-End 100GbE Network to Support Maximum of 32 Tiles
Enabling IP-Broadcast Multiviewer & Signal Analytics
25G & Above is Fastest Growing Ethernet Segment

Mellanox is #1 with 67% share in the Nx25G segment

Key Drivers of 25G+ Networking:
- Cloud, Big Data, Faster Storage, Virtualization, AI & Machine Learning

25G+ Adoption Cycle:
- #1: US Hyperscalers
- #2: BAT in China
- #3: Third wave of cloud, telco, and storage in US, ASIA, and Japan

Source: Crehan Long Range Forecast Server-Class Adapter & Lom Controller (10G and greater), Jan ‘18
Storage Landscape

PRIMARY STORAGE
- Traditional SAN
- Only 20% of capacity

SECONDARY STORAGE
- 80% of capacity
- Rapid growth
- Diverse data types
- Scale-out, Ethernet-based
- Tiered data

File Shares
Archiving
Test/Dev
Backups
Analytics
Cloud
NVMe Growth Projection – $60B by 2021

- NVMe-oF adapter shipments will exceed 1.5M units by 2021
  - $750M @ $500/per

- This does not include ASICs, Custom mezz. cards, etc. inside AFAs and other storage appliances
### Storage Networking Background: Fibre Channel & Ethernet

<table>
<thead>
<tr>
<th>Feature</th>
<th>1997</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>1 G</td>
<td>8/16/32 G</td>
</tr>
<tr>
<td>Supports</td>
<td>Block</td>
<td>Block</td>
</tr>
<tr>
<td>Lossless</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cost</td>
<td>High $$$$$</td>
<td>Medium $$$</td>
</tr>
<tr>
<td>Cloud / HCI</td>
<td>No / No</td>
<td>No / No</td>
</tr>
<tr>
<td>Vendors</td>
<td>Several</td>
<td>Many</td>
</tr>
<tr>
<td>SDS / Scale-out</td>
<td>No / No</td>
<td>Rare / No</td>
</tr>
</tbody>
</table>

#### 1997
- **Feature**
  - Bandwidth: 1 G
  - Supports: Block
  - Lossless: Yes
  - Cost: High $$$$$
  - Cloud / HCI: No / No
  - Vendors: Several
  - SDS / Scale-out: No / No

#### 2017
- **Feature**
  - Bandwidth: 8/16/32 G
  - Supports: Block
  - Lossless: Yes
  - Cost: Medium $$
  - Cloud / HCI: No / No
  - Vendors: 2 / 2
  - SDS / Scale-out: Rare / No

### Yesterday: Storage Network = FC
- Fibre Channel offered best performance
- All interesting storage was tier-1 block
- No cloud or hyperconverged

### Today: Both FC & Ethernet for storage networks
- FC option for Primary Block Storage
- Ethernet only option for all Primary & Secondary Storage (Block, Object, NAS, Cloud, Hyperconverged, Big Data)
Problem Solved With an Ethernet Storage Fabric

Everything a Traditional SAN Offers but ...

Faster, Smarter, & Less Expensive

FAST
• Highest Bandwidth
• Lowest latency
• RDMA and storage offloads
• Native NVMe-oF Acceleration

SMART
• Automated Discovery & Provisioning
• Security & Isolation
• Monitoring, Management, & Visualization
• Storage-aware QoS

EFFICIENT
• Optimized Form Factors
• Just Works Out of the Box
• Flexible: Block, File, Object, HCI
• Affordable: SAN without the $$
Traditional vs. ESF—Where to Draw the Line?

Legacy DC – FC SAN

Modern DC – Ethernet Storage Fabric
Why Not Just Use Your Existing Switches?

Existing switches might not be designed for storage
- Are there enough available ports?
- Will the switches support new speeds – 25, 40, 50, 100GbE?
- Blocking? How many hops? How much latency?
- Are switches too big for storage or HCI clusters?

Need to Look To the Future
- Today’s network purchases must last 3-5 years
- Fast enough for flash
- Flexible for future applications
How to Enable an Ethernet Storage Fabric

Ethernet Storage Fabric needs dedicated ESF switches

- Performance
- High Availability
- Simple
- Automated
- Scalable
- Cost Efficient

- 2 Switches in 1RU
- Storage/HCI port count
- Zero Packet Loss
- Low Latency
- RoCE optimized switches (NVMe-oF)
- NEO for Network automation/visibility
- Native SDK on a container
- Cost optimized
- NOS alternatives
High performance, Multi-Tenant, Scalable Data Centers
Spectrum is Built for ESF

**Congestion Management**

<table>
<thead>
<tr>
<th>Packet size</th>
<th>Max Burst Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>64B</td>
<td>5.2</td>
</tr>
<tr>
<td>512B</td>
<td>8.4</td>
</tr>
<tr>
<td>1.5B</td>
<td>9.6</td>
</tr>
<tr>
<td>9KB</td>
<td>9.7</td>
</tr>
</tbody>
</table>

- **Microburst Absorption Capability**
  - Spectrum: 5.2, 8.4, 9.6, 9.7
  - Tomahawk: 0.3, 0.9, 1.0, 1.1

**Fairness**

- Broadcom
  - Packet Size (Bytes): 64, 128, 164, 200, 1518
- Spectrum
  - Packet Size (Bytes): 64, 128, 164, 200, 1518

**Avoidable Packet Loss**

- Broadcom
  - Packet Loss: 64, 128, 164, 200, 1518
- Spectrum
  - Packet Loss: 64, 128, 164, 200, 1518

**Guaranteed Low Latency**

- Average FIFO Latency (nanoseconds)
  - Spectrum: 3,334ns

**For more information, visit www.zeropacketloss.com**
Ethernet Storage Fabric Must Support RoCE

- RoCE is RDMA over Ethernet
  - Bypass CPU
  - Increase efficiency

- RoCE has growing support
  - All Operating Systems
  - Many Applications
  - More Storage Arrays

- Common Use Cases
  - Storage (incl. NVMe over Fabrics)
  - Big Data
  - Video processing
  - Machine Learning / AI
RoCE Support Done Right!

