FASP 0

Aspera's FASPest FASP

Charles Shiflett

bear @ us . IBM . com

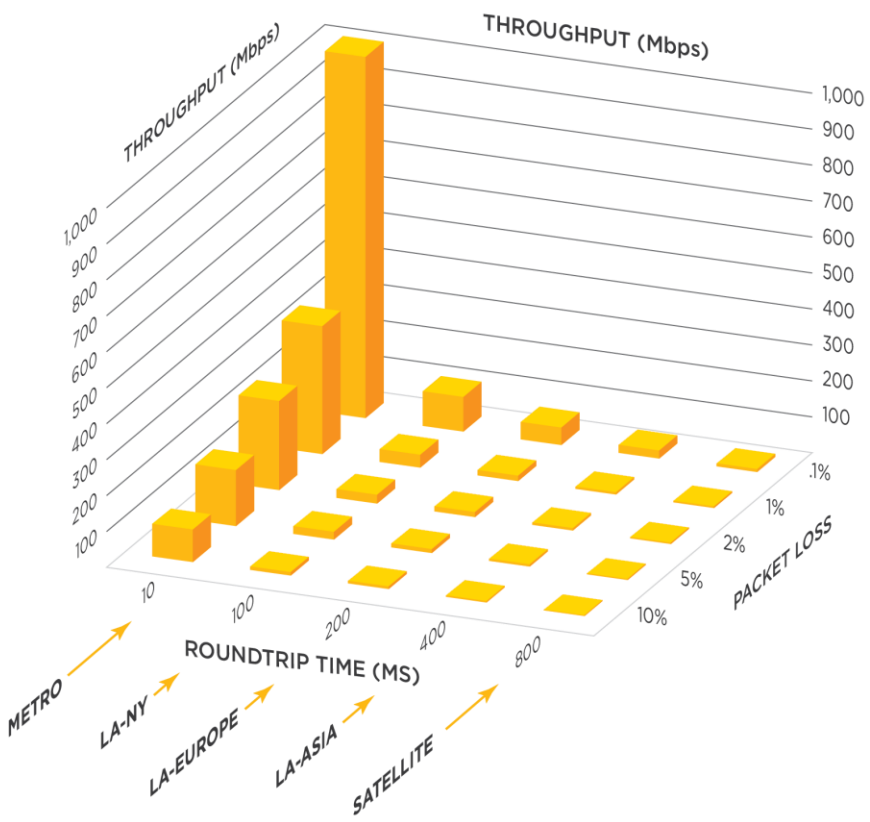
Creating next-generation transport technologies

that move the world's digital assets at maximum speed,

regardless of file size, transfer distance and network conditions

- Founded in 2004 with a focus on solving data transmission across the WAN
- Real networks, Any platform, Any data
- With development branch of FASP 0 performance shown at 100gbit/s using a Single node w/ encryption, regular packets, DPDK acceleration.
- Performance expected to improve as PCIe interconnects improve

CHALLENGES WITH TCP AND ALTERNATIVE TECHNOLOGIES



Note: Table displays throughput degradation of TCP transfers on a 1Gbps network as estimated round trip time and packet loss increases with distance.

Distance degrades conditions on all networks

- Latency (or Round Trip Times) increases
- Packet loss increases
- Fast networks are just as prone to degradation

TCP performance degrades severely with distance

- TCP was designed for LANs and does not perform well over distance
- Throughput bottlenecks are severe as latency & packet loss increase

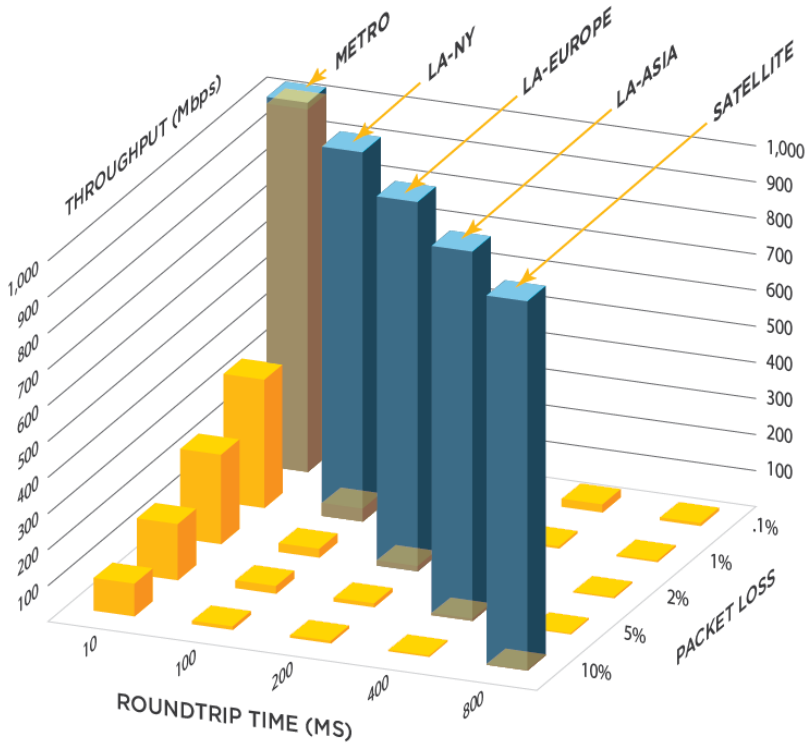
TCP does not scale with bandwidth

- TCP designed for low bandwidth
- Adding more bandwidth does not improve throughput

Alternative technologies

- TCP-based - Network latency & packet loss must be low to work well
- UDP blasters – Inefficient use of bandwidth leads to congestion
- Modified TCP – Does not scale well on high-speed networks
- Data caching - Inappropriate for many large file transfer workflows
- Data compression - Time consuming & impractical for some file types

FASP® – HIGH-PERFORMANCE DATA TRANSPORT



Maximum transfer speed

- Optimal end-to-end throughput efficiency
- Transfer performance scales with bandwidth independent of transfer distance and resilient to packet loss

Congestion avoidance and policy control

- Automatic, full utilization of available bandwidth (fair play)
- On-the-fly prioritization of transfers
- Set caps on bandwidth allocation for transfers

Uncompromising security and reliability

- Secure, SSH user/endpoint authentication
- AES-128 to 256 cryptography of every packet in transit
- Encryption at rest (EAR) requires second password
- FIPS 140-2 compliant, built on the open SSL libraries
- Automatic resume of partial or failed transfers

Scalable management, monitoring and control

- Support highly concurrent transfers
- Real-time progress, performance and bandwidth utilization
- Detailed transfer history, logging, and manifest

Note: The relative bandwidth utilization for FASP transfers over a 1 Gbps network are immune to latency (distance) with very little effect from packet loss.

