Software Defined, Tiered Storage Infrastructure Optimized for Cognitive and Analytics Workloads

Joseph Dain
Storage and SW Defined Systems CTO Office
IBM Master Inventor

Matt Forney
Director of Research and Technology
Ennovar @ Wichita State University

Edge 2016

The Premier IT Infrastructure Conference

Outthink status quo.

Wichita State University
Ennovar

#ibmedge

Ennovar @ Wichita State University Overview

Provides a pipeline of highly qualified and prepared future employees.

- Ennovar at Wichita State University emphasizes education, emerging technology, experiential learning, evolutionary research, exploration and engagement.
- Ennovar employs more than 50 engineering, communication, graphic design, and MBA students and full-time professionals to lead collaborative projects with industry partners
- Solution Reference Architecture lab providing hardware logistics, test plan configurations, testing, data acquisition and verification. Technical marketing including best practices, technical reports, reference architectures, sale collateral and enablement in multiple languages.

Project Objective

- Verify the **benefits/performance of a software-defined tiered storage solution** using **SSD/Flash** disk as a high-performance scratch space for analytics processing, while placing inactive research data on low-cost physical **tape.**

Solution Objectives

- Cost Efficiency
  - Place inactive data on tape to significantly reduce cost
  - Tape TCO on average is over 6 times lower than disk[1]
- Performance
  - Leverage SSD tier in DCS3860 for high performance IO
- Flexibility
  - SSD, disk ,and tape modular building blocks provide scale-out capacity and performance
- Intelligent Management
  - Leverage Spectrum Scale ILM policy engine to place data in the right tier at the right time, transparent to applications.

[1]Clipper Group study 9/2015

IBM Spectrum Scale

+

IBM DCS3860

+

IBM Spectrum Archive

+

IBM Tape

Benefit Analysis

Negligible performance impact on data movement or application crunching when running Spectrum Archive migration and recall in parallel with LS-DYNA simulation.

- *Able to run multiple high-end applications while migrating and recalling data simultaneously.*

Significant advantages of combining SSD based Spectrum Scale file system with LTO-7 Spectrum Archive solution:

- Leverage the performance benefit of SSDs over HDDs for the application space
- Provides less expensive cost/capacity ratio of tape over HDDs for mid- and long-term storage
- Out of the box setup the LTO-7 drives performed at 92% of their max throughput
- Spectrum Scale file system performed 3.5 times faster with SSD versus NL-SAS drives

- Very easy to scale up tape performance by adding additional tape drives

NIAR – National Institute for Aviation Research

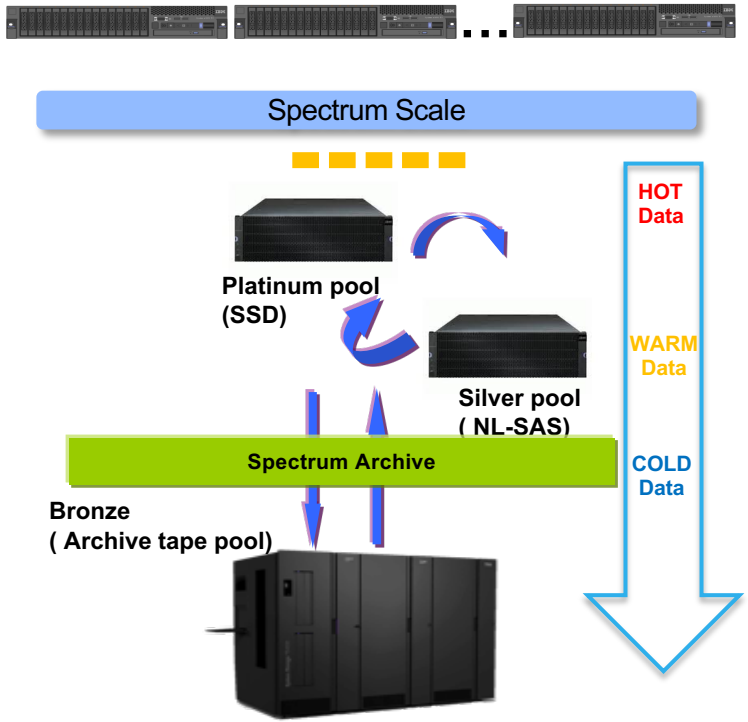
*"The benefit of using this type of technology would help us drive **efficiencies** in how we currently handle our data today. Being able to migrate and recall data sets while running our simulation applications provides a **practical solution** to having this in one location."*

Shawn Ehrstein
Director of the NIAR, CAD/CAM Lab & Virtual Reality Center

Analytics Approach

Store massive amounts of historical or inactive data on tape, place subset of data on platinum pool to be analyzed at fast speeds.

