The Magic and Mystery of In-Memory Apps

Taufik Ma – Industry Insight
Shaun Walsh - Marketeer

Contents

- The Use In Memory Applications?
- Evolution towards & Role of In-Memory Computing
- Role of Storage in In-memory solutions
- Customer Trends
- Emerging Technologies & Some Predictions
- Summary

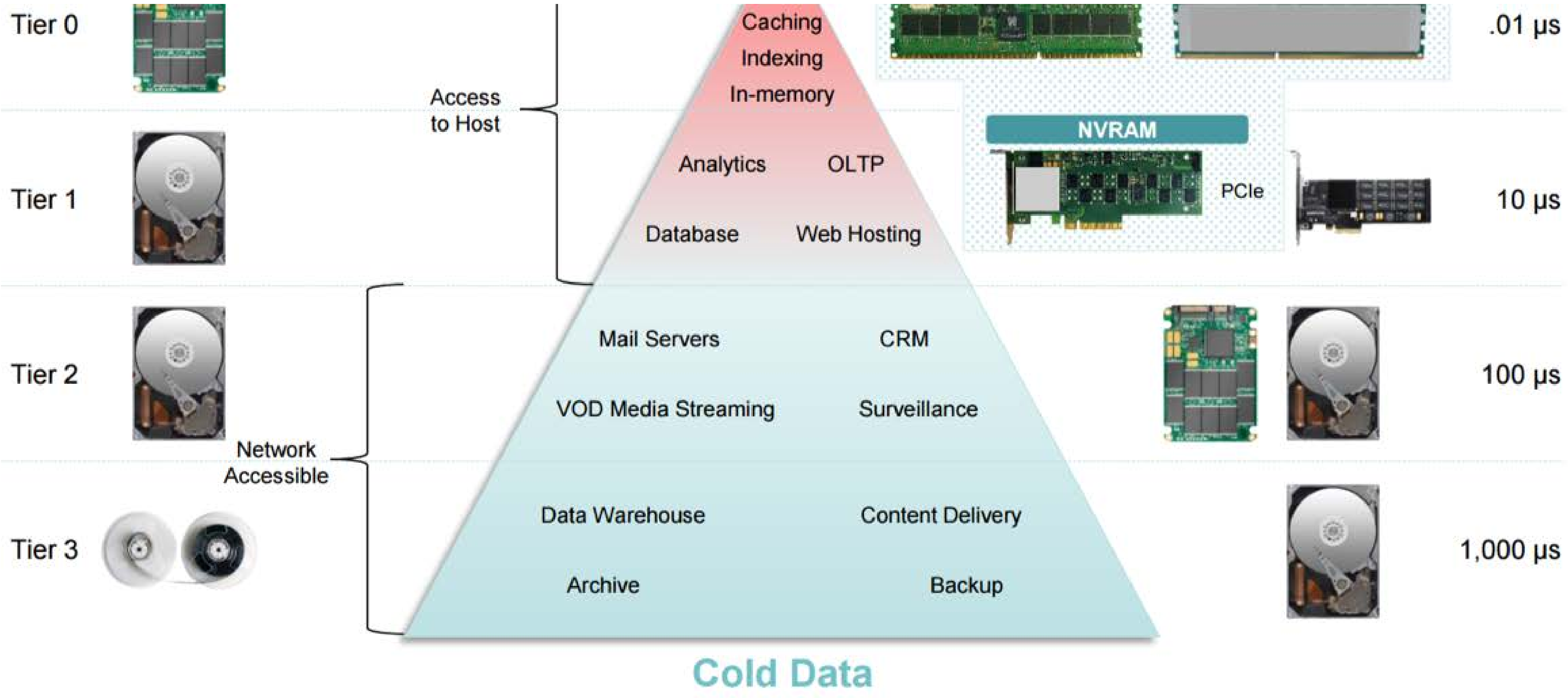
Magic and In-Memory Applications

Shaun Walsh - Marketeer

Any sufficiently advanced technology is indistinguishable from magic

Arthur C. Clarke

The Evolution of Storage Tiers



NVM will Accelerate Both Meta-Data & Application Data

NVDIMM Acceleration Segments

Latency

- In-Memory DBs
- MemCacheD
- RAID
- De-Dupe

NVDIMM-F
NVDIMM-P
3D-XPoint

NVDIMM Type	Presentation	Access Method	Latency
-N	DRAM	Byte	Consistent
-F	Storage	Block	Variable
-P	DRAM and/or Storage	Byte & Block	Variable