FASP® – PERFORMANCE BREAKTHROUGH

MOVING A 10GB FILE		Across US	US - Europe	US - Asia
Legacy Transport	100 Mbps	10-20 Hours	15-20 Hours	Impractical
	1 Gbps			
	10 Gbps			
Aspera FASP®	100 Mbps	14 Min	14 Min	14 Min
	1 Gbps	1.4 Min	1.4 Min	1.4 Min
	10 Gbps	8.4 Sec	8.4 Sec	8.4 Sec

Location Agnostic

FASP transfer speeds remain virtually constant as transfer distances increase while FTP speeds dramatically decrease

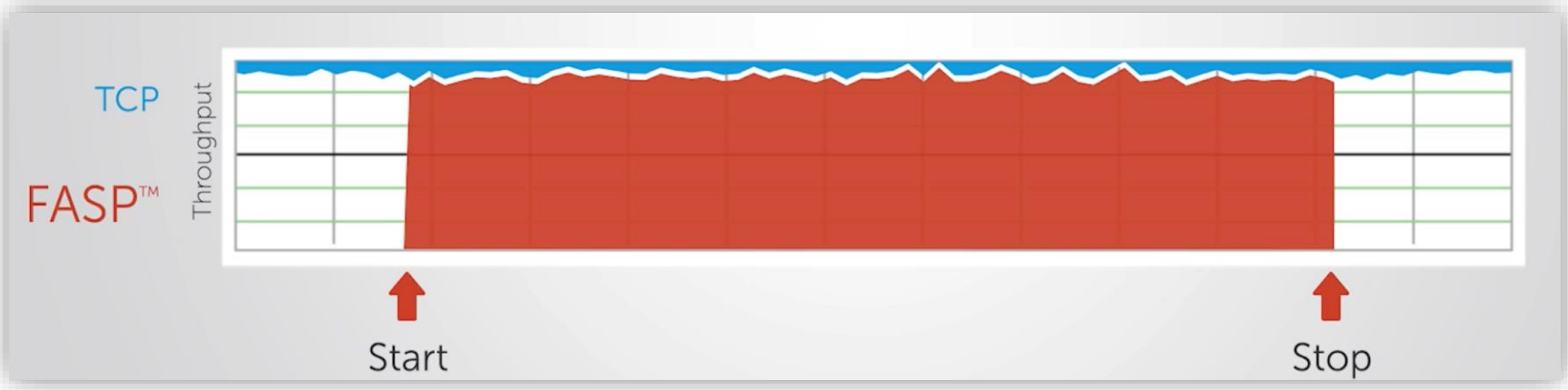
Predictable & Reliable

FASP transfer times decrease linearly as bandwidth increases. However, FTP transfer times don't improve with bandwidth

Versatile

Supports massive file sizes (500 GB+) as easily as very large sets (millions) of small files

FASP – MAXIMUM SPEED WITH NO NETWORK SATURATION



Extraordinary bandwidth control that doesn't saturate the network

- Automatic detection & full utilization of available bandwidth with "fair" policy protection of other network traffic
- Allows "bursts" in TCP traffic and reclaims unused bandwidth as it as it becomes available

FASP ADAPTIVE RATE CONTROL

Real-time prioritization of transfers

- On-the-fly, per flow, user and job prioritization of transfers
- Concurrent transfers adjust bandwidth on the fly, allocating available bandwidth based on transfer priority

System-wide monitoring and reporting

- Real-time progress and performance analysis along with detailed transfer history, logging and manifest

Extraordinary bandwidth control

- Automatic, full utilization of available bandwidth with protection of other network traffic with “fair” policy
- Allows “bursts” in TCP traffic and reclaims unused bandwidth as it as it becomes available

LEGACY FASP – SECURITY & RELIABILITY

Complete Built-in Security

- Secure endpoint authentication, data encryption on-the-fly and at rest, and per-packet integrity verification
- FIPS 140-2 compliant, built on the openssl libraries

Secure User/Endpoint Authentication

- Authentication via secure SSH mechanisms: interactive password or public key
- LDAP, Active Directory user authentication
- Native File System Access Control support across all operating systems

AES-128 Cryptography

- On-the-fly data encryption
- Data encryption in transit and (optionally) at rest (secured storage of transferred content), client and server options

Data Integrity Verification

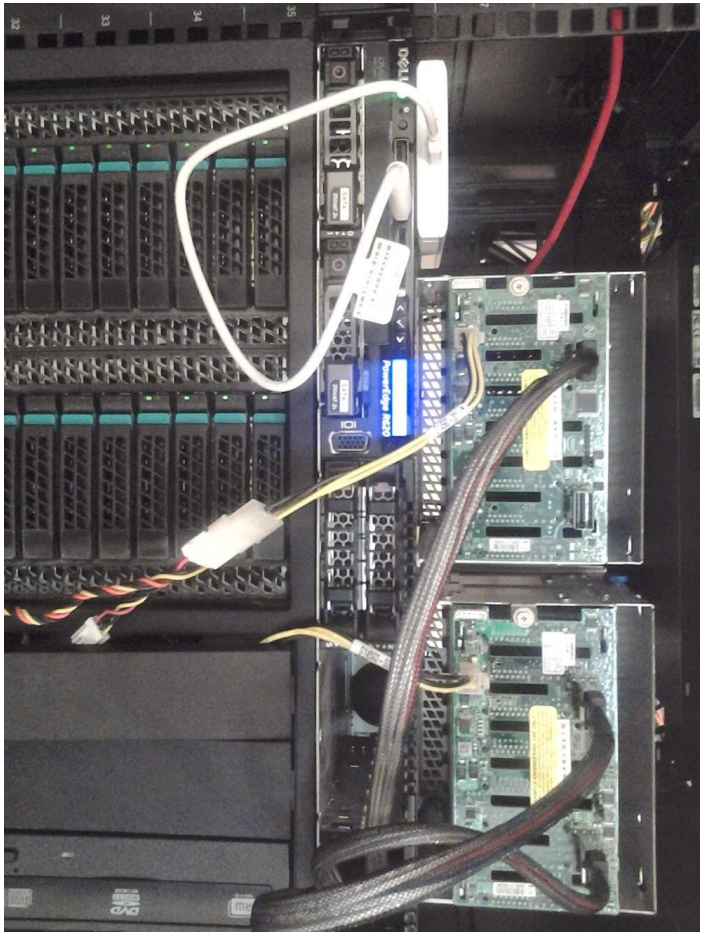
- Each transmitted data block is verified with a cryptographic hash function
- Protects against man-in-the-middle, re-play, and UDP denial-of-service attacks

100% Reliable Data Transmission

- Session semantics guarantee 100% bit-for-bit identical data copy at the destination
- Automatic resume of partial or failed transfers
- Automatic HTTP fallback in highly restrictive networks