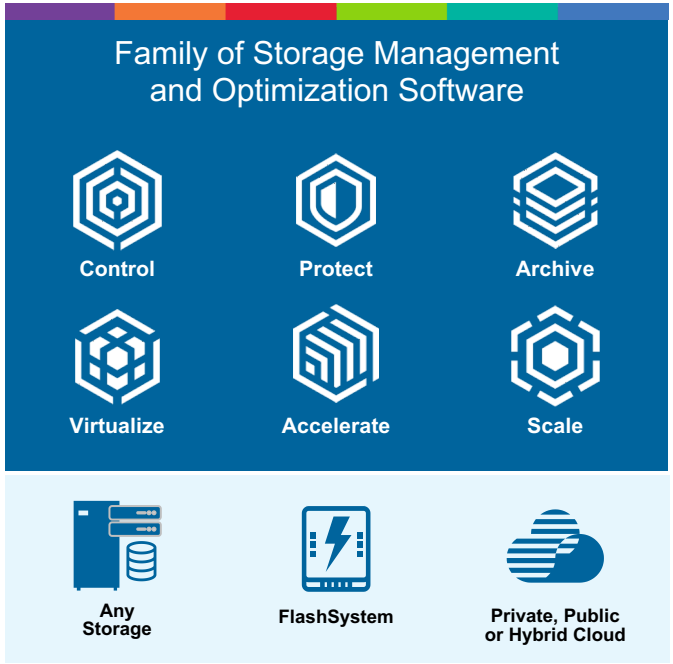
1. Stage dataset to platinum pool by recalling from tape
2. Run analytics job using compute nodes
3. Migrate dataset by optionally transitioning results from platinum pool to tape



- Computes nodes using POSIX interface
- Scratch space for high performance analytics
- Spectrum Scale filesystem metadata
- Spectrum Archive metadata
- Default data placement
- Inactive data sets migrated to tape

IBM Spectrum Storage Family

Proven Technology, Open Standards, Modular Adoption



IBM Spectrum Control	Analytics-driven hybrid cloud data management to reduce costs by up to 73%
IBM Spectrum Protect	Optimized hybrid cloud data protection to reduce backup costs by up to 53 percent
IBM Spectrum Archive	Fast data retention that reduces TCO for active archive data by up to 90%
IBM Spectrum Virtualize	Virtualization of mixed block environments stores up to 5x more data
IBM Spectrum Accelerate	Enterprise block storage for hybrid cloud deployed in minutes instead of months
IBM Spectrum Scale	High-performance, highly scalable hybrid cloud storage for unstructured data driving cognitive applications

IBM Software Defined Infrastructure delivers value across industries

Compute Clusters

Infiniti Red Bull Racing

- Designs and quickly executes multiple complex, interdependent simulations and analysis
- Leverages **20% increase in performance and throughput** to run more simulations is less time on less infrastructure

Global Storage Cloud

Cypress Semiconductor

- Eliminated data access bottlenecks and has **increased performance 10x** using the same hardware
- Continuous data availability across hardware outages

