

# Providing Capacity & TCO to Applications using VMware Software Memory tiering

Presenters:

Sudhir Balasubramanian – Sr Staff Solution Architect, VMware

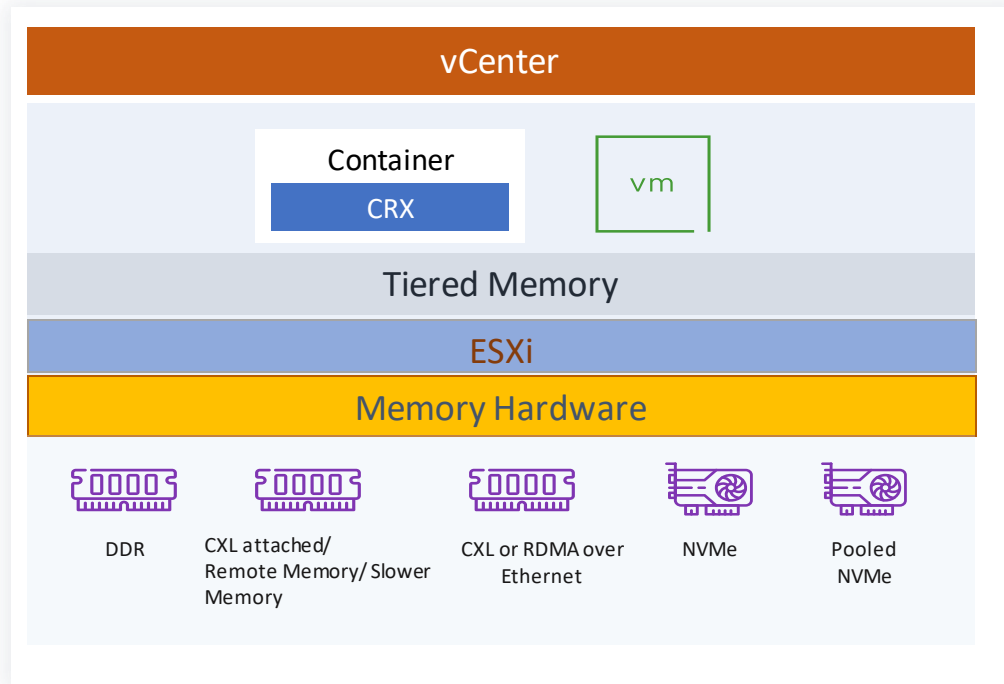
Arvind Jagannath - Sr Product Line Manager, Cloud Platform, VMware

# Agenda

- Memory Tiering and CPU-Memory imbalance
- CXL Use cases
- Deployment Options
- Use Case - Optimizing Oracle Workloads
- Key Takeaways

