



9/15 - 9/18
SANTA CLARA, CALIFORNIA

CREATED BY DEVELOPERS
FOR DEVELOPERS

SNIA 11TH ANNUAL STORAGE DEVELOPER CONFERENCE

[SDC Home \(/events/storage-developer\)](#)

[SDC Agenda \(/events/storage-developer/agenda\)](#)

[New Thinking Track \(/events/storage-developer/agenda/new_thinking\)](#)

[Speakers \(/events/storage-developer/speakers\)](#)

[Registration \(/events/storage-developer/registration\)](#)

[Mobile App \(/events/storage-developer/app\)](#)

[Venue \(/events/storage-developer/venue\)](#)

[Plugfests \(/events/storage-developer/plugfest\)](#)

[2014 Presentations \(/events/storage-developer/presentation\)](#)

[Sponsorship \(/events/storage-developer/sponsorship\)](#)

[Speaker Information \(/events/storage-developer/speaker_info\)](#)

[SDC Archives - 2008 to 2013 \(/events/storage-developer/archive\)](#)

[SNIA Technical Council \(/about/organization/tech_council\)](#)

[Analytics and Big Data Summit \(/events/abds2014\)](#)

[SNIA Home \(\)](#)

Storage Developer Conference Abstracts

Break Out Sessions and Agenda Tracks Include:

Note: This agenda is a work in progress. Check back for updates on additional sessions as well as the agenda schedule.

- [Big Data \(#big_data\)](#)
- [Birds of a Feather \(#bof\)](#)
- [Cloud \(#cloud\)](#)
- [De Dupe \(#de_dupe\)](#)
- [Distributed Storage \(#dist_stor\)](#)
- [etc \(#etc\)](#)
- [File Systems \(#file_sys\)](#)
- [Hardware \(#hw\)](#)
- [iSCSI \(#iscsi\)](#)
- [Key Note and Featured Speakers \(#keynote\)](#)
- [New Thinking \(#new_think\)](#)
- [NFS \(#nfs\)](#)
- [Object Storage \(#obj_stor\)](#)
- [OpenStack \(#openstack\)](#)
- [Open Source Software \(#oss\)](#)
- [Performance \(#performance\)](#)
- [Professional Development \(#pro_dev\)](#)
- [RDMA \(#rdma\)](#)
- [Security \(#security\)](#)
- [SMB \(#smb\)](#)
- [Software Defined Storage \(#soft_def\)](#)
- [Solid State \(#solid\)](#)
- [Storage Architecture \(#stor_arch\)](#)
- [Storage Management \(#stor_man\)](#)
- [Storage Plumbing \(#stor_plumb\)](#)
- [Testing \(#testing\)](#)
- [Testing Performance \(#testing_perf\)](#)
- [Virtualization \(#virt\)](#)

BIG DATA

Big Data Trends and HDFS Evolution

Sanjay Radia, Architect / Founder, Hortonworks

Abstract

Hadoop's usage pattern, along with the underlying hardware technology and platform, are rapidly evolving. Further, cloud infrastructure, (public & private), and the use of virtual machines are influencing Hadoop. This talk describes HDFS evolution to deal with this flux.

We start with HDFS architectural changes to take advantage of platform changes such as SSDs, and virtual machines. We discuss the unique challenges of virtual machines and the need to move MapReduce temp storage into HDFS to avoid storage fragmentation.

Second we focus on real-time and streaming use cases and the HDFS changes to enable them, such as moving from node to storage locality, caching layers, and structure aware data serving.

Finally we examine the trend for on-demand and shared infrastructure, where HDFS changes are necessary to bring up and later freeze clusters in a cloud environment. How will Hadoop and Openstack work together? While use cases such as spinning up development or test clusters are obvious, one needs to avoid resource fragmentation. We discuss the subtle storage storage problems their solutions. Another interesting use case we cover is Hadoop as a service supplemented by valuable data from the Hadoop service provider. Here we contrast a couple of solutions and their trade-offs, including one that we deployed for a Hadoop service provider.

[Back to Top \(#top\)](#)

Hadoop 2 : New and Noteworthy

Sujee Maniyam, Big Data Consultant/Trainer, ElephantScale

Abstract

Hadoop version 2 recently became available. It packs a multitude of features that are important to scalability, inter-operability and

From Terabytes to Exabytes, A paradigm Shift in Big Data Modeling, Analytics and Storage management for Healthcare and Life Sciences Organizations

Ali Eghlima, Director of Bioinformatic, Expert BioSystems

Abstract

Illumina CEO, recently announced availability of whole genome sequencing for just under \$1000. By 2020 whole genome sequencing could cost about \$200. Today, utilizing these technologies, a typical research program could generate from tens of terabytes to petabytes of data for a single study. Within ten years, a large genomic research program may need to analyze many petabytes to Exabyte of data.

Adding patient's genomic data to patient Electronic Health Record (EHR) will increase per patient dataset size from at most a few Gigabytes (today) to several terabytes. So, in a mid to large size hospital computer storage requirements, and associated computing power and network infrastructure performance will need to increase by at least three order of magnitude. Due to patient privacy, regulatory requirements, and issues related to cyber security, healthcare institute such as major hospitals are very reluctant in utilizing public cloud computing, and also, private cloud technology is not appropriate for distributed research collaboration, and large-scale interoperability across many organizations.

Current computing infrastructure of most life sciences research centers, and healthcare organizations/hospitals have not been architected/design to handle "HUGE" Big Data analytics, which is require to manage many Petabytes to Exabyte dataset class size, especially addressing requirements with regard to research collaboration across many organizations.

Learning Objectives

- Review current technology, and common systems architecture used for Big Data Analytics in Health Sciences vs other industries.
- Discuss issues, challenges and potential solutions for real-time and archived data storage managements
- Review, Data integrity/Privacy/Cyber Security concerns of major healthcare/research centers
- Present scalable open source computing platform to manage Exabyte class datasets

Big Data Storage

Apurva Vaidya, Principal Architect, iGate

Abstract

Big Data has emerged as the most booming market in business IT over the past few years. The amount of data to be stored and processed has exploded with a mind boggling degree and speed. Although "Big Data Analytics" has evolved to get a handle on this information content, but emphasize should also be given on appropriately storing data for easy and efficient retrieval.

This paper explains big data characteristics and the storage choices. The paper also discusses the impact of flash and storage tiering on accelerating the performance. The paper concludes by benchmarking and performance analysis and helps make a choice on the right storage platform.

Learning Objectives

- Understand Big Data and its characteristics
- Big Data Storage : Key Requirements, Challenges and Choices
- Analytical comparison of Hyper scale Computing Environments, Scale Out NAS, Object Storage
- Impact of Flash memory and tiered storage technologies
- Benchmarking and performance analysis and Conclusion

Emerging Storage and HPC Technologies to Accelerate Big Data Analytics

Jerome Gaysse, Consultant, Jerome Gaysse Consulting

Abstract

The storage and HPC (High Performance Computing) markets started to use and evaluate a new set of emerging technologies in order to face the big data performances challenges, including lower latency, fast computing and low power consumption. Current technologies include memories (Nandflash, MRAM, RRAM, PCM...), interfaces (PCIe & NVMe, NV-DIMM, CAPI, NVlink, HMC...) and controllers (FPGA, RISC CPU...). This presentation will review all these technologies and explain how it helps big data analytics in

Learning Objectives

- New memories performances
- Interfaces specification and model use
- Alternative to x86 CPUs

[Back to Top \(#top\)](#)

BIRDS OF A FEATHER

The Meaning and Value of Measuring Performance of all Solid State Arrays

Leah Schoeb, Sr. Partner, The Evaluator Group
Peter Murray, Sr. Product Specialist, Load Dynamix

Abstract

All solid state arrays are emerging as an important part in storage infrastructures and solutions. However, measuring performance on these new storage systems accurately is different not only from measuring performance on traditional disk arrays but also hybrid arrays as well. The elements that differentiate all solid state systems are also elements that impact performance behavior as well. Processes build on how we measure solid state devices couple with developing the correct data content and data streams is crucial to accurate performance measurement and reporting.

Learning Objectives

- Learn about the unique characteristics of all solid state and all flash arrays
- Understand a vendor neutral methodology for measure accurate performance
- Learn how these characteristics will set performance expectations with commercial workloads

[Back to Top \(#top\)](#)

The Future Of Cloud Storage: Personal, Ad-hoc, Community-owned Storage Networks

Abstract

The rise of smart phones, tablets, set-top boxes, and ubiquitous network access has created a surge in always-on always-connected commodity devices. These billions of nodes represent one of the greatest untapped resources in storage. Being mobile, wireless, battery-powered, and heterogeneous, it is highly resilient to network and power outages, natural disasters, and computer viruses/worms. With the right software a vast cloud storage system can be created overnight, where users exchange some local storage for storage capacity in the cloud. Yet this system must overcome a number of concerns: guaranteeing reliability, securing privacy, and preventing cheating. In this presentation we outline a path to overcome all these problems, creating the first software-defined cloud.

[Back to Top \(#top\)](#)

Implementing SDS - Developer Experience

Mark Carlson, Senior Staff for Standards, Toshiba
Leah Schoeb, Sr. Partner, The Evaluator Group

Abstract

Software defined storage has emerged as an important concept in storage solutions and management. However, the essential characteristics of software defined storage have been subject to interpretation. This session defines the elements that differentiate software defined storage solutions in a way that enables the industry to rally around their core value. A model of software defined storage infrastructure is described in a way that highlights the roles of virtualization and management in software defined storage solutions.

[Back to Top \(#top\)](#)

SSIF KMIP Testing Program

Wayne M. Adams, SNIA Chairman Emeritus, Senior Technologist, Office of the CTO EMC

Corporation**Abstract**

Pending

[Back to Top \(#top\)](#)**Open Standards vs. Open Source****Mark Carlson, Senior Staff for Standards, Toshiba****Abstract**

There is a debate on the relevance of Industry Standards when faced with Open Source efforts. Yet government bodies still rely on and give preference to ANSI and ISO standards.

At the Storage Developer Conference (SDC) this year we have attendees that participate in the development of standards in various standards bodies. We also have attendees that participate in the Open Source community. A meeting of both groups at SDC presents a unique opportunity to carve a path forward to see if and how both groups can work together.

SNIA's CDMI effort, for example, includes both an ISO standard for cloud storage and an open source reference implementation that provides example code and a running systems with which to interact.

Are these the right ways forward? Can Open Source and Open Standards work together? Are there other paths that may be better? Does documenting an existing Open Source implementation allow Open Source the flexibility to evolve? Does implementing an existing standard represent a viable path? Can simultaneous development really work? Does the standards process need to change? Please join us to discuss these issues.

[Back to Top \(#top\)](#)**SNIA Emerald NAS Power Efficiency Measurement Testing****Wayne Adams, Carlos Pratt, Alan Yoder****Abstract**

SNIA Emerald program is expanding the test tools and taxonomy to include NAS for release in 2015. Participate in a pilot program to validate and refine test methods for power efficiency testing using the SPEC SFS 2014 tool, approved power meter, and your in-house NAS systems.

[Back to Top \(#top\)](#)**Storage for the Internet of Things****David Slik, Technical Director, Object Storage, NetApp****Abstract**

The Internet of Things (IoT) generates data — Lots of data. And like most situations where data is generated, much of it needs to be stored, both transitorily and persistently. This BoF explores emerging data flows in IoT architectures, and explores areas where storage standards can integrate into the emerging ecosystem of capture, transport and analytics.

[Back to Top \(#top\)](#)**SMR, the ZBC/ZAC Standards and the New Libzbc Open Source Project****Jorge Campello, Director of Systems, Architecture and Solutions, HGST Research****Abstract**

Shingled Magnetic Recording (SMR) drives have started to hit the market and the industry is still in the process of determining how to best make use of the technology. The Zoned Block Commands (ZBC) and Zoned ATA Commands (ZAC) standards are in advanced stages of development within T10 and T13 respectively.

In this session, we will explore how to manage SMR drives implementing the ZBC and ZAC standards using the newly introduced libzbc open source project.

[Back to Top \(#top\)](#)

SMB2/SMB3 Protocol Family Test Tools

Mushi Jin, Senior Software Development Engineer, File Services Test Team
Justin Jiang, Software Development Engineer, Interop and Tools Test Team

Abstract

Microsoft has released tools for SMB2&3 protocol family interoperability testing. SMB2&3 file io and services api level tests are available for use at this event as well.

[Back to Top \(#top\)](#)

IP Drives - A New Architectural Partitioning?

Mark Carlson, Senior Staff for Standards, Toshiba

Abstract

A number of scale out storage solutions, as part of open source and other projects, are architected to scale out by incrementally adding and removing storage nodes. Example projects include:

- Hadoop's HDFS
- CEPH
- Swift (OpenStack object storage)

The typical storage node architecture includes inexpensive enclosures with IP networking, CPU, Memory and Direct Attached Storage (DAS). While inexpensive to deploy, these solutions become harder to manage over time. Power and space requirements of Data Centers are difficult to meet with this type of solution. This BOF looks to examine solutions that better meet the requirements by re-partitioning the solutions (drive based storage nodes) and creating points of interoperability.

[Back to Top \(#top\)](#)

Special Microsoft BoF - Title Pending

Abstract

Pending

[Back to Top \(#top\)](#)

Benchmarking with SPEC SFS 2014

Spencer Shepler, Architect, Microsoft

Abstract

An open session/Q&A to discuss the benchmarking features, advanced capabilities, and usage of SPEC SFS 2014, including running custom workloads.

[Back to Top \(#top\)](#)

SNIA Emerald NAS Power Efficiency Measurement Testing

Wayne Adams, SNIA Chairman Emeritus, Senior Technologist, Office of the CTO, EMC

Abstract

To broaden awareness of capable testing engineering services and independent test labs who are SNIA observed competent in performing SNIA Emerald testing in support of SNIA and EPA EnergyStar data storage energy efficiency programs. BoF encourages participation from those performing and or overseeing in-house or contracted services for regulatory requirements.