Content repository

Caris Life Sciences

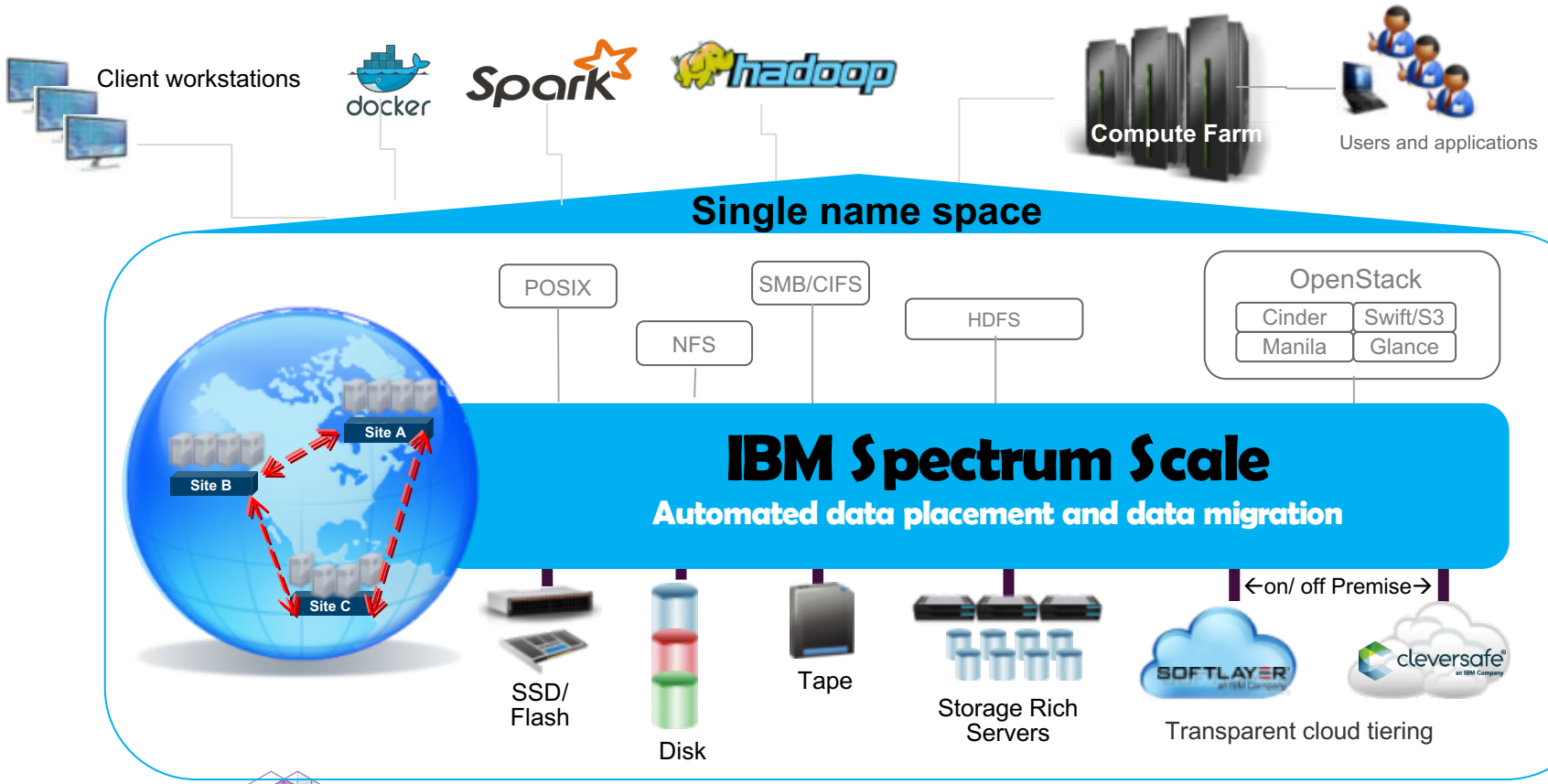
- Correlates molecular data for 65,000 patients and supporting 7,000 oncologists worldwide
- **Manages nearly a terabyte of data** per patient enabling precision cancer treatment

