VMware Memory Vision
for Real World Applications

Arvind Jagannath – Sr Product Line Manager
VMware by Broadcom

Sudhir Balasubramanian – Principal Oracle Architect
(TechBridge BCA Consulting)

MEMORY FABRIC FORUM

OCP GLOBAL SUMMIT

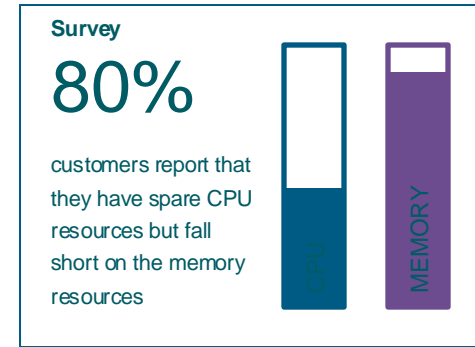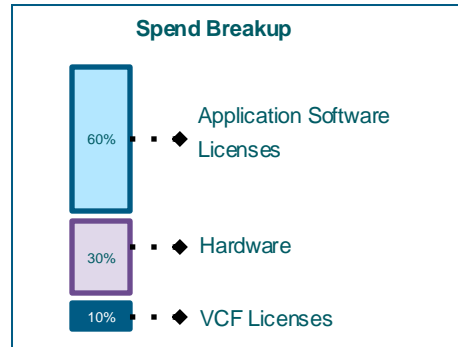OCT 15-17, 2024
SAN JOSE, CA

Disclaimer

- Certain information in this presentation may outline Broadcom's general product direction.

- This presentation shall not serve to (i) affect the rights and/or obligations of Broadcom or its licensees under any existing or future license agreement or services agreement relating to any Broadcom software product; or (ii) amend any product documentation or specifications for any Broadcom software product.

- This presentation is based on current information and resource allocations and is subject to change or withdrawal by Broadcom at any time without notice.

- The development, release and timing of any features or functionality described in this presentation remain at Broadcom's sole discretion.

- Notwithstanding anything in this presentation to the contrary, upon the general availability of any future Broadcom product release referenced in this presentation, Broadcom may make such release available to new licensees in the form of a regularly scheduled major product release.

- Such release may be made available to licensees of the product who are active subscribers to Broadcom maintenance and support, on a when and if-available basis.

- The information in this presentation is not deemed to be incorporated into any contract.

Imbalance between CPU and memory resources

Inefficient memory utilization leading to wastage of spare CPU capacity

Bill of Material	Value
Server Specification	Dell R760 2 Socket/2TB RAM
Server Price	$76865
VCF Per core license	$350
SQL Per Core Licenses	$2000
Total Software price per core	$2350
Cores per server	80
Total HW + SW Cost Per Server	$264865
Server Units	**100**
Total	**26 Million USD**

Spend Breakup

60% ◆ Application Software Licenses

30% ◆ Hardware

10% ◆ VCF Licenses

Survey

80%

customers report that they have spare CPU resources but fall short on the memory resources
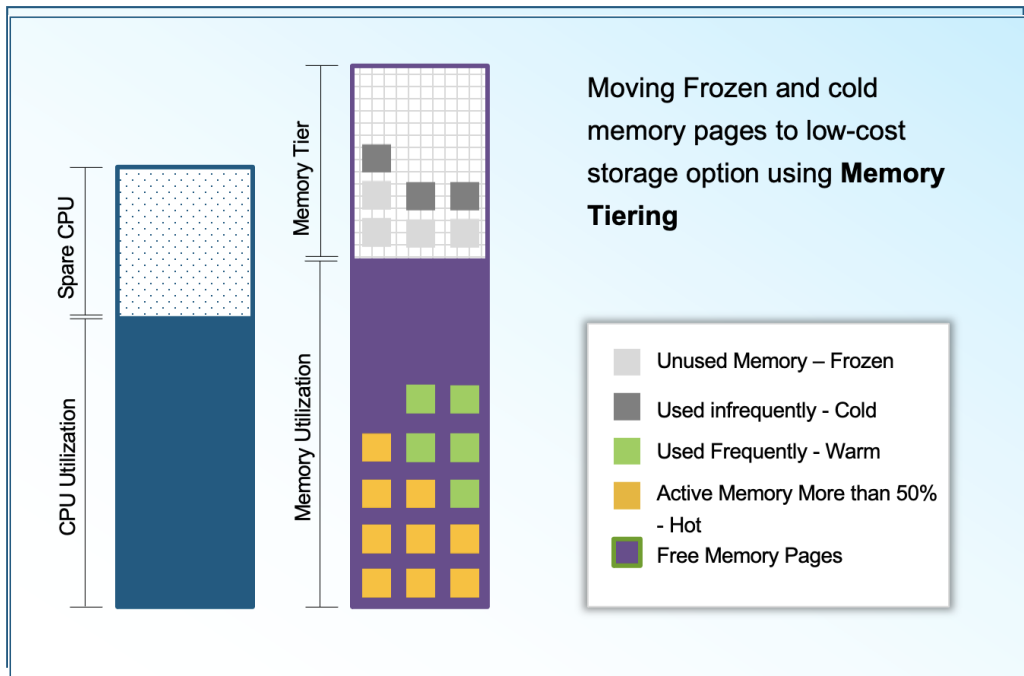
CPU MEMORY

