

**BIG DEMANDS PLACED BY  
BIG DATA ON  
MEMORY TECHNOLOGY  
& THE INDUSTRY**

**Dr. KRS Murthy**

CEO, i3 World

[BigDataExpert@Gmail.Com](mailto:BigDataExpert@Gmail.Com)

(408)-464-3333

Twitter: @BigDataExpert

FMS 2013 Big Data – SSD Session Wednesday Aug 13

# Social Media's Growth

- Social media is coming of age.
- Emergence of the first social media networks some two decades ago.
- Social media has continued to evolve.
- Offers consumers around the world new and meaningful ways to engage with the people, events and brands that matter to them.
- Now, years later, social media is still growing rapidly
- It is becoming an integral part of our daily lives.
- Social networking is now truly a global phenomenon.

# Big Data Growth and Demand

- ❖ Demand for data storage devices has been exponentially increasing with the widespread Internet, mobile devices, social media and networking and cloud computing.
- ❖ This BIG growth is putting BIG demand on all types of memories.
  - Hard Disk Drives (HDDs)
  - Optical Disc Drives (ODDs)
  - Solid-State Drives (SSDs) using NAND and other flash memories.
  - Archival – Tape Memories too.

# Big Data Paradigm and Purpose

- Business Decisions based on Massively Parallel Processing (MPP)
  1. Transactional data stored in relational databases
  2. Potential business value from non-traditional, less structured or unstructured data:
    - ❖ Weblogs
    - ❖ Social Media
    - ❖ Email
    - ❖ Sensors
    - ❖ Machine Generated Data
    - ❖ Audio, Photos, Video etc.

# Big Data Leverages on the Assumption of

- Commodity Processors
- Decreases in the cost of both storage and compute power
- Decreases in the cost of sensors, cameras, machine to machine communication (M2M)
- Increase in embedded data collection devices and systems (E2X)
- Infrastructure embedded Intelligent data collection and interaction (I2X) and (V2X)
- Internet of Things (IoT)
- Server, Processor and Memory Computing Processes Virtualization
- **Supply Chain: HW, SW, Applications, Systems, Installation, Operation, Maintenance, Talent & Capital**

# SSD

## PROS

- High performance IOPS, in MB/sec.
- Low cost per performance (\$/IOPS)
- Fast response, low latency
- No mechanical parts
- Robustness to external vibrations and shocks

## CONS

- ❑ P/E endurance (P/E limitations)
- ❑ Data retention
- ❑ High cost per byte (\$/GB)

# HDD

- ❖ No limitations on write endurance
- ❖ No limitation on data retention
- ❖ Low cost per byte (\$/GB)
- ❖ Low power consumption per byte (W/GB)

# Types of SSDs

- **Client SSDs**, which are mainly used for personal computers, terminal devices and mobile devices, sensor networks, M2M ....
- **Enterprise SSDs**, which are used for servers and storage systems for data centers and corporate systems



# Enterprise SSDs

- ❖ **SSDs used for mission critical systems must feature:**
  - ✓ Higher Reliability
  - ✓ Higher Endurance
  - ✓ High Speed Performance
  - ✓ Guaranteed Sophisticated Data Integrity
- ❑ To ensure 24x7 continuous operations in heavy workload environments

# Use of MLC

- ✓ For enterprise SSDs, SLC (Single Level Cell) NAND has been used due to the severe requirements for write endurance.
- ✓ However, in order to meet the requirement for lower unit price per gigabyte, the use of MLC (Multi Level Cell) NAND has been promoted.

# MLC NAND Adoption

- Adoption of MLC NAND requires increased parallel operations of the NAND flash memories to improve:
- Read/Write performance
- Optimization of ECC Mechanism
- Digital Signal Processing (DSP)
- Error Recovery Algorithm to
- Secure Effective Write Endurance
- Data Retention Characteristics
- Implementation of Dynamic Write Performance throttling mechanism to manage the Endurance and Product Life.