[Back to Top \(#top\)](#)

Non-Volatile DIMMs: Memory or Storage?

**Arthur Sainio, Co-Chair, SNIA NVDIMM SIG, SMART Modular Systems
Mario Martinez, SNIA NVDIMM SIG member, Netlist**

Abstract

NVDIMMs are gaining momentum with industry standardization efforts, but questions remain on what they are and how organizations and development staffs can best take advantage of them. This session will discuss how NVDIMMs function in server and storage systems and how they can be integrated into a standard server platform. The new SNIA SSSI NVDIMM SIG will also share their latest projects on NVDIMM taxonomy and welcome all those interested in NVDIMM and NVM topics for a discussion on a closer relationship between the SIG and the NVM Programming TWG.

[Back to Top \(#top\)](#)

Linux Kernel Storage Developers

Abstract

This BOF is focused on discussion of current and planned work in the Linux kernel in the storage space. The BOF will be lead by Christoph Hellwig and Martin Petersen and is open to anyone interested in current activity in the Linux file and storage stack.

[Back to Top \(#top\)](#)

Azure Cloud Storage

Abstract

Pending

[Back to Top \(#top\)](#)

Cloud

Introducing CDMI 1.1

David Slik, Technical Director, Object Storage, NetApp

Abstract

Subsequent to the SNIA Cloud Data Management Interface (CDMI) becoming adopted as an international standard (ISO 17826:2012), there has been significant adoption and innovation around the CDMI standard. This session introduces CDMI 1.1, the next major release of the CDMI standard, and provides an overview of new capabilities added to the standard, major errata, and what CDMI implementers need to know when moving from CDMI 1.0.2 to CDMI 1.1.

Learning Objectives

- Learn about the adoption of CDMI and how this drives improvements to the standard
- Learn what new capabilities were added to CDMI 1.1
- Learn about the major errata corrected in CDMI 1.1
- Learn what changes you need to make as a CDMI server or client implementer to move to CDMI 1.1

[Back to Top \(#top\)](#)

LTFS Bulk Transfer Standard

David Slik, Technical Director, Object Storage, NetApp

Abstract

LTFS tape technology provides compelling economics for bulk transportation of data between enterprise locations and to and from clouds. This session provides an update on the joint work of the LTFS and Cloud Technical Working Groups on a bulk transfer standard that uses LTFS to allow for the reliable movement of data and merging of namespaces. This session introduces the use cases for inter and intra-enterprise data transport, and cloud data transport, and describes the entities and XML documents used to

Learning Objectives

- Learn about how LTFS Bulk Data Transport reduces the cost of bulk data transport
- Learn about how the LTFS Bulk Data Transport standard works
- See a demonstration of the LTFS Bulk Data Transport used to bulk retrieve data from a cloud

[Back to Top \(#top\)](#)

Stratus to Cirrus: Avoiding Nose-Bleeds During Upgrades of Cloud Storage Systems

Tom Cocagne, Senior Software Developer, Cleversafe

Abstract

Implementing zero-downtime upgrades of live cloud storage systems is a surprisingly complex problem that has proven difficult to completely automate. Beyond merely preventing availability outages, the upgrade process must proactively detect and repair errors, prevent cascading failures from leading to data loss, be resilient in the face of transient network communication errors, and gracefully handle disk failures that occur during device upgrades. At the scale of today's deployments, occasional human intervention to help the process along is tolerable. With hundreds of thousands of devices comprising multi-exabyte, single-system deployments on the horizon though, completely automated solutions are required. Please join us as we discuss the challenges inherent to upgrading cloud storage systems and how those challenges may be overcome at scale.

Learning Objectives

- The challenges involved in cloud storage upgrades
- Techniques to address those challenges
- Ramifications at scale

[Back to Top \(#top\)](#)

Introduction and Evaluations of a Wide Area Distributed Storage System

Hiroki Kashiwazaki, Assistant Professor, Osaka University

Abstract

In recent years, much attention has been paid to wide area distributed storages to backup data remotely and ensure that business processes can continue in terms of disaster recovery. In the "distcloud" project, we have been involved in the research of wide area distributed storage by clustering many computer resources located in geographically distributed areas, where the number of sites is more than 2 ($N > 2$). The storage supports a shared single POSIX file system so that long distance live migration (LDLM) of virtual machines (VMs) works well between multiple sites. We introduce the concept and basic architecture of the wide area distributed storage and its technical improvement for LDLM. We describe the result of our experiment, that is;

1. Nation Wide Live Migration (about 500Km) in Japan, and
2. Transpacific Live Migration (over 24,000Km)

We show the technical benefit of the current implementation and discuss suitable applications and remaining issues for further research topics.

Note: This presentation will be jointly-conducted with SNIA-J.

Learning Objectives

- Live migration
- Distributed storage
- POSIX file system

[Back to Top \(#top\)](#)

One Ring Cannot Rule Them All

Gary Ogasawara, VP Engineering, Cloudian

Abstract

Implementing an object storage system with user definable data protection models (replicas, erasure coding, NOSQL DB) on a per workload basis. I will discuss the Cloudian architecture that allows per-container and even per-request selection of storage type.

Learning Objectives

- Why one storage technique is not sufficient
- Important components of a traffic model for cloud/object storage
- Architecture for a multiple storage type system
- Challenges for an intelligent virtual storage system

[Back to Top \(#top\)](#)

Is CDMI and Non CDMI Operations Inter-Operable in Conformance Testing?

Sachin Goswami, Solution Architect, TATA Consultancy Services
Ankit Agrawal, Solution Developer, TATA Consultancy Services

Abstract

Is CDMI and Non CDMI operations interoperable in Conformance testing: Addressing Challenges, Approach & best practice?

With the rapid growth of the cloud market, today there are a slew of vendors offering multiple cloud solutions for cloud migration, data management and cloud security. Multiple cloud solutions put end-users in qualm about the best solution. The Cloud Data management Interface specification has adopted CDMI, NON CDMI as well as Profile based categories to resolve end user muddle.

TCS has been concentrating on implementing the Conformance Test Suite as well as contributing to SNIA CDMI Conformance Test Specification, which is focusing towards incorporating CDMI, NON CDMI and profile based scenarios.

In this proposal we will share the approach and challenges for testing of interoperability of CDMI, Non CDMI specification as well as profile based scenarios of the cloud products. Also, we will share additional challenges / learning's gathered from testing of CDMI Products for conformance. These learning's will help as a ready reference for organizations developing CDMI, NON CDMI and profile based cloud storage products.

Learning Objectives

- Understanding CDMI, Non CDMI and Profile based Specification
- Understanding the CDMI, Non CDMI and Profile based interoperability approach
- Understanding of existing gaps that are present in CDMI, Non CDMI and Profile based specification

[Back to Top \(#top\)](#)

DE DUPE

Building Efficient All-Flash Scale-Out Block Storage with Deduplication Support

Doron Tal, Chief Architect, Kaminario

Abstract

Scale-out architectures have emerged in recent years as a way to address increasing capacity and performance needs. In this presentation we will discuss about the importance of combining scale-out and scale-up in a flexible way to gain the best possible TCO. Adding inline deduplication into the mix of scale-out and scale-up makes it much more challenging, but also more interesting. We will deep dive into the architecture details of the Kaminario scale-out block storage architecture to understand how global deduplication can be implemented in an efficient way which allows significant scalability while providing high performance.

Learning Objectives

- Flexible scale-out and scale up
- Variable size deduplication
- Implementing variable size deduplication on a scale-out architecture

[Back to Top \(#top\)](#)

DISTRIBUTED STORAGE

Next Generation Erasure Coding Techniques

Wesley Leggett, Senior Software Architect, Cleversafe

Abstract

Straight forward erasure coding methods can offer significant improvements to reliability, availability and storage efficiency. But these improvements are not nearly as optimal as they could be. Recently, we discovered a new method for storing erasure coded data in spite of an arbitrary number of failures. We call this technique Adaptive Slice Placement. This technique dispenses with old assumptions and redefines familiar concepts, and in the process yields a storage system with substantial benefits over traditional erasure coded systems. Adaptive Slice Placement can reduce overhead by a third while at the same time substantially improving reliability, availability, and performance. Please join us for our first public presentation of this new erasure coding technique.

Learning Objectives

- Some of the limitations and drawbacks of first-generation erasure coding
- How second-generation erasure coding works, and how it leads to such improvements
- Why adaptive slice placement necessary to realize second-generation erasure coding

[Back to Top \(#top\)](#)

Taming the Flood: Massively Scalable Semi-P2P Content Distribution

Yogesh Vedpathak, Software Developer, Cleversafe

Abstract

Sometimes demand for popular data exceeds the capability of storage servers to deliver it, but new protocols offer a solution. Protocols that leverage client resources can scale capacity to meet any level of demand—a famous example being BitTorrent. Yet there are many challenges with dynamically creating, tracking and seeding torrents to satisfy millions of users, accessing petabytes of data, across an enterprise-class storage system. Problems compound when peers are untrusted, potentially malicious and sometimes uncooperative. In this presentation we consider whether the problem can be solved for the general case, and evaluate the benefits of predictive caching, machine learning, and modifications to the BitTorrent protocol to create a storage system of truly unlimited capacity for content distribution.

Learning Objectives

- Massively scalable enterprise-class content distribution using torrent-based protocols
- Intelligent, predictive server-side caching based on past behavior and current state of clients
- How client trust models affect the efficiency of a P2P file distribution scheme

[Back to Top \(#top\)](#)

Benchmarking Cloud Storage through Standard Approach

Yaguang Wang, Sr. Software Engineer, Intel

Abstract

This session will cover the design and implementation details on supporting SNIA CDMI standard in COSBench (a benchmarking tool for cloud storage) support SNIA CDMI standard, and the CDMI implementations which are verified, and the recipes on making tests against CDMI server. Finally, an update of major enhancement will be shared.

Learning Objectives

- How is CDMI supported in COSBench
- What CDMI implementations are supported
- How to make tests to different CDMI servers
- Any other new enhancements included in the tool

[Back to Top \(#top\)](#)

Swift Object Storage: Adding Erasure Codes

Paul Luse, Sr. Staff Engineer, Intel

Kevin Greenan, Staff Software Engineer, Box

Abstract

This session will provide insight into this extremely successful community effort of adding an Erasure Code capability to the OpenStack Swift Object Storage System by walking the audience through the design and development experience through the eyes of the developers from key contributors. An overview of Swift Architecture and basic Erasure Codes will be followed by design/implementation details.

Learning Objectives

- Introduction to Swift Object Storage
- Introduction To Erasure Codes
- Design overview of Erasure Codes in Swift

[Back to Top \(#top\)](#)

FILE SYSTEMS

Data Deduplication for Distributed Segmented Parallel Filesystem

Boris Zuckerman, Distinguished Technologist, HP

Oskar Batuner, Expert Software Engineer, HP

Abstract

This presentation explores the design ideas behind de-duplicating of data in the distributed segmented parallel file systems (Ibrix). There are special challenges related to the large scale of our file system. Many entry point servers generate new content simultaneously; meta-data and directories are widely distributed; the system can grow both in capacity and performance by adding new storage segments and destination storage servers. While adding ability to de-duplicate the data content we have to preserve flexibility and scalability of the original design. This presentation shows the key points of our design for de-duplication: how to achieve the balance between efficiency of de-duplication and the size of indexes, how to use RAM efficiently, how to preserve parallelism and efficiency of I/O streams, how to avoid bottlenecks and scale linearly by adding more storage and servers.

Learning Objectives

- Expose fundamentals of the highly distributed segmented parallel file system architecture
- Review the challenges of associated with data de-duplication in such environment
- Explore details of the design: indexes, data containers, representative indexing, evolution of index
- Review effectiveness of data placement and parallelism of I/O streams
- Review the basis for scalability and parallelism

[Back to Top \(#top\)](#)

Samba and Btrfs - A Snapshot of Progress

Jim McDonough, Consulting Software Engineer, SUSE and Samba Team

Abstract

With strong community backing and an impressive array of features, Btrfs is widely regarded as _the_ next generation filesystem for Linux.

This talk will outline some of the features offered by Btrfs - namely snapshots, compression and clones - and demonstrate at how they can be exposed to Windows clients via Samba. In addition to the demonstration, this talk will also cover some Samba implementation details.

[Back to Top \(#top\)](#)

Reliable, Scaling and High Performance Storage System: LeoFS

Yosuke Hara, Lead Technologist, Rakuten, Inc.

Abstract

LeoFS is an unstructured data storage for the web and a highly available, distributed, eventually consistent storage system. Organizations are able to use LeoFS to store lots of data efficiently, safely, and inexpensively. I will present the design and architecture of LeoFS how we realized high reliable, high scalability and high performance ratio as well as demonstrate how developers are easily able to run and manage LeoFS in their environments. Also, I will introduce how we administrate LeoFS at Rakuten, Inc.

Learning Objectives

- To give a deep understanding of LeoFS
- To share how we realized high reliability, high performance and high scalability storage system
- Architects and project managers who want to discover highly reliable S3-compatible object storage sy

[Back to Top \(#top\)](#)

Toward High-Performance Shadow Migration

Youngjin Nam, Principal Software Engineer, Oracle

Aaron Dailey, Senior Manager, Oracle Storage

Abstract

Shadow migration has been widely used for migrating existing file systems to Solaris servers and the Oracle ZFS Storage Appliance (ZFSSA). Shadow migration is an interposing technology making the data in the old file system immediately available and modifiable in a new file system that "shadow's" the old. In this talk, we will explain how the technology works at a high level and discuss some of the challenges we've faced in optimizing and improving its performance.