NVDIMM-N

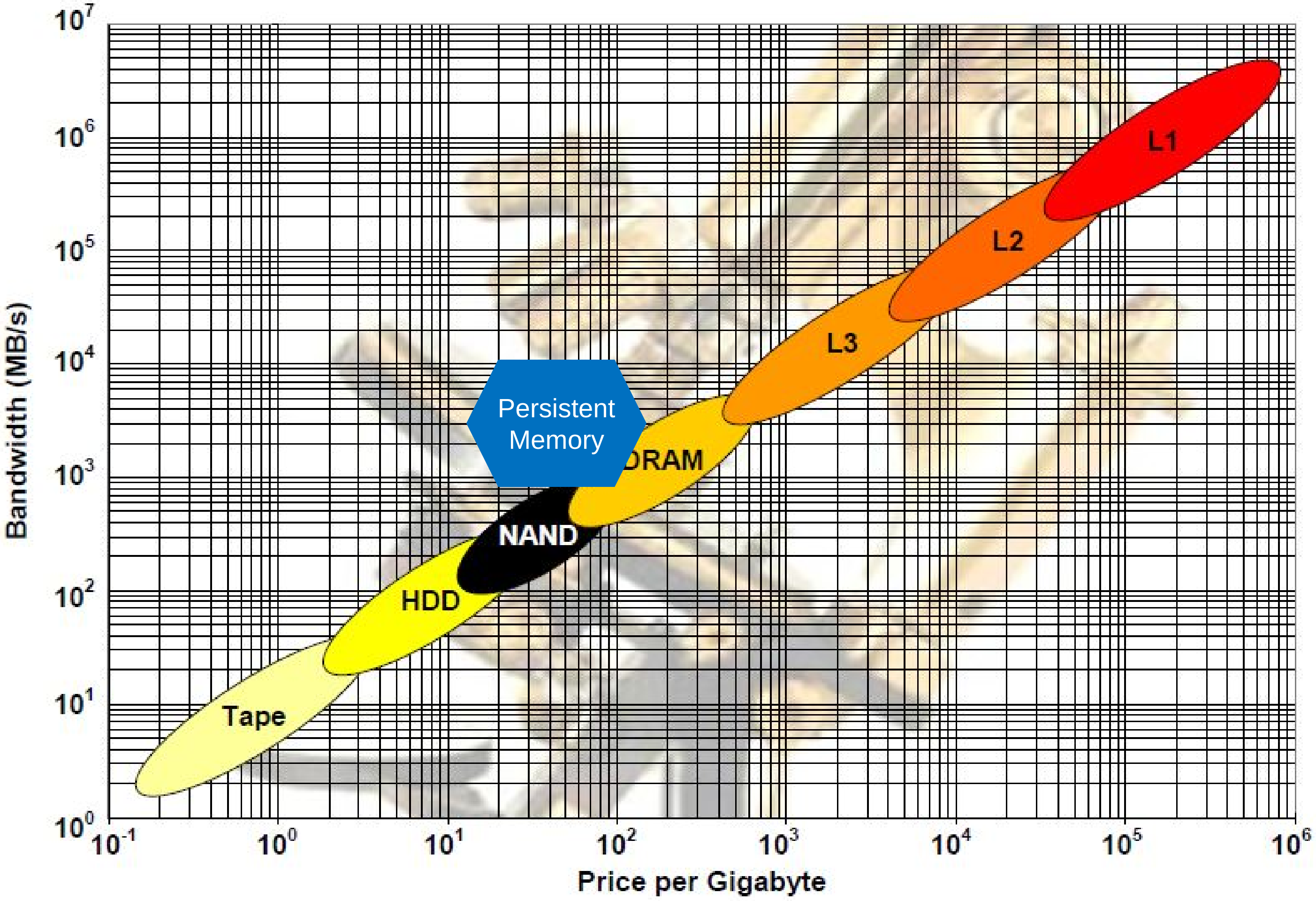
- Data Base Log Files
- Clustering
- Cache Synchronization

Meta Data Acceleration

Meta Data Acceleration

NVM-DIMM – fills growing DRAM-NAND gap

- In Memory Applications are driving a new class of Storage Class Memory (SMC)
- Latency and persistence are as important as absolute bandwidth
- Byte and Block address flexibility is vital to scaling In-Memory Applications (IMA)



Source: Objective Analysis, 2015

The Future of Business Intelligence

Bandwidth & Capacity

- Old performance was data rates (GB/s) & capacity (TB)
- Store Everything, Sort Later
- Higher Cost, Slow Decisions

Latency & Persistence

- Real-Time is Business Critical
- Major Players Driving NMV
- Store the Vital & Analyze now