FASP 0

40 GBIT/S - 2013

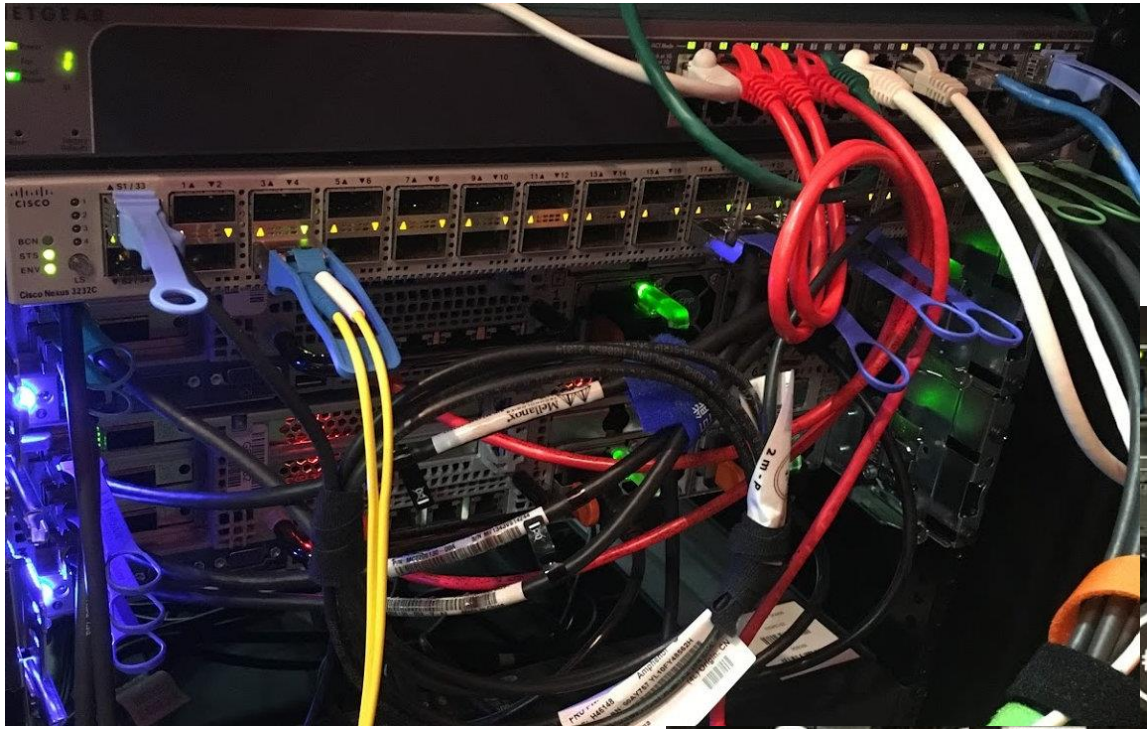


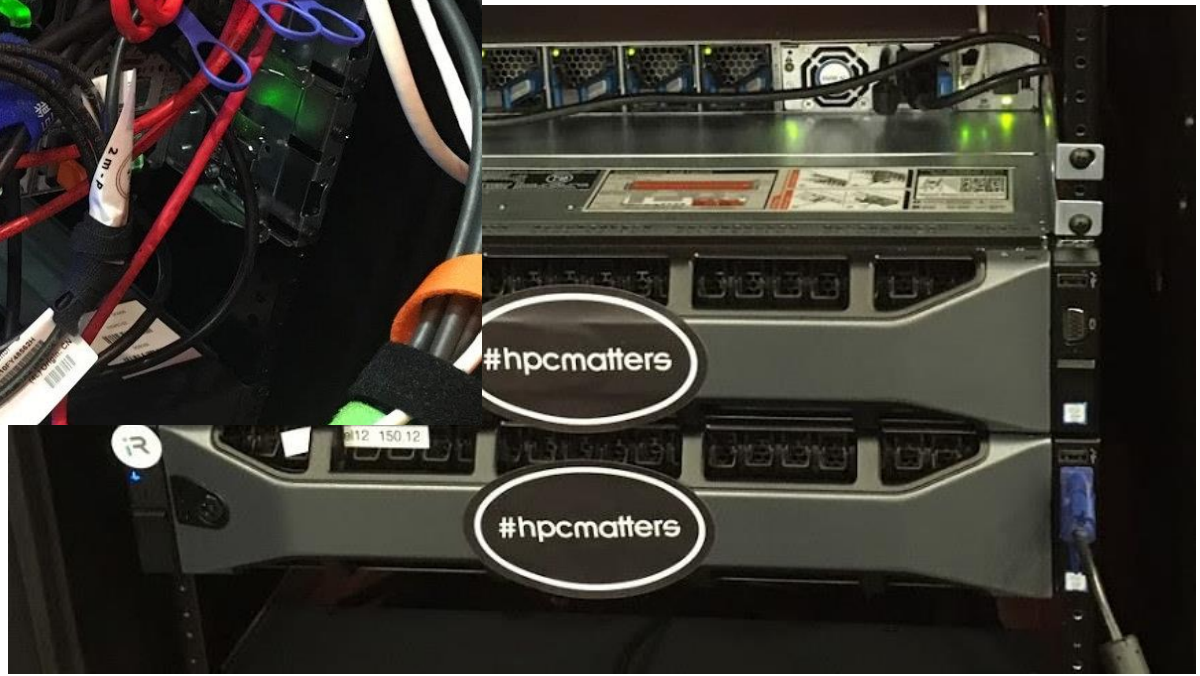
- Initial Xeon Ivy Bridge based DTN
- Used Hardware RAID and 4x10 gbit Intel NICs
- Single PCIe v2 x16 Hardware RAID card
- 12 Intel SATA SSDs
- Storage was bottleneck
- Hacked together using a combination of Intel Grizzly Pass systems and Dell Server. Note the internal SAS cables routing outside the Dell chassis along with power fed from other system.
- Whitepaper:
  http://asperasoft.com/fileadmin/media/Asperasoft.com/Resources/White_Papers/Big_Data_Transfer_Phase_2_WP_FINAL.pdf