Application Blocked by the Switch

Pause time (Microseconds)

Time/Seconds

Application runs Smoothly

Other Switches
Summary: Choosing an ESF Switch

- Better Performance
- Right Form Factors
- Easy Setup
- Better Visibility
- Tested End-2-End
Maximizing Efficiency of Software Video Streaming

- Lower CPU Utilization: 80% Less CORES*
- Less Servers & Network: 40% Less*
- Lower Power, Cooling & Space: 30% Less*
- Highest Throughput & Quality: 100 Gb/s PER PORT 0% Packet LOSS

2110-21 Compliance
Packet Paced in NIC hardware

* Comparing to Kernel based application performance.
Ultra HD Resolution Demands Hardware Performance

- Raw uncompressed video bandwidth requirement exceeds 100Gb/s
  - Bandwidth for 1 video stream with 8K UHD high frame rate
- Packet pacing in software is not SMPTE 2110-21 compliant
  - Sub 1us Inter Packet Gap (IPG) is not achievable with software based solutions

<table>
<thead>
<tr>
<th>Resolution</th>
<th>bpp</th>
<th>Frame Rate</th>
<th>IPG (nSec)</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full HD</td>
<td>20</td>
<td>25fps</td>
<td>11,000</td>
<td>Can be achieved with software packet pacing</td>
</tr>
<tr>
<td>4K UHD</td>
<td>20</td>
<td>60fps</td>
<td>1,150</td>
<td>Cannot be achieved with software packet pacing</td>
</tr>
<tr>
<td>4K UHD</td>
<td>24</td>
<td>120fps</td>
<td>469</td>
<td>Cannot be achieved with software packet pacing</td>
</tr>
<tr>
<td>8K UHD</td>
<td>36</td>
<td>60fps</td>
<td>156</td>
<td>Cannot be achieved with software packet pacing</td>
</tr>
</tbody>
</table>

Inter Packet Gap (IPG) Requirements for UHD

Bandwidth Requirements for UHD
Rivermax Key Features

**Packet Pacing**
- Leverages ConnectX-5 hardware based Packet Pacing
- SMPTE ST 2110-21 compliance at any bit rate
- No dependency on CPU Strength, OS interrupt level or Application

**Kernel Bypass**
- Reduced Kernel overhead with direct network adapter access
- Selective bypass – enables to select traffic bypasses and which flows to kernel
- Reduced latency
- Reduced CPU utilization
- Increased throughput

**Packet Aggregation**
- Application at Frame/Line(s) level
- Receive: fully assembled frame/lines(s) in memory
- Transmit: synchronously transmit packet paced full frames/lines (/chunks)

Based on ConnectX-5 Technology
Thank You
NEXT-GEN CREATIVITY: BEYOND THE TECH
PRESENTED BY NICK ANDERSON
MAKING 4K WITH HD TECH

STORAGE LIMITATIONS
• 1 & 10 GbE
• Non-realtime Image Sequences
• Compressed RAW for real-time

WORKSTATION LIMITATIONS
• Intermediate & Proxy Codecs
• Prores, Cineform, and DNx
• Optimized hardware acceleration
THE GREAT PLATEAU

HOW FAR IS ENOUGH?

• 2K-8K Resolution (based on viewing distance)
• Physiological Color Gamut
• 21 Stops of Dynamic Range
  7 stops visible at one time
• 12-bit
SO... WHAT CAN WE DO NOW?

- End-to-end Online Workflows
- Seamless Collaboration
- Artificial Intelligent Tools
- Masters as Deliverables
- …and then it gets spooky
END-TO-END ONLINE

• No waiting for transcodes for source workflows
• Source to master quality instead of proxy (16-bit OpenEXR)
• More eyes looking at content to catch issues
• Non-destructive processing throughout the pipeline
• ACES Color Transforms
SEAMLESS COLLABORATION

• NLEs with collaboration
  - DaVinci Resolve – Edit, Color, Audio, and VFX in one application
  - Adobe – Dynamic Linking with Premiere, After Effects, and Audition
  - AVID – Bin Sharing and Fusion Connect to DaVinci Resolve

• Entire team working in Parallel with the same assets
• No conforms, relinking, or project version tracking
ARTIFICIAL INTELLIGENCE FOR POST

- AI-friendly storage enables new workflows
  1. AI-based auto-tagging and transcription
  2. Creating new pixels and bits (Nvidia)
  3. Machine learning for automation

- AI designed to enable non-technical creatives
MASTERS AS DELIVERABLES

- High-bandwidth Fibre and 5G for content delivery
- IMF with encoding on the fly
- Metadata driven decode
- Display transforms
THE COMING CREATIVE REVOLUTION

• Brain to computer interfaces
• P2P streaming
• Integrated AI creative tools
• Blockchain asset management
• Augmented, virtual, and mixed reality
• Femto-photography
THANK YOU

NICK ANDERSON
PRODUCT MANAGER, DIGITALGLUE
nick.anderson@digitalglue.com