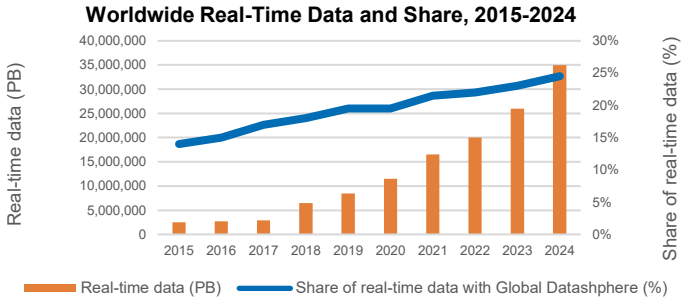
MemVerge

Persistent Memory + Data Services = Big Memory

September 2020

Industry Trends

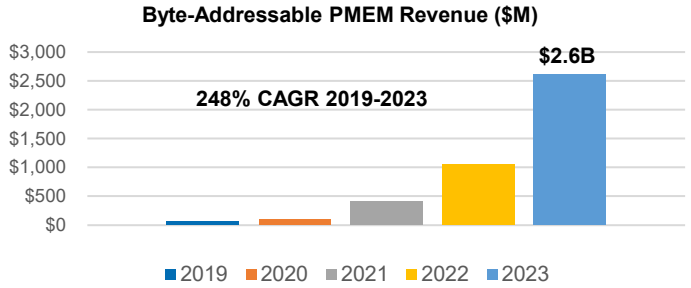
Real-Time Data
as % of All Data is Growing



- Growth of Real-Time Data & Applications Demand Memory-Centric Infrastructure:
  - By 2024, ¼ of world's data will be real time data
  - By 2021, 2/3 of Global 2000 companies will have at least one mission critical app needing to process real time data

Big Memory Software
Required to Unlock the Potential of Real-Time Data on Big Memory

Persistent Memory
Emerging…

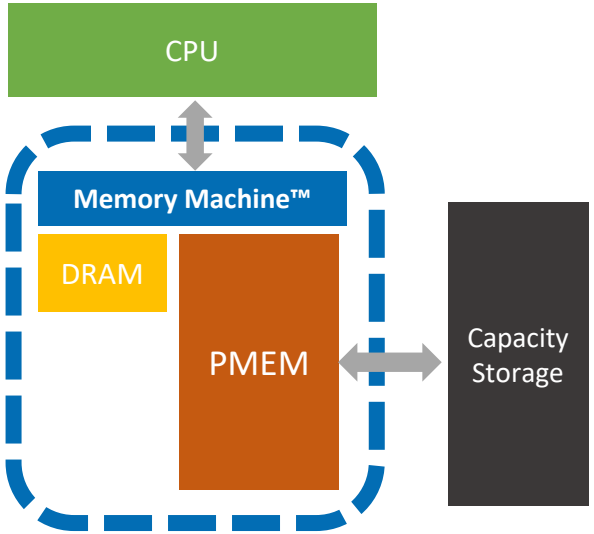


- Intel Shipped Optane Persistent Memory (PMEM) in April 2019
  - Larger capacity and lower price than DRAM
  - Persistence built-in
- Other companies (Micron, SK Hynix and others) are expected to follow

MemVerge

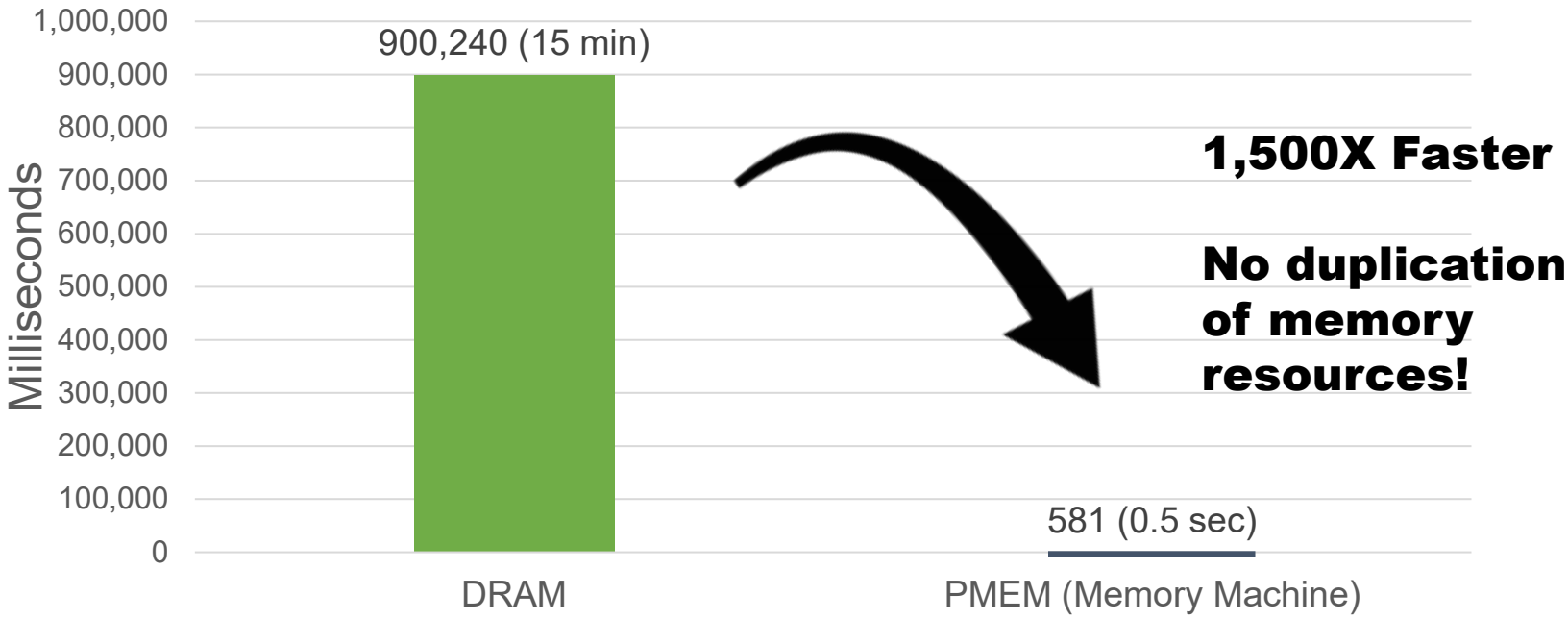
MemVerge Memory Machine™ is the World's 1st Big Memory Software