Learning Objectives

- Understanding use cases of shadow migration in ZFSSA & Solaris
- Understanding shadow migration at a high level
- Understanding optimization issues for higher performance

[Back to Top \(#top\)](#)

Hadoop on Scality RING

Paul Speciale, Sr. Director of Product Management, Scality

Abstract

Scality RING is by design an object store but the market requires a unified storage solution. Why continue to have a dedicated Hadoop Cluster or an Hadoop Compute Cluster connected to a Storage Cluster ? With Scality, you do native Hadoop data processing within the RING with just ONE cluster. Scality leverages its own file system for Hadoop and replaces HDFS while maintaining HDFS API. Scality Scale Out File System aka SOFS is a POSIX parallel file system based on a symmetric architecture. This implementation addresses the Name Node limitations both in term of availability and bottleneck with the absence of meta data server with SOFS. Scality leverages also CDMI and continues its effort to promote the standard as the key element for data access.

Learning Objectives

- Illustrate a new usage of CDMI
- Learn Scality SOFS design with CDMI
- Address Hadoop limitations with CDMI

[Back to Top \(#top\)](#)

LustreFS and its ongoing Evolution for High Performance Computing and Data Analysis Solutions

Roger Goff, Senior Product Manager, DataDirect Networks

Abstract

The Lustre file system was born out of the need to deliver huge amounts of data to the fastest high performance computing systems in the world. Research institutions adopted it quickly and helped to create the fastest and most widely adopted parallel file system for research computing systems. From its early days in environments accepting a fair amount of instability, Lustre deployments are now found in commercial research labs across oil and gas, manufacturing, rich media, and finance sectors. This presentation will cover the elements and architecture of Lustre and several current technology developments helping to continue its evolution as a solid foundation for High Performance Computing and Data Analysis (HPDA).

Learning Objectives

- High level understanding of the Lustre filesystem architecture and elements
- Lustre recent and upcoming improvements for small file IO including architectural changes and the use of flash caches
- An approach to extending Lustre access beyond the data center and into the cloud for tiering, collaboration and disaster recovery

[Back to Top \(#top\)](#)

Multi-Protocol Support in GlusterFS

Ira Cooper, Principal Software Engineer, Red Hat

Abstract

With NFS, FUSE, SMB, SMB2, SMB2.1, SMB3, Cinder, and Swift all accessing one file system, you have the recipe for a nightmare for most companies, or a dream for a community charged up and ready to take the challenge on!

This presentation will cover how the flexible architecture of GlusterFS will be used to solve the problems of cache coherency (oplocks/leases/delegations), locking, ACLs, and share modes. The presentation will also show how similar approaches can be taken in the future to help extend our solutions, and solve future problems.

Learning Objectives

- Why multi-protocol access is a big problem
- A basic understanding of how GlusterFS is architected
- How we will use the architecture of GlusterFS to solve the multi protocol problems in a novel and efficient manner

[Back to Top \(#top\)](#)

BorgFS File System Metadata Index Search

**Stephen Morgan, Senior Staff Research Engineer, Huawei Technologies and
Masood Mortazavi, Distinguished Engineer, Huawei Technologies**

Abstract

This talk describes the design, implementation, and evaluation of the metadata index search component of a highly scalable, cost-efficient file system. Our system maintains traditional file system interfaces (e.g., POSIX) because they are used by many enterprise and consumer applications. However, storing hundreds of millions or billions of files in such a file system makes it difficult for a user to keep track of files and their status. Hierarchical naming is helpful up to a point, but does not solve the whole problem of managing files, which can easily be "lost." Therefore, in such large file systems, a search facility is required. Searching for a file by a combination of file name and metadata makes it easier to find files. A POSIX file system already stores metadata such as file owner, group, creation date, change date, and size. Here, we focus on facilities that we have built to maintain file metadata indices and to service file meta-data search queries. Our metadata search subsystem uses open-source components, an OS-level file system notification system and an index partitioning and distribution mechanism that allows for fast searches over billions of files. In a typical installation, typical queries, including those that touch all file indices, respond within reasonable delays. We also discuss future work.

Learning Objectives

- Metadata Index Search for a Scalable POSIX File System
- Use of Open Source Software in a Prototype System
- Overview of a Scalable Distributed POSIX File System

[Back to Top \(#top\)](#)

Updated Implementation of Hadoop Distributed File System Protocol on OneFS

Tanuj Khurana, Senior Software Engineer, EMC Isilon

Abstract

SDC 2012 introduced the concept of HDFS as a protocol on top of OneFS. Since then, Apache HDFS as well as the OneFS HDFS implementation have come a long way forward. This talk is a follow-up to that presentation and focuses on how the Isilon HDFS implementation has evolved to support some of the newer features in HDFS such as Kerberos authentication and WebHDFS. There is also talk of how our HDFS implementation integrates with Access zones which are the basic constructs for supporting multi-tenancy on Isilon clusters.

Learning Objectives

- Isilon implementation of HDFS Namenode, Datanode and WebHDFS protocols
- Isilon implementation of HDFS authentication
- Leveraging a single Isilon cluster to create multiple virtual HDFS clusters

[Back to Top \(#top\)](#)

HARDWARE

Advanced Barium Ferrite Tape Technologies

**Osamu Shimizu, Research Engineer, FUJIFILM Corporation and
Hitoshi Noguchi, General Manager, FUJIFILM Corporation**

Abstract

Barium ferrite tapes offer high capacity and long-term stability and are therefore set to replace the metal particulate tapes currently in widespread use. We recently developed an advanced barium ferrite magnetic tape medium that is more than comparable to the media to be used in the 128-TB cartridge expected to be launched in 2022 based on the 2012–2022 Roadmap from the Information Storage Industry Consortium. The limitations of enhancing the capacity by using metal particulate tapes are presented, followed by the details of the key features of advanced barium ferrite tapes, which include the mechanism for realizing high capacity with high reliability. In addition, the future prospects for tape media are discussed.

Learning Objectives

- Magnetic tape media technology
- Difference between metal particulate tape and barium ferrite tape
- How to realize higher capacity with barium ferrite tape
- Reliability and long-term archiveability of barium ferrite media
- Future prospects for magnetic tape

[Back to Top \(#top\)](#)

Storage Systems Can Now Get ENERGY STAR Labels and Why You Should Care

Dennis Martin, President, Demartek

Abstract

We all know about ENERGY STAR labels on refrigerators and other household appliances. In an effort to drive energy efficiency in data centers, the EPA announced its ENERGY STAR Data Center Storage program in December 2013 that allows storage systems to get ENERGY STAR labels. This program uses the taxonomies and test methods described in the SNIA Emerald Power Efficiency Measurement specification, which is part of the SNIA Green Storage Initiative. In this session, Dennis Martin will discuss the similarities and differences in power supplies used in computers you build yourself and in data center storage equipment, 80PLUS ratings, and why it is more efficient to run your storage systems at 230v or 240v rather than 115v or 120v. Dennis will share his experiences running the EPA ENERGY STAR Data Center Storage tests for storage systems and why vendors want to get approved.

Learning Objectives

- Learn about power supply efficiencies
- Learn about 80PLUS power supply ratings
- Learn about running datacenter equipment at 230v vs. 115v
- Learn about the SNIA Emerald Power Efficiency Measurement
- Learn about the EPA ENERGY STAR Data Center Storage program

[Back to Top \(#top\)](#)

Next Generation Storage Networking for Next Generation Data Centers

Dennis Martin, President, Demartek

Abstract

With 10GigE gaining popularity in data centers and storage technologies such as 16Gb Fibre Channel beginning to appear, it's time to rethink your storage and network infrastructures. Learn about futures for Ethernet such as 40GbE and 100GbE, 32Gb Fibre Channel, 12Gb SAS and other storage networking technologies. We will touch on some technologies such as NVMe, USB 3.1 and Thunderbolt 2 that may find their way into datacenters later in 2014. We will also discuss cabling and connectors and which cables NOT to buy for your next datacenter build out.

Learning Objectives

- What is the future of Fibre Channel?
- What I/O bandwidth capabilities are available with the new crop of servers?
- Share some performance data from the Demartek lab

[Back to Top \(#top\)](#)

Shingled Magnetic Recording – SMR Models, Standardization, and Applications

**Mary Dunn, Technologist, Seagate and
Timothy Feldman, Technologist, Seagate**

Abstract

Shingled Magnetic Recording (SMR) is a new technology that allows disk drive suppliers to extend the areal density growth curve with today's conventional components (heads, media) as well as providing compatibility with future evolutionary technologies. While the shingled recording subsystem techniques are similar amongst SMR device types, there are a variety of solutions with respect to the drive interface and resultant host implications across the various market segments utilizing HDDs. This presentation will provide insight into the primary SMR models—Drive Managed and Zoned Block Devices; provide an update on the progress in the Interface Committee Standards (ZBC/ZAC); provide insight regarding alignment of the SMR models to a variety of applications/workload models; and provide insights regarding other industry ecosystem infrastructure support.

Learning Objectives

- Gain an understanding of the 3 SMR Models—Drive Managed (Autonomous) direct access devices, and Host Aware & Host Managed Zoned Block devices
- Interface Committee Standardization for Zoned Block Devices: T10 (SCSI) ZBC and T13 (SATA) ZAC: New Commands and Best Practices
- Where to use Drive Managed SMR drives and where to use Zoned Block Devices
- Update on the State of the Storage Stack: File systems, Device drivers, Host bus adapters and port expanders

[Back to Top \(#top\)](#)

iSCSI

Next Generation iSCSI Enterprise Grade Data Integrity and Performance

Wael Noureddine, Vice President of Technology, Chelsio

Abstract

This session will explore the latest developments related to iSCSI, and discuss the attributes that differentiate iSCSI as a foundation for next generation storage platforms. Developed to enable SAN convergence, iSCSI has garnered broad industry support with native support in all major operating systems and hypervisors. Today, mature offloaded iSCSI implementations offer high performance, advanced data integrity protection, and leverage a robust TCP/IP foundation that allows operation over Wireless, LAN and WAN networks without the need for specialized equipment, switches, or forwarders. iSCSI today enables true network convergence and currently ships at 40Gbps with a roadmap to 100Gbps and beyond.

[Back to Top \(#top\)](#)

iSCSI Protocol Advancements from IETF Storm WG

**Mallikarjun Chadalapaka, Principal Program Manager, Microsoft and
Frederick Knight, NetApp**

Abstract

IETF Storm Working Group has just finished a major round of iSCSI protocol standardization. This session jointly presented by the co-authors of just-published RFC 7143 and 7144 will provide an overview of what to expect in the new iSCSI RFCs, and the architectural/design considerations behind the new protocol semantics.

Learning Objectives

- Overview of iSCSI standards landscape
- Architectural context of iSCSI in the SCSI protocol stack
- Appreciation of RFC 7143
- Appreciation of RFC 7144

[Back to Top \(#top\)](#)

KEY NOTE AND FEATURED SPEAKERS

RAMCloud and the Low-Latency Datacenter

John Ousterhout, Professor, Stanford University

Abstract

Datacenter computing has driven many of the innovations in computer systems over the last decade. The first phase of datacenter computing focused on scale (harnessing thousands of machines for a single application), but the next phase will focus on latency

(taking advantage of the close proximity between machines). In this talk I will discuss why low latency matters in datacenters and how it will be achieved over the next 5-10 years. I will also introduce RAMCloud, a storage system that keeps all data in DRAM at all times in order to provide 100-1000x faster access than existing storage systems. Low-latency datacenters, combined with infrastructure such as RAMCloud, will enable a new class of applications that manipulate large datasets more intensively than has ever been possible.

[Back to Top \(#top\)](#)

New Directions for NAND Memory in the Changing Data Center Era

Bob Brennan, Senior Vice President, Samsung Memory Solutions Lab

Abstract

Memory innovators, led by Samsung, are blazing new directions in memory components and memory architecture to help advance digital communications and the data center infrastructure behind them. Most importantly, NAND flash is taking on an ever-greater role in increasingly interconnected computing markets. Samsung's recently introduced V-NAND technology and other NAND advances will enable much more responsiveness in handling clouds and big data. Vertically integrated and TCO-optimized NAND solutions, including 3-bit technology, will have greater impact in hyperscale data centers. In addition, enterprise SSDs, such as V-NAND drives and PCIe NVMe SSDs, offer compelling benefits for data center managers. Furthermore, advances in NAND architecture will play an increasingly important role in propelling the cloud to new horizons, while major improvements in NAND performance, reliability, power efficiency and endurance reduce the data center TCO.

[Back to Top \(#top\)](#)

Leveraging Software-Defined Storage to Meet Today and Tomorrow's Infrastructure Demands

Molly Rector, Executive Vice President, Product Management, DataDirect Networks

Abstract

Optimizing the data driven businesses requires balancing data and budget growth. A Software-Defined Storage strategy enables users to get all the value they desire from Cloud-type infrastructures with the flexibility to optimize for individual cost, security and access requirements. This session focuses on the key interfaces, capabilities and benefits Software-Defined Storage platforms enable to bring web-scale infrastructure flexibility and capabilities to any data center.

[Back to Top \(#top\)](#)

Software Defined Storage: Changing the Rules for Storage Architects

Ric Wheeler, Director of Red Hat Storage Engineering, Red Hat

Abstract

Software Defined Storage, to those of us who have been writing storage software for years, sounds like yet another marketing term. In effect, software defined storage changes the model for how our users do storage - they buy the hardware and storage architects write the software. This talk will give an overview of how that impacts storage architects and also discuss how open source software plays an important role in making SDS viable for both storage designers and storage consumers.

[Back to Top \(#top\)](#)

Big Data Storage Challenges for Industrial Internet

**Shyam Nath, Principal Architect, GE and
Diwakar Kasibhotla, Principal Architect, GE**

Abstract

Industrial and Machine data is pushing the storage paradigms to new limits. With the Internet of Things connecting 26 billion new "things" by 2020, data centers will go through complete transformation to handle the Big Data. In order to make a large scale economic impact on country level infrastructure, the sensor data from the industrial machines such as jet engines, locomotives, power generation equipment and utilities have to be analyzed with very little latency. Such sensor data has to be married with other Enterprise data typically stored in Asset Management and other ERP and CRM systems. This session will showcase such challenges in context of Industrial Internet of Things. The framework for data ingestion, transformation, analysis as well as persistence of data, in context of streams and near-real-time batches will be discussed. We will address the different dimensions of Big Data namely volume, velocity and variety from storage and retrieval perspective.