Challenges

- Software licenses are based on per core and contribute significantly to customers IT spendings
- 80% customers reports that they have spare CPU cores that can not be utilized because of unavailability of memory resources
- The typical response to such situation is adding more DRAM but its not ideal solution due to limiting factors such as cost and limited DIMM slots on hardware

Introducing Memory Tiering

Introducing memory tiering with page classification

Moving Frozen and cold memory pages to low-cost storage option using **Memory Tiering**

Legend:
- Unused Memory – Frozen
- Used infrequently - Cold
- Used Frequently - Warm
- Active Memory More than 50% - Hot
- Free Memory Pages

NVMe Memory Tiering

1. Recommended for optimizing VDI, web services, and IT management workloads
2. Uses PCIe and software interface
3. 4:1 ration for DRAM to secondary Tier
4. Local memory tiering only

Peaberry accelerator

1. Recommended for optimizing memory intensive workloads such as databases, SAP, and Redis
2. CXL and hardware co-designed solution
3. Recommended 1:1 ration for DRAM to secondary Tier
4. Local memory tiering, remote tiering/pooling, vMotion, Offloads

Unlocking the Potential Value

Effective utilization of CPU cores unlock the potential

Bill of Material	Peaberry	NVMe
Total HW + SW Cost Per Server	$264865	$264865
Additional Cost of Storage tier	$5000 (+)	$1 /GB (+)
Number of cores unlocked per server	28	12
VCF license saving per host	$9800 (-)	$4200 (-)
SQL Per Core Licenses	$56000 (-)	$24000(-)
Total ROI per unit	$60800	$28200
Server Units	100	100
Total ROI	~$6 Million	~$2.8 Million

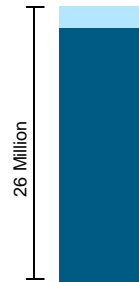Planned VCF 9U1

26 Million
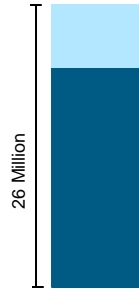
$ 35%

PEABERRY ACCELERATOR

By using CXL based accelerators, customers can unlock up to 35% of total cores

$6M

Worth of value unlocking using Peaberry Accelerator

$1M — VCF Licensing value unlocked

$5M — Software Licensing value unlocked

Tech Preview vSphere 8U3

26 Million

$ 15%

NVMe Based Memory Tiering

By using NVMe based memory tiering, customers can unlock up to 15% of total cores
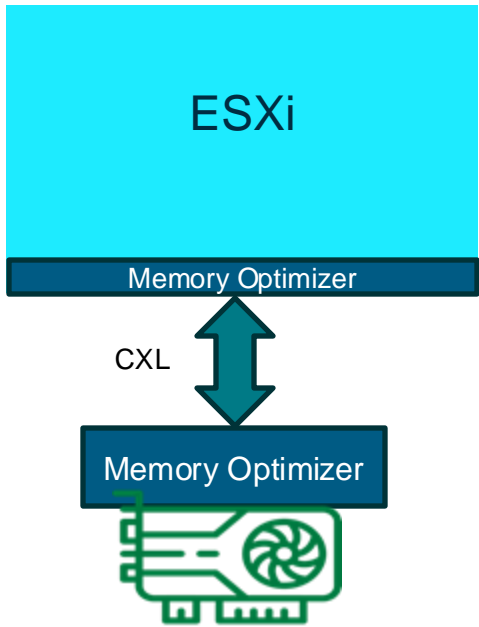
$2.8M