Big Data Analytics

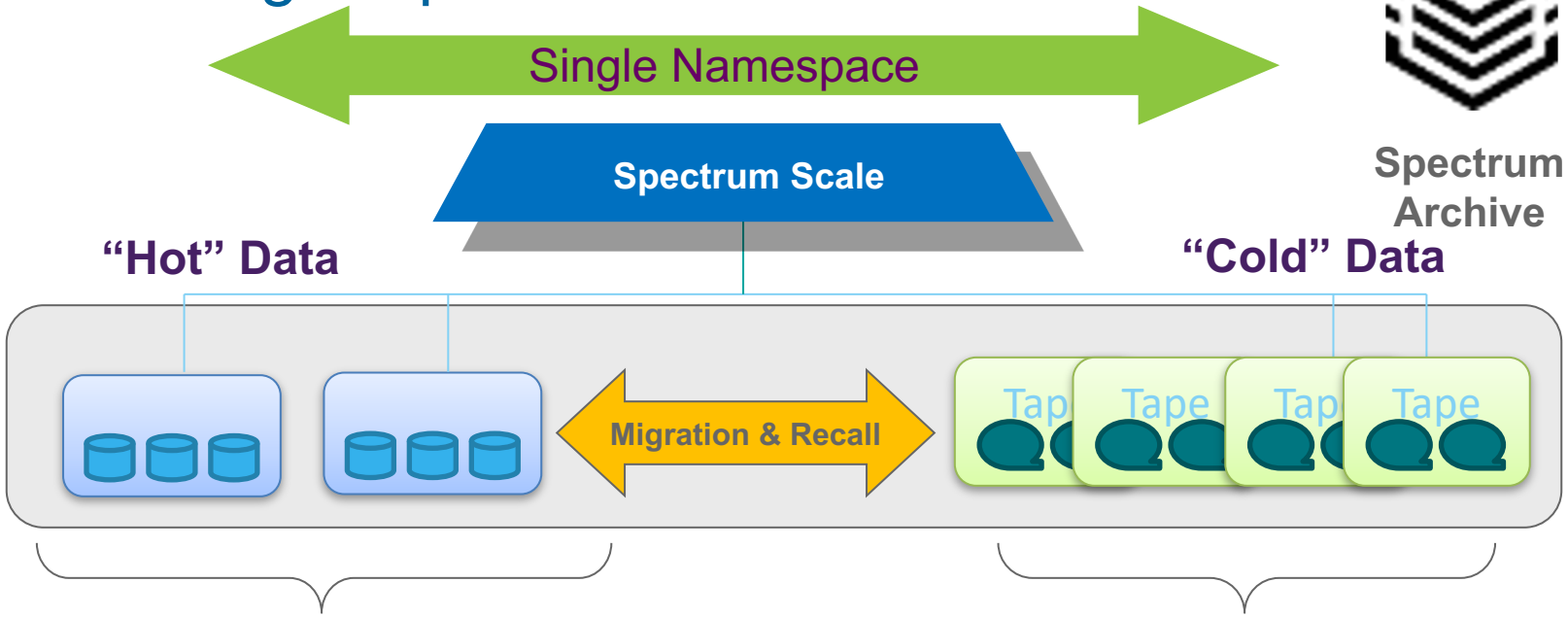
Citi

- **100X performance improvement** combined with on-demand access to compute power drastically speeds time to results
- Computing resources used more effectively with **hardware utilization increasing from 20% to 80%**

Spectrum Scale: Data Ocean for File, Object, and Analytics



Tiered Storage: Spectrum Archive



Fast and high available storage

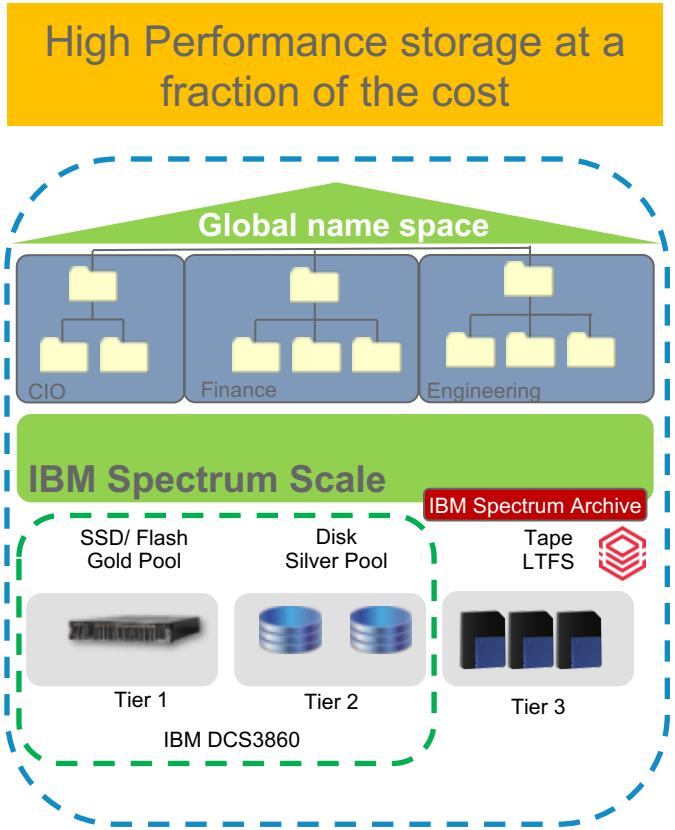
- Time critical applications and data
- Protected by Spectrum Scale RAID, Snapshots and replication

Lower cost, scalable storage

- Protected by multiple tape copies
- Export and import of tapes

Spectrum Scale + Spectrum Archive + DCS3860: Intelligent Tiering Solution

- Cost Efficiency
  - Place inactive data on tape to significantly reduce cost
- Performance
  - Leverage SSD tier in DCS3860 for high performance IO
- Flexibility
  - SSD, disk ,and tape modular building blocks provide scale-out capacity and performance
- Intelligent Management
  - Leverage Spectrum Scale ILM policy engine to place data in the right tier at the right time, transparent to applications



Outthink status quo.

