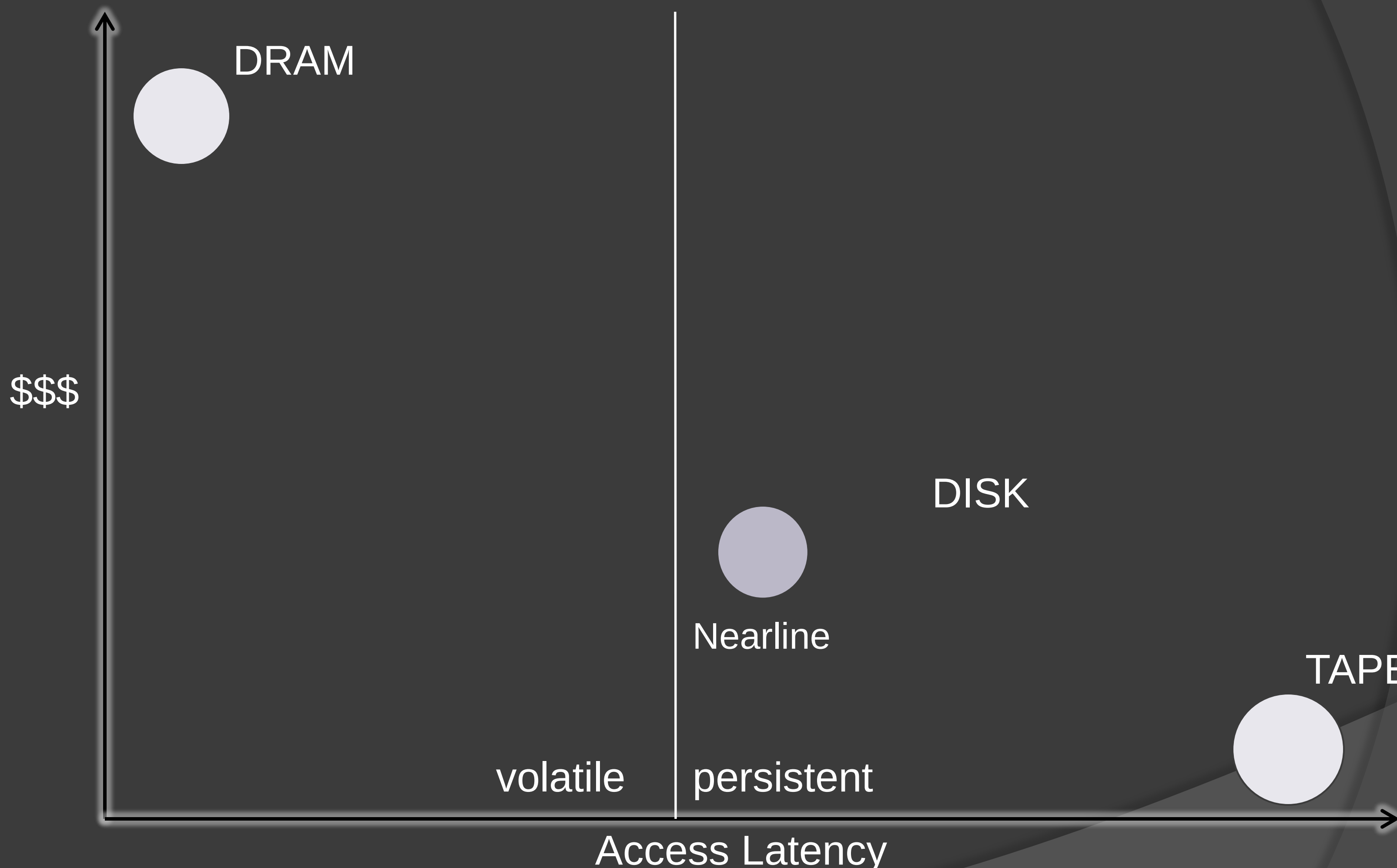
FLASH MEMORY SUMMIT

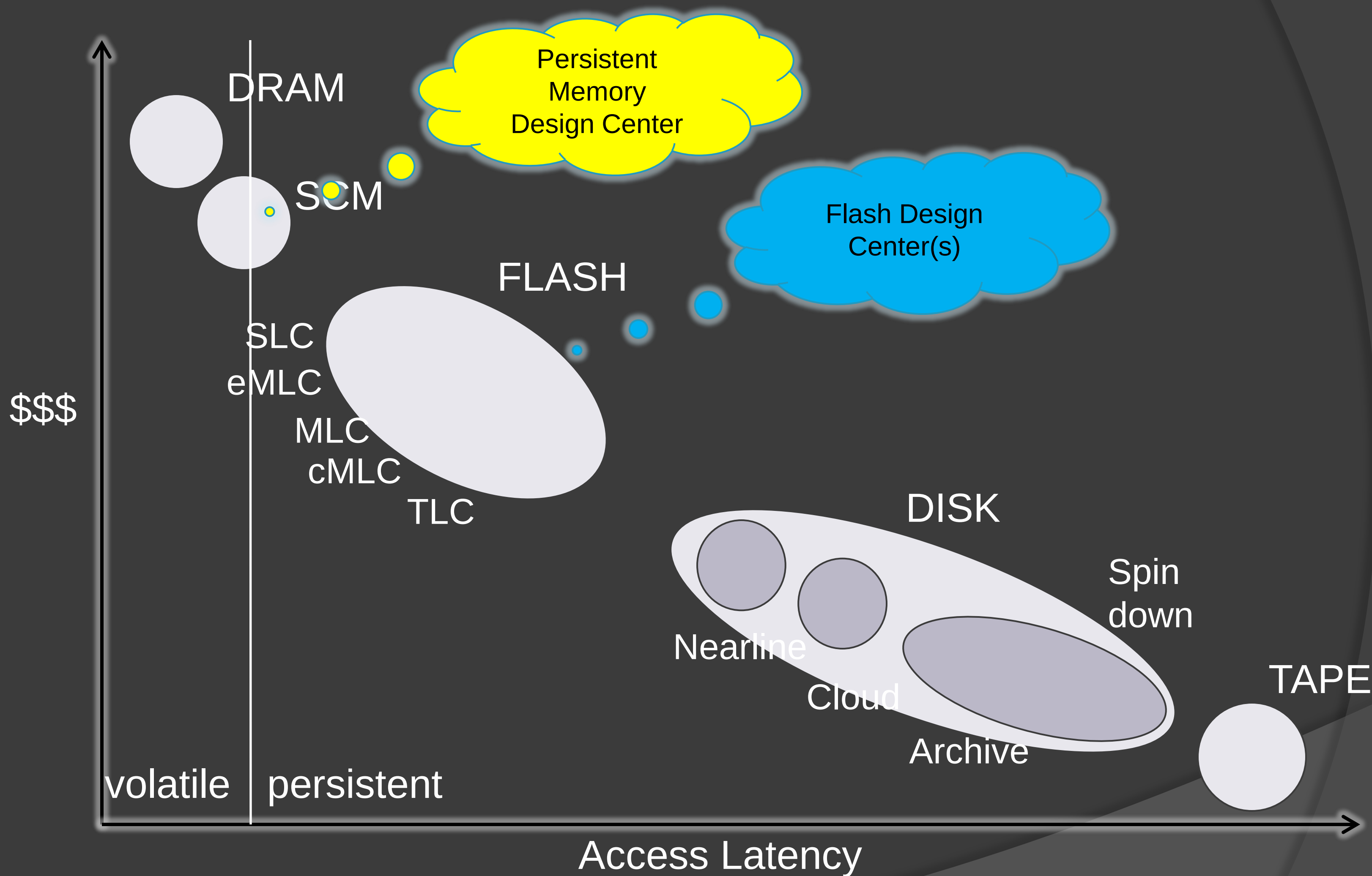
PERSISTENT MEMORY APPLICATIONS TRACK
AUGUST, 2014

Kaladhar Voruganti
Senior Technical Director
NetApp, CTO Office

Storage Used to Be Simple



Talk Focus: Persistent Memory Design Center



Confluence of Trends

Real-Time Analytics

High Transaction Rates
Cannot Tolerate
Disk/Network Latencies
Middleware using KV Interface

Storage Class Memories (SCM)

-Better Density than DRAM
Servers will be shipped
With SCMs
-Writes to SCM faster than
Writes to Flash and over
the wire writes

Persistent Memory Design Center

Primary Data Management
Moving to Host
(Key Control Point)