Worth of value unlocking using NVMe Memory Tiering

$.3M — VCF Licensing value unlocked

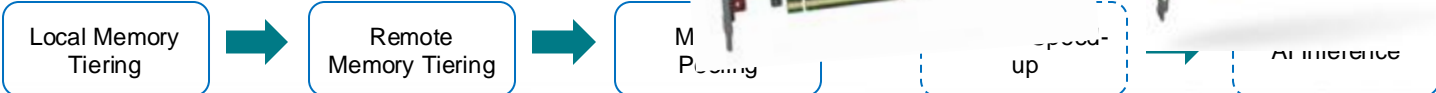$2M — Software Licensing value unlocked
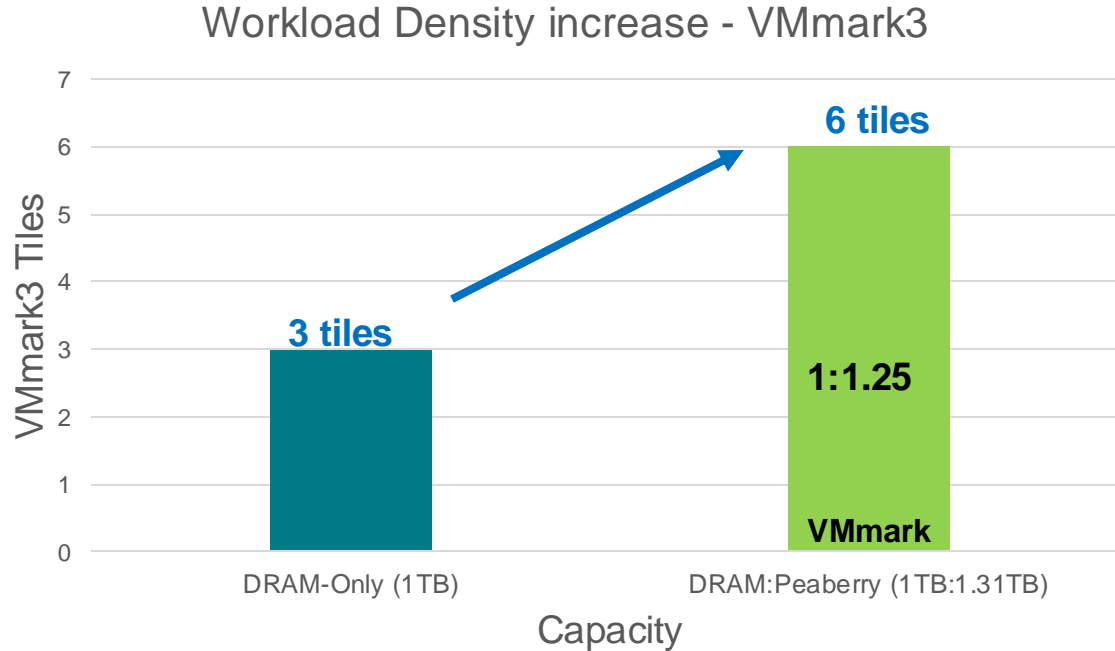
Project Peaberry - CXL-based Accelerator

GA Target VCF 9.1

ESXi

Memory Optimizer

CXL

Memory Optimizer

SAMSUNG

SAMSUNG CMM-H

1. CXL/Hardware Co-design
    1. 0 page-... ...cking
2. Low App...
3. Perform... ...ds
4. Auto-managed lifecycle
5. No add-on license
6. 4TB Capaci...