# VMware's Big Memory and Disaggregation Vision starts with CXL and Tiering

# VMware Memory Tiering



## Benefits

- Higher Density - core utilization
- Lower TCO
- Larger bandwidth

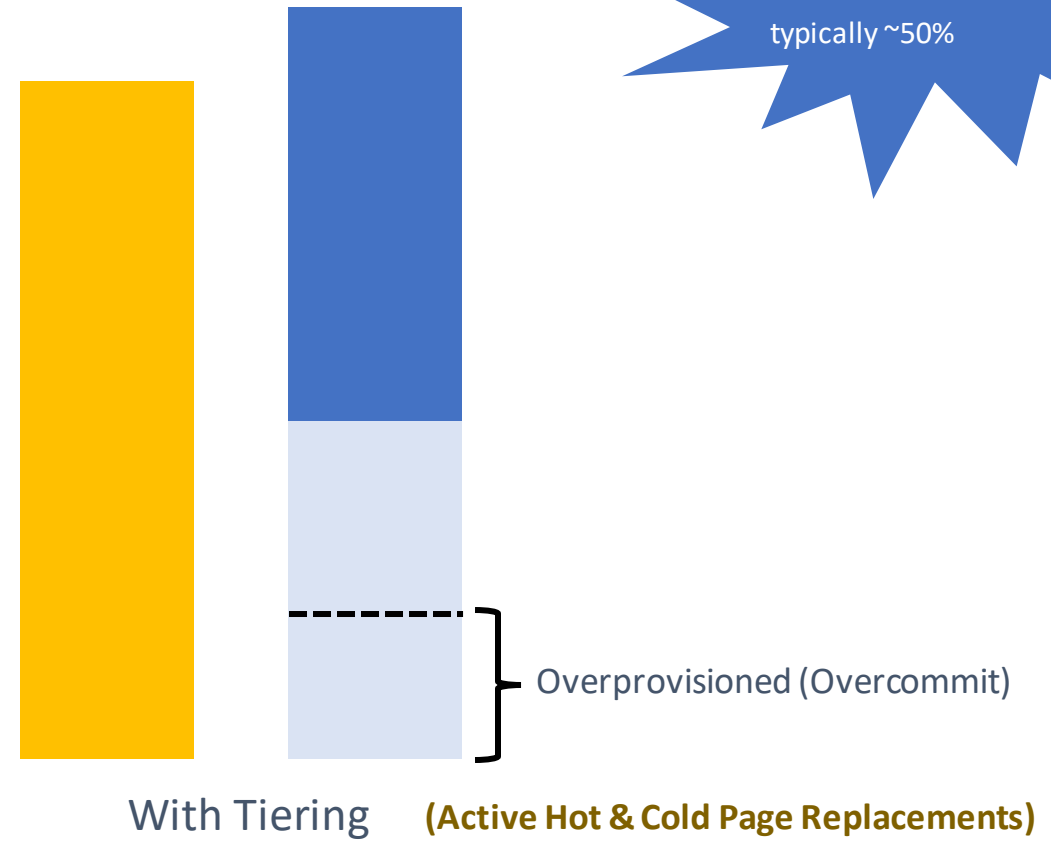
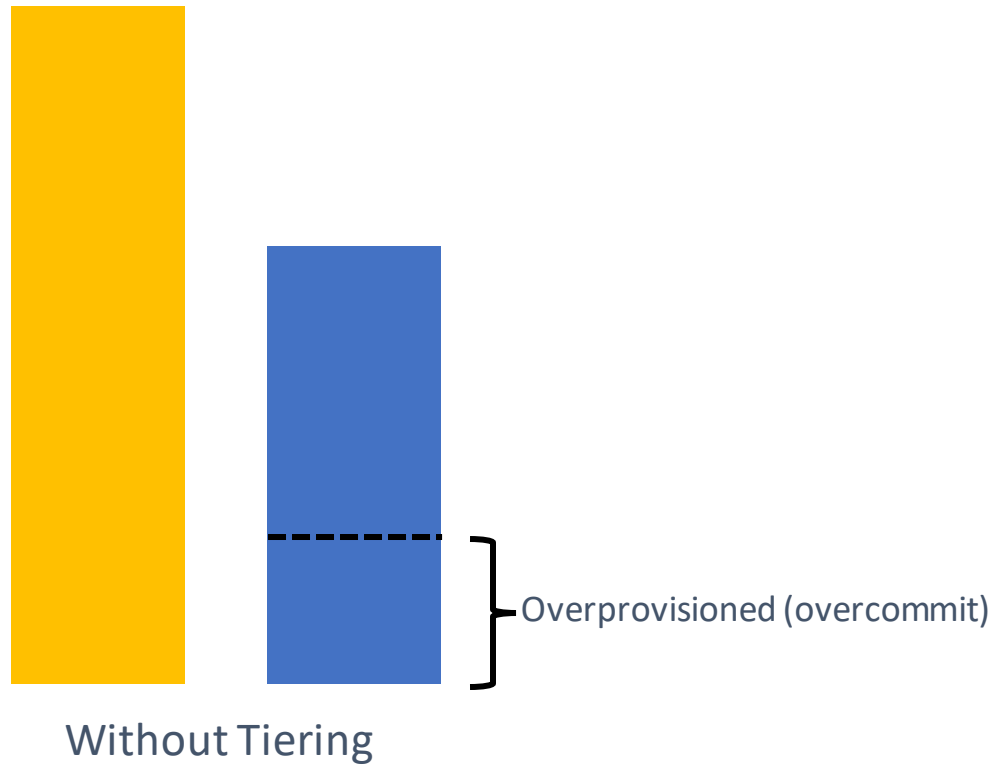
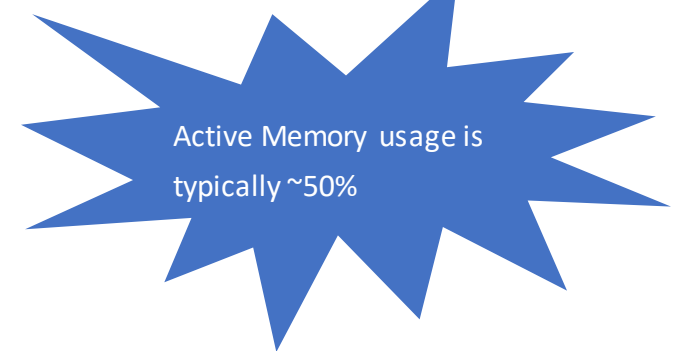
## Value over traditional approaches