Testing Environment and Results

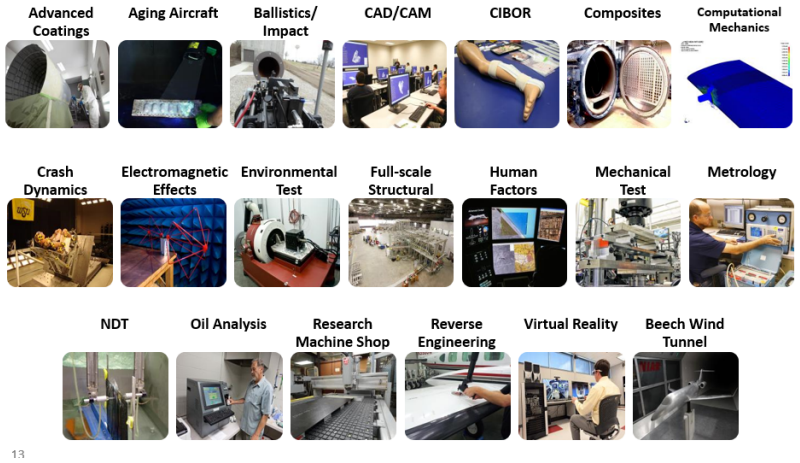
Edge 2016

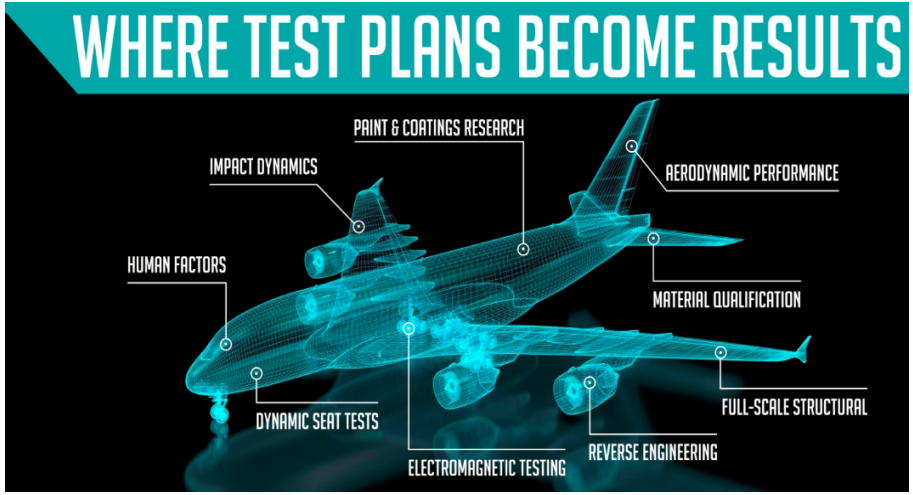
The Premier IT Infrastructure Conference

#ibmedge

NIAR – National Institute for Aviation Research

Since its inception in 1985, National Institute for Aviation Research (NIAR) at Wichita State University has made a name for itself as the most capable university-based aviation research center in the United States, providing research, design, testing, certification and training to the aviation manufacturing industry, government agencies, educational entities and other clients that can benefit from our services. “multiphysics simulation”

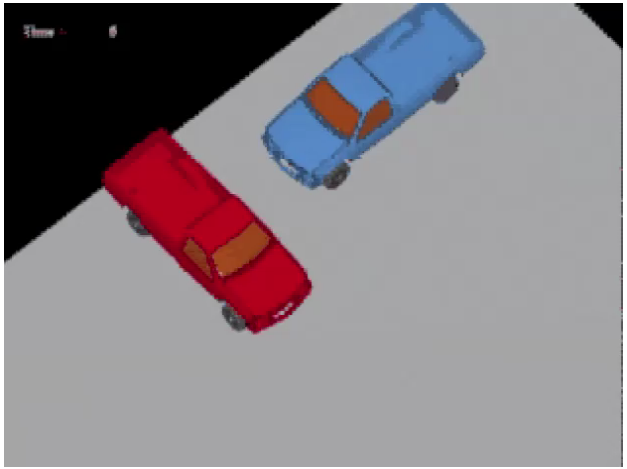




What is LS-DYNA

An advanced general-purpose Multiphysics simulation software package developed by the Livermore Software Technology Corporation.