Latency and Persistence are the new value currency for real-time applications & storage

Procter & Gamble - Real-Time Reporting & Business Decisions

400%

Increase in decision support systems performance

35,000

Retail, supply chain and business users supported

55%

Reduced database from 36TB to 16TB all in memory

P&G achieved faster, more reliable reporting and analytics

https://hana.sap.com/abouthana/customer-stories/pg.html

G2M COMMUNICATIONS

McLaren Group – Faster Formula 1

- Faster and more consistent lap times
- Improved down force for better grip
- Real-time telemetric analysis
- More World Championships

The Art and Science of In Memory Applications

- Taufik Ma
  Industry Insight

Evolution of Databases & Analytics

1980s

RDBMS

1990s

Operational
(OLTP, ERM)

RDBMS

Oracle, MS SQL, Sybase

Data Warehousing
(Data mining, DSS, Analytics)

EDW/OLAP

Teradata, Oracle,
SAS, etc

2000s-2015

RDBMS

MySQL
Postgres

NoSQL

MongoDB
Cassandra

Hadoop

MapReduce
HBase

EDW/OLAP

IBM Netezza
EMC Greenplum

Ongoing Evolution & Specialization…

Real-time, Online Operations	**RDBMS**	**NoSQL**	OLTP, ERM Purchases, clicks User profiles, reviews Content Management
Batch, Offline Analytics	**EDW/OLAP**	**Hadoop**	User Segmentation Daily offer recommendation Ad serving engine Fraud Detection
	Structured Data, Relational	Unstructured, Schema-less	

Ongoing Evolution & Specialization…

	Structured Data, Relational	Unstructured, Schema-less	
Real-time, Online Operations	**RDBMS**	**NoSQL**	OLTP, ERM Purchases, clicks User profiles, reviews Content Management
Real-time analytics	**In-Memory Database** *Hana, Exalytics, MemSQL, etc*	**In-Mem Data Processing** *Spark, Hadoop in-mem*	Financial risk/value analysis Fraud Prevention Real-time recommendations Profitability analysis
Batch, Offline Analytics	**EDW/OLAP**	**Hadoop**	User Segmentation Daily offer recommendation Ad serving engine Fraud Detection

Multiple Tools Within A Customer

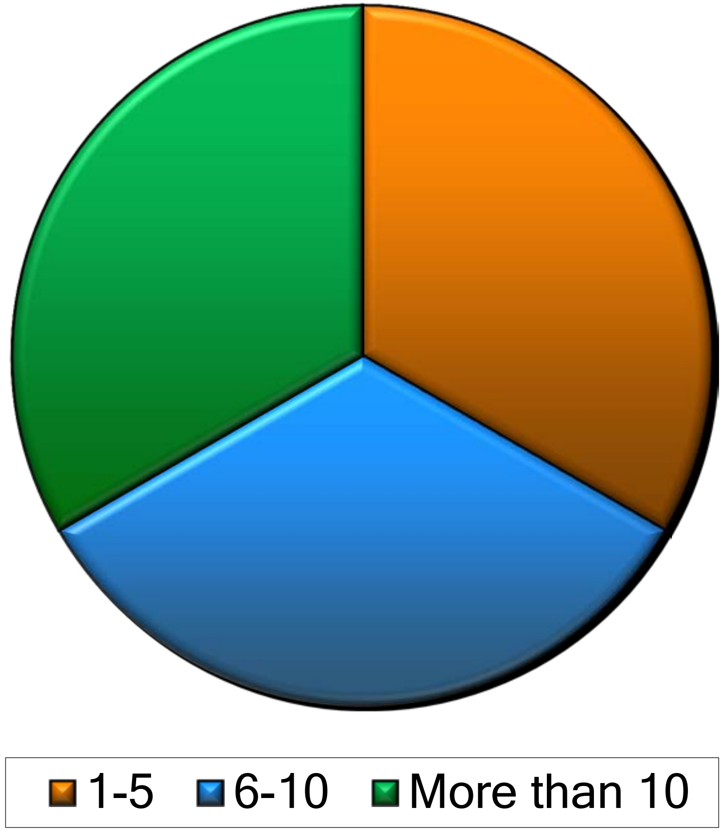
	Customer Profiles (G2M Survey)					
	$500M+ Retail	$500M+ Pharma	$1B+ Manufacturing	$1B+ Pharma	$1B+ SaaS	$250M+ Healthcare
Hadoop	Yes	Yes	Yes	Yes	Yes	Yes
MongoDB	Yes	No plans		Yes	Yes	No plans
Spark	Yes	No plans	Considering	Yes, in 6 months	Yes	Yes, in 6 months
SAP HANA	No plans	Yes	Considering	Yes	No plans	Considering
Microsoft Hekaton	No plans	No plans	Considering	Yes, in 6 months	No plans	Yes, in 12 months
memSQL	No plans	No plans	Considering	Yes, in 6 months	No plans	Yes, in 12+ months
Oracle Exalytics	No plans	No plans	Yes	Yes	No plans	Yes, in 12+ months

“Specialized Tools for Specific Needs”

(Or “Too Many Data Islands”?)

