

# SanDisk®

## At-Scale Data Centers & Demand for New Architectures

Allen Samuels

Software Architect, Software and Systems Solutions

August 12, 2015

# Forward-Looking Statements

During our meeting today we may make forward-looking statements.

Any statement that refers to expectations, projections or other characterizations of future events or circumstances is a forward-looking statement, including those relating to market growth, industry trends, technology transitions and developments in flash pricing and capacity.

Actual results may differ materially from those expressed in these forward-looking statements due to factors detailed under the caption “Risk Factors” and elsewhere in the documents we file from time to time with the SEC, including our annual and quarterly reports.

We undertake no obligation to update these forward-looking statements, which speak only as of the date hereof.

# A Global Leader in Storage Solutions



## Produce Close to Half of Industry Bit Output

Joint Ventures between SanDisk and Toshiba \*

## Computing: Enterprise, Client and Retail SSDs

Qualified at  
**5 largest server & 4 largest storage OEMs**



Approved Supplier to  
Leading PC Manufacturers

- **SanDisk Supports Open Source Innovation**
- **SanDisk is committed to Open Source**
- **FOSS is a key part of all of our mega markets of *compute, mobile and consumer***
- **More on the possibilities of storage at [bigdataflash.sandisk.com/opensource](http://bigdataflash.sandisk.com/opensource)**

\* Gartner NAND Flash Supply & Demand, WW 1Q'13 – 4Q'15, March 2014.

# The Data Tsunami is Still Getting Bigger

## CONTENT REPOSITORIES



- Large containers for long periods with on-demand rapid access
- Mixed media container, active-archiving, backup, locality of data

## BIG DATA ANALYTICS



- Hadoop, NoSQL
- Time-to-Value and Time-to-Insight

## MEDIA STREAMING



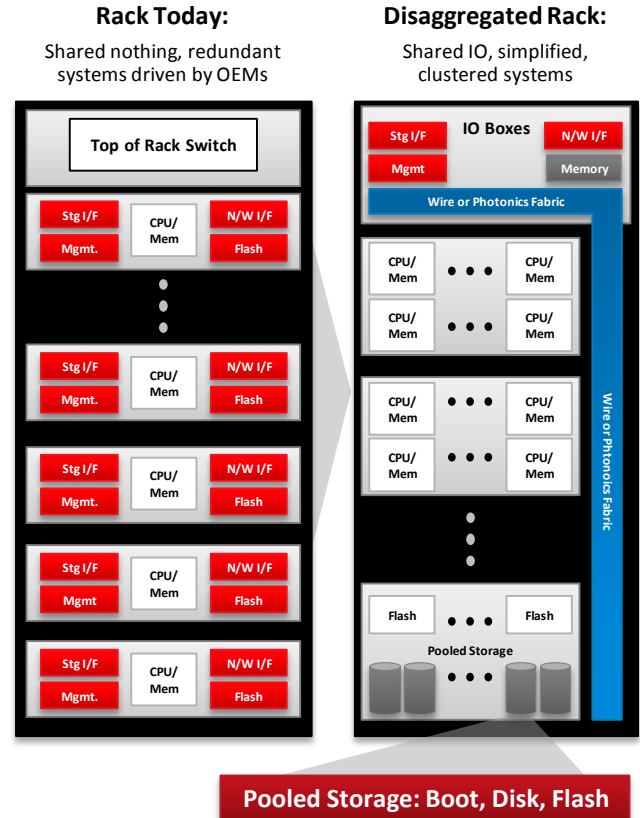
- High read intensive access from billions of edge devices
- Hi-def video driving even greater demand for capacity and performance

# We are in the Middle of a Massive Change

- Big Data is changing our Data Center Architectures
  - Compute-Centric → Dataflow-Centric
- Scalability is the new paradigm
- Price/Performance trumps Performance
  - You can always get performance from scale

# Scalable Hardware

- Rethinking the Rack – Rack Scale Architecture
  - Decompose, aggregate and recompose
- High performance dense fabrics are key enablers
- Independent scaling of compute, memory, storage and networking
- Deployment flexibility and configuration
- Lifecycle management, operation efficiency