- **Standard Edition: The BEST way of using PMEM**
  - Software-Defined Memory Service compatible with DRAM
    - No application rewrite required
  - Expand memory capacity by tiering DRAM and PMEM
    - Up to 4.5TB per CPU socket
  - Lower TCO of memory due to lower cost of PMEM
    - 30-50% cost savings on memory
  - DRAM-like performance
    - Software optimization in memory allocation and tiering



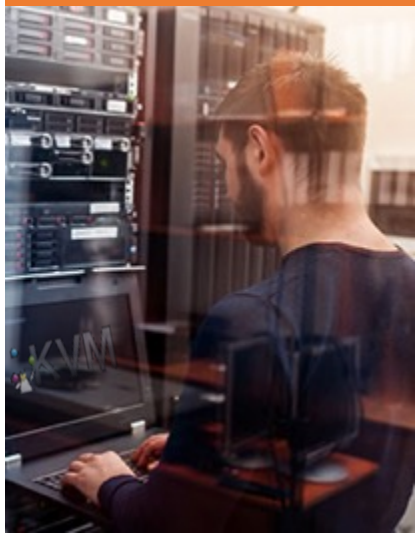
- **Advanced Edition with ZeroIO™ Snapshot : The BEST way to persist onto PMEM**
  - No application rewrite required
  - Auto-save feature protects against application crashes
  - Thin-clone feature enables agile and resource-efficient deployment

Restore a Redis Database with 300M Keys



Use Case Examples

Lower Cost of Cloud Memory Infrastructure



Increase Studio Business Continuity

Increase Agility with Thin Cloning

Memory Machine Use Case

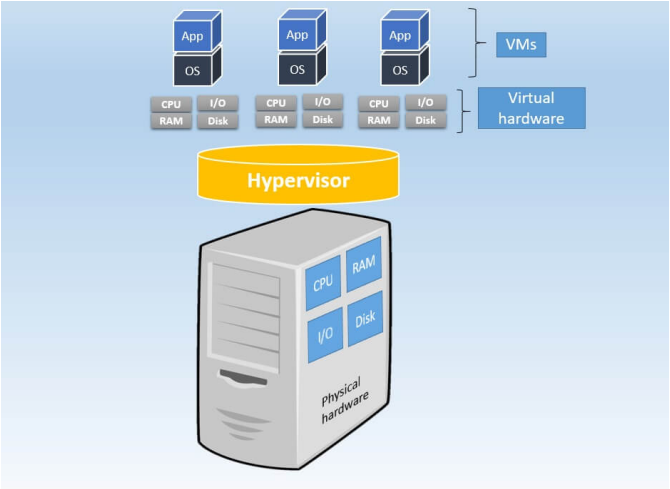
KVM Virtual Machines

Problem

- Increase VM density
- Reduce per-VM cost

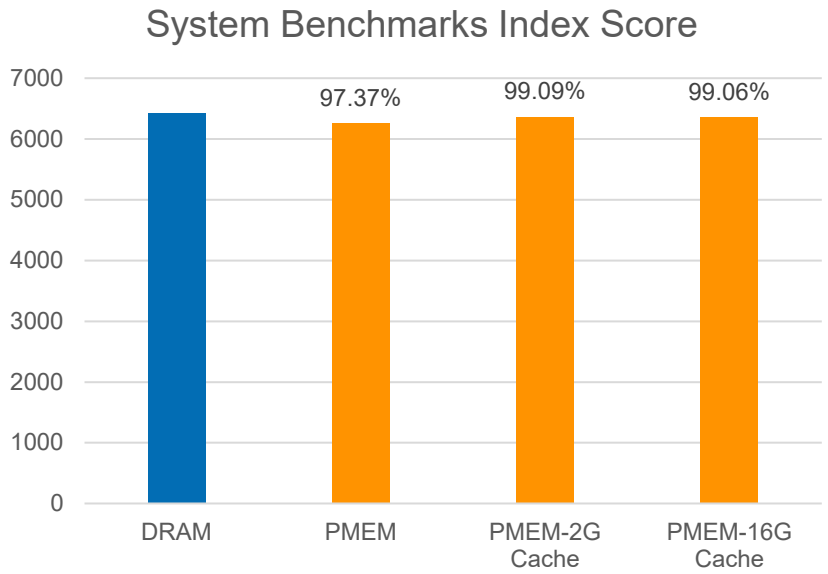
Solution

- Memory Machine™ software delivers software-defined Big Memory service, multiply amount of memory available to the hypervisor
- Adjustable DRAM:PMEM ratio, enabling memory performance similar to DRAM