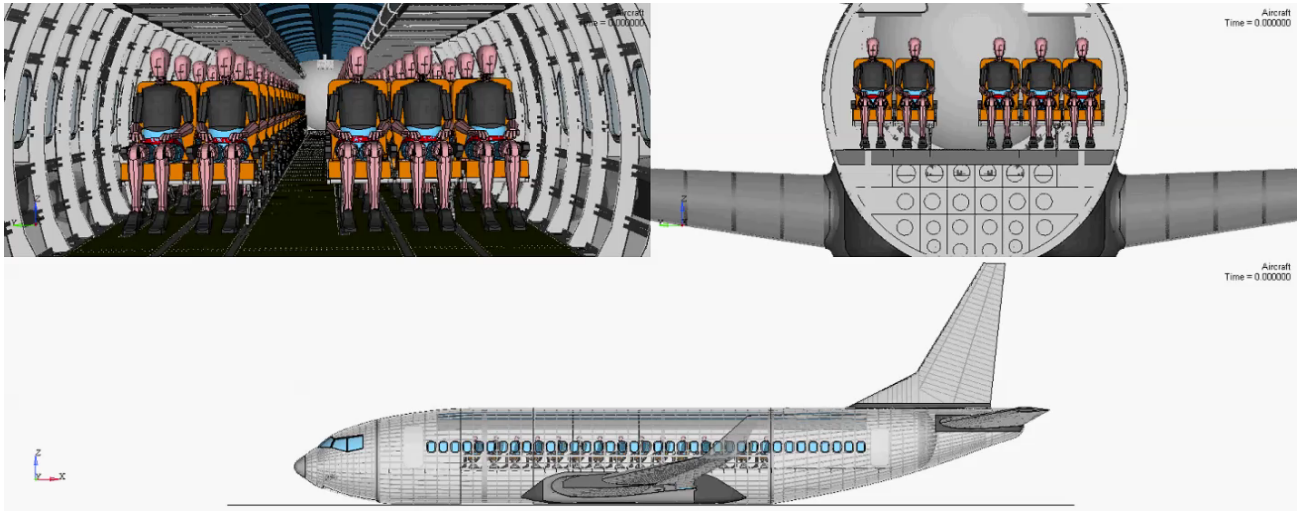
- "Transient dynamic" analysis of high speed, short duration events where inertial forces are important.
    - Crash simulations (deformation of chassis or structure, airbag inflation, seatbelt tensioning)
    - Explosions (underwater Naval mine, shaped charges)
    - Manufacturing (sheet metal stamping)



http://www.lstc.com/

LS-DYNA Use Case

- Simulation represents a 30ft drop of a single aisle commercial aircraft full of passengers.
- The model has over 10 million elements including the ATDs and the Aircraft structure. File size grows to 335GB during render process and that takes more than 46 hours of computational process time.



This information was provided by the NIAR computational mechanics laboratory located at Wichita State University.

NIAR Wichita State University National Institute for Aviation Research

Lab Overview

Mellanox MCX353A-FCBT

compute1 compute2 compute3 compute4 compute5 compute6

6 nodes x 2 FDR IB HCA per node on PCIe Gen2 (x8) = 72 GB/s

Mellanox SX6036 FDR Infiniband

2 nodes x 2 x 2-port FDR IB HBA per node on PCIe Gen3 (x8) = 56 GB/s

2 nodes x 2 LSI-SAS9300-16e per node on PCIe Gen3 (x16) = 24 GB/s

Lenovo x3650 M5

Lenovo x3650 M5

DCS3860 - 1

DCS3860 - 2

DCS3860 - 1
2x quad-port 12Gbs SAS
(1 connection per controller)
60x NL-SAS 6TB 7200RPM

DCS3860 - 2
2x quad-port 12Gbs SAS
(1 connection per controller)
60x NL-SAS 6TB 7200RPM
20x 800GB SSD per system

Brocade 6505

TS4500

LT06

LT07

Spectrum Scale
- Version 4.2.0.3
- NSD volume – 10 disk RAID6
- 1MB file system metadata block size
- 4MB file system data block size