# Scaling the WorkFlow

- Data Center scale operating system
- APIs, resource allocation, access control
  - Same problems, different scale
  - OpenStack, VMWare, Mesos, Kubernetes, Microsoft CPS
- Managing, monitoring, configuring and provisioning at scale
- Storage
  - Block, file, object, key/value
  - Cinder, Ceph, Gluster, Redis, Cassandra, MongoDB, Couchbase

# Scalable Storage Systems

- Many storage scaling paradigms are already proven and deployed
  - Object stores (Swift, etc.)
  - Scale-out block (Ceph, Gluster, SheepDog, etc.)
  - NoSQL/KV store (Cassandra, MongoDB, redis, memcached, etc.)
  - Hadoop file system



# Trends in Storage Components

- Performance oriented HDD vanishing
- Capacity optimized HDD thrives (SMR, Cloud Drives, etc.)
  - BW/bit shrinks as capacities grow
- Flash cost reductions continue (3D)
- Flash diversifies and drives new standards and form factors
- New NVM technologies (PCM, STT, etc.) still waiting in the wings

# HDD Based Computation

- HDD BW roughly constant into the future
- CPU processing power increasing into the future
- HDD / CPU ratios increases 2-3x each generation *just to stand still*
- HDD-based systems becoming constant \$ / transaction
  - *No longer cheaper, faster or smaller. Just more capacious.*

# Flash to the Rescue

- Flash BW/bit is an economic choice
- BW/\$ continues to decrease into the future
- Flash \$ / bit continues exponential decrease
- \$ / Computation for Flash decreases exponentially
  
- How can I afford to obtain the benefits of Flash?

# Best of Both Worlds -- Tiering

- Primary Tier of erasure coded Flash
- Secondary Tier of cloud-speed capacity-optimized HDD
- Easy to sustain BW during inter-tier migrations

# Primary Tier Design

- Data Replication with Flash is not best practice
  - Sufficient BW without RAID-0
  - Overkill for data protection, Flash AFR  $\ll$  HDD AFR
  - Wasteful of \$
- Erasure Coding is best practice with Flash
  - High coding rates provide comparable protection to replication
  - Overhead rates can be easily be less than 20% ( $< 1.2x$ )
  - Recovery times drastically reduced, MTDDL reduced accordingly
- EC BW increase easily handled @ modern LAN speeds
  - Rack level failure model directly obtainable

# Building the Secondary (HDD) Tier

- Ideal use-case for SMR drives
  - SMR Optimized file systems are under development
- Only 2x replication needed for many situations
  - Some data sets don't justify high MTTDL values
  - Some data sets don't justify mission critical availability
- Remote placement of one replica satisfies many DR needs

# Open Source Support for EC and Tiering

- HDFS doesn't natively implement EC – yet
  - HDFS-RAID (Facebook) is available
- Ceph Object (RGW) does EC today
- Swift Erasure Coding in beta today
- Ceph direct HDFS connector does EC with local tiering

# What's the Status?

- Current scale-out OSS software are \$ inefficient on flash
  - SanDisk Ceph contributions have achieved 7x-10x improvement<sup>1</sup>
  - SanDisk work on Cassandra, MongoDB, and Redis shows 2-3x boost<sup>2</sup>
- Performance gap translates into significant \$ at datacenter scale
- Proprietary and in-house solutions are filling the gap
  - Fragmentation impedes portable application development

<sup>1</sup> Based on internal testing comparing Emperor to Hammer releases using vdbench and fio on a two node cluster (Dell R720 2.8GHz, 2xE5-2680, 64GB DRAM), krbd client systems on Dell R620 (2x E5-2680, 32GB DRAM), 40Gbe interconnect

<sup>2</sup> <http://www.sandisk.com/assets/docs/sandisk-zetascale-whitepaper.pdf>



# Summary

- Flash as primary storage is here today, NOT tomorrow
- We want to work with you!
- Let's talk!

# Thank you

© 2015 SanDisk Corporation. All rights reserved. SanDisk is a trademark of SanDisk Corporation, registered in the United States and other countries. Other brand names mentioned herein are for identification purposes only and may be the trademarks of their respective holder(s).

The SanDisk logo is displayed in a bold, red, sans-serif font. The word "SanDisk" is written in a slightly stylized manner, with the "S" and "D" being larger and more prominent. A registered trademark symbol (®) is located at the top right of the "k".

**SanDisk®**