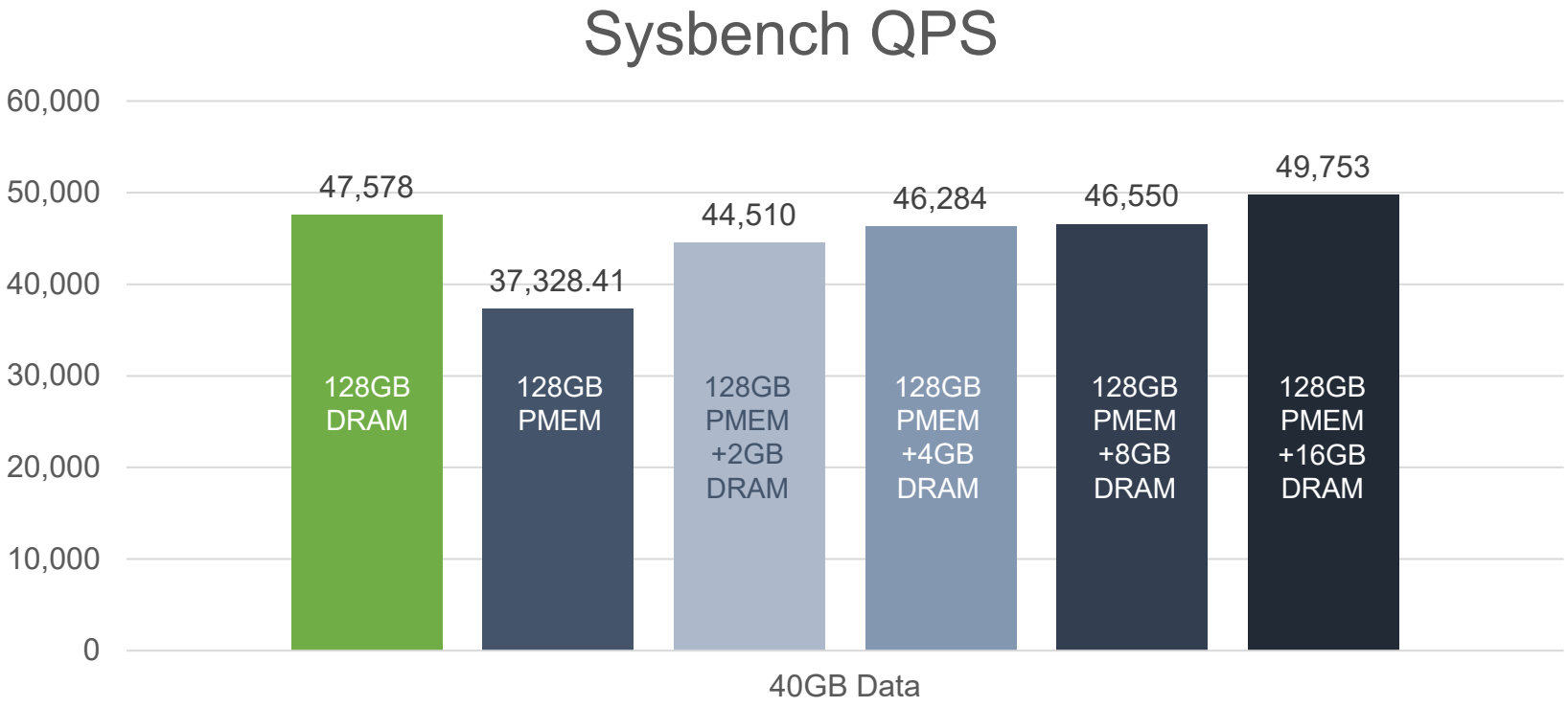


Unix-Bench Performance

- Dhrystone 2 using register variables
- Double-Precision Whetstone
- Execl Throughput
- File Copy 1024 bufsize 2000 maxblocks
- File Copy 256 bufsize 500 maxblocks
- File Copy 4096 bufsize 8000 maxblocks
- Pipe Throughput
- Pipe-based Context Switching
- Process Creation
- Shell Scripts (1 concurrent)
- Shell Scripts (8 concurrent)
- System Call Overhead



MySQL Performance on Memory Machine™



Memory Machine Use Case

Crash Recovery

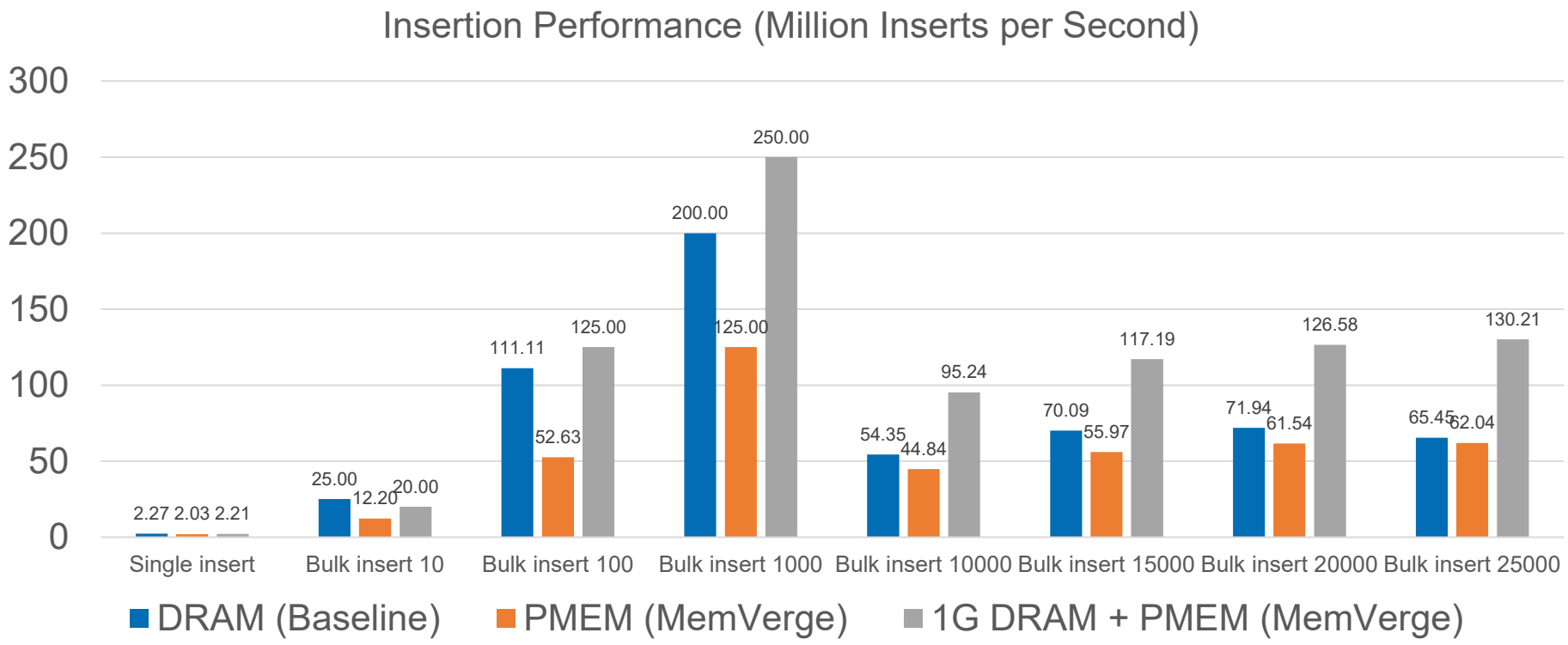
Problem

- Memory-intensive applications take a long time to restart after crash or planned shutdown