80 GBIT/S @ SC14

@ SC15

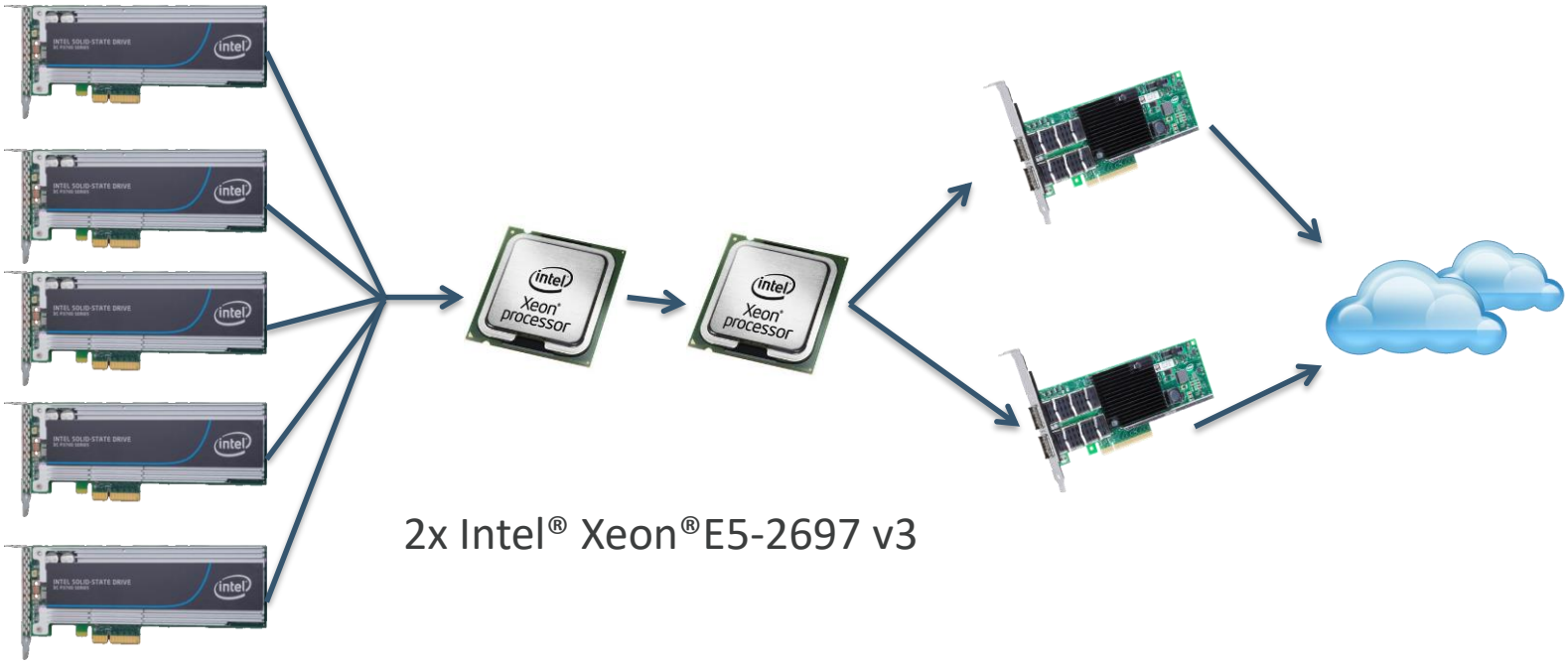
100 GBIT/S @ SC16





TODAY! 2X100GBIT NIC + 20X NVME

HARDWARE CONFIGURATION



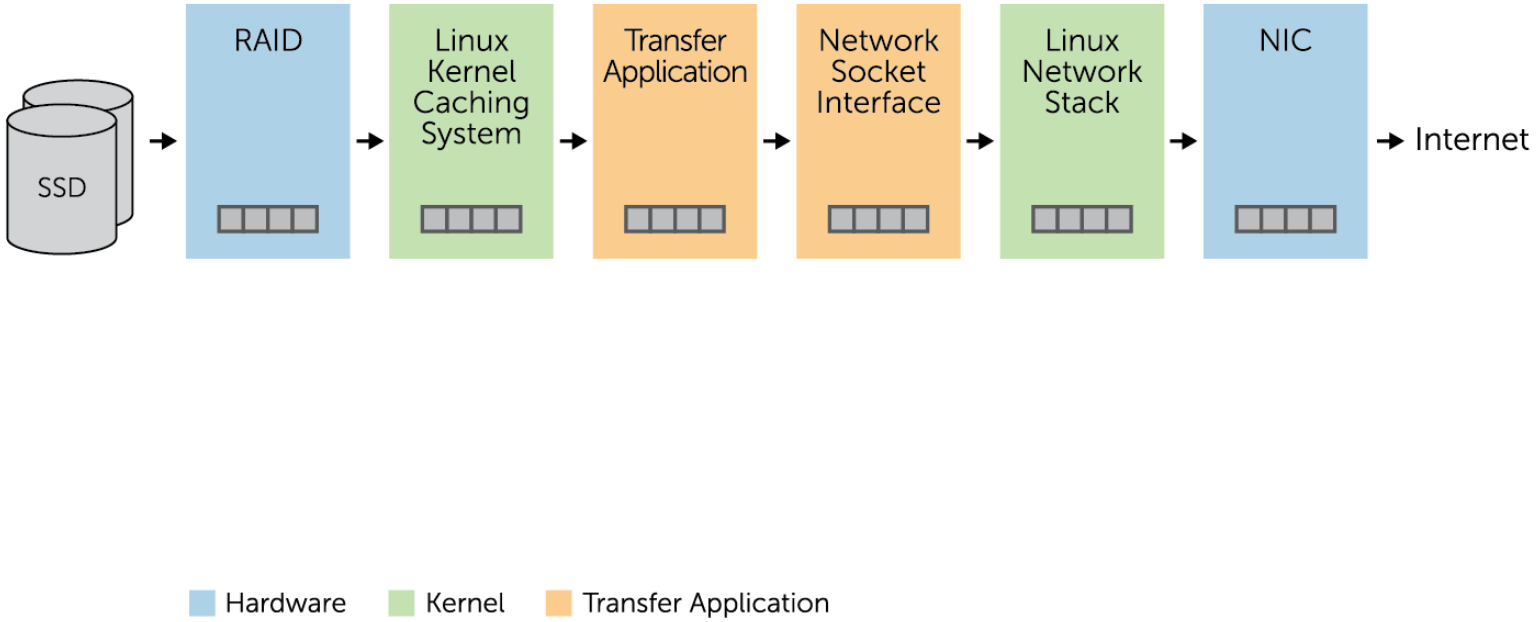
2x Intel® Xeon®E5-2697 v3

5x Intel® DC P3700 NVMe SSD

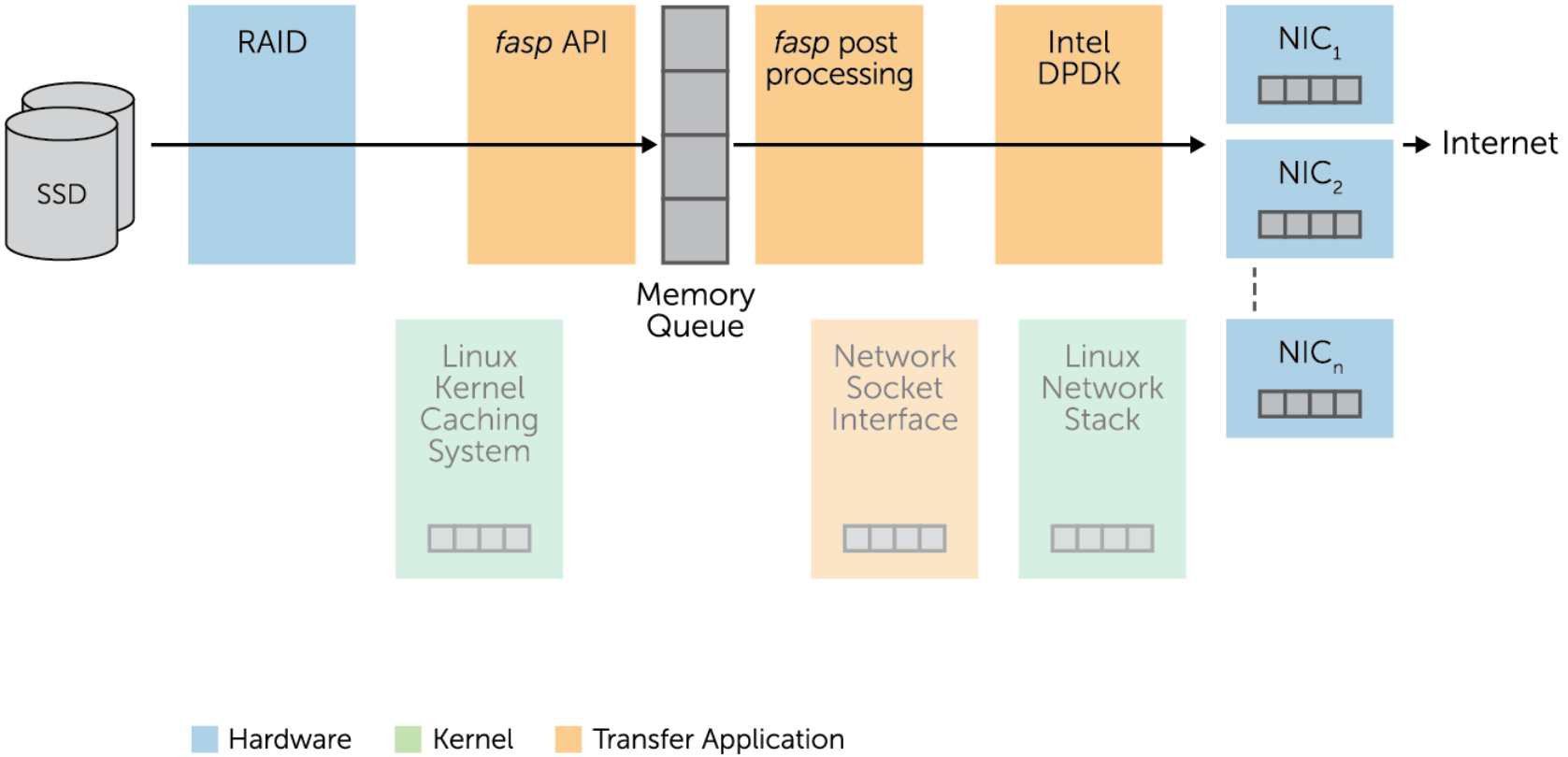
2x Intel® XL710 40 GbE Ethernet QSFP+

TYPICAL TRANSFER APPLICATION

BEFORE

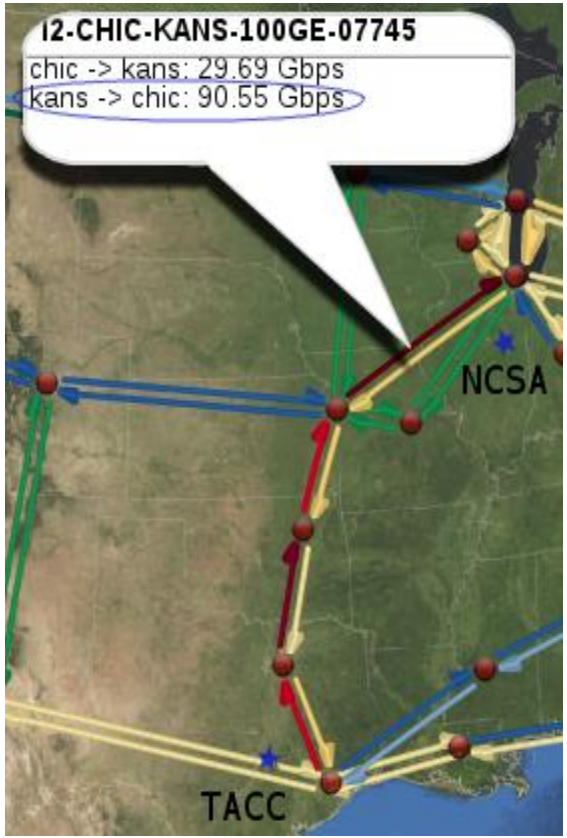


AFTER


RAID
fasp API
fasp post processing
Intel DPDK
NIC1
NIC2
NICn
Internet
SSD
Memory Queue
Linux Kernel Caching System
Network Socket Interface
Linux Network Stack
Hardware
Kernel
Transfer Application

FASP 0 TRANSFER BETWEEN TACC AND NCSA (~2016)

- Performance of in Development FASP 0 Protocol
  - 65 Gbit/s Wire Rate transfer, TACC to NCSA
  - 61 Gbit/s Effective rate, over 90% of available bandwidth utilized for transfer
- Eliminate traditional bottlenecks which impede the efficient transmission of data
- Single Stream Single Node Transport Solution
- 91 Gbit/s effective throughput within LAN environment with single Mellanox ConnectX®-4
  - 1 PB of data transferred every day
  - 675 GB per minute



THREE MAJOR REVISIONS SINCE PRESENT

Version 1: Optimized around major bottleneck areas

- Minimize memory usage (Zero Copy Transfer Solution)
- Optimize Network (DPDK Based Network Stack)
- Optimize IO (Direct I/O to disk, memory aligned around Page Boundaries)

Version 2: Optimize memory layout

- Threading performance strongly correlated to how processors share data
- Improved throughput from ~40 to ~70 GBPS