Learning Objectives

- Nature of Data from Internet of Things
- Storage Challenges posed by Machine and Industrial Data
- Nature of Industrial Internet of Things
- Storage Paradigm for Data Analysis
- Marrying Machine Data with Human Data

[Back to Top \(#top\)](#)

Software Defined Storage - Moving Beyond the Hype

Greg Scott, Chief Cloud Storage Strategist, Intel - Platinum Sponsor

Abstract

Deployment of virtualization technologies has had a significant impact on the way companies view and manage their IT infrastructure. Virtualization has enabled both enterprises and cloud service providers to improve utilization and drive down cost. This is resulting in a fundamental change in the way Information Technology (IT) is viewed and managed as it shifts from a being a cost center to an efficient service. To fully realize this change, the Datacenter must be viewed as a single system, not a collection of isolated pools of hardware. As server virtualization leads to the separation of the application from the physical HW, Software Defined Infrastructure (SDI) will further separate the compute, networking, and storage functions from the physical data center environment, to provide better utilization and lower costs. Over the last year, Software Defined Storage (SDS) has become a popular marketing term, however there is much less agreement on SDS framework and collective thinking in the industry to take advantage of this disruptive trend in the industry. In this keynote, we explore SDS framework within the context of Software Defined Infrastructure, how it enables flexible service delivery model, role of open interfaces and call to action for broader adoption in the industry.

Learning Objectives

- Understand Software Defined Infrastructure and how SDS fits into this broader framework
- Role of SDS in addressing storage management challenges
- Understand existing open standards role as well as need for new standards
- Industry wide initiatives needed to ramp SDS solutions and promote interoperability

[Back to Top \(#top\)](#)

Who Moved My Bits?

Val Bercovici, Big Data and Cloud Czar, Office of the CTO, NetApp

Abstract

Data drives our modern economy with a spectacular variety of sources, apps and consumption models expanding on a regular basis. A virtuous circle has emerged starting with the consumerization of technology driving dramatic supply chain roadmap updates for the storage industry. This talk will explore the new data processing trends and their unexpected impacts on new classes of storage which will emerge to address these trends. Apps, Databases, Media and Data Center design will all be impacted. Will these impacts change the way you plan to create or consume storage by 2020?

[Back to Top \(#top\)](#)

Is It Really All Going into the Cloud?

Geoff Barrall, Chief Executive Officer, Drobo

Abstract

Storage administrators find themselves walking a line between meeting employees demands to use public cloud storage services, and their organizations need to store information on-premises for security, performance, cost and compliance reasons. However, as file sharing protocols like CIFS and NFS continue to loose their relevance, simply relying only a NAS-based environment creates inefficiencies that hurt productivity and the bottom line. IT wants to implement cloud storage it can purchase and own like NAS (NTAP) but that works like traditional public cloud storage services like Dropbox, Box and Google Drive. This talk will look at what's really happening with file protocols and explore the truth behind Silicon Valley's cloud dreams

[Back to Top \(#top\)](#)

NEW THINKING

In Search of an Understandable Consensus Algorithm

Diego Ongaro, PhD, Stanford University

Abstract

Consensus is a fundamental building block for fault-tolerant systems, but it's poorly understood. We struggled to build a real system with Paxos, the most widely used consensus algorithm today. As a result, we developed Raft to be easier to understand. In this talk, I'll give an overview of how Raft works. More info on Raft can be found at <http://raftconsensus.github.io>

[Back to Top \(#top\)](#)

Failure-Atomic Msync(): A Simple and Efficient Mechanism for Preserving the Integrity of Durable Data

Stan Park, Research Engineer, HP Labs

Abstract

Preserving the integrity of application data across updates is difficult if power outages and system crashes may occur during updates. Existing approaches such as relational databases and transactional key-value stores restrict programming flexibility by mandating narrow data access interfaces. We have designed, implemented, and evaluated an approach that strengthens the semantics of a standard operating system primitive while maintaining conceptual simplicity and supporting highly flexible programming: Failure atomic msync() commits changes to a memory-mapped file atomically, even in the presence of failures. Our Linux implementation of failure-atomic msync() has preserved application data integrity across hundreds of whole-machine power interruptions and exhibits good microbenchmark performance on both spinning disks and solid-state storage. Failure-atomic msync() supports higher layers of fully general programming abstraction, e.g., a persistent heap that easily slips beneath the C++ Standard Template Library. An STL built atop failure-atomic msync() outperforms several local key-value stores that support transactional updates. We integrated failure-atomic msync() into the Kyoto Tycoon key-value server by modifying exactly one line of code; our modified server reduces response times by 26–43% compared to Tycoon's existing transaction support while providing the same data integrity guarantees. Compared to a Tycoon server setup that makes almost no I/O (and therefore provides no support for data durability and integrity over failures), failure-atomic msync() incurs a three-fold response time increase on a fast Flash-based SSD—an acceptable cost of data reliability for many.

[Back to Top \(#top\)](#)

Aerie: Flexible File-System Interfaces to Storage-Class Memory

Haris Volos, Research Engineer, HP

Abstract

Storage-class memory technologies such as phase-change memory and memristors present a radically different interface to storage than existing block devices. As a result, they provide a unique opportunity to re-examine storage architectures. We find that the existing kernel-based stack of components, well suited for disks, unnecessarily limits the design and implementation of file systems for this new technology.

We present Aerie, a flexible file-system architecture that exposes storage-class memory to user-mode programs so they can access files without kernel interaction. Aerie can implement a generic POSIX-like file system with performance similar to or better than a kernel implementation. The main benefit of Aerie, though, comes from enabling applications to optimize the file system interface. We demonstrate a specialized file system that reduces a hierarchical file system abstraction to a key/value store with fewer consistency guarantees but 20–109% higher performance than a kernel file system.

[Back to Top \(#top\)](#)

Tango: Distributed Data Structures Over a Shared Log

Mahesh Balakrishnan, Microsoft

Abstract

We argue that highly available, strongly consistent distributed systems can be realized via a simple storage abstraction: the shared log. In this talk, we describe Tango objects, a new class of replicated, in-memory data structures (maps, lists, queues, etc.) backed by a shared log. Replicas of a Tango object are synchronized via simple append and read operations on the shared log instead of complex distributed protocols. The shared log is the source of durability and consensus in the system, subsuming the role of protocols such as Paxos. In addition, it enables ACID transactions across multiple Tango objects. Distributed systems such as ZooKeeper and BookKeeper can be replaced by Tango objects comprising a few hundred lines of code. In turn, these Tango objects are used to harden the meta-data components of larger systems (e.g., the HDFS name-node).

[Back to Top \(#top\)](#)

NFS

Dynamic Placement Layouts in pNFS

**Adam C. Emerson, Software Engineer, CohortFS, LLC and
Marcus Watts, Software Engineer, CohortFS, LLC**

Abstract

The current NFS layout types are insufficient for next generation distributed file systems. Pseudorandom structure traversals and distributed hash tables are too complex to be captured by rectangular striping arrays, and distributed systems that use them thus have difficulty making full use of pNFS. We are developing a dynamic placement layout that includes executable code that can be evaluated to find the correct location for data. Dynamic placement layouts are better for vendors and users, as their support on clients will allow many distributed filesystem system to be exported through NFS, rather than having a layout type implemented for each strategy.

Learning Objectives

- Learn about latest developments in NFSv4
- Review distributed data placement strategies in NFSv4 context
- Keep up with distributed file system standardization proposals

[Back to Top \(#top\)](#)

Customer-Oriented Storage Performance Management

Dany Felzenswalbe, Senior Systems Engineer, Intel

Abstract

End-users always want to get the "best" performance for their project activities, regardless of which file servers they use.

The monitoring of the NAS performance on file servers in Intel has transformed in the last few years and we have gone from no visibility on performance to good visibility and control over it. The user-experience monitoring, that we have been using, generally provides good correlation to when users are being slowed down because of the file-servers but we also miss performance problems with this method, some of which have a major impact to the projects' progress.

This presentation will discuss the different methodologies used to monitor NFS file-services performance in Intel's design data-centers and the performance is controlled and managed.

[Back to Top \(#top\)](#)

NFS-Ganesha for Clustered NAS

**Poornima Gupte, Senior Staff Software Engineer, IBM and
Venkateswararao Jujuri, Sr. Software Engineer, IBM**

Abstract

With the storage requirements growing at exponential rate and the industry shift towards cloud computing and software defined architecture, the need for bigger, reliable and centralized storage servers is increasing. Industry is making use these centralized clustered storage units to serve the storage needs. NFS-Ganesha, a user space NFS server has been gaining popularity and has become central part of multiple industry leaders in serving NFS side of the NAS needs. NFS-Ganesha abstracts out various file systems interfaces through its unique File System Abstract Layer (FSAL). Because of this, NFS-Ganesha is able to support various types of file systems like GPFS, Ceph, Gluster, Lustre, VFS, ZFS and more. As the enterprise users are demanding clustered NAS, we have introduced a new clustering framework, called Cluster Manger Abstraction Layer (CMAL) which allows NFS-Ganesha to work with various cluster mangers seamlessly.

In this presentation we intent to describe the CMAL interface and how it can be used to implement Clustered Duplicate Reply Cache(cDRC), Cluster wide distributed Lock manager and Lock Recovery.

Learning Objectives

- Understand how NFS-Ganesha can be used in clustered NAS
- Overview of CMAL interface for NFS-Ganesha
- Understand some of the use-cases of CMAL, like Clustered DRC

[Back to Top \(#top\)](#)

Introducing FedFS on Linux

Chuck Lever, Linux Kernel Architect, Oracle

Abstract

FedFS provides a standard way to create a network file namespace that crosses server and share boundaries (similar to autofs). Presenter will introduce the FedFS standard, illustrate storage administration benefits, and walk through the FedFS implementation on Linux.

Learning Objectives

- What is FedFS?
- Scaling storage administration with FedFS
- How to install and configure on Linux

[Back to Top \(#top\)](#)

OBJECT STORAGE

Best Practice on Distributed Intelligent Storage with NVMe-SSDs and Fast Interconnect

Dieter Kasper, CTO Data Center Infrastructure, Fujitsu

Abstract

The open source distributed storage solution Ceph is designed for highest scalability while maintaining performance and availability. RADOS as a self-healing reliable, autonomous, distributed object store is the base for the unified front-end layers File, Block and Object.

Best practice hints will demonstrate how NVMe-SSDs can help to achieve highest throughput for block, file while keeping Storage efficiency. Fast interconnects like Infiniband in combination with tuning of the code can help to reduce latency in an distributed environment with commodity server in a significant manner.

Learning Objectives

- Get an overview on NVMe-SSDs and fast interconnect protocols
- Learn how NVMe-SSDs can help to get highest throughput
- Learn how tuning of the Ceph code can reduce IO latency
- Learn about the influence of fast interconnect protocols

[Back to Top \(#top\)](#)

Kinetic Open Storage Platform

Mayur Shetty, Senior Solutions Architect, Seagate

Abstract

The Kinetic Open Storage platform reduce the inefficiencies of traditional datacenters whose legacy architectures are not well-adapted to highly distributed and capacity-optimized workloads. Kinetic Open Storage platform is a new class of key/value Ethernet drives + an open API and series of libraries designed to provide the simplest semantic abstraction and enable applications through an easy-to-use, minimalist API that allows the application direct access to the storage.

Learning Objectives

- Understand Key/Value Storage
- Understand the value to applications
- Understand the requirements that a key/value system provides on an infrastructure

[Back to Top \(#top\)](#)

OPENSTACK

Delivering Standards Based SDS Framework with OpenStack SDS Controller Implementation

Anjaneya Chagam, Principal Engineer, Intel

Abstract

Software Defined Storage (SDS) has significant impact on how companies deploy and manage public and private cloud storage solutions to deliver on-demand storage services while reducing the cost. Similar to Software Defined Networking (SDN), SDS

promises to simplify management of diverse storage solutions and ease of use. However in order to deliver this promise, there is a need to define SDS framework with specific focus on abstracting control plane functionality that paves the way for using distributed storage solutions on standard high volume servers as well as traditional storage appliances. This presentation explores standards based SDS framework for north bound and south bound interaction as well as working prototype using Openstack Cinder with SNIA standards (SMI-S, CDMI) based integration.

Learning Objectives

- Learn Software Defined Storage framework for managing cloud wide storage services
- Understand SDS controller control plane abstraction, application and storage interfaces using open standards (SMI-S, CDMI)
- Understand the paradigm shift in application integration using Service Level Objectives and how SDS controller abstracts underlying storage system implementations
- Learn about the emerging OpenStack shared file service, Manila
- Learn about new developments in OpenStack storage

[Back to Top \(#top\)](#)

OpenStack Cloud Storage

Sam Fineberg, Distinguished Technologist, HP

Abstract

OpenStack is an open source cloud operating system that controls pools of compute, storage, and networking. It is currently being developed by thousands of developers from hundreds of companies across the globe, and is the basis of multiple public and private cloud offerings. In this presentation I will outline the storage aspects of OpenStack including the core projects for block storage (Cinder) and object storage (Swift), as well as the emerging shared file service. It will cover some common configurations and use cases for these technologies, and how they interact with the other parts of OpenStack. The talk will also cover new developments in Cinder and Swift that enable advanced array features, QoS, new storage fabrics, and new types of drives.