Real-Time Analytics Applications

Applications	Characteristics	Middleware	Products
Reservation Banking Financial	1. Moderate Transaction Rate 2. ACID 3. Batch OLAP Queries on copy of data 4. Moderate amount of data	SQL	DB2, Oracle, SQLServer
Document Systems, User Preference Data Machine Generated Logs/Data	1. Eventual Consistency 2. Very large amount of data 3. Key-Value Access Model 4. Both Batch/Real-time queries	NoSQL	Hbase, Cassandra CouchDB MongoDB
Supply-Chain Apps Fraud Detection Stock Trading Mobile Location Services	1. Very High Transaction Rate and cannot tolerate disk latencies (and in cases network latencies) 2. Very Large Amount of Data 3. Mostly ACID Semantics 4. Real-Time Queries on same data	NewSQL	NuoDB,MemSQL, SQLFire, VoltDB SAP Hana, Oracle TimesTen

Call for persistent Memory Architectures

Different Types of Persistent Memory

	PCM	STT-MRAM	ReRAM /Memristor	N-RAM (Nantero)	DRAM	NAND (SLC/MLC) 2D-planar
Storage Mechanism	Phase change – amorphous and crystalline	Magnetization of ferromagnetic layer	Ion transport and redox reactions	Carbon nanotube-based resistance change elements	Charge on a capacitor	Floating Gate
Feature Size F	45nm	65 nm	30nm	22nm	36nm	22nm
Cell Size	$4.8F^2$	$20F^2$ [$14F^2$ (54nm) (Grandis)]	$4F^2$	$6F^2$	$6F^2$	$4F^2$ /$2F^2$
Read latency	12 ns	[illegible]	[illegible]	[illegible]	10 ns	50us/100us
Write / Erase latency	[illegible]	[illegible]	[illegible]	[illegible]	10 ns	500us/ 1ms
Write endurance	1E9	1E12	[illegible]	>1E9	1E16	1E5/ 1E4
Data Retention	>10yrs	> 10 yrs	> 10 yrs	>10 yrs	64 ms	>10yrs (Fn of writes)
Write Voltage (V)	3	1.8	0.6	3.5	2.5	15
Read Voltage (V)	1.2	[illegible]	[illegible]	[illegible]	1.8	1.8
Write Energy (pJ/bit)	[illegible]	[illegible]	[illegible]	[illegible]	4E-15	2E-9
Idle Power	[illegible]	[illegible]	[illegible]	[illegible]	Med. (refresh)	Low
MLC (bits/cell)	2	Difficult	Anticipated	Possible		2/3 (max till 4)
3D solution	Possible	At higher cost (vertical MOSFET)	Feasible at higher cost	?	Hybrid Memory Cube applies TSV	3D (yes)
Scalability Prospects	F < 10nm	F = 10-45nm	F < 10nm	F < 5nm	F=10nm	F = 10nm
Addressability	byte	byte	?	?	byte	Page (r/w) Block (erase)

Better Endurance & Speeds than Flash
Byte Addressable

Lower speeds but better densities than DRAM

Two Persistent Memory Design Centers

- Transparent

- Disruptive

“Transparent” NVM Adoption Drivers

- SCM/NVM treated as storage or DRAM
- Every app benefits from ‘faster flash’
- Every app benefits from more ‘memory’
  - DRAM cache in front of SCM
- Likely that mobile use cases will drive cost of SCM technologies down

"Transparent" NVM Applications

- Application characteristics
  - Memory-based or Flash-based today
  - Real-time latency / Hi Tx rate Databases
  - Performance at any cost

Disruptive Persistent Memory Design Center

Current Design Center

- POSIX Access Model
- Disk Optimized data structures (e.g. Column store) try to localize updates
- Coarse Grained file/LUN/Volume level Snapshots
- Coarse Grained SLOs (File, LUN/Volume) [specified out of band]
- Replication needed for both durability and HA
- Client-Network Storage Server Model

Persistent Memory Design Center

- Fine-grained Load/Store Access From User Space
- Processor Optimized Graph Data Structures that don't try to localize updates
- Data Structure level Versioning and recovery
- Data Structure level SLOs [specified in-band]
- Replication need for HA
- Peer to Peer Storage Model will gain traction

“Disruptive” NVM Applications

- Memory-based or Flash-based today
- Real-time latency / Hi Tx rate
- Larger datasets than today’s in-memory apps
- Apps that need zero downtime and instant recovery

Disruptive NVM Applications

Medical/Wearable/Embedded

Conclusion: Asks

- Need end to end solutions (i.e. cannot develop persistent memory design center in isolation)
  - Need to move checkpoints, versions to capacity storage
- Need New Hardware Service Models
- Need data structure level management semantic standards (e.g. persistence, consistency, SLOs)
- Need Rapid Re-Build Algorithms from slower media during failures