Version 3: Improve Memory Locality

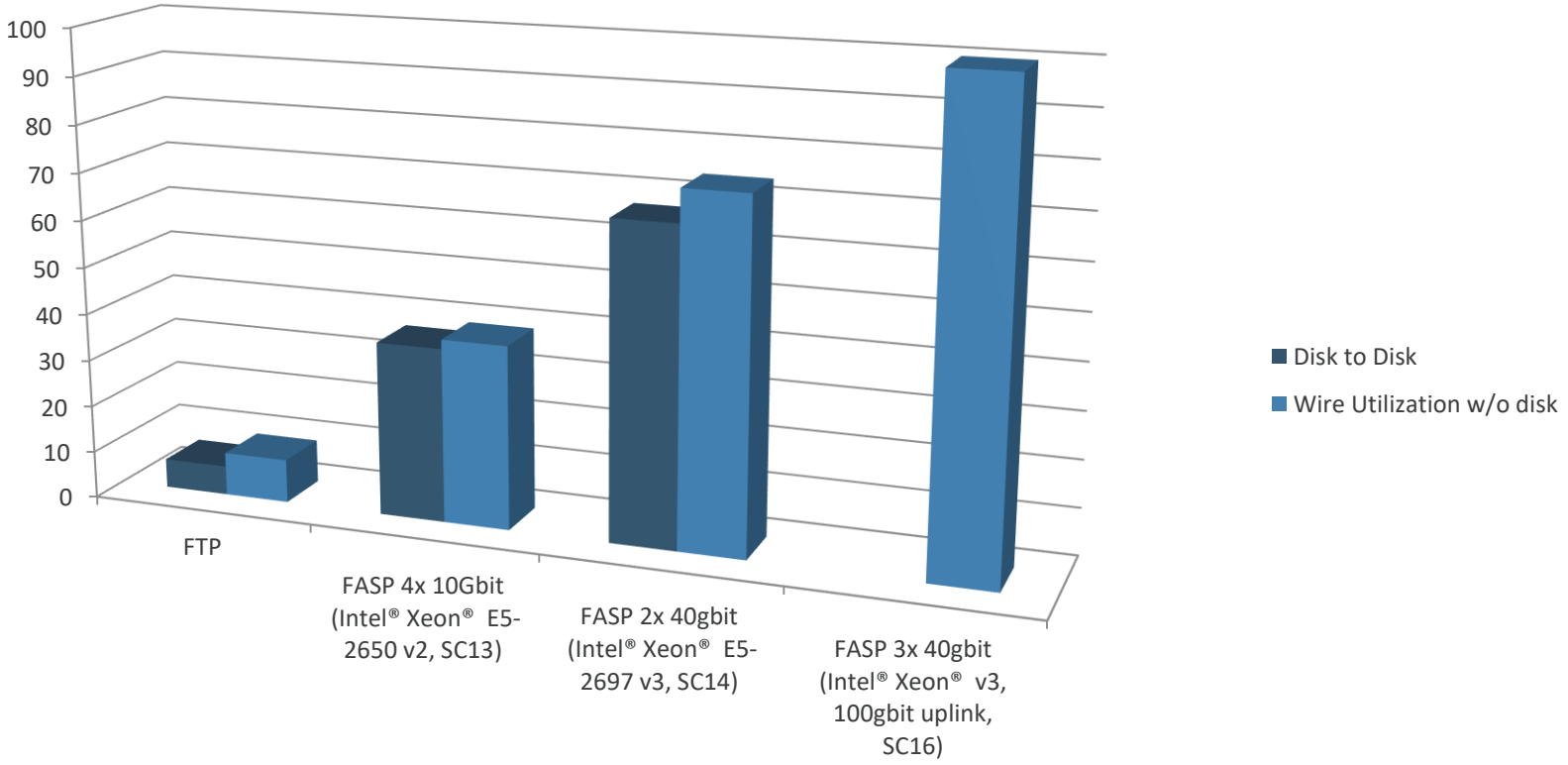
- Reduce how often threads need to synchronize
- Refine “Block” structure and how memory pages are owned by threads
- Improved Performance (120 GBPS) & Stability (> 1.2 PB transferred single session)

PLUS CURRENT WORK

- Support for regular sockets
- Better negotiation of traffic flows
  - Dynamically change ports to better take advantage of Receive Side Scaling (RSS), a way of distributing network load across cores
  - Simplified session negotiation
- Integration into ASCP 4
  (Two GA's of ASCP; *ASCP standard* and *ASCP 4*)
- Support stand alone DTLS authentication

Performance & Security

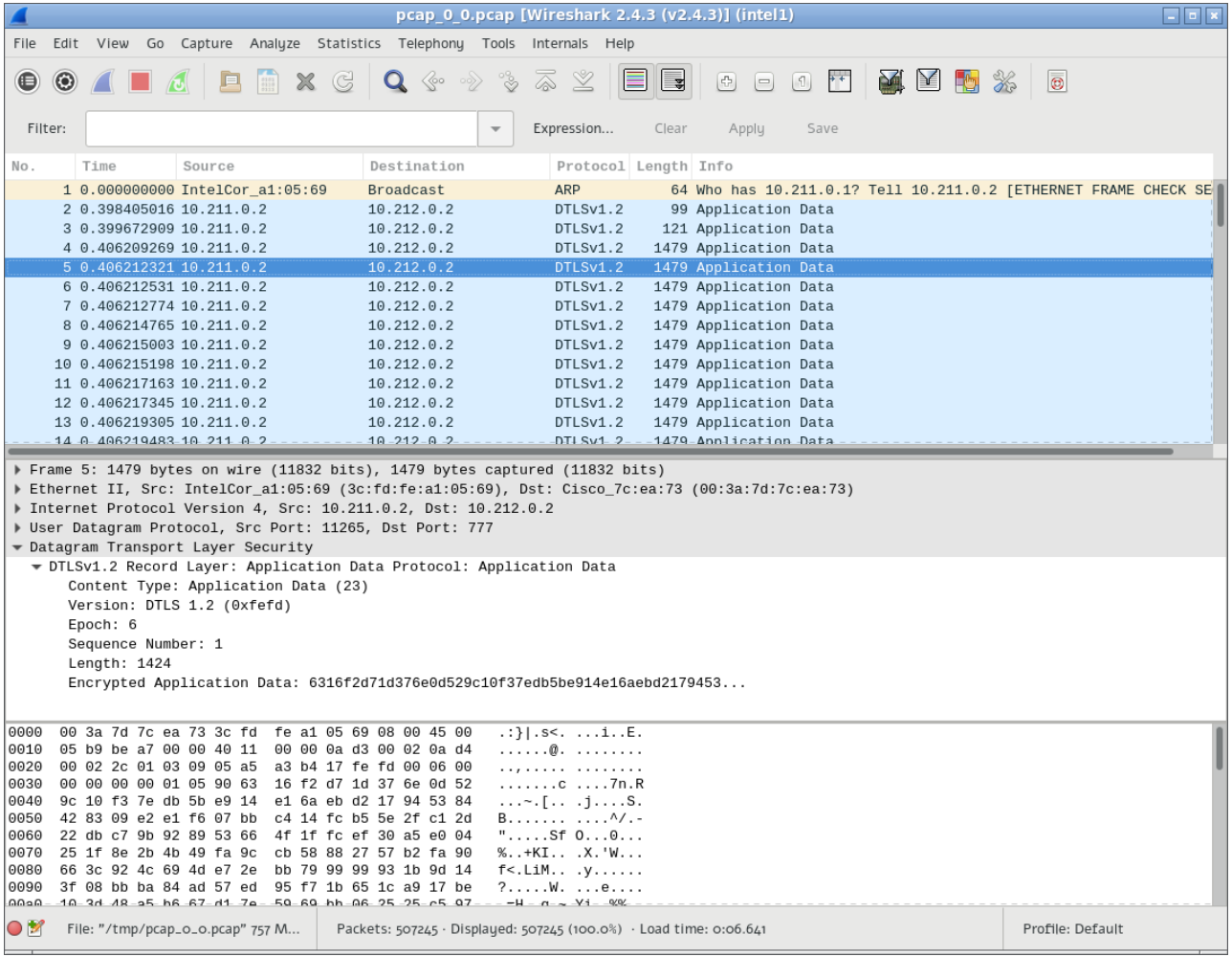
IT'S REALLY FASP!