Inventec

BittWare
a molex company

altera
An Intel Company

Inventec 3680m CXL Memory Accelerator

| Local Memory Tiering | → | Remote Memory Tiering | → | M... P...ng | Speed-up | → | AI Inference |

Project Peaberry – Increase workload density

Workload Density increase - VMmark3
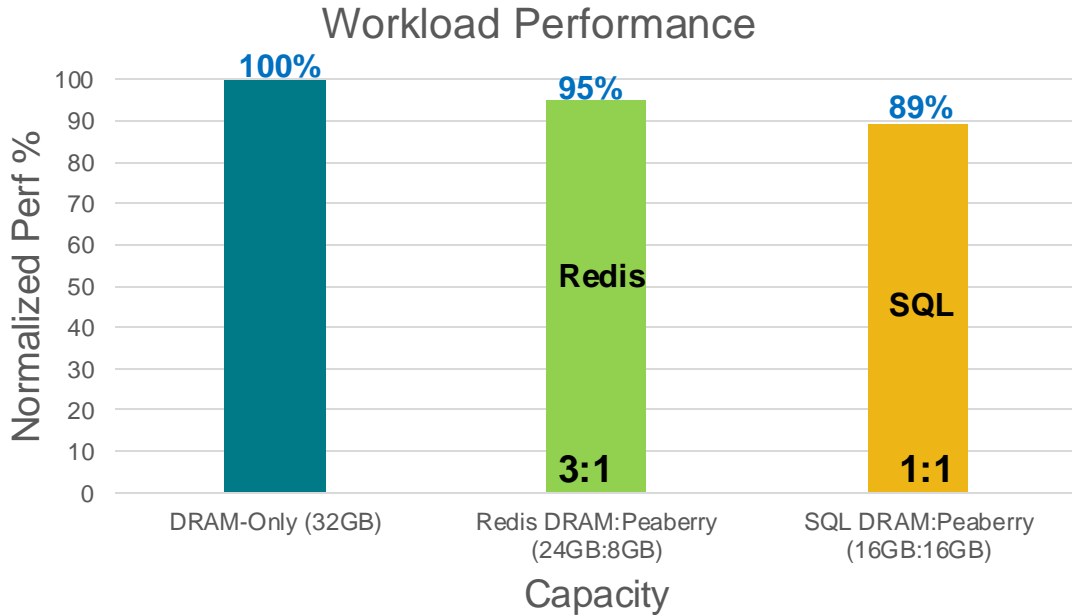

	Host
System under test (SUT)	DELL PowerEdge R750 with 2x Intel Xeon Platinum 8380 (40 cores per socket)
SUT: Memory DDR4 Total Host DRAM Capacity	1TB (16x64GB DIMMs)
SUT: CXL Accelerator Total Capacity	4TB PCIe card

100% more workload density at a 40:60 ratio with <5% perf impact

Project Peaberry - Redis and SQL with No Performance Impact

Workload Performance

Bar chart titled "Workload Performance" with Y-axis "Normalized Perf %" (0 to 100) and X-axis "Capacity":
- DRAM-Only (32GB): **100%**
- Redis DRAM:Peaberry (24GB:8GB): **95%** — Redis — 3:1
- SQL DRAM:Peaberry (16GB:16GB): **89%** — SQL — 1:1
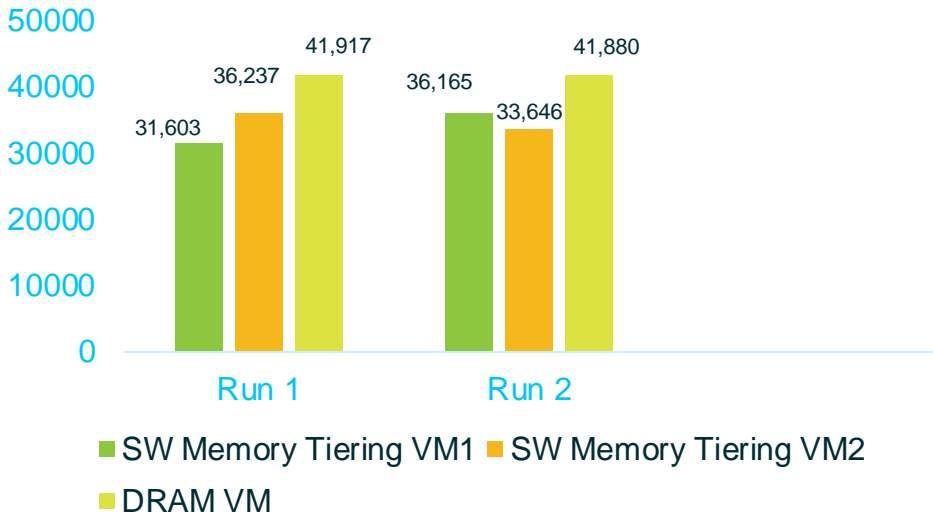
| Host	
System under test (SUT)	Intel EMR platform Windows VM (8vCPUs)
SUT: Memory DDR4 DRAM Total Capacity (VM Total Memory Capacity)	1TB (16x64GB DIMMs) VM Memory (32GB)
SUT: CXL accelerator (PCIe)	4TB Peaberry card