- **Virtualization**
  - Independent underlying hardware changes
- **Transparent – Single volatile memory address**
  - No Guest or Application changes
    - Run any Operating System
  - ESX internally handles page placement
- **DRS and vMotion to mitigate risks**
  - Tiering/device heuristics fed to DRS
- **Ensure Fairness across workloads**
  - Consistent performance
- **Min Configuration changes**
  - No special tiering settings
- **Minimum Performance Degradation**
- **Processor specific monitoring**
  - vMMR monitors at both VM- and Host-levels

# Bridging the CPU-Memory imbalance



Flash Memory Summit



CPU



DRAM



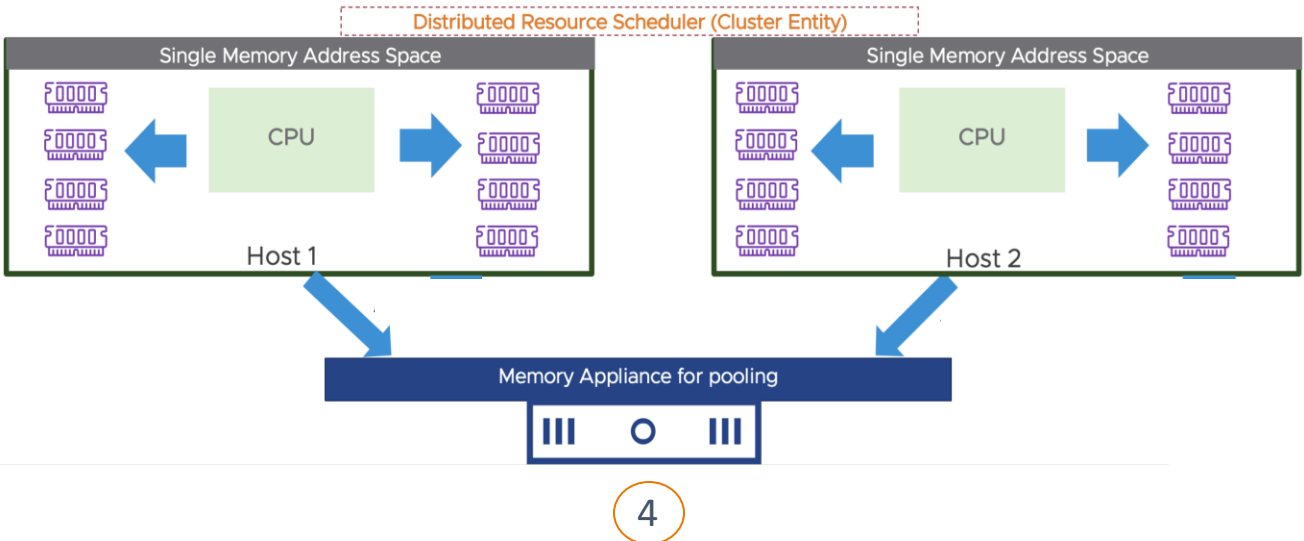
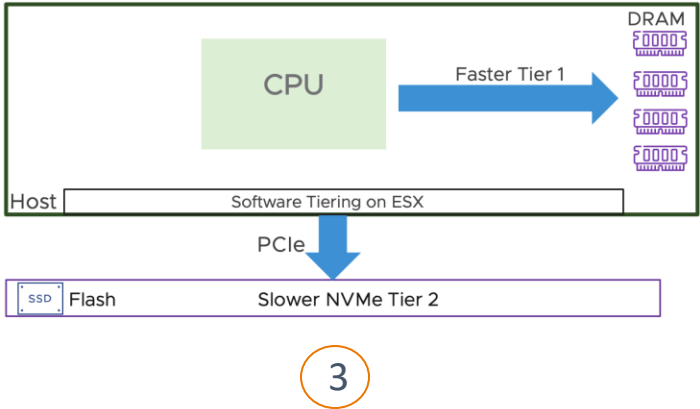
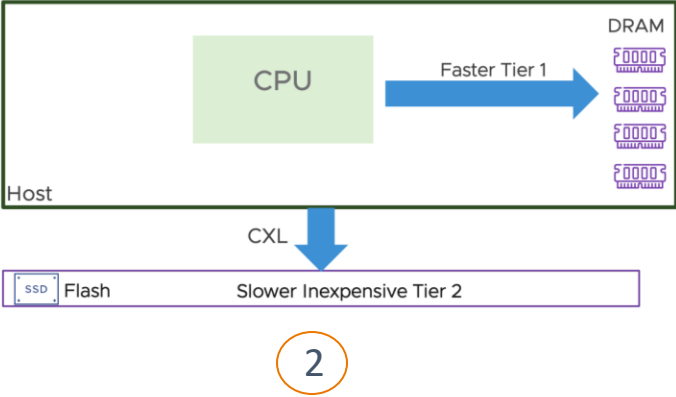
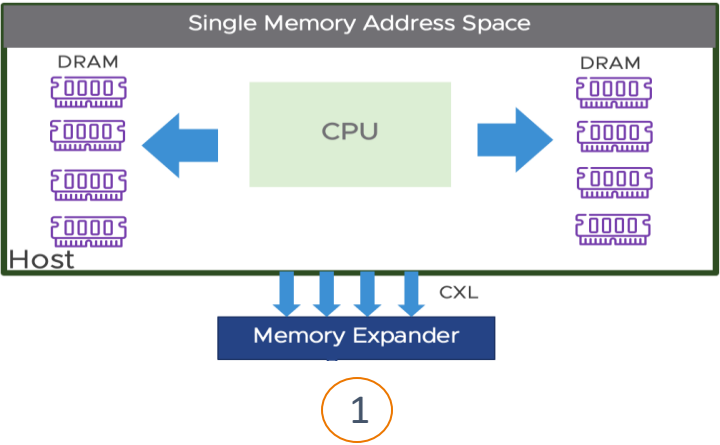
Tier-2