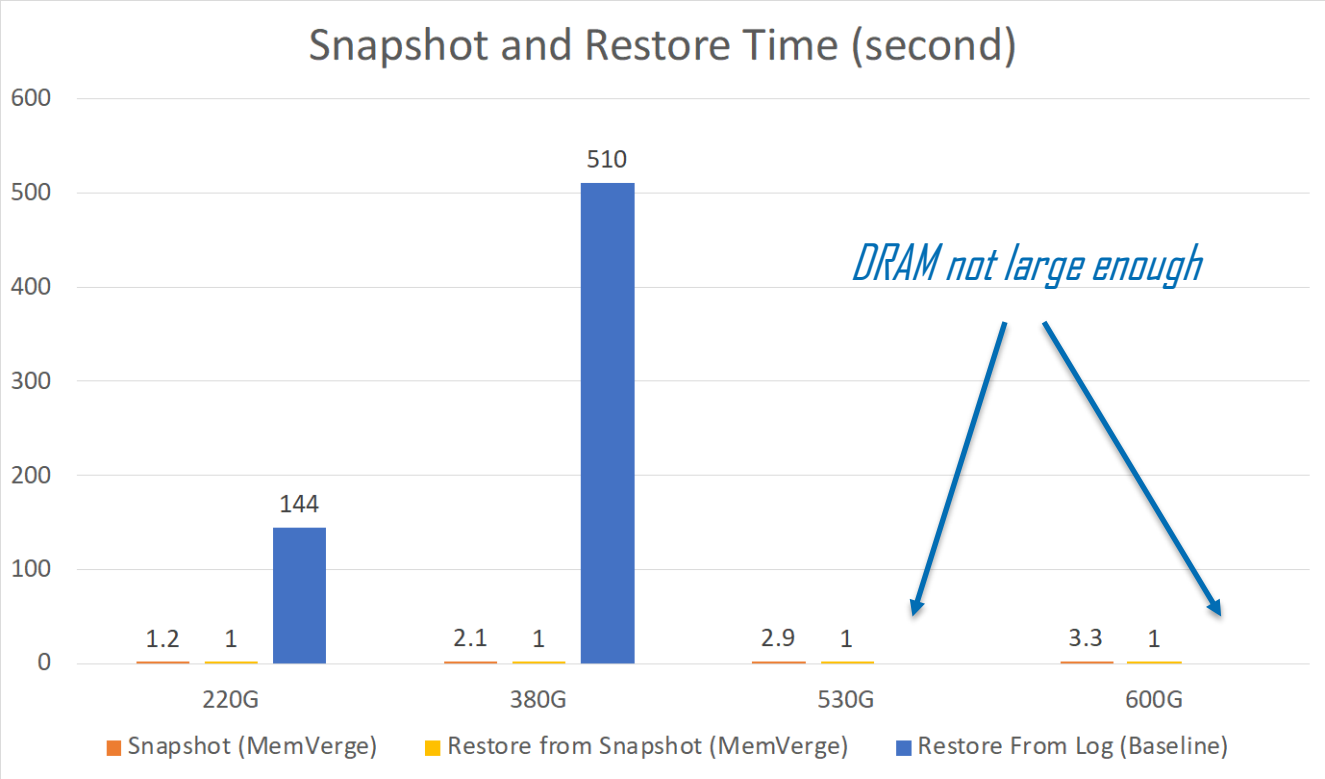
Solution

- Memory Machine takes instant snapshot, as frequently as every 1 minute
- Fast restart of the application from the last snapshot