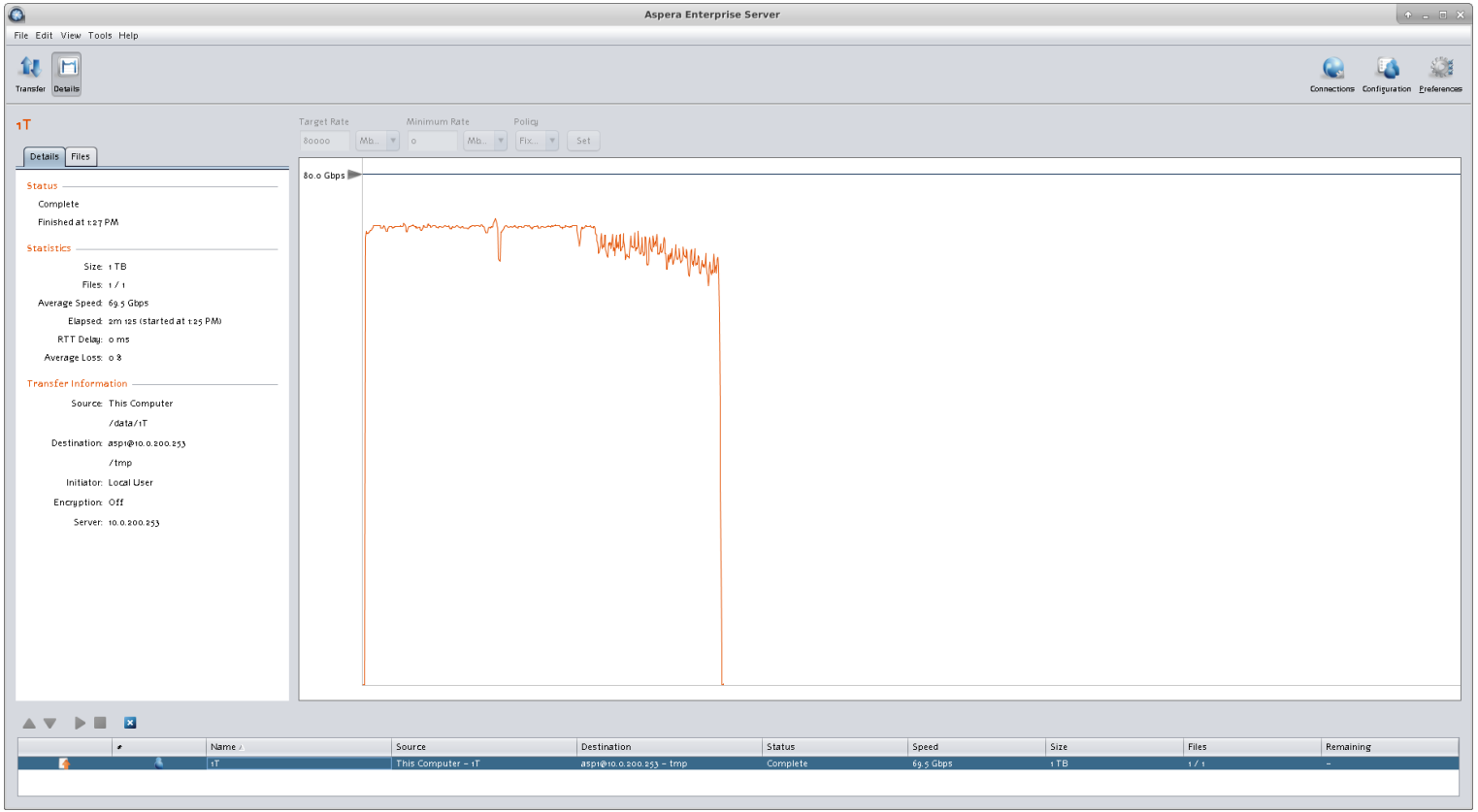
Remove Bottlenecks Typically Associated with Network Transfers

aspera
an IBM® company

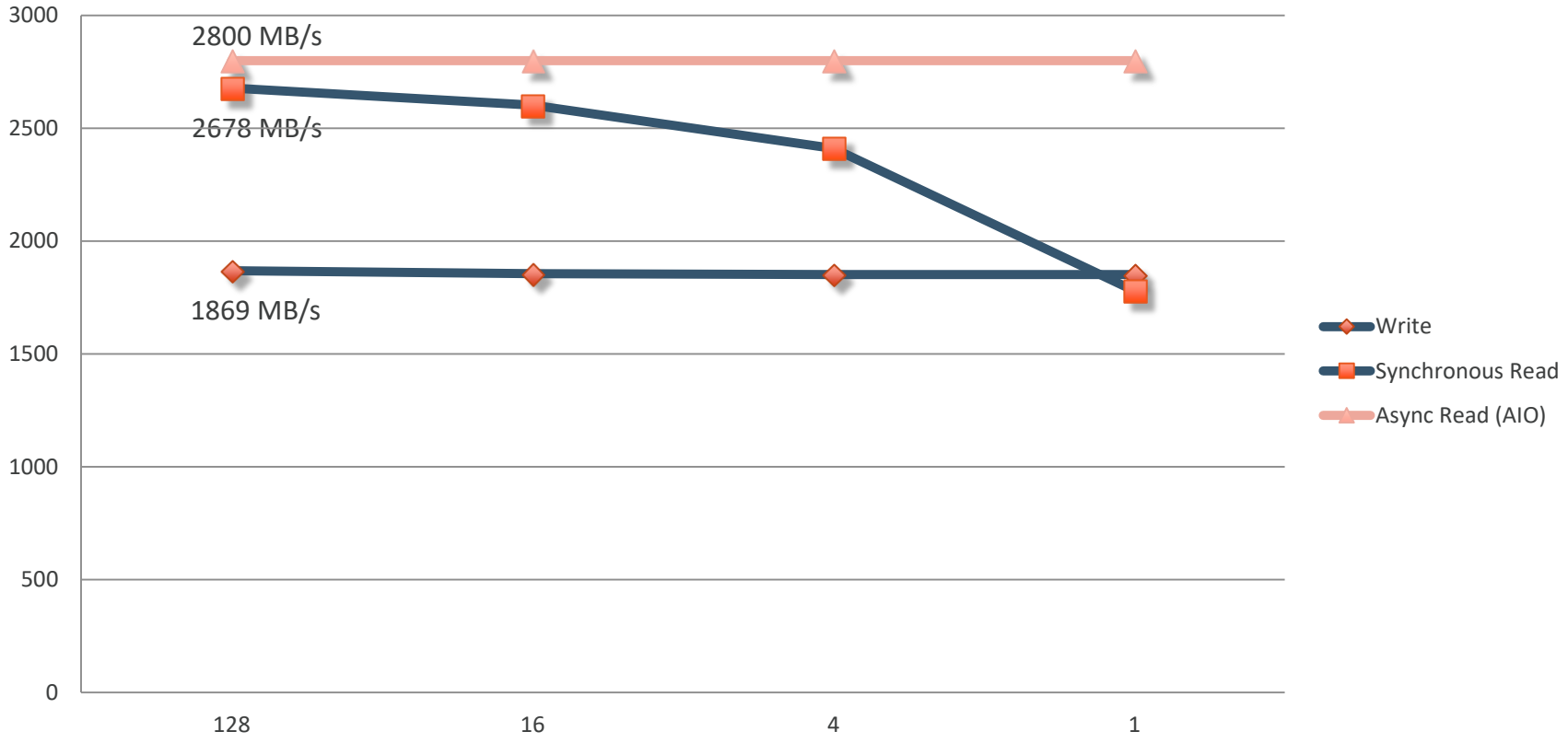
END-TO-END DTLS!