# Key Use Cases emerging with CXL

Memory Expansion with NUMA-like latencies	Memory Tiering	Memory pooling across hosts on a cluster using memory appliances	Memory sharing	CXL switching and shared access (future)
<ul style="list-style-type: none"><li>-Increase capacity/scale</li><li>-Flat (non-tiered) expansion</li><li>-Consolidate server memory</li><li>-Improve bandwidth</li><li>-Improve core utilization</li></ul>	Lower TCO – combinations of lower cost memory with DRAM	Consolidate memory usage on a cluster	Utilize stranded memory on hosts	Disaggregation and Composability

# Deployment Options



# CapEx and OpEx improvements with CXL



---

## Reduce CapEx

- Provide lower TCO by DRAM substitution
- Improve capacity and bandwidth
- Improve host CPU core utilization
- CPU Savings from page tracking/transfer offloads

---

## Reduce OpEx

- Improve OpEx by reducing migration times
- Reduce failure probability
- Reduce host evacuation time for maintenance

---

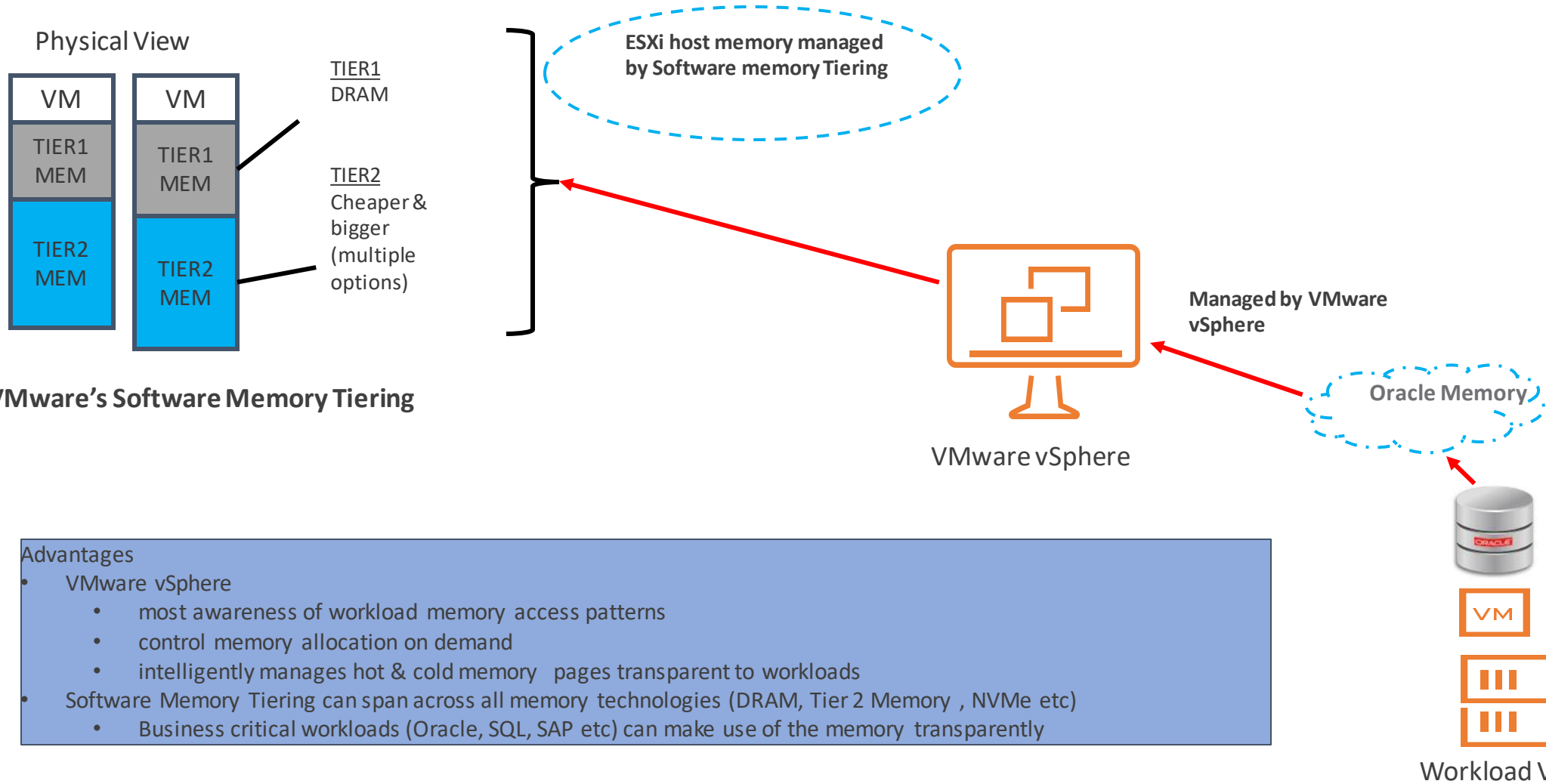
## Other future use-cases

- Pooling/sharing
- Remote mem tiering (stranded mem)
- Memory tracking/scrubbing & proactive maintenance
- App-specific acceleration (AI/ML, Analytics)
- Encryption/dedupe/compression)