Learning Objectives

- Learn what OpenStack is, and what storage support is available in OpenStack
- Learn about the OpenStack block storage service, Cinder
- Learn about the OpenStack object storage service, Swift
- Learn about the emerging OpenStack shared file service, Manila
- Learn about new developments in OpenStack storage

[Back to Top \(#top\)](#)

OpenStack Manila - File Storage

Robert Callaway, Reference Architect NetApp's Cloud Solutions Group, NetApp

Abstract

This page documents the concept and vision for establishing a shared file system service for OpenStack. The development name for this project is Manila. We propose and are in the process of implementing a new OpenStack service (originally based on Cinder). Cinder presently functions as the canonical storage provisioning control plane in OpenStack for block storage as well as delivering a persistence model for instance storage. The File Share Service prototype, in a similar manner, provides coordinated access to shared or distributed file systems. While the primary consumption of file shares would be across OpenStack Compute instances, the service is also intended to be accessible as an independent capability in line with the modular design established by other OpenStack services. The design and prototype implementation provide extensibility for multiple backends (to support vendor or file system specific nuances / capabilities) but is intended to be sufficiently abstract to accommodate any of a variety of shared or distributed file system types. The team's intention is to introduce the capability as an OpenStack incubated project in the Juno time frame, graduate it and submit for consideration as a core service as early as the as of yet unnamed "K" release.

[Back to Top \(#top\)](#)

How to Manage your Swift Cluster Using Swift Metrics

Sreedhar Varma, Director, SW Development, Vedams, Inc

Abstract

Redundancy is built into OpenStack Swift at various levels that I/O operations are capable of riding through failures happening in the cluster. The failures could be disk faults, services stopping or failing, node failures, etc. This presentation talks about building a monitoring system that is constantly receiving and analyzing the Swift metrics and reporting the status of the cluster to the Administrator. We present techniques to baseline a Swift cluster, figure out variances in metrics during failures and ways to report appropriate metrics and errors in a dashboard to the administrator. We will also be covering setup of configuration files in order to enable reporting of StatsD metrics and Swift-Informant metrics in the Swift cluster.

Learning Objectives

- Monitoring and managing a Swift cluster
- Receiving StatsD metrics from Swift cluster
- Receiving Swift-Informant metrics from Swift cluster
- Tuning a swift cluster

[Back to Top \(#top\)](#)

OPEN SOURCE SOFTWARE

Finding the Right Open Source Storage

Ric Wheeler, Director of Red Hat Storage Engineering, Red Hat

Abstract

In the open source storage world, there is a wealth of options to choose from. This presentation will give a technical overview of several of these technologies and go into the current challenges that we face in the upstream development communities. We will also present some guidance on typical use cases for specific technologies and try to give guidance on how to choose.

[Back to Top \(#top\)](#)

Storage Tiering and Erasure Coding in Ceph

Sage Weil, CTO, Inktank

Abstract

Ceph is designed around the assumption that all components of the system (disks, hosts, networks) can fail, and has traditionally leveraged replication to provide data durability and reliability. The CRUSH placement algorithm is used to allow failure domains to be defined across hosts, racks, rows, or datacenters, depending on the deployment scale and requirements.

Recent releases have added support for erasure coding, which can provide much higher data durability and lower storage overheads. However, in practice erasure codes have different performance characteristics than traditional replication and, under some workloads, come at some expense. At the same time, we have introduced a storage tiering infrastructure and cache pools that allow alternate hardware backends (like high-end flash) to be leveraged for active data sets while cold data are transparently migrated to slower backends. The combination of these two features enables a surprisingly broad range of new applications and deployment configurations.

This talk will cover a few Ceph fundamentals, discuss the new tiering and erasure coding features, and then discuss a variety of ways that the new capabilities can be leveraged.

[Back to Top \(#top\)](#)

PERFORMANCE

NFS over 40Gbps iWARP RDMA

Wael Noureddine, Vice President of Technology, Chelsio

Abstract

NFS over RDMA is an exciting development that allows storage managers unprecedented Ethernet performance and efficiency. The Internet Wide Area RDMA Protocol (iWARP) is the IETF standard for RDMA over Ethernet and it is particularly well suited for storage protocols like NFS (SMBDirect is another example) because of their characteristics and performance requirements.

iWARP builds upon the proven TCP/IP foundation, and lets users preserve their investments in network security, load balancing and monitoring appliances, and switches and routers. pNFS is another significant advance for NFS, removing scaling bottlenecks through parallelizing client access to storage.

This presentation goes over the trends motivating the use of RDMA in storage applications, and discusses the latest developments on the NFS/RDMA front. It presents new benchmark results for NFS running over 40Gb Ethernet with iWARP RDMA, showing the benefits in performance and efficiency it brings in.

Learning Objectives

- The audience will gain an understanding of the benefits of RDMA and iWARP
- The audience will gain an understanding of the performance comparisons of 40GE vs FDR IB
- The audience will gain an understanding of the performance comparisons of NIC/40GbE, iWARP/40GbE and FDR IB

[Back to Top \(#top\)](#)

Impact of Hypervisor Based Data Acceleration Tier on Virtualized Applications

Chethan Kumar, Member of Technical Staff, PernixData

Abstract

Hypervisor-based I/O acceleration tiers built using server side high-speed I/O media promise to scale the I/O performance of virtualized data centers to new heights. This talk quantifies the core hypothesis of an acceleration tier – liberating I/Os from the clutches of data access latency, and achieving linear storage scale. I/O profiles of enterprise applications with varying I/O requirements, from low latency to high throughput are used to study the impact of the acceleration tier on the applications' reads and writes both without and with fault tolerance across a hypervisor (vSphere) cluster. This talk will also analyze the effectiveness of various low latency IO media as hardware building blocks for said acceleration tier.

Learning Objectives

- Performance implications of accelerating reads and writes of virtualized applications
- High-speed server side I/O media and their impact on I/O acceleration
- Scaling I/O acceleration across a vSphere cluster

[Back to Top \(#top\)](#)

Storage Performance Council: Comprehensive, Industry Standard Benchmarks for Storage Products

Walter E. Baker, CEO/Founder, Gradient Systems, Inc

Abstract

The presentation will provide insight into the complex, comprehensive aspects of the SPC workloads, including the use of "hot spots", implementation of data re-reference, the use of multiple types of sequential access patterns and applicability to various application types. A more detailed review of various measured SPC data for both "internal" and "external" use will also be presented, leading to discussion of the technology 'agnostic' nature of SPC workloads/benchmarks, which provides applicability for multiple storage technologies such as HDDs, SSDs and evolving storage technologies. The presentation will conclude with an announcement and review of the new SPC-1 Toolkit, which includes a number of enhancements such as a defined content mix and support of compression. The announcement will also include the initial presentation of new SPC-1 Results from multiple SPC member companies using the new SPC-1 Toolkit

[Back to Top \(#top\)](#)

Deploying Ceph with High Performance Networks, Architectures and Benchmarks for Block Storage Solutions

John Kim, Director of Storage Marketing, Mellanox Technologies

Abstract

As data continues to grow exponentially storing today's data volumes in an efficient way is a challenge. Many traditional storage solutions neither scale-out nor make it feasible from Capex and Opex perspective, to deploy Peta-Byte or Exa-Byte data stores. A novel approach is required to manage present-day data volumes and provide users with reasonable access time at a manageable cost.

This paper summarizes the installation and performance benchmarks of a Ceph storage solution. Ceph is a massively scalable, open source, software-defined storage solution, which uniquely provides object, block and file system services with a single, unified Ceph Storage Cluster. The testing emphasizes the careful network architecture design necessary to handle users' data throughput and transaction requirements. Benchmarks show that a single user can generate read throughput requirements able to saturate a 10Gbps Ethernet network, while the write performance is largely determined by the cluster's media (Hard Drives and Solid State Drives) capabilities. For even a modestly sized Ceph deployment, the usage of a 40Gbps Ethernet network as the cluster network ("backend") is imperative to maintain a performing cluster.

Learning Objectives

- Challenges of big data storage
- Ceph as an open source storage solution
- Network requirements for large-scale Ceph deployments
- Using 40GbE to increase Ceph performance
- Benchmark test results

[Back to Top \(#top\)](#)

SPEC SFS 2014 - The Workloads and Metrics an Under-the-Hood Review

**Spencer Shepler, Architect, Microsoft and
Nick Principe, Senior Software Engineer, EMC**

Abstract

Historically, the SPEC SFS benchmark and its NFS and CIFS workloads have been the industry standard for peer reviewed, published performance results for the NAS industry.

The current generation of SPEC SFS benchmark suffers from two major flaws. The first is that it generates the workload directly by constructing the NFS and CIFS protocol requests directly and thus limits itself to what portions of the file system and storage can be measured. The second is a result of the first in that creating new workloads on top of this framework is difficult and truly synthetic.

With the new version of SFS, both of these issues are addressed. First, the SPEC SFS 2014 benchmark generates the workloads using traditional operating system APIs. Second, the workload definitions are easy to define and thus the benchmark framework provides for measurement of multiple workload types.

This presentation will provide an in-depth review of the workloads being delivered in SPEC SFS 2014 and the methods used to develop them. The attendee will leave with the knowledge to effectively understand reported SPEC SFS 2014 results or to start using the benchmark to measure and understand their own file systems or storage system.

[Back to Top \(#top\)](#)

PROFESSIONAL DEVELOPMENT

Leverage Agile Project Management to Foster Collaboration in Distributed Teams

Hasnain Rizvi, CIO and Agile Coach, AAA University

Abstract

Rapid growth of global markets is forcing organizations to become more flexible and responsive. Effective Project Management with distributed teams is a critical success factor, with many working towards Agile-centric frameworks. Yet many organizations today face a 'crisis' in projects across distributed teams. Introduction of total quality management, continuous improvement programs and the drive to radically redesign business processes requires an alignment with strong project management skills that can successfully lead distributed resources. Successful and effective implementation of change employs specific skills, which are no more the domain of a few technical professionals. Proficiency in these skills is a prerequisite to managing change and growth at all levels. Agile and distributed team models seem to be at odds with each other. One is about close communication and short feedback loops, while the other is about being effective with resources in different locations. Yet Agile Project Management can provide a structured and organized way to successfully leverage the strengths of distributed teams to foster collaboration.

Practical insights from prior projects managed by Hasnain will be covered in this presentation. Hasnain's talk provides insight into patterns common for setting up Agile Distributed teams and will show the results that can be achieved once teams cross the initial evolutionary bumps of establishing a distributed agile culture.

[Back to Top \(#top\)](#)

RDMA

Running SMB3.x Over RDMA on Linux Platforms

**Mark Rabinovich R&D Manager, Visuality Systems
John Kim, Director of Storage Marketing, Mellanox Technologies**

Abstract

Enabling high-end storage solution for a LINUX-based CIFS solution requires an advanced transport layer. We will speak about the methodology and the architecture of the Mellanox network acceleration platform which enables a powerful and scalable transport layer. Being coupled with the Storage version of Visuality's NQ CIFS Server, this combination grants an enterprise-level SMB traffic. SMB3.x, indeed, is the must to make this happen.

Learning Objectives

- How an effective network acceleration platform for Linux should look like
- How to benefit from a file-oriented RDMA on Linux
- How to combine SMB3.x solution and RDMA together

[Back to Top \(#top\)](#)

Enhancements to the iSER iSCSI Protocol

Sagi Grimberg, Senior Software Engineer, Mellanox Technologies**Abstract**

As customers deploy more flash storage and storage vendors support more 40Gb and 56Gb connections, faster block protocols are needed for server-storage connections. iSER is iSCSI over RDMA and provides significantly higher throughput and lower latency than other block protocols while still supporting the availability and management features of iSCSI. Sagi Grimberg will cover the latest enhancements to iSER.

Learning Objectives

- Enhancements to iSER
- iSER added to VMware
- OpenStack and iSER
- Latest iSER performance

[Back to Top \(#top\)](#)**RDMA Requirements for High Availability in the NVM Programming Model****Doug Voigt, Distinguished Technologist, HP****Abstract**

The SNIA NVM programming model includes actions that assure durability of data in persistent memory. New work is now in progress on remote access to NVM in support of high availability for persistent memory. While RDMA is an obvious choice of technologies for this purpose, certain challenges arise with single microsecond latencies for remote synchronization of small byte ranges. This presentation describes the work the SNIA NVM Programming Technical Working group on RDMA requirements for highly available persistent memory.

Learning Objectives

- Understand the role of RDMA in highly available persistent memory
- Understand how this use of RDMA differs from typical usage in HPC and storage systems
- Understand new requirements placed on RDMA and the potential roles of hardware and software in addressing them

[Back to Top \(#top\)](#)**SECURITY****Samba's AD DC: Samba 4.2 and Beyond****Andrew Bartlett, Samba Developer, Samba Team****Abstract**

This will be a overview of where Samba's AD DC is in Samba 4.2, where we are going with it for future releases, and a discussion of how the tools it contains it can be leveraged by storage, cloud and identify industry vendors.

Learning Objectives

- Understand the current state of Samba 4.2 and the AD DC
- Understand how to apply the Samba 4.2 AD DC in your product

[Back to Top \(#top\)](#)**Practical Secure Storage: A Vendor Agnostic Overview****Walt Hubis, Storage Standards Architect, Hubis Technical Associates****Abstract**

This presentation will explore the fundamental concepts of implementing secure enterprise storage using current technologies. This tutorial will focus on the implementation of a practical secure storage system, independent of any specific vendor implementation or

methodology. The high level requirements that drive the implementation of secure storage for the enterprise, including legal issues, key management, current technologies available to the end user, and fiscal considerations will be explored in detail. In addition, actual implementation examples will be provided that illustrate how these requirements are applied to actual systems implementations.

This presentation has been significantly updated to include emerging technologies and changes in the international standards environment (ISO/IEC) as related to storage security.

Learning Objectives

- Understand what is driving the need for secure storage

[Back to Top \(#top\)](#)

Best Practices for Cloud Security and Privacy

Eric Hibbard, CTO Security & Privacy, Hitachi Data Systems

Abstract

As organizations embrace various cloud computing offerings it is important to address security and privacy as part of good governance, risk management and due diligence. Failure to adequately handle these requirements can place the organization at significant risk for not meeting compliance obligations and exposing sensitive data to possible data breaches. Fortunately, ISO/IEC, ITU-T and the Cloud Security Alliance (CSA) have been busy developing standards and guidance in these areas for cloud computing, and these materials can be used as a starting point for what some believe is a make-or-break aspect of cloud computing.