kdb+ Performance Stress Test - Insertion



kdb+ Snapshot and Restore Performance



Memory Machine Use Case

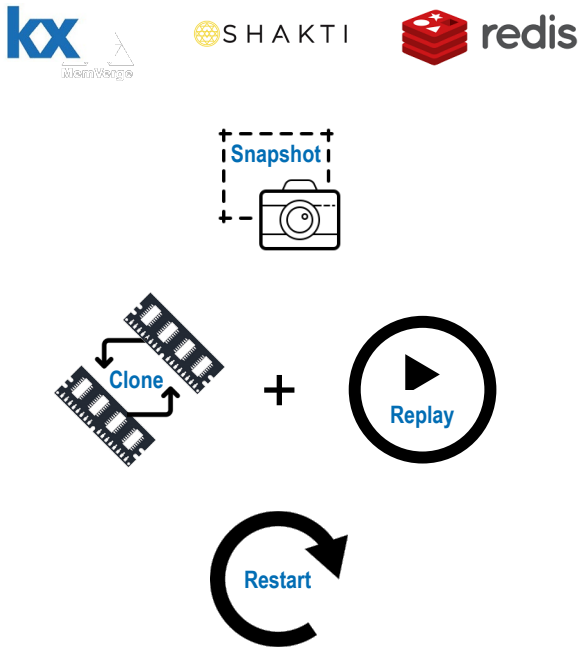
Cloning

Problem

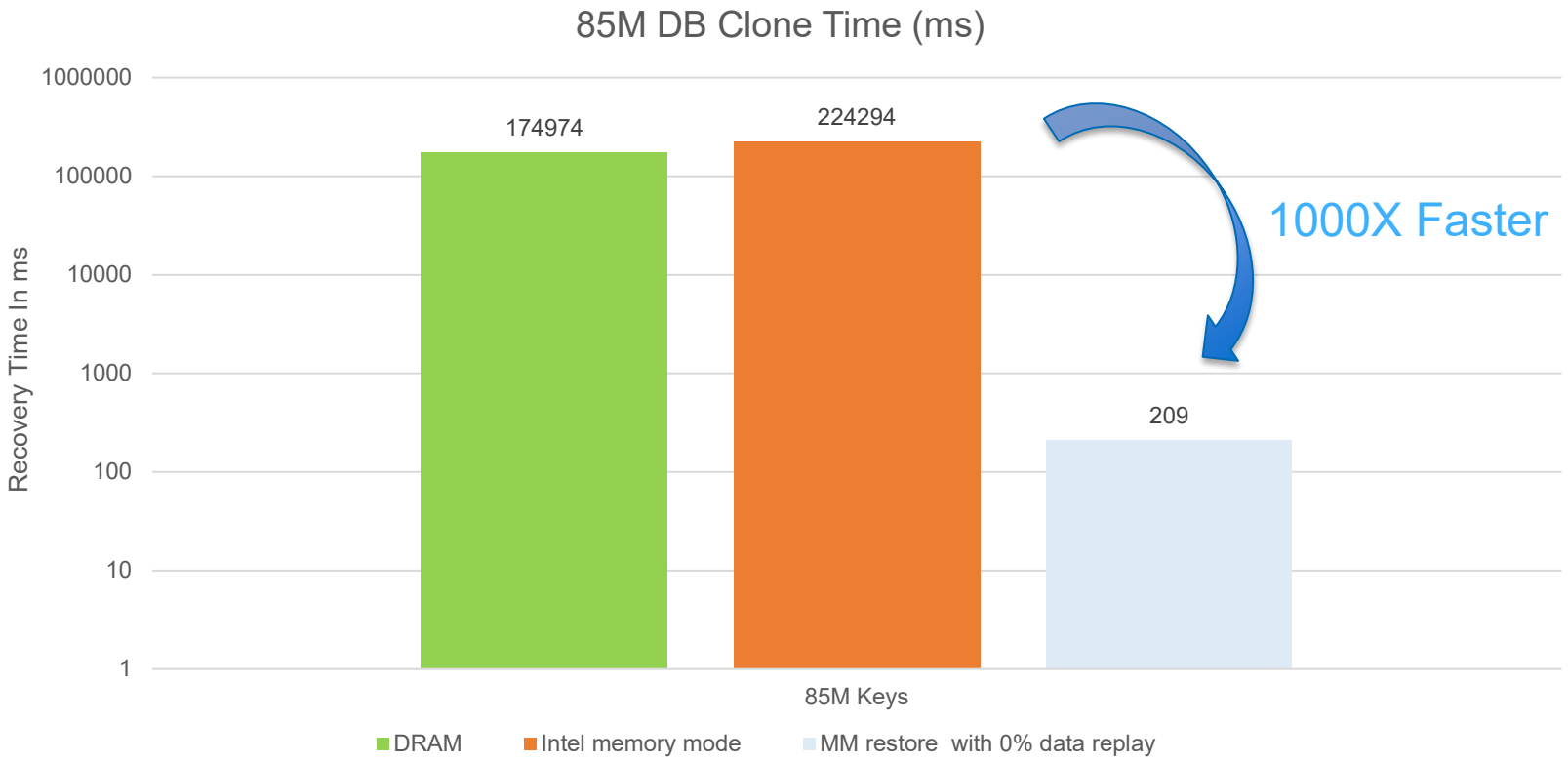
- Need to set up additional instances of memory-intensive applications, but it requires too much memory or takes too long.

Solution

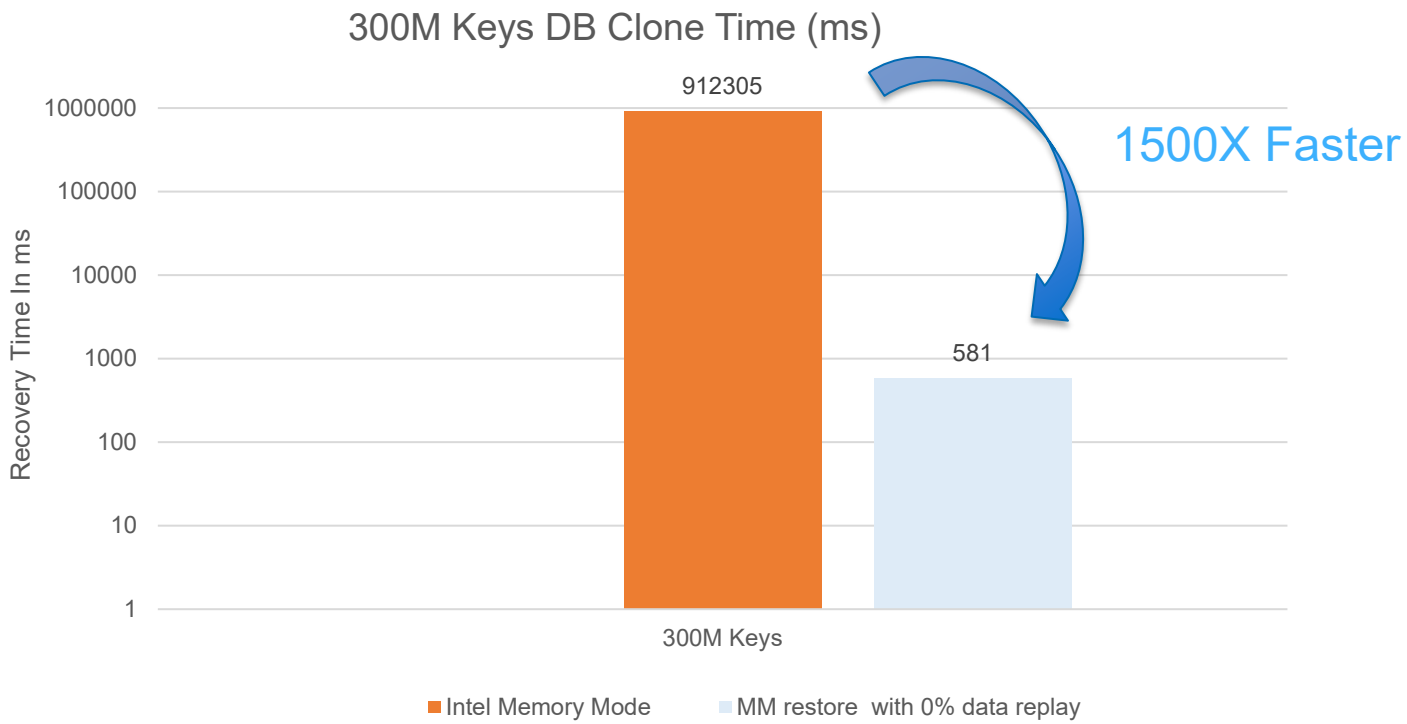
- Memory Machine takes instant snapshot, as frequently as every 1 minute
- A new application instance can be created from any snapshot quickly, without memory copy.