Multiple In-Memory Applications within a Customer

How many in-memory applications do you (or will you) run?

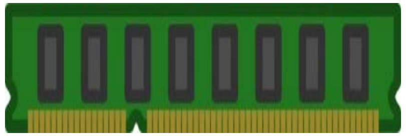


Key Enabler of In-Memory Computing: Today's Technologies

	Time to get data
CPU L1 cache	0.001 usec
DRAM	0.01 usec
NAND	100 usec
HDD	10,000 usec

On a human scale…

If I complete 50 operations in 50 seconds, then have to wait for data…



DRAM = getting food from the fridge (10's of seconds)

NAND = taking the day off

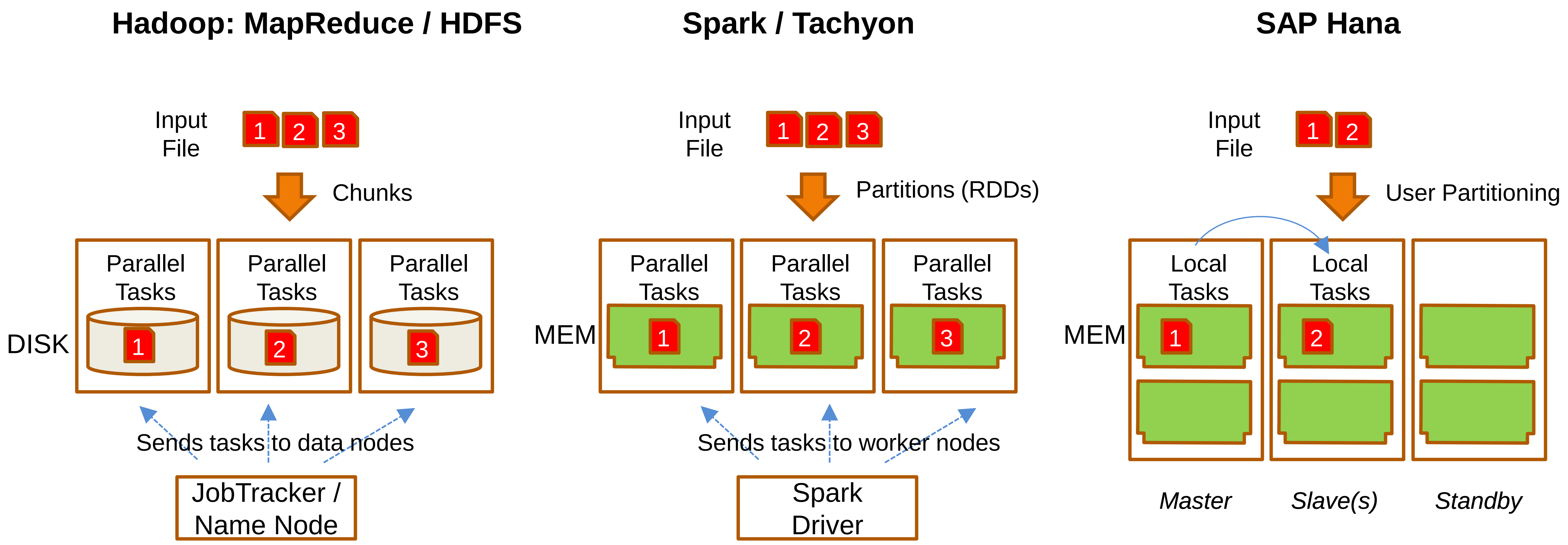
HDDs = hiking the Pacific Coast Trail (months)