Flash vs Disk – IOR HPC Benchmark

	Spectrum Scale block allocation type: Cluster				Spectrum Scale block allocation type: Scatter			
Drives	Write (MB/sec)	Normalized (MB/sec)	Read (MB/sec)	Normalized (MB/sec)	Write (MB/sec)	Normalized (MB/sec)	Read (MB/sec)	Normalized (MB/sec)
20 SSD	3579	179.0	9426	471.3	2557	127.9	9124	456.2
20 NL-SAS	1963	98.15	2756	137.8	930	46.5	1480	74
60 NL-SAS	2221	37.02	5789	96.5	2271	37.85	4731	78.9
Normalized SSD:NL-SAS (20drives)		1.82		3.42		2.75		6.16

The average (20 drive) normalized performance of SSD was 3.54 times faster than NL-SAS over all tests

Cache settings on the IBM 3860 storage system:

- Read Cache, Write Cache and write Cache w/Mirroring = ON
- Read Cache w/Prefetch = Off

- *Normalized based on single drive calculation*
- *Test optimized for throughput*
- *Better performance with Scatter*

FLAPE Test Results

	6 TB 7200 RPM NL-SAS		800 GB SSD	
	Migrate (sec)	**Recall (sec)**	**Migrate (sec)**	**Recall (sec)**
LTO-6 (4 drives)	594/9min	703/11min	405/6min	576/9min
LTO-7 (4 drives)	355/6min	507/8min	311/5min	491/8min
Performance Increase LTO-6 to LTO-7	67.3%	38.66%	30.23%	17.31%

LS-DYNA simulation was running in parallel while migrating and recalling data in the back ground. (46 hours, 29 minutes, 28 seconds)

Use case data used to determine migrate and recall rates was 320GB of LS-DYNA key and output data from an actual Boeing 737-800 crash landing simulation.

Minimal tuning results with true end user data.

- Emulex HBA settings modified according to best practices
- GPFS files system tuned for throughput

Actual results will vary according to number and size of files. use case data was a mix of both large and small files.

- *No performance lose during migrate and recall.*
- *Better performance running your applications on SSD's then NL-SAS drives.*
- *Cost affective solution for MB per Sec*
- *Energy Efficient*

Thank You

Edge 2016

The Premier IT Infrastructure Conference

Outthink status quo.

#ibmedge

Notices and Disclaimers

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IN NO EVENT SHALL IBM BE LIABLE FOR ANY DAMAGE ARISING FROM THE USE OF THIS INFORMATION, INCLUDING BUT NOT LIMITED TO, LOSS OF DATA, BUSINESS INTERRUPTION, LOSS OF PROFIT OR LOSS OF OPPORTUNITY. IBM products and services are warranted according to the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law

Notices and Disclaimers Con't.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. IBM EXPRESSLY DISCLAIMS ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.

The provision of the information contained h erein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com, Aspera®, Bluemix, Blueworks Live, CICS, Clearcase, Cognos®, DOORS®, Emptoris®, Enterprise Document Management System™, FASP®, FileNet®, Global Business Services ®, Global Technology Services ®, IBM ExperienceOne™, IBM SmartCloud®, IBM Social Business®, Information on Demand, ILOG, Maximo®, MQIntegrator®, MQSeries®, Netcool®, OMEGAMON, OpenPower, PureAnalytics™, PureApplication®, pureCluster™, PureCoverage®, PureData®, PureExperience®, PureFlex®, pureQuery®, pureScale®, PureSystems®, QRadar®, Rational®, Rhapsody®, Smarter Commerce®, SoDA, SPSS, Sterling Commerce®, StoredIQ, Tealeaf®, Tivoli®, Trusteer®, Unica®, urban{code}®, Watson, WebSphere®, Worklight®, X-Force® and System z® Z/OS, are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