# Use Case - Optimizing Oracle Workloads

# Oracle Workloads Using VMware Software Memory Tiering - Concept



# VMware Software Memory Tiering - ESXi Details



sc2esx64.vslab.local ACTIONS

Summary Monitor Configure Permissions VMs Resource Pools

### Host Details

**Hypervisor:** VMware ESXi

**Model:** SYS-2049U-TR4

**Processor Type:** Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz

**Logical Processors:** 192

**NICs:** 10

**Virtual Machines:** 2

**Memory Tiering:** Software

**State:** Connected

**Uptime:** 10 days

### Hardware

**CPU** 192 CPU(s) x Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz

**Memory** 4557.15 GB

**Virtual Flash Resource** 43.2 GB / 119.75 GB

**Networking** 10 Network(s)

**Storage** 5 Datastore(s)

## Software Memory Tiering Server

- 4 sockets, 24 cores / socket
- Total Memory 4.5 TB

- 1.5 TB DRAM
- 3 TB Tier 2 Memory

sc2esx65.vslab.local ACTIONS

Summary Monitor Configure Permissions VMs Resource Pools

### Host Details

**Hypervisor:** VMware ESXi

**Model:** SYS-2049U-TR4

**Processor Type:** Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz

**Logical Processors:** 192

**NICs:** 10

**Virtual Machines:** 1

**State:** Connected

**Uptime:** 18 days

### Hardware

**CPU** 192 CPU(s) x Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz

**Memory** 1534.66 GB

**Virtual Flash Resource** 43.19 GB / 119.75 GB

**Networking** 10 Network(s)

**Storage** 5 Datastore(s)

## DRAM Mode

- 4 sockets, 24 cores / socket
- Total Memory 1.5 TB

- 1.5 TB DRAM

# VMware Software Memory Tiering - NUMA and Memory Details



sc2esx64.vslab.local ACTIONS

Summary Monitor Configure Permissions VMs Resource Pools

Host Details

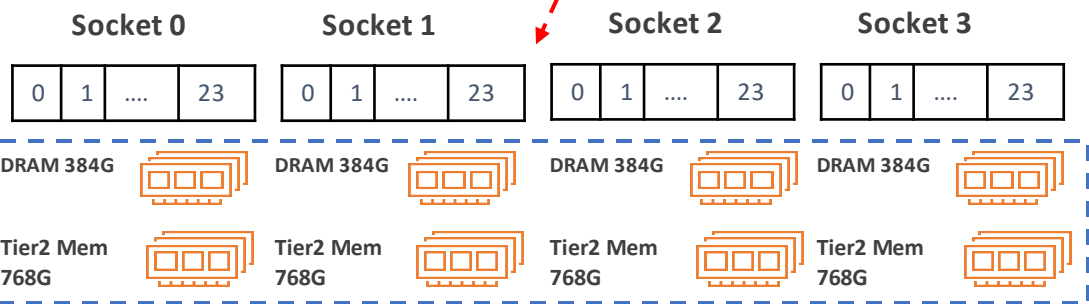
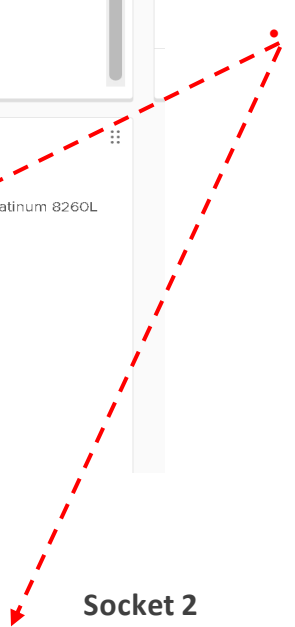
**Hypervisor:** VMware ESXi  
**Model:** SYS-2049U-TR4  
**Processor Type:** Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz  
**Logical Processors:** 192  
**NICs:** 10  
**Virtual Machines:** 2  
**Memory Tiering:** Software  
**State:** Connected  
**Uptime:** 10 days

Hardware

**CPU** 192 CPU(s) x Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz  
**Memory** 4557.15 GB  
**Virtual Flash Resource** 43.2 GB / 119.75 GB  
**Networking** 10 Network(s)  
**Storage** 5 Datastore(s)

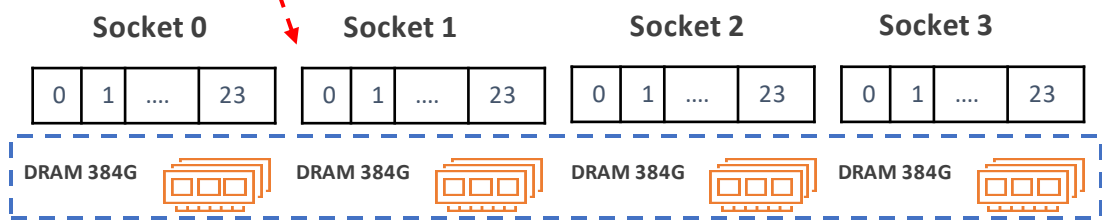
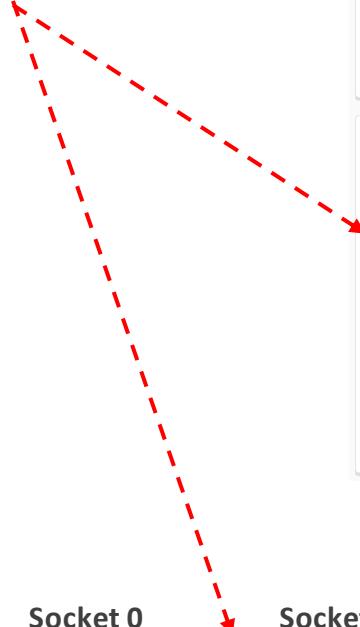
## Software Memory Tiering Server