Performance Comes at a Price

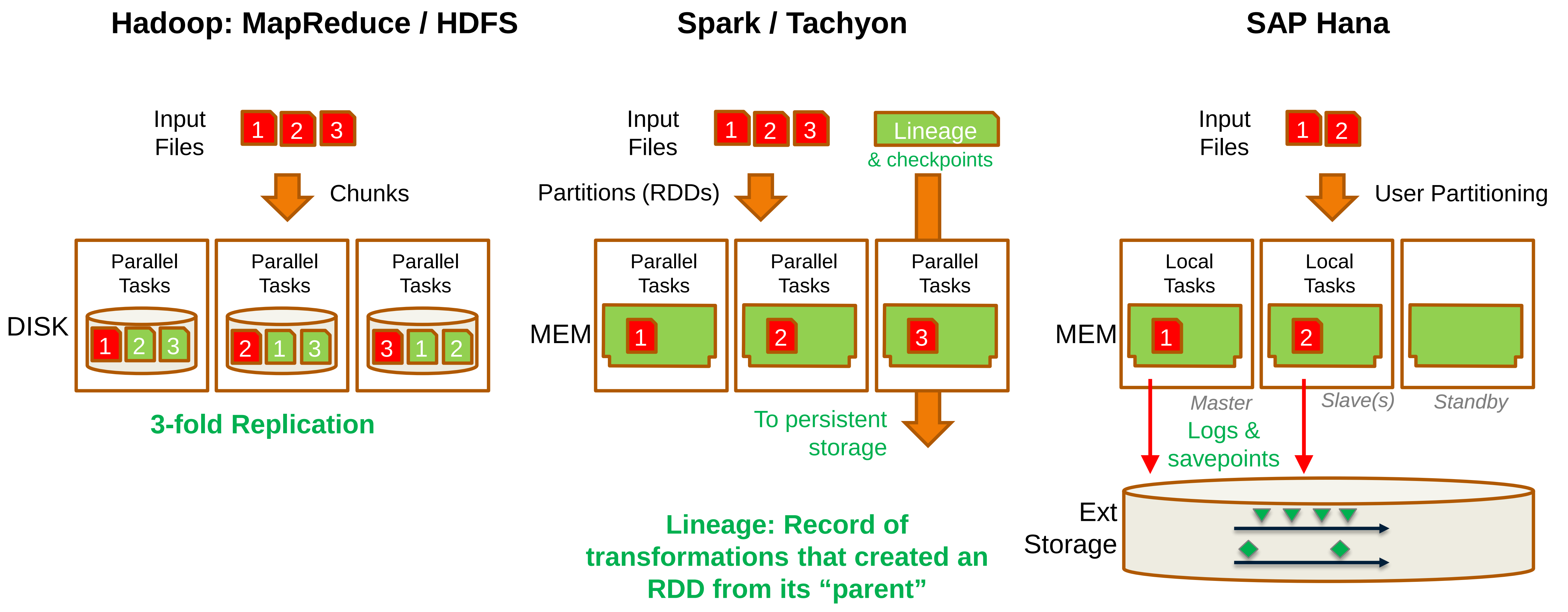
Storage	Time to get data	Price / GB	Cost for 100TB	# 2U Servers Req'd to Hold 100TB*
DRAM	0.01 usec	$5.60 32G DIMM for $179 ea, Samsung Registered DDR4, M393A4K40BB0-CPB0	$560,000 3125 x 32G DIMMs	130
NAND	100 usec	$0.35 2.5" 1TB SSD, $350 ea, Intel 540S	$35,000 100 x 2.5" 1TB SSD	5
HDD	10,000 usec	$0.03 3.5" 4TB SATA HDD for $120 ea, Seagate ST4000DM000	$3,000 25 x 3.5" 4TB SATA HDD	2-3

