In-Memory Computing:
Facts & Myths

Jon Webster
VP of Product
Management

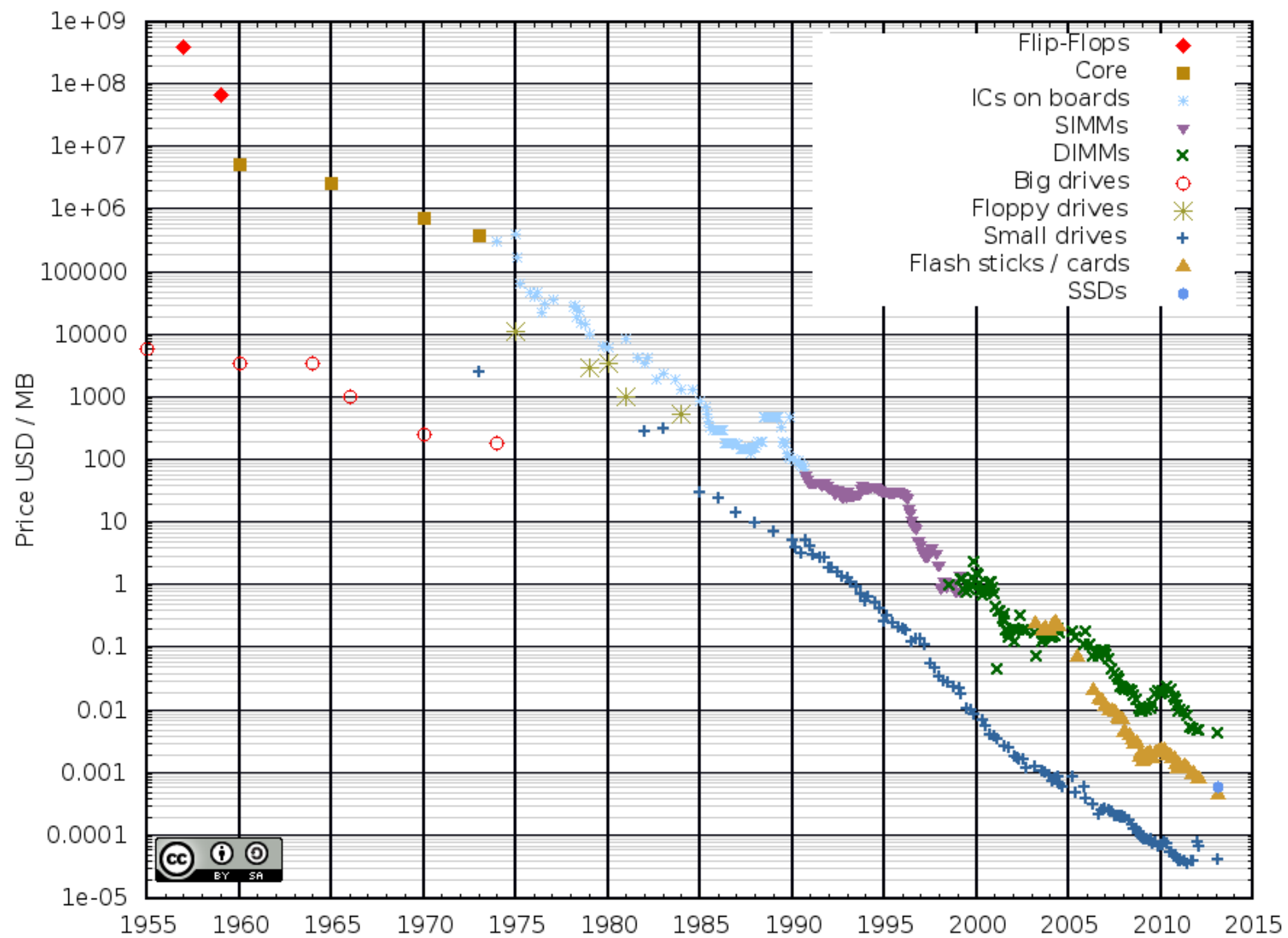GridGain
IN-MEMORY COMPUTING

What is In-Memory Computing?

In-Memory Computing uses high-performance, integrated, distributed memory systems to compute and transact on large-scale data sets in real-time - orders of magnitude faster than legacy disk-based systems.

Why Now?

Paradigm Shift à la 1970s

1970s:

> IBM released "Winchester" IBM 340 disk - **Era of HDD**
> Tapes start to decline

> SQL

> Era of Structured Data

2010s:

> 64-bit CPUs + DRAM prices drop 30% YoY - **Era of Memory**
> HDDs start to decline

> NoSQL + SQL

> Era of Unstructured Data

Gartner **RAM is a new disk, disk is a new tape.**

GridGain
IN-MEMORY COMPUTING

Memory First vs. Disk First

> **Memory First** Architecture:

Memory is primary storage, disk for backups

Reading Record: API call <-> pointer arithmetic

Latency: nanoseconds

> **Disk First** Architecture:

Disk as primary storage, memory for caching

Reading Record: API call <-> OS I/O <-> I/O controller <-> disk

Latency: milliseconds

Bring Computations To Data

> **Client-Server, J2EE, SMP** 1990s-2005s

  > Data is moved to application layer for processing:
    Data not-partitioned and stored in central RDBMS
    Data sizes are relatively small
    Technically **impossible** to distributed computations to central RDBMS

> **In-Memory Computing, Hadoop, MPP** 2005s+

  > Computations are moved to data:
    Data is partitioned and stored in distributed systems
    Data overall sizes are massive
    Technically **possible** to distribute computations for distributed data