This session provides an introduction to cloud computing security concepts and issues as well as identifying key guidance and emerging standards. Specific CSA materials are identified and discussed to help address common issues. The session concludes by providing a security review of the emerging ISO/IEC and ITU-T standards in the cloud space.

Learning Objectives

- General introduction to cloud security threats and risks
- Identify applicable materials to help secure cloud services
- Understand key cloud security guidance and requirements

[Back to Top \(#top\)](#)

SMB

The Rewards of Jealousy: An SMB2 Toolkit in Python

**Christopher Hertel, Storage Architect, Samba Team / Red Hat and
Jose Rivera, Software Engineer, Red Hat**

Abstract

Python is a widely popular object-oriented programming language, useful for rapid prototyping and integration of disparate software modules into a cohesive whole. The 2013 Storage Developer's Conference saw the announcement of a new SMB2 test framework written in Python, which generated both interest and a bit of code envy among some in the audience. This presentation will cover the fruits of that envy, an as-yet-unnamed SMB2 rapid development toolkit written in Python. Imitation is the sincerest form of flattery.

Learning Objectives

- Open Source Toolkit
- Rapid SMB2 development
- SMB2 internals

[Back to Top \(#top\)](#)

A New DCERPC Infrastructure for Samba

Stefan Metzmacher, Developer, Samba Team / SerNet

Abstract

There are currently 4 independent DCERPC implementations (2 servers and 2 clients). They work fine, but they're missing some important features.

The new infrastructure will combine all 4 implementations and add important new features: full async client and server support, support for association groups, multiplexing of security contexts, multiplexing of presentation contexts, support for DCERPC pipes and maybe DCERPC callbacks.

- SMB Witness support for file server clusters
- SYSVOL replication for domain controllers
- the new async printing of Windows 8
- remote filesystem snapshot support
- windows search protocol support
- maybe better DCOM and WMI support

It will also be possible to use this new infrastructure by external projects like OpenChange and maybe others in future.

Learning Objectives

- Why is a new infrastructure needed
- How is the new infrastructure designed
- How this impacts Samba's file server (cluster) support
- Obtain more information about the state of this project

[Back to Top \(#top\)](#)

Scalable CHANGE_NOTIFY

Volker Lendecke, Developer, Samba Team / SerNet

Abstract

Samba's implementation of the SMB CHANGE_NOTIFY request has seen a few iterations. The first implementation part of Andrew Tridgell's NTvfs effort in Samba 4 created the first understanding of the semantics of that request before the SMB documents were published by Microsoft. Since then, the Samba Team has made significant modifications to the internal algorithm and data structures, in particular to make CHANGE_NOTIFY scale well in a clustered environment. This talk will cover the history of our CHANGE_NOTIFY.

Implementation and describe how Samba now implements a very well scalable and low-overhead implementation of recursive CHANGE_NOTIFY.

- CHANGE_NOTIFY is difficult to make scalable, in particular in a cluster environment
- Samba implements CHANGE_NOTIFY with very little overhead
- Samba offers a simple interface for other protocols like NFS to interoperate for CHANGE_NOTIFY
- Objective4

[Back to Top \(#top\)](#)

Beyond SMB3: New Developments in the Linux SMB3 Implementation

Steven French, Principal System Engineer, Samba Team / Primary Data

Abstract

With another year of development, the Linux SMB3 kernel implementation continues to improve. This presentation will summarize the current status of the Linux SMB3 implementation, and newly implemented features such as compression and server copy offload. It will also discuss the status of the Unix Extensions for SMB3 which will help provide more strict POSIX semantics over SMB3 mounts, and also other proposals for improving SMB3 in Linux environments. Performance and stability of the SMB3 implementation have been much improved and the progress on multicredit support will be described. Newly implemented features will also be demonstrated.

Learning Objectives

- What new SMB3 features are now available on Linux and how are they enabled and configured?
- When should you use SMB3 for Linux?
- Are the SMB3 Unix Extensions important for your workload?
- How has the performance of SMB3 on Linux improved?
- How do we test the Linux SMB3 implementation?

[Back to Top \(#top\)](#)

smb[3]status

Michael Adam, Team Lead Samba, Samba Team / SerNet

Abstract

Version 3 added a whole set of new features to the SMB protocol, notably all-active clustering capabilities. The two most prominent consequences are that it is now possible to run Hyper-V instances off an SMB share and that SMB supports RDMA as a transport in the guise of the so called SMB Direct.

Samba supports basic SMB3 since version 4.0, but without these more advanced features. Designs and plans have been developed for implementing SMB Direct and all that is needed for Hyper-V-support, and meanwhile, development has really begun.

This talk describes the advanced SMB3-features from the perspective of a Samba developer. The new concepts create quite a few challenges for Samba. Especially interesting is the result of our research regarding the relation of the SMB3 clustering concepts with the existing CTDB-clustering of Samba. The talk then describes the status of development of the SMB3 features in Samba.

Learning Objectives

- SMB3 features as seen from a Samba perspective
- Challenges imposed by SMB3 for Samba's architecture
- Status of SMB3 development in Samba

[Back to Top \(#top\)](#)

Directory Write Leases in the MagFS Distributed File System

**Deepti Chheda, Staff Engineer, Maginatics and
Nate Rosenblum, Architect, Maginatics**

Abstract

Typical metadata-heavy workloads incur significant network round trip latencies for each namespace-modifying operations. Leases or delegations found in traditional network file systems like SMB or NFSv4 allow clients to cache directory-level information on a "read-only" basis and hence force a network round trip on every create, rename, or delete operation. This talk will focus on the concept of Directory Write Leases, a protocol-level enhancement made in the Maginatics File System (MagFS) to considerably speed up small file metadata-heavy workloads.

Directory Write Leases allow the client to act on behalf of the server for all file system operations in that directory. This is a powerful concept because it enables the client to locally serve namespace-modifying operations within that directory, and asynchronously propagate these operations to the server. We will talk about the semantics of this new lease state needed to preserve strong consistency guarantees in a distributed file system like MagFS. Finally, we will demonstrate that using Directory Write Leases we were able to hide a significant fraction of the network latency and bottlenecks for build workloads when compared to NFS or SMB, and were able to achieve a significant performance boost when compared to traditional leasing mechanisms.

Learning Objectives

- Bottlenecks in small file metadata-heavy workloads e.g. build workloads
- Consistency semantics and guarantees of Directory Write Leases
- Implementation details and challenges of caching namespace modifying operations on the client

[Back to Top \(#top\)](#)

Evolution of Message Analyzer and Windows Interoperability

Paul Long, Senior Program Manager, Microsoft

Abstract

Today, Message Analyzer lets you find deviations in your implementation by exposing differences from Microsoft's documented protocols. Focused on interoperability, we provide new techniques for correlating multiple logs from varied operating systems. Learn how these techniques can lead you to discover interoperability problems and sift through a hayfield of log files to pinpoint the haystack you need to look in. See how you can integrate Message Analyzer into your varied operating system environments and let us show you a new integrated way to decrypt traces, providing you have the private key.

Learning Objectives

- Discover new analysis techniques with Message Analyzer
- Find deviations in windows interoperable solutions
- Message Analyzer analysis with varied operation system environments

[Back to Top \(#top\)](#)

Implementing Witness Service for Various Cluster Failover Scenarios

Rafal Szczesniak, Principal Software Engineer, EMC Isilon

Abstract

The witness service came as part of cluster enabled SMB3 protocol implementation. For the first time SMB clients have become

aware of the server state. They can be notified immediately when the state changes, so they can make their own decision or be suggested where should they reconnect. The information coming from the network subsystem is a natural candidate for the data feed to such service. As network interfaces on the server can be configured and tuned according to the current load and system state, providing some of that information to the clients facilitates in less disruptive experience of client-server connection. Other sources could include monitoring of critical system services or even administrator-driven actions. The talk covers the design concepts behind the witness service employing modules realising several different data feeds combined together to provide a better failover experience.

Learning Objectives

- Sources of data that can be useful for Witness (and SMB3) clients
- Extendable witness service design
- Practical experience from implementing some of the modules

[Back to Top \(#top\)](#)

Introduction to SMB 3.1

**David Kruse, Software Developer, Microsoft and
Greg Kramer, Sr. Software Engineer, Microsoft**

Abstract

The SMB3 ecosystem continues to grow with the introduction of new clients and server products, and a growing deployment base. This talk will look at some potential upcoming changes to the SMB3 protocol and what is driving their need, and how it will affect both protocol implementers as well as customers of those solutions.

[Back to Top \(#top\)](#)

Why SMB3 and How to Implement SMB 3 on Unix/Linux Storage Platforms

Dilip Naik, Software Engineer, HvNAS Pty

Abstract

This talk starts with a quick overview of the significant enhancements in SMB2 and SMB3 over CIFS. The focus is on the features that make the protocol a must have for modern data centers. From there, the presentation examines 5 different architectural ways to have Windows/Hyper-V 2012 access their workloads residing on Linux/Unix storage. Finally, the presentation does a quick survey of the SMB 3 implementations readily available for storage OEMs and cloud gateway vendors.

Learning Objectives

- SMB 3
- Portability of code
- Solutions available in market
- Technology overview

[Back to Top \(#top\)](#)

Comparing SMB Direct performance Using RoCE, InfiniBand, and TCP Networking

Anand Rangaswamy, Sr. System Engineer, Mellanox Technologies

Abstract

SMB 3.0 has been growing in popularity for critical applications in Windows environments. It can be deployed on standard or RDMA networks. Previous presentations have shown excellent performance for SMB Direct on InfiniBand but have not directly compared InfiniBand to 40Gb RoCE Ethernet or non-RDMA networks. Mellanox presents SQL Server benchmark results that compare SMB 3.0 performance on FDR 56Gb InfiniBand with 10 and 40Gb Ethernet using both RoCE (RDMA on Converged Ethernet) and non-RDMA TCP/IP, including IOPS, throughput, latency, and CPU utilization.

Learning Objectives

- SMB Direct performance on FDR InfiniBand
- SMB Direct on 40GbE, RoCE and TCP
- SMB Direct on 10GbE, RoCE and TCP
- IOPS, throughput, latency, and CPU utilization comparisons

[Back to Top \(#top\)](#)

SOFTWARE DEFINED STORAGE

ViPR - Scalable Distributed Storage Built on Software Defined Storage Fundamentals

**Shashwat Srivastav, Director Software, EMC and
Kamal Srinivasan, Product Manager, EMC**

Abstract

Storage customer needs have evolved from client server applications to more of web scale, mobile first applications. A recent IDC study indicates, Enterprises and hosters alike have billions of users and millions of apps within their environment referred to as 3rd platform in the industry. We need to rethink storage while building for this scale out 3rd platform. In this talk, we'll provide an overview of ViPR - A Scalable Distributed Storage platform built from commodity Hardware with ability to meet web scale application needs. ViPR is a Software Defined Storage platform that fits into the Software Defined Datacenter. The control and data plane abstraction help decouple the key storage engine from both Hardware as well as the Application. This enables ViPR to both manage Arrays + Commodity as a single pane and to store objects on them as a unified pool of resources.

ViPR also has an unique approach to Geo distribution. Most scale out deployments have multiple sites and our platform optimizes for storage efficiency while meeting the scale out needs. Providing multi-protocol access over the storage engine enables multitude of enterprise applications to be supported on the platform. Furthermore, support for standards based REST access like S3 and Swift make it a platform that enables standards based access of data in this environment.

Learning Objectives

- Object platform
- Software Defined Storage
- Geo replication
- Commodity storage
- Industry trends in storage

[Back to Top \(#top\)](#)

SNIA Technical Council: SDS Automation and Orchestration

**Mark Carlson, Senior Staff for Standards, Toshiba
Leah Schoeb, Sr. Partner, The Evaluator Group**

Abstract

Software Defined Storage (SDS) has been proposed as a new category of storage software products. SDS can be an element within a Software Defined Data Center but can also function as a stand-alone technology. This talk will present the definitional work of the SNIA Technical Council in this area that may lead to new technical work.

[Back to Top \(#top\)](#)

Horizontal and Elastic Orchestration of Software-Defined Storage

Cheng Wu, Senior Director, ProphetStor

Abstract

Ever since the term “software-defined storage”, or SDS, was coined a few short years ago, the industry has seen a long list of SDS offerings in all shapes and sizes, promising scale-out solutions to handle the explosive amount of data that is said to grow 40–50% per year. While many of these new breeds bring convincing value propositions to the marketplace, few considered themselves as an extension to the good old legacy systems such as EMC or NetApp. However, whether the data of yesterday should be entirely isolated from the new (big) data of today is a question best answered by the application users, not the storage vendors. Lacking a horizontal and elastic storage orchestration platform to simultaneously manage both legacy storage arrays and scale-out x86 hardware from a single pane of glass probably explains why SDS is still primarily a vendor-driven technology.

Learning Objectives

- Horizontal federation for heterogeneous storage environments
- Separation of control and data planes defines SDS
- Pluggable storage docking station architecture
- Policy based self-service provisioning and automatic, predictive storage allocation is holy grail

- Automatic deployment of the storage cloud is key to adoption

[Back to Top \(#top\)](#)

SOLID STATE STORAGE

Flash Data Reduction Techniques and Expectations

Doug Dumitru, CTO, EasyCo LLC

Abstract

Data reduction holds the promise most sought after for solid-state storage: lower cost. In some cases, data reduced SSD arrays can cost less than hard disk drive solutions. We will discuss data reduction techniques including compression, block de-duplication, thin provisioning and how these techniques tax host resources, impact performance and flash wear, and can enhance your storage capacity and save you money. The solutions discussed will include not only pre-built appliances, but also software-only solutions that can be deployed on existing data center storage arrays.