* Assuming 24 DIMM slots, 24x 2.5" drives or 12x 3.5" drives

Location of Data & Tasks

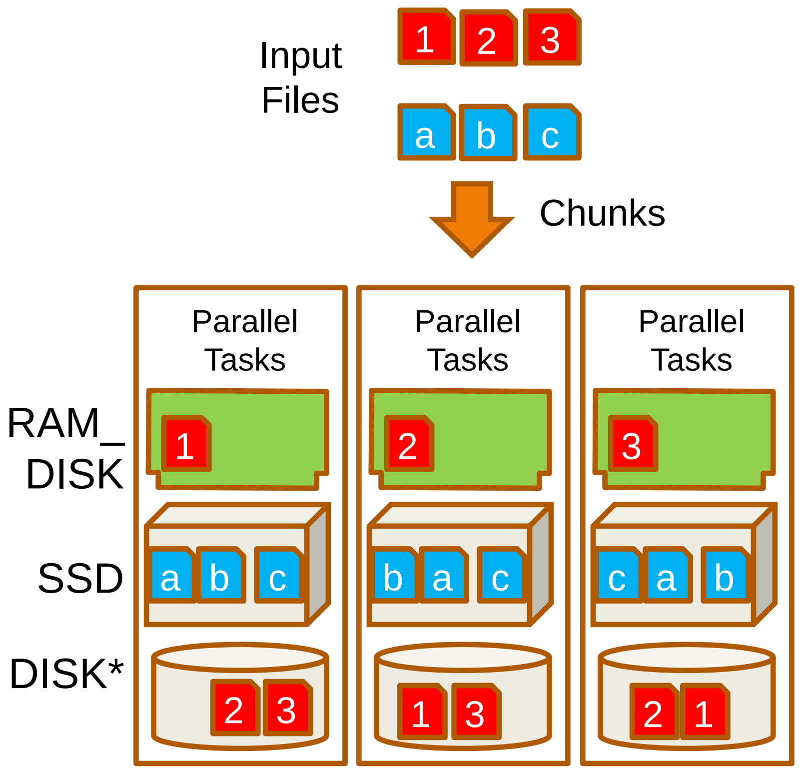


Surviving Failures



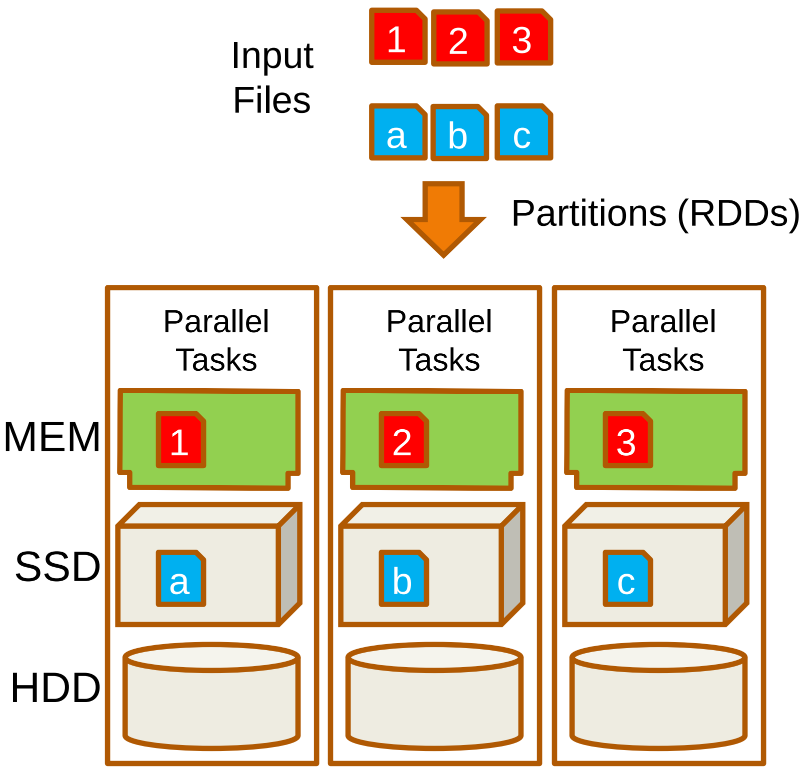
No such thing as 100% In-Memory

Hadoop: MapReduce / HDFS



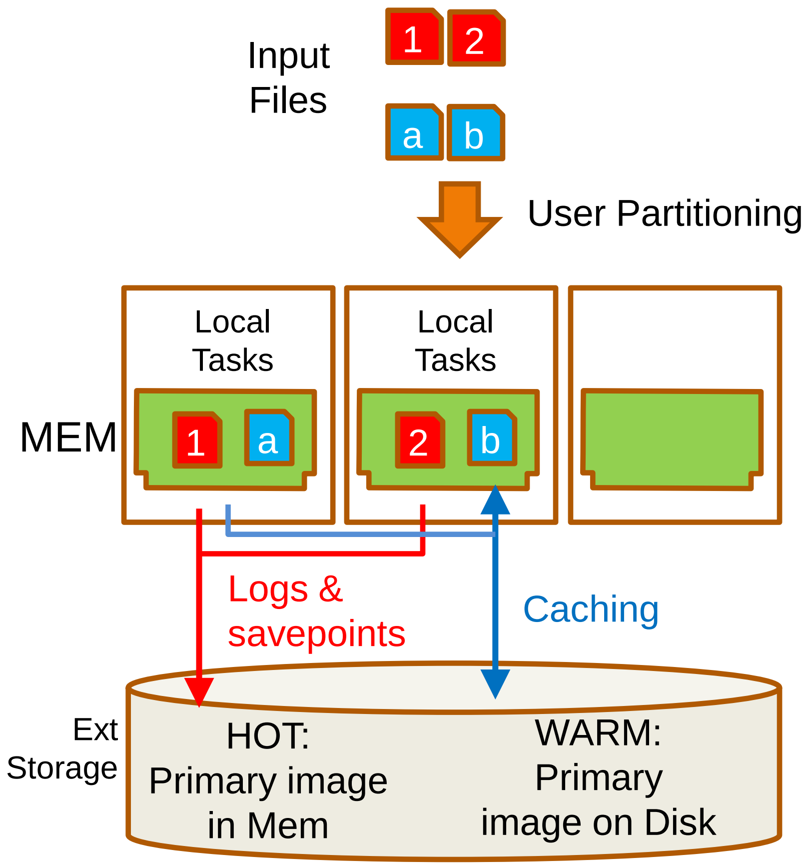
HDFS2.0 Heterogeneous Storage
Storage Types & Policies
Files/directories assigned policies
(e.g. Lazy_persist, All_SSD)