- **Total Memory 4.5 TB**
  - 1.5 TB DRAM
  - 3 TB Tier 2 Memory



## DRAM Mode Server

- **Total Memory 1.5 TB**
  - 1.5 TB DRAM



sc2esx65.vslab.local ACTIONS

Summary Monitor Configure Permissions VMs Resource Pools

Host Details

**Hypervisor:** VMware ESXi  
**Model:** SYS-2049U-TR4  
**Processor Type:** Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz  
**Logical Processors:** 192  
**NICs:** 10  
**Virtual Machines:** 1  
**State:** Connected  
**Uptime:** 18 days

Hardware

**CPU** 192 CPU(s) x Intel(R) Xeon(R) Platinum 8260L CPU @ 2.40GHz  
**Memory** 1534.66 GB  
**Virtual Flash Resource** 43.19 GB / 119.75 GB  
**Networking** 10 Network(s)  
**Storage** 5 Datastore(s)

mmit

# VMware Software Memory Tiering & DRAM only VM Details



Oracle21C-OL8-Capitola

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.64

Encryption

Not encrypted

Guest OS

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

SMT1

VM Hardware

CPU12 CPU(s), 622 MHz used

Memory256 GB, 3 GB memory active

Hard disk 1 (of 7)80 GB | Thin Provision

Network adapter 1APPS-1614 (connected) | 00:50:56:ad:82:76

CD/DVD drive 1Disconnected

CompatibilityESXi 7.0 U2 and later (VM version 19)

EDIT

Related Objects

Hostsc2esx64.vslab.local

NetworksAPPS-1614

StorageSC2-Pure-Oracle

Oracle21C-OL8-Capitola2

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.164

Encryption

Not encrypted

Guest OS

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

SMT2

VM Hardware

CPU12 CPU(s), 622 MHz used

Memory256 GB, 3 GB memory active

Hard disk 1 (of 7)80 GB | Thin Provision

Network adapter 1APPS-1614 (connected) | 00:50:56:80:a0:23

CD/DVD drive 1Disconnected

CompatibilityESXi 7.0 U2 and later (VM version 19)

EDIT

Related Objects

Hostsc2esx64.vslab.local

NetworksAPPS-1614

StorageSC2-Pure-Oracle

Oracle21C-OL8-DRAM

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.65

Encryption

Not encrypted

Guest OS

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

DRAM Only

VM Hardware

CPU12 CPU(s), 406 MHz used

Memory256 GB, 3 GB memory active

Hard disk 1 (of 7)80 GB | Thin Provision

Network adapter 1APPS-1614 (connected) | 00:50:56:80:4a:3f

CD/DVD drive 1Disconnected

CompatibilityESXi 7.0 U2 and later (VM version 19)

EDIT

Related Objects

Hostsc2esx65.vslab.local

NetworksAPPS-1614

StorageSC2-Pure-Oracle

Software Memory Tiering VM's – SMT1 and SMT2

DRAM Only VM

# Oracle Database – Details



- OEL 8.5 UEK ,Oracle 21.5
- Oracle Standalone DB , ASM & ASMLIB
- Oracle on VMware Best Practices Followed

Oracle Memory



- VM
  - vCPUS=12, RAM = 256G
- Oracle
  - SGA=192G, PGA=30G

Oracle Memory



- VM
  - vCPUS=12, RAM = 256G
- Oracle
  - SGA=192G, PGA=30G

Oracle Memory



- VM
  - vCPUS=12, RAM = 256G
- Oracle
  - SGA=192G, PGA=30G

Oracle21C-OL8-SMT1

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.64

Encryption

Not encrypted

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

Guest OS

VM Hardware

CPU

12 CPU(s), 622 MHz used

Memory

256 GB, 3 GB memory active

Hard disk 1 (of 7)

80 GB | Thin Provision

Network adapter 1

APPS-1614 (connected) | 00:50:56:ad:82:76

CD/DVD drive 1

Disconnected

Compatibility

ESXi 7.0 U2 and later (VM version 19)

Related Objects

Host

sc2esx64.vslab.local

Networks

APPS-1614

Storage

SC2-Pure-Oracle

SW Memory Tiering VM – SMT1

Oracle21C-OL8-SMT2

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.164

Encryption

Not encrypted

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

Guest OS

VM Hardware

CPU

12 CPU(s), 622 MHz used

Memory

256 GB, 3 GB memory active

Hard disk 1 (of 7)