[Back to Top \(#top\)](#)

Experiences Designing a Persistent Memory SDK

Paul Von Behren, Software Architect, Intel

Abstract

The SNTA NVM Programming Model specification defines a programming model for byte-addressable persistent memory (including certain types of NVDIMMs). The programming model's recommended way to use persistent memory (PM) is as a new storage tier, rather than as a replacement for either volatile memory or disks. Intel is developing an open source software development kit (SDK) that eases the effort in developing applications to use PM optimally. This SDK will initially be available on Linux (utilizing emerging file system support for PM); support for other operating systems will be added over time.

Learning Objectives

- The basics of the persistent memory programming model
- The use of persistent memory as a new storage tier in applications
- How applications are expected to use features of the SDK

[Back to Top \(#top\)](#)

How Persistent Memory will Change Our Approach to Computing

Jim Handy, Director, Objective Analysis

Abstract

The data processing industry is approaching a point in which the line of demarcation between storage and memory will become blurred if not eliminated. Memory chips will begin to offer persistent storage. Certain functions of storage will need to be offloaded to this memory. Current coherency models will no longer meet the needs of performance architectures. How will the industry deal with these changes? This presentation explores SNTA's efforts to produce persistent memory programming standards to create an environment in anticipation of this change.

Learning Objectives

- Learn why memory architectures are about to undergo fundamental changes
- Understand how today's memory/storage delineation will eventually disappear
- See how SNTA's nonvolatile memory initiatives will help to solve the problems posed by persistent
- Understand how you will need to adapt your approach to computing to accommodate these changes

[Back to Top \(#top\)](#)

From ARIES to MARS: Transaction Support for Next-Generation, Solid-State Drives

Joel Coburn, Software Engineer, Google

Abstract

Transaction-based systems often rely on write-ahead logging (WAL) algorithms designed to maximize performance on disk-based storage. However, emerging fast, byte-addressable, non-volatile memory (NVM) technologies (e.g., phase-change memories, spin-transfer torque MRAMs, and the memristor) present very different performance characteristics, so blithely applying existing algorithms can lead to disappointing performance. This work presents a novel storage primitive, called editable atomic writes (EAW), that enables sophisticated, highly-optimized WAL schemes in fast NVM-based storage systems. EAWs allow applications to safely access and modify log contents rather than treating the log as an append-only, write-only data structure, and we demonstrate that this can make implementing complex transactions simpler and more efficient. We use EAWs to build MARS, a WAL scheme that provides the same features as ARIES (a widely-used WAL system for databases) but avoids making disk-centric implementation decisions. We have implemented EAWs and MARS in a next-generation SSD to demonstrate that the overhead of EAWs is minimal compared to normal writes, and that they provide large speedups for transactional updates to hash tables, B+trees, and large graphs. Finally, MARS outperforms ARIES by up to 3.7x while reducing the software complexity of database storage managers.

[Back to Top \(#top\)](#)

StorScore: SSD Qualification for Cloud Applications

**Laura Caulfield, Firmware Dev. Engineer 2, Microsoft and
Mark Santaniello, Sr. Performance Engineer, Microsoft**

Abstract

Deploying SSDs in the cloud requires testing under many workloads. Drives must be fungible between applications with flexible and inflexible workloads so we need both clarity into SSDs' strengths and assurances that the drive will not fail in any corner case.

Current SSD testing tools severely limit the number of workloads which are practical to study. Preconditioning, for example, requires either heavy user interaction or running for a worst-case of many hours when most workload transitions require only minutes.

StorScore combines existing standards and tools to automate the testing, increasing the number of workloads we can study. It implements concepts from SNIA standards to automatically detect steady state. It can easily call and parse results from any scriptable performance testing tool. The parser extracts performance (BW, throughput, high-percentile latency, etc.) and endurance metrics (wear distribution and write amplification) per workload. Finally, StorScore simplifies the thousands of metrics into one score.

Learning Objectives

- Automated testing, enabling many workloads
- Challenges of measuring performance and endurance of TB-scale drives
- Cloud scale needs from performance and endurance testing

[Back to Top \(#top\)](#)

New Fundamental Data, Storage and Device Technologies

Page Tagizad, Senior Product Marketing Manager, SanDisk

Abstract

As data sets continue to grow, IT managers have begun seeking out new ways for flash to be deployed in the data center in order to take greater advantage of the performance and latency benefits. With traditional interfaces such as SAS, SATA and PCIe already taking advantage of flash, the focus has shifted to non-traditional interfaces in order to further penetrate current infrastructure. This has led to the emergence of new solutions that leverage the DDR3 interface and are deployed via existing DIMM slots in server hardware, creating vast pools of flash and enabling it to be deployed on the edges of the data center.

In this tutorial, Page Tagizad of SanDisk, will provide an overview of various DIMM-based approaches that have emerged, including the ULLtraDIMM, Hybrid DIMM, SATA DIMMs and NVDIMMs, as well as discuss the advantages of each approach and what applications are best addressed by them.

Learning Objectives

- How the DIMM form factor closes the gap between storage devices and system memory
- How DIMM-based infrastructures meet the demands of data-intensive workloads by reducing processing time compared to HDDs and SSDs leveraging other interfaces
- How DIMM solutions meet the needs of time-sensitive workloads such as high-frequency trading (HFT), Big Data, Analytics, Virtualization and Virtualized Desktop Infrastructure (VDI)

[Back to Top \(#top\)](#)

Phase Change Memory and Its Positive Influence on Flash Algorithms

Rajagopal Vaideeswaran, Principal Software Engineer, Symantec Software

Abstract

NAND based Flash have scaling difficulties as chip lithography shrinks. Each burst of voltage across the cell causes degradation and Flash memory leaks charges which causes corruption and data loss. Also, repeated writes and rewrites of data blocks on Flash without giving it time to perform garbage collection and cleaning can overwhelm the Flash controller's ability to manage free blocks and can lead to low observed performance.

The presentation focuses on Phase Change Memory (PCM) based solutions and its algorithms which has key advantages over Flash (NAND/NOR) as Memory element can be switched more quickly. Also enduring 100 million write cycles, handling 85C working temperatures, retaining data for 300+ years and exhibiting resistance to radiation can make PCM a compelling option.

Learning Objectives

- Evaluate strategies for designing storage solutions that can benefit from Phase Change Memory (PCM)
- Recognize the efficiency and benefits of PCM
- Determine the limitations of NAND based storage and transition plan to PCM
- Thoughts on changes to key Flash Algorithms that will be required

[Back to Top \(#top\)](#)

Method to Enhance the Performance of a Storage Array System with SSD Drives

**Dr. M. K. Jibbe, Director of Quality Architect Team, NetApp and
Kuok Hoe Tan, Senior QA Engineer, NetApp**

Abstract

As the enterprise adoption of SSD in the datacenter continues to accelerate, the need to further understand SSD characteristics with various enterprise workload has become more critical to the success of an all flash storage platform. For the NetApp EF-Series of all SSD storage arrays, we have been studying and analyzing different methods on how our customers can fully realize overall application performance gains after deploying EF-Series arrays into their datacenters by enhancing how we approach array performance optimization to guarantee consistency low latency I/O performance.

Implementing a hybrid solution to optimize the high traffic component of a customer configuration In this presentation we will be covering the various workload that was benchmarked and the results of the various optimization done and their impact to the overall performance of the NetApp E-series arrays

Learning Objectives

- Improve performance by dynamically disabling of full stripe writes
- Improve performance by disabling of write caching under certain conditions
- How to design the volume to establish a high I/O rate with a low latency
- Implement a hybrid solution to optimize high traffic component of a customer CF
- How the mix of array features can optimize and impact overall performance

[Back to Top \(#top\)](#)

Thanks for the Memories: Emerging Non-Volatile Memory Technologies

**Tom Coughlin, Coughlin Associates and
Ed Grochowski, Computer Storage Consultant**

Abstract

While Flash and DRAM devices based on silicon as well as magnetic hard disk drives will continue to be the dominant storage and memory technologies in 2014, this trend is expected to be impacted through 2016 and beyond by new and emerging memory technologies. These advanced technologies are based on emerging non-volatile memory technologies in combination with existing silicon cells to create high density, lower power and low cost products with higher storage capacities. These technologies include MRAM, RRAM, FRAM, PRAM and others. The promise of terabyte devices appearing in the near future to replace existing memory and storage products is based on a continued improvement in processing techniques. The rise of non-volatile, high endurance, fast solid-state memory will change the fundamental design of microprocessor devices and the software that runs on them. This talk will include estimates for the replacement of volatile with non-volatile memory and the eventual replacement of flash memory with a new and scalable storage technology. It will also give some guidance on how non-volatile memory will change electronic device architectures.

Learning Objectives

- Technology advancing to identify new and emerging memory and storage devices
- Potential candidates MRAM, RRAM, Advanced Flash
- New storage mechanisms
- Shift in production equipment priorities
- Changes in computer design and software

[Back to Top \(#top\)](#)

STORAGE ARCHITECTURE

Defining The Software-Defined Technology Market

Mario Blandini, Sr. Director, Marketing Cold Storage Solutions Group, HGST

Abstract

Software-defined storage is driving new requirements in hardware and opening up new opportunities for innovation. HGST is building on decades of drive quality and reliability with new drive technology that connects over Ethernet. In this session, HGST will cover how CPU and memory resources residing on these storage devices can be leveraged to run storage services as close to the data as possible. The technology will be demonstrated using open software like Ceph and Swift, which run in the standard Linux environment without modification. Additionally, HGST will share its SDK and thoughts around taking advantage of this new architecture.

Learning Objectives

- The new software defined disk drive solution from HGST
- This will be the first time enterprises can run distributed storage/applications directly onto storage media for next generation big data, analytics and research
- How CPU and memory resources residing on these storage devices can be leveraged to run storage services as close to the data as possible

[Back to Top \(#top\)](#)

Building Next-Gen CDN with Extremely Low Power and High IO Architecture

Hong Cai, CTO of Cloud Computing, ZTE TX

Abstract

Content Delivery Network (CDN) has never been so important when multimedia files such as video, image, and other document explode on Web. CDN infrastructure plays a key role to improve end user experience while control the overall delivery cost by pushing content to the edge network near the end users. In this presentation, we describe an optimized CDN solution using innovation lower power high IO architecture. Unlike traditional x86 commodity servers, this solution offers far more higher IO (2X higher) with much less power (1/7) and much less form factor (1/4).

Value: Saving space, lower power, lower TCO and offer extremely high IO that cannot be achieved by other SOC architecture.
Audience: engineers having interest in future technologies in CDN system.

Learning Objectives

- CDN architecture
- New system architecture to enable next generation CDN
- Low power/high IO architecture

[Back to Top \(#top\)](#)

Application Agnostic, Analytical Modeling of End-To-End Memory Hierarchy

Pradip Mukhopadhyay, Member of Technical Staff, NetApp

Dhishankar Sengupta, Test Architect, Dell

Krishanu Dhar, Software Engineer Manager, Dell

Abstract

Analytical modeling of memory hierarchy is typically discussed from the perspective of CPU architecture. From the perspective of service tiers Storage-Network-Compute paradigm, memory hierarchy plays a pivotal role for serving data and meeting stringent SLA for applications.

All applications today evolve over time, this appears as a completely different workload to the service tiers. All the tools-and-techniques available allow us to measure and determine the memory requirement in pockets of the overall configuration while provisioning initially. The standard practice of over-provision results in sub-optimal resource allocation and usage.

The solution is to stitch together the end to end memory requirement and provide knobs to try out various parametric manipulations. This approach allows tweaking various parameters and come up with a deterministic requirement of memory at various levels. E.g.: keeping block size constant, change the pattern (sequential-random-access), mix (Read-Write percentage) and the IO working set, cache-agnostic eviction policy, IO concurrency etc.

Learning Objectives

- Looking at memory requirement from an end to end perspective
- How do we account for changing memory requirement while the application is evolving over time
- What solution is the most cost effective and serves my application needs over its lifetime
- A step towards making my application to be cloud-fit

[Back to Top \(#top\)](#)

Deployment Planning and Optimization for Big Data & Cloud Storage Systems

Bianny Bian, Engineering Manager, Intel

Abstract

With the rise of big data analytics systems, IT spending on storage system is increasing. In order to minimize costs architects must optimize system capacities and characteristics. Current capacity planning is mostly based on trial and errors as well as rough resource estimations. With increasing hardware diversity and software stack complexity this approach is not efficient enough. This session presents a novel modeling framework, built with Intel® CoFluent™ Studio, that can be used before system provisioning for cluster capacity planning, performance evaluation and optimization. The methodology uses a top-down approach to model behavior of a complete software stack and simulates the activities of cluster components including storage, network and processors. In addition, simulations can scale to a large number of server nodes while attaining good accuracy and fast simulation speeds (even faster than native execution).

Learning Objectives

- Storage system modeling technology for Big Data & Cloud
- HDFS and Swift simulation vs. measurements
- Real use case: the planning and optimization of a video streaming cluster

[Back to Top \(#top\)](#)

Using Reinforcement Learning to Optimize Storage Decisions

Ravi Khadiwala, Software Developer, Cleversafe

Abstract

Effective use of distributed storage systems requires real-time decision making: what nodes to read from, where to write new data, and when to schedule maintenance operations to name a few. Effectively using available resources is everyone's goal, but in systems as complex and dynamic as distributed storage, the number of variables makes it impossible for any developer to work out every possible situation in advance. Therefore, making optimum decisions requires building intelligent logic into the storage application. But optimizing the logic and getting the information to base decisions on is not easy. In this talk we show that many decision problems in distributed storage are solved by the "multi-armed bandit" model, a well researched approach in reinforcement learning. We also explain how we've put multi-armed bandits to use in our product, to create adaptive agents that make performance optimizing storage decisions in real-time.