Spark / Tachyon



Tachyon Tiered Storage
(for Off_heap Spark RDDs)
Auto or manual

SAP Hana



SAP HANA Dynamic Tiering
Data spec'd as either Hot or Warm

* ARCHIVE tier not shown

Customer In-Memory Computing Trends (based on G2M survey)

SIZE

- **Cluster sizes similar to big data solutions**
  - ½ respondents > 500 servers, 1/3 at >50
  - And not just for Spark
- **With datasets that fit available DRAM capacity**
  - 1/3 at >100TB, 1/3 at >10TB

GROWTH

~Half with 10-20%+/yr dataset growth

Majority use/want tier-ing when dataset > DRAM

Only minority would rely on scale-out only

Mixed on whether tier-ing should be transparent or not

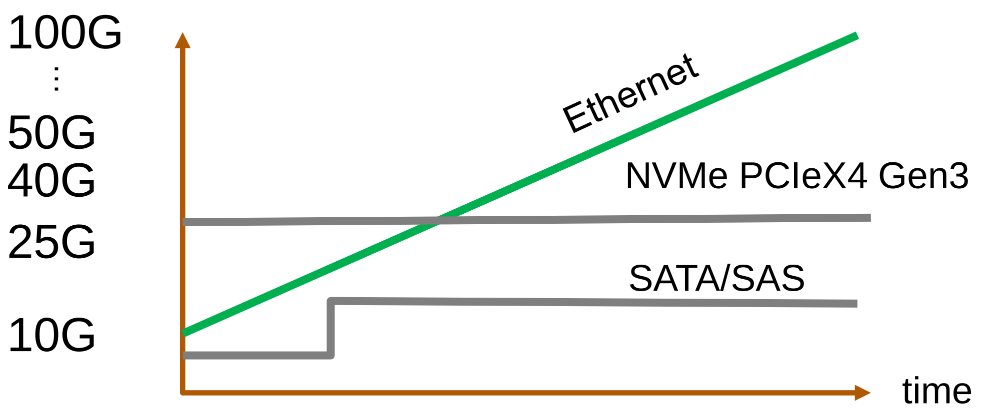
Some want it transparent to developer; Rest want developer to have control via policy