FASP MANAGEMENT

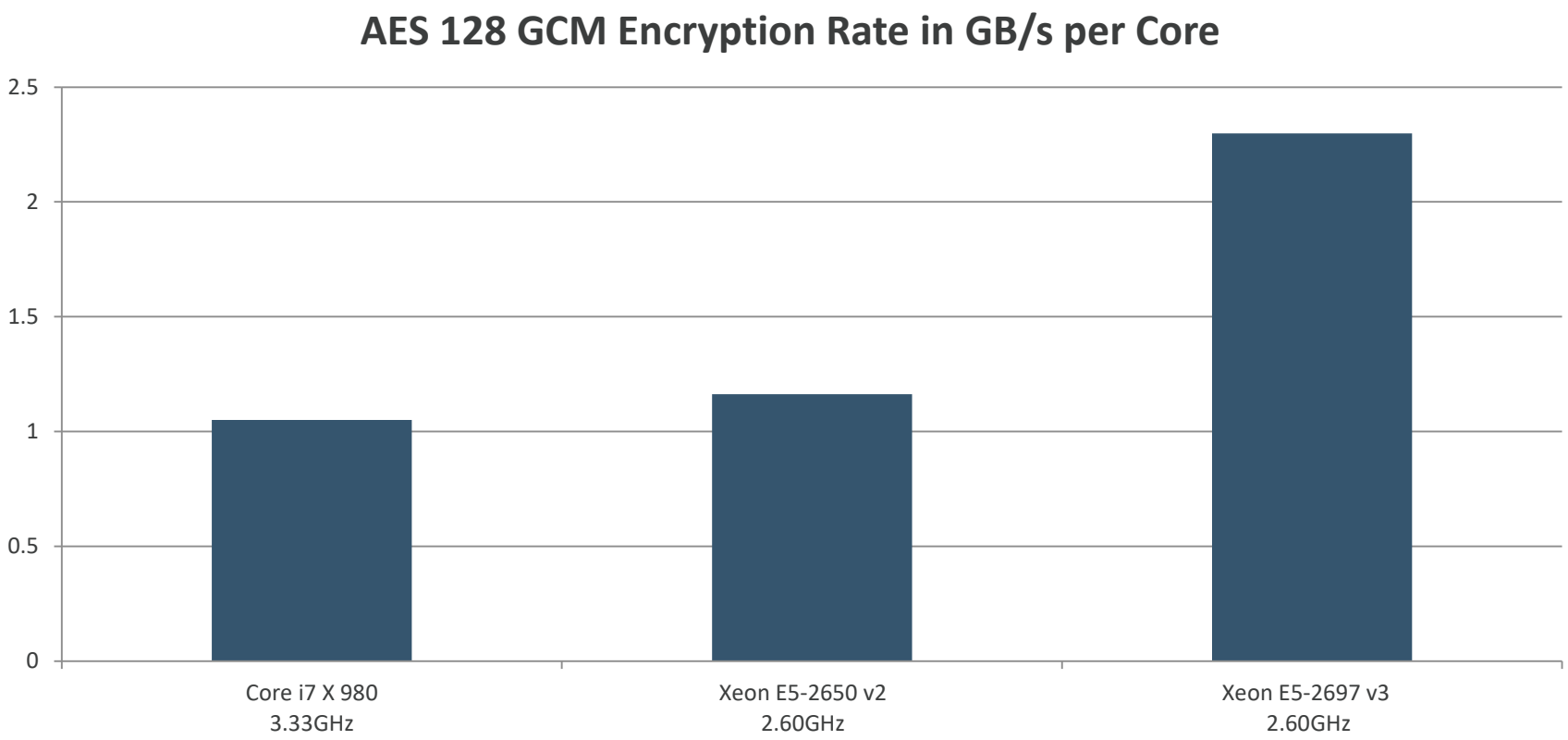


DC P3700 PERFORMANCE (SINGLE DRIVE)



Performance relative to block size (SI Units in MB)

ENCRYPTION IS NOT A BOTTLENECK WITH HW AES-GCM



USING ASCP

- Originally developed as a faster version of SCP.
  - Uses SSH for Authentication and Authorization.
  - Command line interface originally derived from SCP, but significantly different feature set to better support customer use cases
  - Default configuration tries to provide a good mix of security and usability.
- Language bindings for most languages (C, C++, Java, Python, .Net, Ruby, Javascript, Go, …)
  - Persistent Sessions
  - Possible to bypass SSH Authentication, useful for propagating a session over TLS
  - Data transfer still over encrypted UDP. Internally encrypted UDP changing to DTLS 1.2 as FASP engine is upgraded across Aspera products.
  - Fully transparent; Language bindings fully expose transfer functionality
  - Streaming Capable
  - Cloud enabled (Direct to object store, SAAS, Cloud VMs)
- Web based collaboration Suite
  - AoC, Formerly Files, Aspera's SAAS offering
  - Faspex & Shares (Web based file collaboration), Console (Management), Orchestrator

FASP 0

Aspera's FASPest FASP

Charles Shiflett
bear @ us . IBM . Com

Thank you for your interest!