# SSD Classifications & Requirements

## ▣ Operating Condition At Power-On

- **Client:** 8 hours/day at 40°C
- **Enterprise:** 24 hours/day at 55°C

## ▣ Operating Condition At Power Off

- **Client:** 1 year at 30°C
- **Enterprise:** 3 months at 40°C ≤ 10-16

## ▣ UBER (Uncorrectable Bit Error Rate):

- **Client** ≤ 10 to the power of -15
- **Enterprise** ≤ 10 to the power of -16

# Comparison of Specifications

## ENTERPRISE SSDS

## CLIENT SSDS

- |  |  |  |
|--|--|--|
| ❖ <b>Application:</b> Servers, storage systems<br>PCs, notebooks, etc. |  | ▣ PCs, Notebooks,<br>Terminal & Mobile |
| ❖ <b>Interface</b> SAS, (PCI Express), (SATA) SATA                     |  | ▣ (PCI Express), (USB)                 |
| ❖ <b>NAND Type</b> SLC, (MLC) MLC                                      |  | ▣ (MLC)                                |
| ❖ <b>Performance</b>   |  |  |
| ✓ Read/write (IOPS) High   |  | ▣ Low                                  |
| ✓ Read/write (MB/sec.) High  |  | ▣ Low                                  |
| ✓ Endurance (Bytes) High   |  | ▣ Low                                  |
| ✓ Data protection in the event of power<br>failure Yes                 |  | ▣ Low                                  |
| ✓ End-to-End Data Protection Yes                                       |  | ▣ No                                   |
| ✓ Life (Years) 5 (SLC)   |  | ▣ No                                   |
|  |  | ▣ 3 (MLC)                              |

# SSD Problems - Solutions

- ▣ Data protection
- ▣ Robust ECC
- ▣ Internal RAID
- ▣ In-flight data protection (SECDED, chipkill, 8B/10B, etc.)
- ▣ DIF
- ▣ Wear leveling
- ▣ Thermal protection
- ▣ Power loss protection
- ▣ Partial page program avoidance
- ▣ Intelligent and granular memory retirement
- ▣ Block size tuning
- ▣ Reserve space tuning

# Stages of Data Movement

- ✓ Data Acquisition Stage:
  - Acquiring data from the point of data creation:
    - ❖ Mobile Devices – Position, Activity, History
    - ❖ M2M – Machine to Machine environment
    - ❖ I2X – Infrastructure to X Environment
    - ❖ V2X – Vehicle to X Environment
    - ❖ S2S - Sensor to Sensor Mesh Environment

# Stages of Data Movement - Aggregation

- ❖ Aggregation of Related Data or Similar Context Data:
  - Mesh Gateways
  - Intermediary Processing
  - Prioritization based on the Source & Content Type
  - Prioritization based on Criticality
  - Buffering
  - Upload Frequency
  - Data Filtering



# Stages of Data Movement

## Data Organization

- ❖ Function performed close to or adjunct to the Big Data Server and Storage Cluster in the Data Center or Cloud
- ❖ Part of the typical ETL and different sequencing and combinations of E, T and L
- ❖ Buffering
- ❖ Sequencing
- ❖ Preprocessing based on Data Formats and Content

# Big Data MPP

- ❖ Three copies of the data will be needed for MPP Map Reduce
- ❖ Depends on the Data Center Server and Storage Architecture
- ❖ Most of the time, three work horses are busy and occupied even though processed results from only one is used to go to the next stage of processing or storage in the meta data.
- ❖ It is like asking three separate kitchens with its own chefs to cook the same menu and as you are super hungry, and finally you pick the one who cooks first. Then throw away the other two, as you needed only one

# Meta Data

- ❖ The Meta Data is created
- ❖ To be accessed for query for business intelligence by the data scientists
- ❖ On behalf of the different decision makers
- ❖ The different decision makers have their specialized requirements:
  - CEO
  - CFO
  - COO
  - CMO
  - Different Executive and Management levels

# Multiple Stages with Multiple Copies

- ❖ Multiple copies of data requires multiples of memory capacity over all
- ❖ Multiple stages of data requires multiples of memory capacity over all
- ❖ Some are client, mix and enterprise level processing requiring both SLC, MLC SSDs
- ❖ Hybrid: SSD-HDD
- ❖ Archived Data – Tapes
- ❖ Enterprise Class DRAM

# Demand on Talent

- ❖ Professionals of different categories, specialization and experience levels are required.
- ❖ The CAGR rate of growth of memory, servers, data centers, networking, cloud, infrastructure deployment of sensors, meters, equipment etc. require architects, design from device to large system levels, engineering, manufacturing, test, field engineering, sales / marketing, customer support..... Has to be matched by talent availability.
- ❖ The talent supply chain comes colleges, universities, training companies
- ❖ The talent supply chain for the above comes from K-12 schools

# Demand of Capital Investment

- ❖ Capital Intensity to meet all aspects and phases of
  - Device Manufacturing
  - Components Manufacturing
  - Subsystems and Systems Manufacturing
  - All aspects of Supply Chain Management
  - Even capital for colleges, universities and K-12 schools
  - Capital availability for these have to compete with all other capital requirements for other industry verticals
  - All at favorable rates, terms and conditions

**BIG DATA**

Is a Hungry  
Monster

# BIG THANKS FROM THE BIG DATA EXPERT

**Dr. KRS Murthy**

CEO, i3 World

[BigDataExpert@Gmail.Com](mailto:BigDataExpert@Gmail.Com)

(408)-464-3333

Twitter: @BigDataExpert