Both Redis and SQL show performance close to DRAM-Only system

Oracle on NVMe Memory Tiering – Run 2 VMs instead of 1

Executes (SQL) per second

Logical Reads (blocks) per second
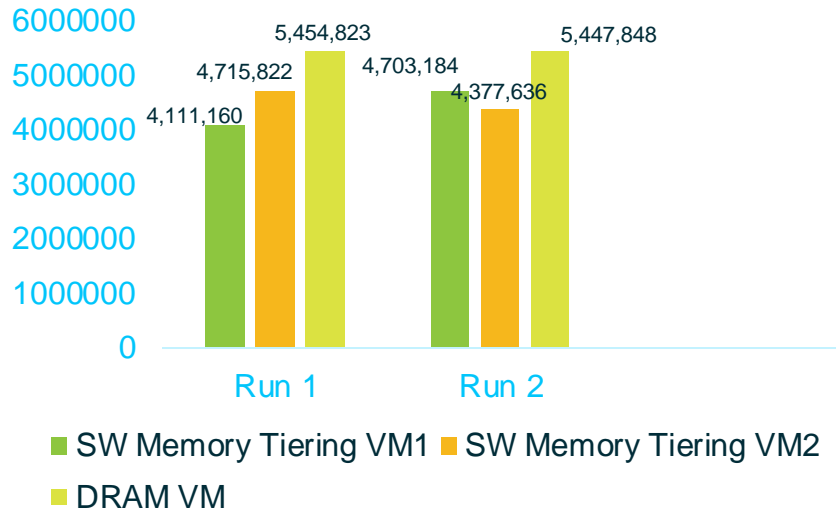
- **Load Generator chosen as SLOB 2.5.4.0**
  - UPDATE_PCT=0 - READ only test
  - RUN_TIME=1200 secs (20mins)
- **Test Results**
  - Executes(SQL) / second
  - Run 1
    - **Aggregate SW Tier VM1+VM2 = 69,841/sec V/S DRAM Mode VM** - 41,917/sec
  - Run 2
    - **Aggregate SW Tier VM1+VM2 = 69,811/sec V/S DRAM Mode VM** - 41,880/sec

- **Test Results**
  - Logical Reads (blocks) per second
  - Run 1
    - **Aggregate SW Mem Tier VM1 + VM2 = 8,826,982/sec**
    - **DRAM Mode VM - 5,454,823/sec**
  - Run 2
    - **Aggregate SW Mem Tier VM1 + VM2 = 9,080,820/sec**
    - **DRAM Mode VM - 5,447,848/sec**

Summary of preliminary workload performance with each solution

NVMe Tiering

1. LoginVSI (VDI Benchmark):
   - 1:1 ratio: Compared to DRAM baseline –
     - 100% Increase in VM density:128->256 VMs with 1-2% performance drop, 5-6 additional cores usage
2. VMmark (Mixed DB, In-mem DB, MySQL, Web services, varied VM sizes (2-256GB), vCPUs (2-8)):
   - 4:1 ratio: (4 DRAM & 1 NVMe) Compared to DRAM baseline –
     - 33% increase in tiles (Tile=Mixed bundle) 3 tiles -> 4 tiles. Score (Measure of performance) Not affected.
     - More workloads are being planned

Project Peaberry - CXL Accelerator

1. HammerDB with Oracle TPC-H profile (End-to-End solution with NVMe usage)
   - Running on a 48GB VM – (1:1 - DRAM:CXL Accelerator)
     - 90+% performance of DRAM
2. VMMark and LoginVSI 1:1 ratio – 2X density

Call to Action

Please contact Arvind.Jagannath@broadcom.com for more details on VMware memory tiering and Project Peaberry, and to learn more about VMware customers and pain-points

Oracle Collateral on VMware (or Oracle & Business Critical Applications Expertise) – Contact sudhirb@techbridgebca.com/sudhirbala@gmail.com

https://blogs.vmware.com/apps/2017/01/oracle-vmware-collateral-one-stop-shop.html

Thank you!