Learning Objectives

- The importance of adaptive intelligence in large scale distributed storage systems
- The basics of the exploration/exploitation trade off
- Statistical approaches to the "Multi-armed Bandit" and "Thompson sampling"
- How to implement distributed on-line and real-time learning

[Back to Top \(#top\)](#)

Green and Energy Efficient Big Data Processing at the Software Level

**Da Qi Ren, Staff Research Engineer, Huawei Technologies and
Zane Wei, Director, Huawei Technologies**

Abstract

With the explosive growth of big data applications, energy efficiency is at the forefront of evaluating the performance of a data center, to deliver green solutions in analyzing both structured information and unstructured big data. Focusing on handling the critical design constraints at the software level in a distributed system composed of huge numbers of power-hungry components, in this proposal, an optimized program design approach is introduced in order to achieve the best possible power performance in big data processing. Methodologies to model and evaluate large scale big data computer architectures with multi-core and GPU are introduced. The model allows obtaining design characteristic values at the early design stage, thus benefits programmers by providing the necessary environmental information for choosing the best power-efficient alternative. The energy efficiency improvements from the new designed approach have been validated by real measurements on a multiproCESSOR system.

Learning Objectives

- Introducing energy efficient software design methodologies for big data processing, power performance metrics and measurements
- Modeling and evaluating large scale computer architectures with multi-core and GPU
- Global optimization for choosing the best power-efficient alternative based on data characteristics and quantitative performance analysis
- Validation of the energy efficiency improvements from the new designed approach

[Back to Top \(#top\)](#)

Making Storage Smarter

Jim Williams, Product Manager, Oracle

Martin Petersen, Linux Storage Architect, Oracle

Abstract

Abstraction of storage through protocols, such SCSI and increased storage virtualization, while having the effect of improving compatibility and flexibility of storage, has the adverse effect of isolating the intelligence in storage products from the awareness of how storage is used by applications. It would be extremely useful to penetrate the walls brought about by protocol and abstraction, between storage and application, to enable storage to better accommodate the actions and events at the application. This is not a new concept and has sometimes been called "Intelligent Storage." Perhaps a better term is "Storage Intelligence Coupling" because it enables storage to make better decisions predicated on desired behavior at the application level. This talk discusses some of the many challenges to achieving this objective, a few proprietary approaches that have been taken, and the outlook for the future.

Learning Objectives

- Understand the negative impact of storage abstraction for Quality of Service
- Learn why Intelligent Storage has been difficult to build and bring to market
- Learn of proprietary approaches to Intelligent Storage
- Learn of possible changes in the storage stack that might help make Intelligent Storage possible

[Back to Top \(#top\)](#)

Cold Data Archival - Cellular Biology as Storage Engines

Sanjay Joshi, CTO Life Sciences, EMC Isilon

Abstract

Plant and animal cells (with and without nuclei) have used DNA as information engines for millennia the source code, compiler, executable and application, all rolled into one small and power efficient package. This presentation will focus on the mechanics and details of using synthetic DNA as the primary engine of storage systems: the technology to read, write and archive. The concepts of a Data Center in this context and the notion of Privacy in the DNA age will also be discussed.

Learning Objectives

- Cells as Storage Engines
- Writing to DNA
- Reading from DNA
- Future Data Centers
- What does Privacy mean in a DNA age

[Back to Top \(#top\)](#)

STORAGE MANAGEMENT

Gamification Approach for Next Gen Storage and Infrastructure Management UI

Abhinav Jawadekar, Sr. Director, Symphony Teleca and

Vishal Kirpalani, Director, Symphony Teleca

Abstract

Traditionally storage and IT infrastructure products were managed by tech-savvy administrators using cryptic command line interfaces or complex GUI consoles which resembled fighter jet dashboards. The advent of cloud and mobility has completely changed the personas of the users and managers of these products. There is a need to support self serviceability by end users as well as the needs of a business savvy CIO who is focused on managing SLAs to ensure that IT is being run as a profit center. It is also very important to keep in mind the expectations of the smart phone generation forming an increasing percentage of the users and

Learning Objectives

- Understanding the paradigm shift for management UI on IT infrastructure products due to the advent
- Persona based user experience design for storage and IT infrastructure products
- Approach to uplift the user experience for older products

[Back to Top \(#top\)](#)

Private Cloud Storage Management using SMI-S, Windows Server, and System Center

Hector Linares, Principal Program Manager, Microsoft
Rajesh Balwani, Software Engineer, Microsoft

Abstract

Service providers and enterprises deploy private cloud infrastructure offering powerful capabilities that reduce costs, streamline management, and deliver new value.

With SMI-S, Windows Server and System Center manages SAN, NAS, and Fibre Channel Fabrics for virtualized environments.

Learning Objectives

- Use SMI-S and SAPI to invoke CIM directly for advanced management scenarios
- Determine differentiating value-add using SMI-S
- Design cloud tiering models that take into account storage capabilities
- Rapid VM provisioning using snapshot/clone technology

[Back to Top \(#top\)](#)

Storage Quality of Service for Enterprise Workloads

Tom Talpey, Architect, Microsoft
Eno Thereska, Researcher, Microsoft

Abstract

Storage Quality of Service (QoS) is an increasingly critical aspect of modern datacenter workloads, such as virtualization and cloud deployments. Storage resources are in high demand, and highly scaled and deeply layered contention presents many interrelated challenges, all of which must be addressed to meet service level agreements, and to provide predictable response.

This talk presents IOFlow, a new architecture for classifying and queuing storage traffic at dataplane "stages", and a controller to implement policies. Policies such as minimum guarantee, bandwidth limit, and fairness are supported, with diverse classification of storage traffic. Real-world application of the technology to several virtualization scenarios are explored. The IOFlow implementation is independent of the physical storage technology and also the storage interconnect, and therefore can apply equally for block, file and cloud storage.

This effort is joint work between Microsoft Windows Server and Microsoft Research.

Learning Objectives

- Understand new techniques for storage traffic control
- Learn how Storage QoS can manage virtual machine workloads
- Explore compelling real-world storage management scenarios

[Back to Top \(#top\)](#)

STORAGE PLUMBING

Multiqueue Block Storage in Linux

Christoph Hellwig, Consultant

Abstract

This presentation gives an overview over the problems of the existing Linux storage stack when dealing with low-latency and high IOPS devices, and explains how these are addressed by the blk-mq and scsi-mq frameworks. The blk-mq framework provides a replacement for the lower half of the Linux block layer and allows drivers to be written in a way that allows them to handle low-latency I/O and a high number of IOPS, as well as scale better to large number of CPUs. The scsi-mq framework uses blk-mq to speed up access to common SAN and directly attached storage that uses various SCSI protocols. This presentation will explain the architecture of blk-mq and scsi-mq, and show performance data comparing it to the older block layer and SCSI implementations.

SCSI Standards and Technology Update

Marty Czekalski, President, SCSI Trade Association

Abstract

SCSI continues to be the backbone of enterprise storage deployments and continues to rapidly evolve by adding new features, capabilities, and performance enhancements. This presentation includes an up-to-the-minute recap of the latest additions to the SAS standard and roadmaps, the status of 12Gb/s SAS deployment, advanced connectivity solutions, MultiLink SAS™, SCSI Express, and 24Gb/s development. Presenters will also provide updates on new SCSI features such as atomic writes and Zoned Block Commands (ZBC) for shingled magnetic recording.

Learning Objectives

- Attendees will learn how Express Bay improves how slot-oriented Solid State Drives
- The latest development status and design guidelines for 12Gb/s SAS

TESTING

Cloud Scale Testing Infrastructure – Cloud Simulation, Fault Injection and Capacity Planning

Tanmay Waghmare, Principal Test Manager, Microsoft and Anitha Adusumilli, Senior Test Lead, Microsoft

Abstract

Storage backend for today's cloud deployments require integration of several discrete software and hardware components which need to interoperate correctly with each other. They need to meet the high standards of reliability and availability that end customers expect from a cloud platform. Cloud deployments scaling to 1000s of VMs are comprised of interesting elements like workloads, migrations, admin actions, software faults, hardware faults, planned/unplanned failovers etc. This talk goes over how we simulate cloud deployment reliability aspects with fault injection to meet customer expectations. Besides having reliable cloud platform it is important to perform capacity planning to determine tipping points or targets for acceptance performance at scale. This talk also explains the methodology, success criteria and learnings from implementing cloud scale reliability and performance testing infrastructure.

Learning Objectives

- Understand how the test team modeled end to end solutions for testing cloud deployments
- Understand elements of cloud simulation and applying customer driven success metric for cloud reliability and performance

iSCSI Protocol Test Suite

Tejas Bhise, Director of Engineering, Calsoft

Arshad Hussain, Lead Developer, Calsoft

Abstract

This paper presents iSCSI Protocol Test Suite (ITS). iSCSI is widely used with multiple open and close sourced implementations. Since there is no easy way to test these implementations for protocol conformance and interoperability with myriad initiator implementations, hence the relevance of ITS. Existing software that tests iSCSI (e.g. sahlberg/libiscsi), focuses more on SCSI block command (SBC) rather than iSCSI. ITS is portable, written in C, and compiled as a user space binary under Linux. Each test is a self-contained binary that can be integrated in any existing test suite. With around 200 test cases in iSCSI login, Full Feature Phase (FFP) and Errors, we have successfully run our test suite against multiple iSCSI target implementations and detected several errors. Roadmap for ITS includes additional cases for Login, FFP, ERROR and including auxiliary RFCs like CHAP, SLP, iSER and Boot etc.

Learning Objectives

- Demystify iSCSI protocol compliance testing challenges
- Integrate ITS into existing iSCSI target test suite
- Understand the Calsoft ITS solution

TESTING PERFORMANCE

Dedupe, Compression and Pattern-Based Testing for Flash Storage: Getting It Right

Peter Murray, Sr. Product Specialist, Load Dynamix and

Leah Schoeb, Sr. Partner, The Evaluator Group

Abstract

Measuring flash array storage performance involves more than just measuring speeds & feeds using common IO tools that were designed for measuring single devices. Many of today's flash-based storage arrays implement sophisticated compression, deduplication and pattern reduction processing to minimize the amount of data written to flash memory in order to reduce storage capacity requirements and extend the life of flash memory. Effectively measuring the performance and capacity of flash-based storage now requires more than the usual steps documented by the SNIA SSS1 specification. Instead, testing now requires the inclusion of complex data patterns that effectively stress data reduction technologies like pattern recognition, compression and deduplication. Measuring performance without including these important features or by using tools that offer a limited set of data patterns falsely overstates modern flash array performance. Only by evaluating with these inline data reduction capabilities enabled, based on real-world application workloads, can vendors and customers truly understand the performance, capacity and effectiveness of a particular flash storage array offering these advanced features.

Learning Objectives

- How today's flash memory storage arrays implement pattern recognition
- What pattern-based performance measurement involve
- How an effective performance measurement and validation solution

[Back to Top \(#top\)](#)

Synthetic Enterprise Application Workload Testing

Eden Kim, CEO, Calypso Systems, Inc.

Abstract

Introduction of Synthetic Application Workloads and new testing that focuses on IOPS and Response Times as demand intensity increases. Discussion of SSD performance states, PTS test methodology for drive preparation and test, definition of synthetic application workloads and a case study showing IOPS and Response Time saturation as demand intensity increases. Particular focus on response time Quality of Service and confidence levels up to 99.999% (or 5 nines) of confidence.

Learning Objectives

- Understanding Demand Intensity, Confidence Levels and Quality of Service
- Defining Synthetic Application Workloads
- Developing a test plan using SNIA SSS PTS test methodologies
- Evaluate Response Time QoS as demand intensity increases - selecting optimal OIO operating points

[Back to Top \(#top\)](#)

VIRTUALIZATION

The Ultimate Storage Virtualization

Felix Xavier, Founder, CloudByte

Abstract

The concept of virtualization is not new for the storage. Logical Volume Manager is the earliest form of storage virtualization which allowed to create virtual volumes abstracted from the underlying hardware. Subsequently, scale-out NAS virtualized the NFS mount point through virtual IP address where NFS mount point is floating above the group of hardware storage nodes. Virtual Server or commonly known as VM in the server world triggered the new form of virtualization in the storage world, it is storage virtual machines. This is in fact reverse of scale-out NAS problem. Due to the arrival of SSDs, storage systems are becoming larger and larger in terms of storage performance, hence this hardware storage system needs to be shared among multiple applications. This brings out the need for storage virtual machines, which abstract the physical characteristics of storage such as IOPS, throughput, latency and capacity into software and credibly share the larger storage system with multiple applications. This is the ultimate form of storage virtualization.

[Back to Top \(#top\)](#)

Solid State of Affairs: How to Evaluate the Benefits (or not) of SSDs for VMware

Irfan Ahmad, CTO and Co-Founder, CloudPhysics

Abstract

Evaluating the benefits and challenges of Flash SSDs in VMware virtualized systems: from best practices to data science

A lot of industry buzz surrounds the value of SSDs. New flash-based products have entered the server and storage market in the past few years. Flash storage can do wonders for critical virtualized applications, but most VMware shops are still on the sidelines, not yet sure of the value. The key question being asked is: how can I figure out the benefit of SSDs to my datacenter and is it worth the cost?

Our experience has shown that SSDs aren't a silver bullet nearly as often as you'd think. Different products do well for different workloads. If you don't know the workload and don't match I/O patterns to the capabilities that SSDs bring to the table--you'll end up spending money where you don't need it, not spending money where you do need it, or spending the right money for the wrong application.

Irfan Ahmad, the tech lead of VMware's own Swap-to-SSD, Storage DRS and Storage I/O Control features will share his experiences working with SSDs in virtualization systems. Irfan will demonstrate the techniques for precise prediction of SSD benefits and choosing the right solution for the right workload.

Learning Objectives

- How to use data science to evaluate production workload caching benefits
- How workload traces can be captured and replayed
- How new sampling techniques can be used for production workloads

[Back to Top \(#top\)](#)

Copyright © 2014 SNIA

[Site Map \(/sitemap\)](#) | [Contact Us \(/about/contact_us\)](#) | [Privacy Policy \(/about/privacy_policy\)](#)