Redis 85M Keys DB Clone



Redis 300M Keys DB Clone



Memory Machine Use Case

AI/ML Inference

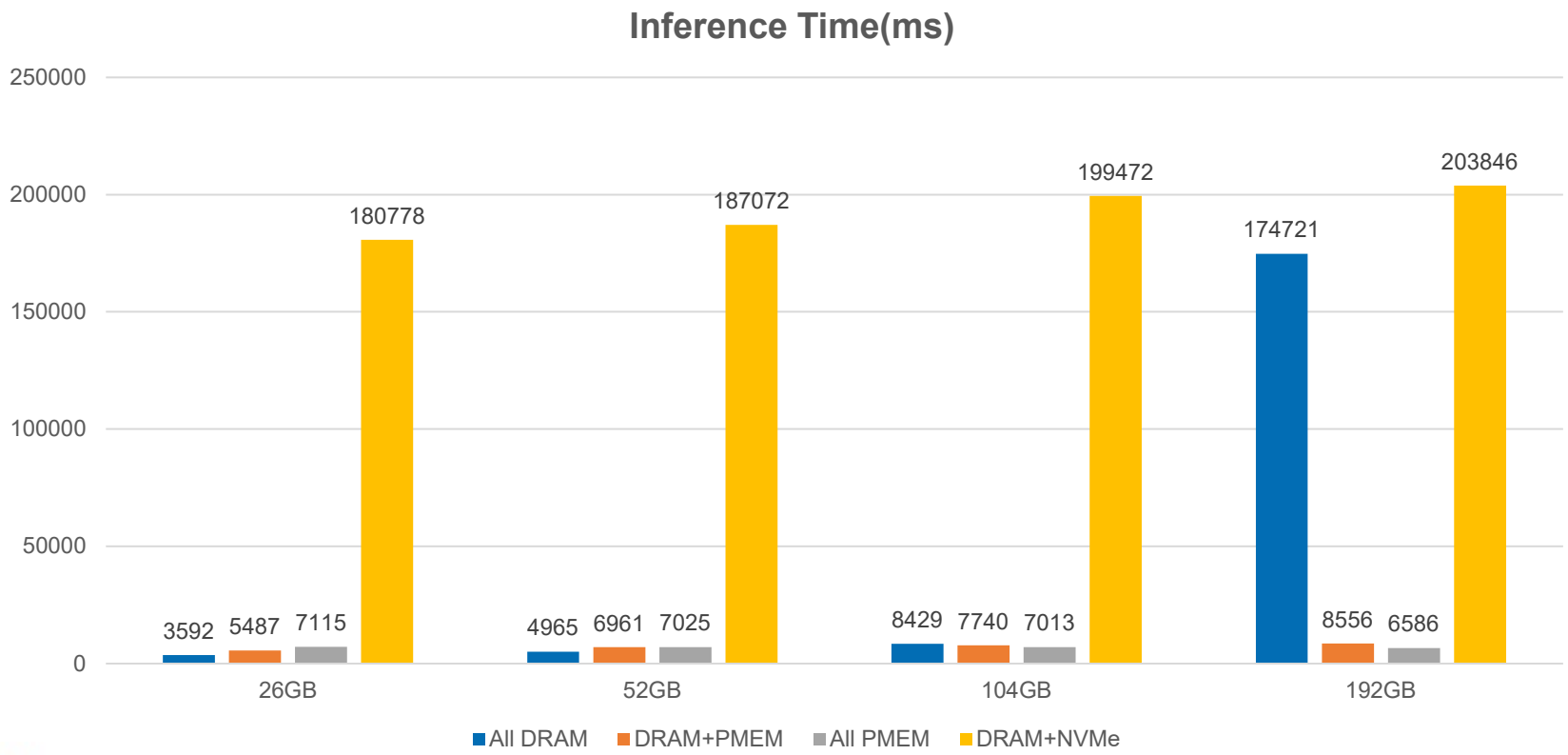
Problem

- When data is greater than the size of DRAM, AI/ML performance slows down dramatically
- Memory-intensive inference jobs take a long time to load and restart