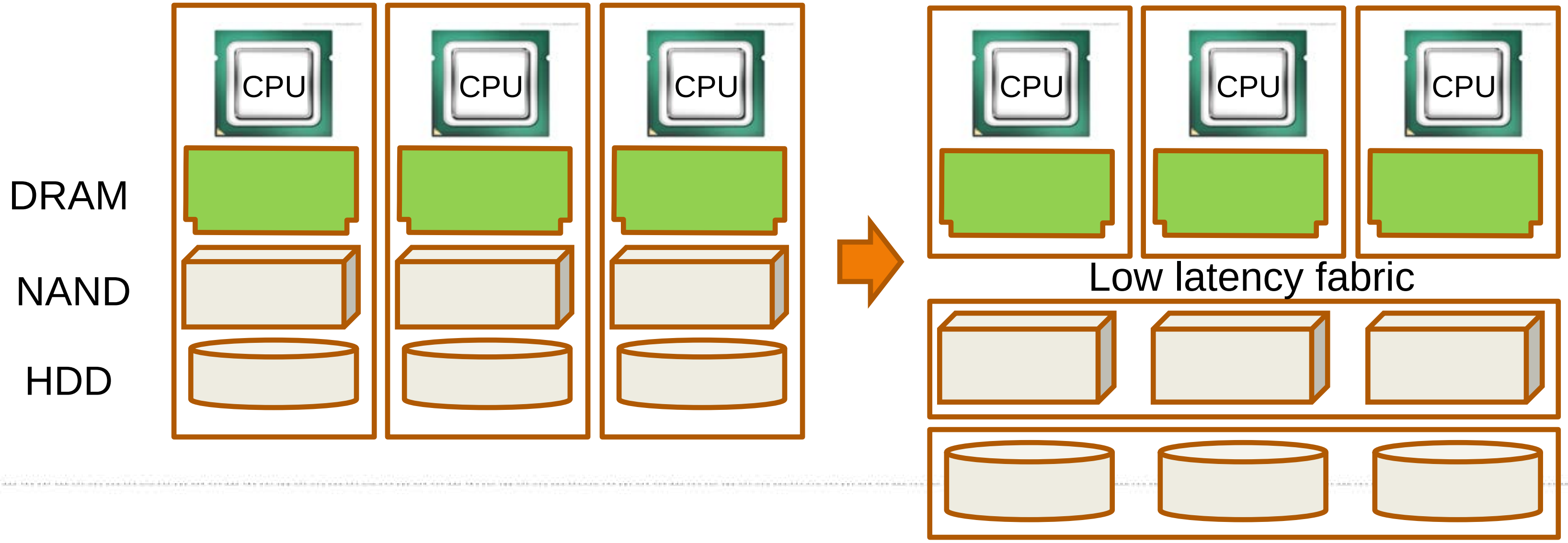
EFFICIENCY

- **~Half believe *"my storage capacity forces me to have more compute capacity then I need"***
- **Majority have or have plans for consolidated data silos**
  - OLTP+IMDB, Spark+Hadoop, NoSQL+Hadoop

Emerging Technologies: High-speed Fabrics & Disaggregated Storage

- Data Center Ethernet speeds ramping faster than drive speeds: 10/25/40/50/100G
- RDMA-over-Ethernet technologies
- Multi-host PCIe fabrics emerging (e.g. OCP Lightning) albeit w/ less scalability





- Ethernet or PCIe based fabric
- DAS-like performance Local or SAN
- Map any drive to any host
- Scale each storage tier separately from compute
- Early proof points: EMC DSSD, SanDisk InfiniFlash, DriveScale

Emerging Technologies: Storage Class Memory

Storage	Persist-ence	Time to access data	Price / GB	Cost for 100TB	# 2U Servers Req'd to Hold 100TB*
DRAM	N	10ns+	$5.60	$560,000 3125 x 32G DIMMs	130
NV-DIMM -N	Y	10ns+	$10+ If 2X+ DRAM	$1,000,000+	260 16G NVDIMM, supercap
3DXP DIMM		100ns Rd 500ns Wr	$2+ If 1/3+ DRAM	$190,000+	~50 assuming 96 or 128GB DIMMs
NAND	Y	100 usec	$0.35 2.5" 1TB SSD, $350 ea, Intel 540S	$35,000 100 x 2.5" 1TB SSD	5
HDD	Y	10,000 usec	$0.03 3.5" 4TB SATA HDD for $120 ea, Seagate ST4000DM000	$3,000 25 x 3.5" 4TB SATA HDD	2-3

* Assuming 24 DIMM slots, 24x 2.5" drives or 12x 3.5" drives

In-Memory Computing Predictions / Trends

1. 3DXP DIMMs used for **“Jumbo Memory”** – value in lower $/GB vs DRAM, not persistence
   - Mix of 3DXP & DRAM DIMMs in server nodes
   - Tier-ing will be tuned to accommodate slower writes & reads
   - Spark, In-mem Hadoop, MemSQL, Hana, etc
   - NV-DIMM –P might have similar adoption but predictable latency is a concern
2. Increasing use of NVMe SSDs as **“Far Memory”** – as next tier (below DRAM/3DXP)
   - Priority on $/TB, not persistence. Resiliency still via Lineage, logs, etc
   - Remove ”last-inch” of latency via BLKB (block-layer/kernel bypass) stacks (e.g. EMC libflood, SPDK)
   - Implemented as a fabric-disaggregated cluster to enable efficiency & independent scalability
   - Longer-term, HW-based paging of near-memory to far-memory
3. Use of **“Persistent Memory”** for In-Mem computing will evolve
   - For 3DXP & NV-DIMM –N
   - Industry progress on pmem file systems (Linux, Windows)
   - Does persistence replace or complement lineage/logs?
   - Need low latency replication across nodes (PMoF)

Summary

- In-memory solutions growing in adoption – driven by real-time analytics
- Co-existence of structured (e.g. Hana) and unstructured frameworks (e.g. Spark)
- Confluence of big-data & real-time analytics drives increasing adoption of tier-ing
- Newer technologies on horizon will continue to create disruptions to in-memory computing architectures

G2M COMMUNICATIONS

G2M

COMMUNICATIONS