Backup

Lab Configuration – more details

- Lenovo System X 3650-M5
  - Dual E5-2650 v4 12 cores each @ 2.2GHz “Broadwell”
  - 3x LSI-SAS9206-16e quad-port 6 Gb SAS adapter > PCIe Gen3 x8
  - 3x IBM OEM Mellanox dual-port FDR Connect-IB, PCIe Gen3 x16
  - OS - RHEL 7.1
- Intel-Based NSD Server
  - Single Intel E5-1650 v3 6-core 3.5 GHz (Haswell) processor, 64GB RAM
  - 1x Avago quad-port 9302-16e 12 Gb SAS HBA, PCIe Gen3 x16
  - 1x Avago dual-port 9302-8e 12 Gb SAS HBA, PCIe Gen3 x8
  - 1x IBM OEM Mellanox dual-port FDR Connect-IB, PCIe Gen3 x16
  - OS – CentOS 7.2
- Spectrum Scale
  - Version 4.2.0.3
  - NSD volume – 10 disk RAID6
  - 1MB file system metadata block size
  - 4MB file system data block size

Spectrum Scale TS4500 Tape Library

- High density enterprise tape library
- Integrates seamlessly into Spectrum Scale topology
- Reduced setup and configuration time with Spectrum Archive
- Single-pane-of-glass management console and easy to use GUI
- Automated tiering for fast data migration

LTO7 Tape Drives

- 2.4 times more capacity
  - single cartridge supports of up to 15TB*
  - TS4500 supports up to 526.5PB** of storage
- 87% higher performance
  - Up to 96GBpS** in a single TS4500 Library
  - 300MB/s throughput for greater RTO/RPO
- Full height and half height form factors
- Spectrum Archive enabled for File System based storage of data
  - Incorporates Long Term File System™ format

LTO Gen to Gen	Capacity Increase	Performance Increase
4 to 5	87.5%	16.7%
5 to 6	66.7%	14.3%
6 to 7	140%	87.5%

IBM DCS3860

Expandable	▪ Up to 360 drives	▪ Scale up to meet requirements
Drive Options	▪ NL-SAS: 4/6TB 7.2k ▪ SSD: 800GB/1.6TB	▪ Flexibility to right size for performance and capacity
I/O Interface Options	▪ 8 12Gb SAS ▪ 8 16Gb FC ▪ 8 10Gb iSCSI	▪ Complete set of high-performance and cost-effective interconnect options for SAN fabrics and direct attach
Storage Software	▪ FW 8.20 ▪ DCS Storage Manager 11.20	▪ Performance-optimized OS ▪ SAN management simplified
Premium Features	▪ Performance Read Cache ▪ Disaster Recovery Option ▪ Copy Services Option ▪ Super Key Option	

SAS host card

- 4-port 12Gb/s wide-port SAS

FC / iSCSI host card*

- 4-port 16Gb/s FC or
- 10Gb/s iSCSI optical SFP+

Configuration

	Balanced-Workload NL-SAS Building Block	Balanced-Workload SSD Building Block	Balanced-Workload Mixed Building Block
Compute Workload Servers	6 IBM System X 3690-X5 Dual Intel(R) Xeon(R) CPU E6540 6 Cores Each @ 2.00GHz “Beckton”	6 IBM System X 3690-X5 Dual Intel(R) Xeon(R) CPU E6540 6 Cores Each @ 2.00GHz “Beckton”	6 IBM System X 3690-X5 Dual Intel(R) Xeon(R) CPU E6540 6 Cores Each @ 2.00GHz “Beckton”
Controllers/ NSD Servers	2 Lenovo System X 3650-M5 Dual E5-2650 v4 12 cores each @ 2.2GHz “Broadwell”	2 Lenovo System X 3650-M5 Dual E5-2650 v4 12 cores each @ 2.2GHz “Broadwell”	2 Lenovo System X 3650-M5 Dual E5-2650 v4 12 cores each @ 2.2GHz “Broadwell”
Storage Subsystem	1 DCS3860-G2 systems 60 - 6TB NL-SAS per system 12Gb SAS connections max. 60 disks, 262 TB usable capacity	1 DCS3860-G2 systems 20 - 800GB SSD per system 12Gb SAS connections max. 20 disks, 12 TB usable capacity	2 DCS3860-G2 systems 20 - 800GB SSD per system 100 - 6TB NL-SAS per system 12Gb SAS connections max. 100 disks, 436 TB usable capacity max. 20 disks, 12 TB usable capacity (MetaData resides here)
Spectrum Scale Tape Library	TS4500 Library, Tape Drives 4 - LTO7 (300MB sec) and 4 – LTO6 (160MB sec)	TS4500 Library, Tape Drives 4 - LTO7 (300MB sec) and 4 – LTO6 (160MB sec)	TS4500 Library, Tape Drives 4 - LTO7 (300MB sec) and 4 – LTO6 (160MB sec)