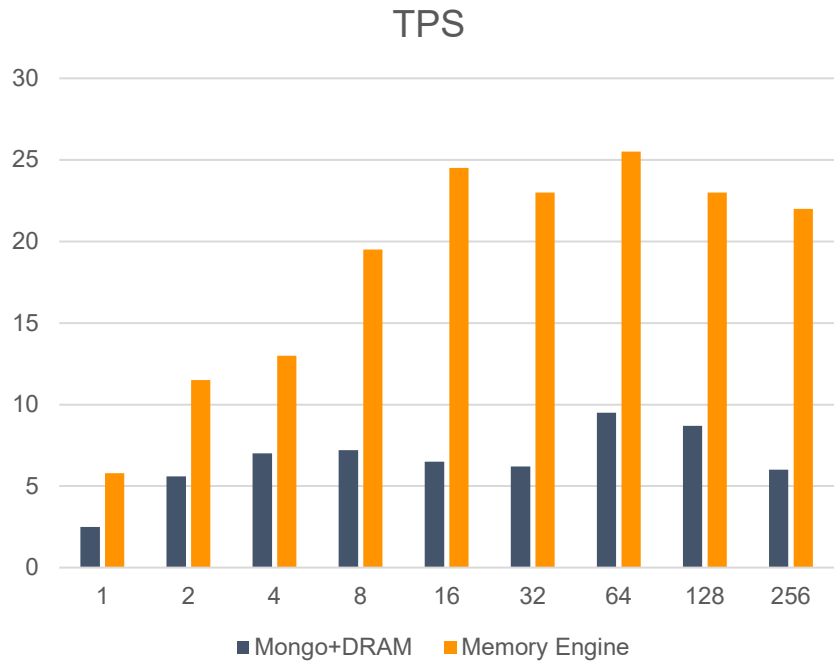
Solution

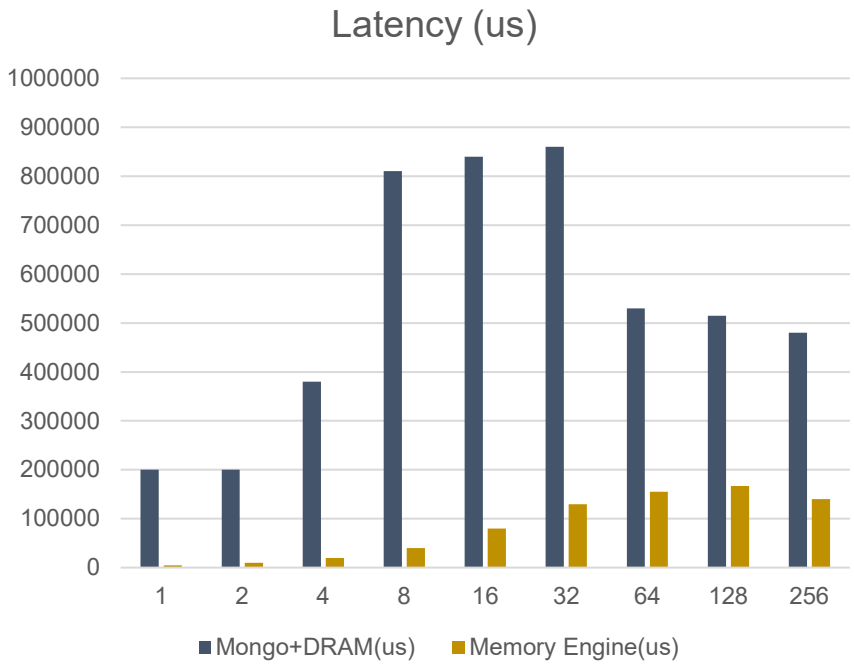
- Create big memory lakes consisting of DRAM and PMEM to provide capacity needed for all data including models and embeddings
- Fast data recovery and restart by using in-memory data snapshot

DLRM Inference Performance



Facial Recognition Inference Performance





Memory Machine Use Case

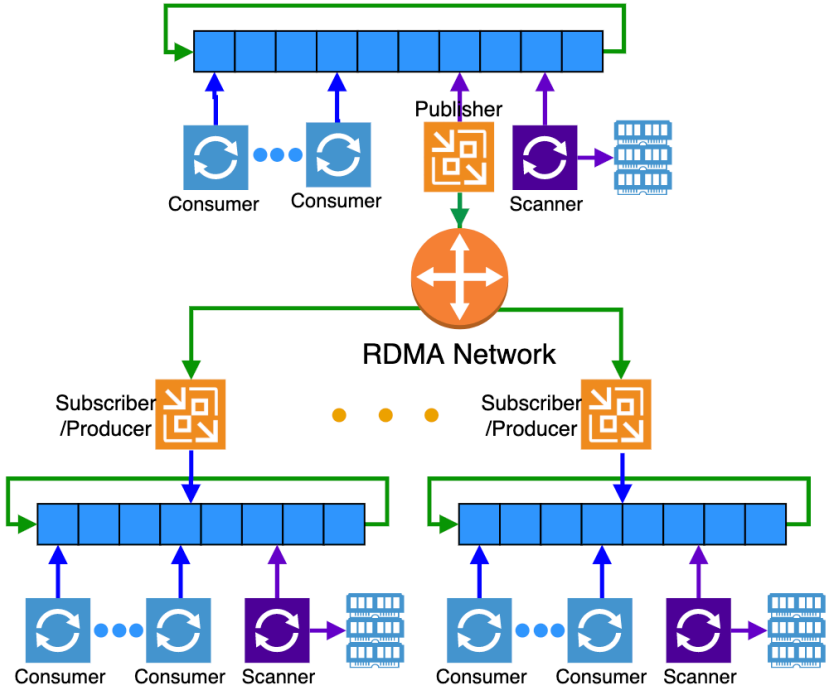
Real Time Market Data Pub/Sub

Problem

- Market data event stream published to multiple subscribers with the lowest latency and jitter