GridGain
IN-MEMORY COMPUTING

Myth #1: Too Expensive

Facts:

> 2013: 1TB DRAM cluster ~$25K

> 2015: 1TB DRAM cluster ~$10K
30% reduction YoY

> Memory Channel Storage (MCS)
NAND in DRAM form factor, 2x speed of flash, same price as flash

> Storage Class Memory (SCM)
~10x slower than DRAM, Flash price, non-volatile

Myth #2: Not Durable

Facts:

> IMC have durable backups and disk storage
> Active or passive replicas, transactional read-through and write-through

> Mature IMC provide tiered storage
> DRAM - Local Swap - RDBMS/HDFS

> Operational vs. Historical datasets
> 99% of operational datasets < 10TB

GridGain
IN-MEMORY COMPUTING

Myth #3: Flash Is Fast Enough

Facts:

> Flash on PCI-E is still... **a block device, i.e. disk.** A faster one.
> Still going through OS I/O, I/O controller, marshaling, buffering.
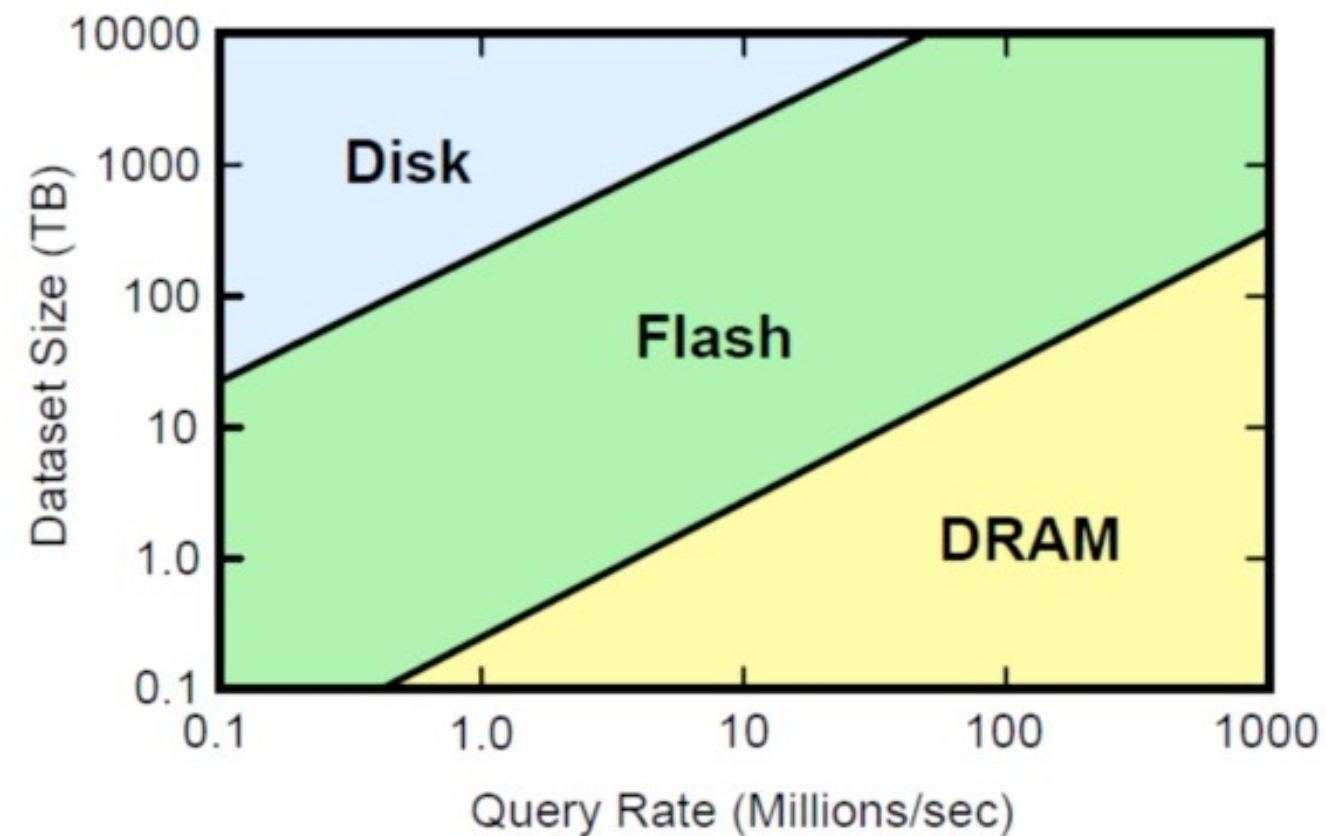
Myth #4: Only In-Memory Databases

Facts:

> **IMC is not a product - it is a technology**
> It is applied to different products and payloads

> **In-Memory Database is important use case for today**
> Easiest adoption and a "low-hanging fruit"

> **Streaming is an ideal use case for IMC going forward**
> Streaming CAN ONLY be supported on IMC

> **Vertical and PnP products are the future**
> Minimal integration, maximum benefit

GridGain
IN-MEMORY COMPUTING

Four Use Cases

> **Real-Time Risk Analytics**
Able to grow book of business while reducing latencies

> **Railroad Logistics**
Able to ingest and process sensor data for instant logistics

> **Energy Generation**
Able to decide on trade vs. generate in real-time as demand spikes

> **Oil & Gas Drilling**
Able to provide real-time safety monitoring during fracking

GridGain
IN-MEMORY COMPUTING

Thank you!

GridGain
IN-MEMORY COMPUTING

www.gridgain.com   [#gridgain](#)