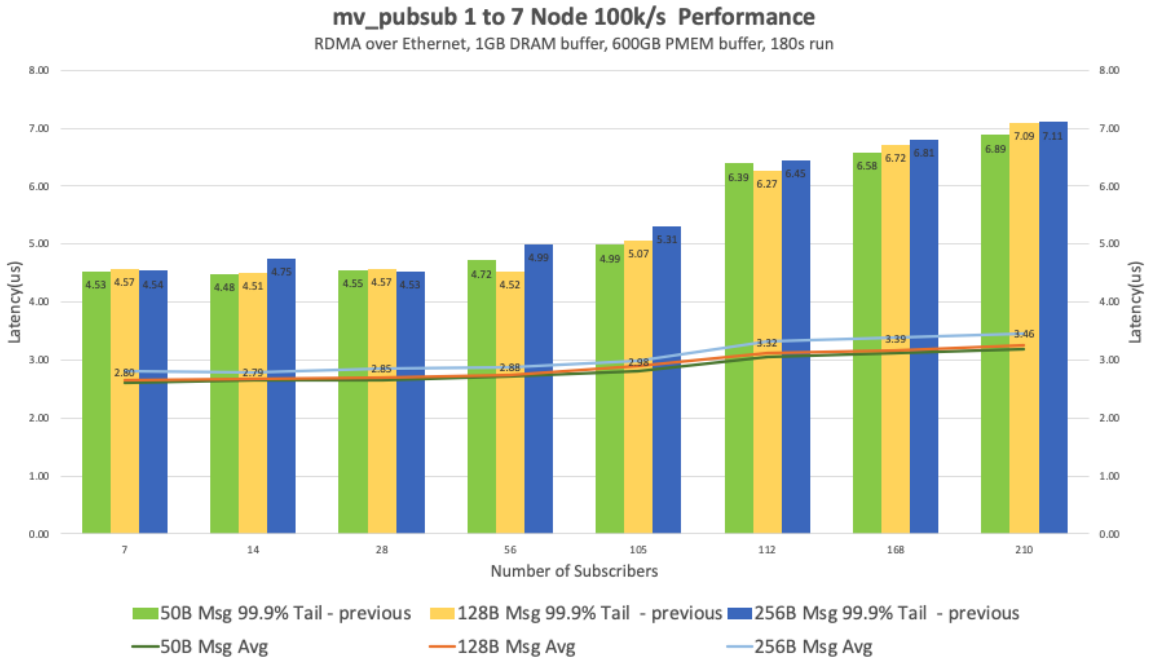
Solution

- Memory Machine™ software writes market data event stream to in-memory bus
- The events are written to memory of remote servers via RDMA
- Subscriber processes across all servers read the event stream with low latency



Real Time Market Data Pub/Sub Performance

Results with 210 Subscribers



Memory Machine™

Big Memory Software

Key Benefits to Real-Time Apps

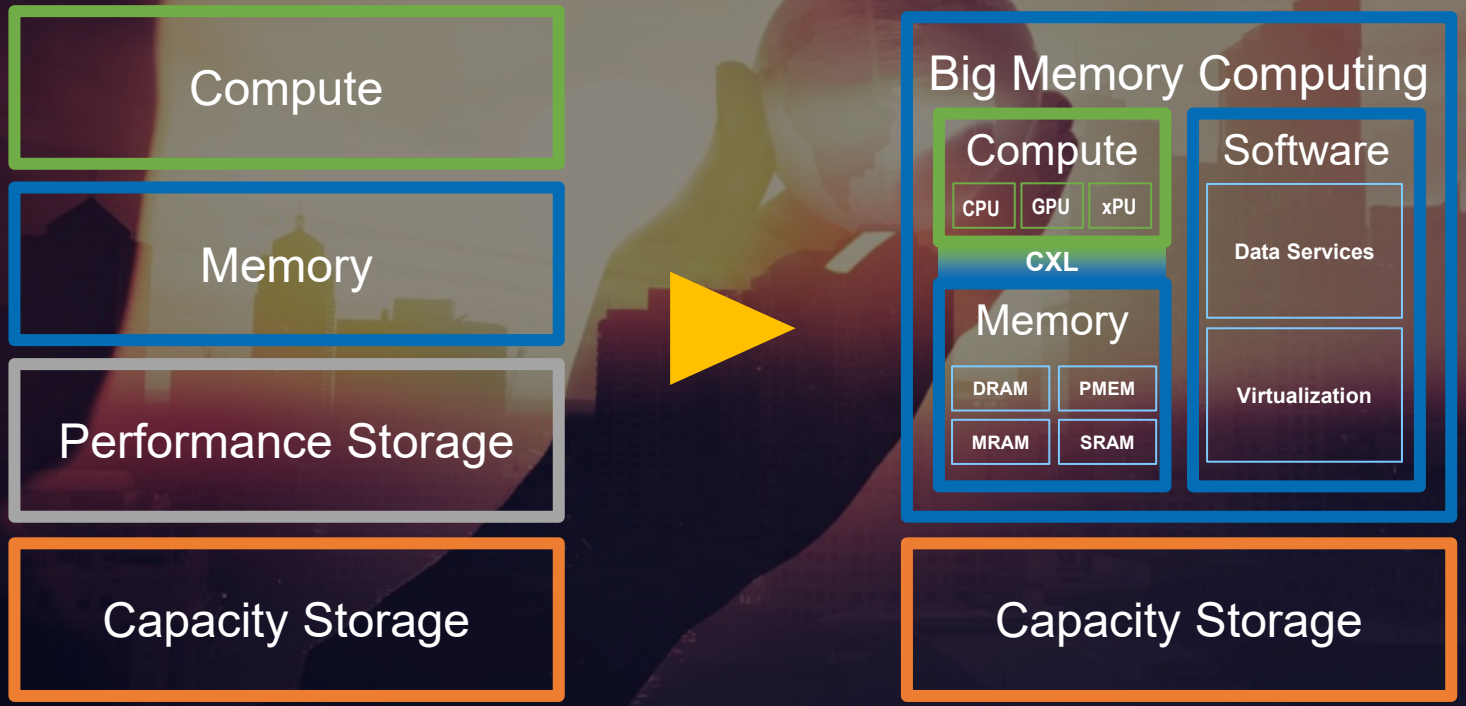
Scalability
Terabytes of DRAM-speed memory available to memory-intensive apps

Availability
ZeroIO™ Snapshot enables faster crash recovery

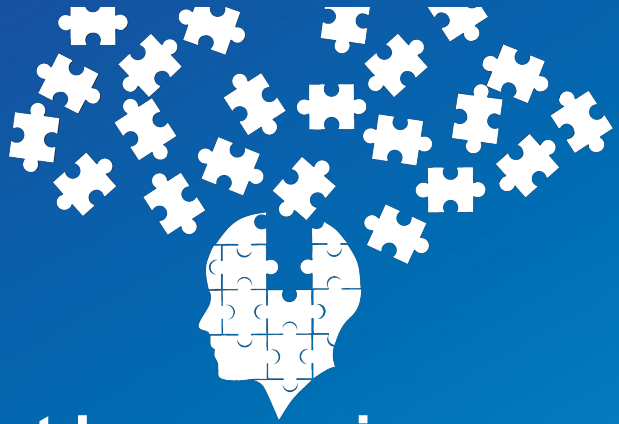
Agility
Instant Clones without duplicating of physical memory

Compatibility
No changes to application necessary

Big Memory Software Vision:
The New Software Platform for Next-Gen Computing



MemVerge



What happens in memory stays in memory…