80 GB | Thin Provision

Network adapter 1

APPS-1614 (connected) | 00:50:56:80:a0:23

CD/DVD drive 1

Disconnected

Compatibility

ESXi 7.0 U2 and later (VM version 19)

Related Objects

Host

sc2esx64.vslab.local

Networks

APPS-1614

Storage

SC2-Pure-Oracle

SW Memory Tiering VM – SMT2

Oracle21C-OL8-DRAM

Summary Monitor Configure Permissions Datastores Networks Snapshots Updates

Virtual Machine Details

Power Status

Powered On

Guest OS

Oracle Linux 8 (64-bit)

VMware Tools

Running, version:11333 (Guest Managed)

DNS Name (1)

oracle21c-ol8.vslab.local

IP Addresses (1)

172.16.14.65

Encryption

Not encrypted

LAUNCH REMOTE CONSOLE

LAUNCH WEB CONSOLE

Guest OS

VM Hardware

CPU

12 CPU(s), 406 MHz used

Memory

256 GB, 3 GB memory active

Hard disk 1 (of 7)

80 GB | Thin Provision

Network adapter 1

APPS-1614 (connected) | 00:50:56:80:4a:3f

CD/DVD drive 1

Disconnected

Compatibility

ESXi 7.0 U2 and later (VM version 19)

Related Objects

Host

sc2esx65.vslab.local

Networks

APPS-1614

Storage

SC2-Pure-Oracle

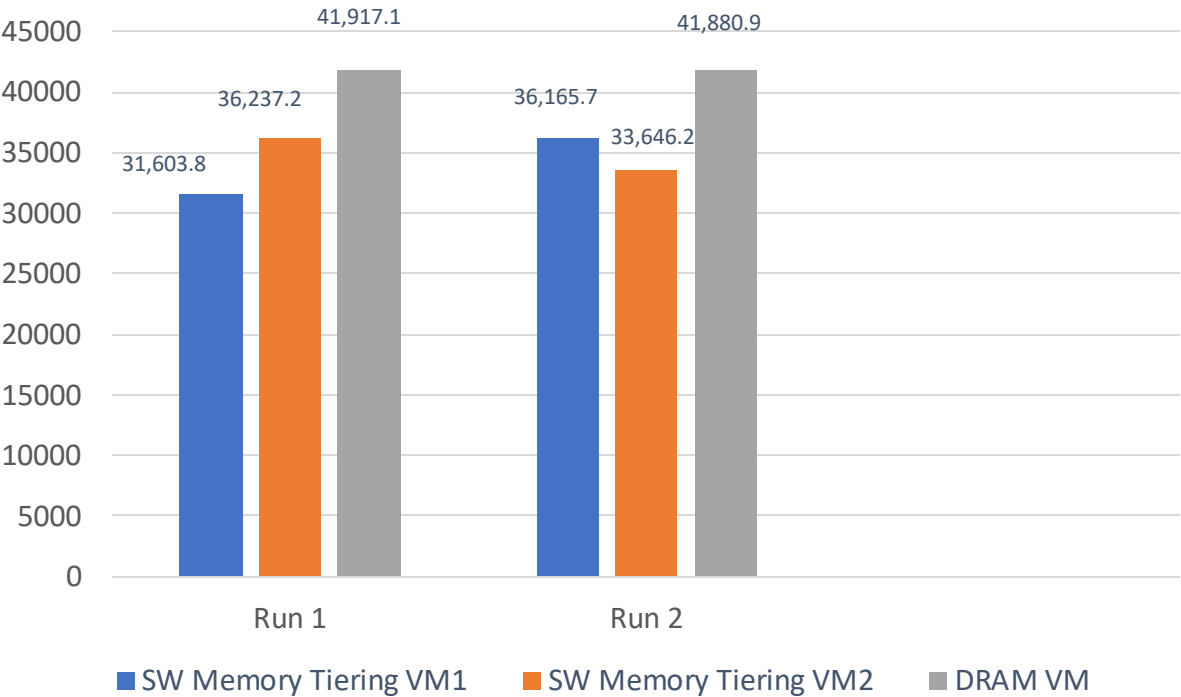
DRAM Mode VM1

Goal – Run ‘2 SW Memory Tiering’ VM’s on SMT Server on 1 NUMA node v/s ‘1 DRAM VM’ on DRAM only Server on 1 NUMA node – Can we double our workload performance with lower TCO ?

# Oracle Workload on SW Memory Tiering & DRAM Mode - Metrics

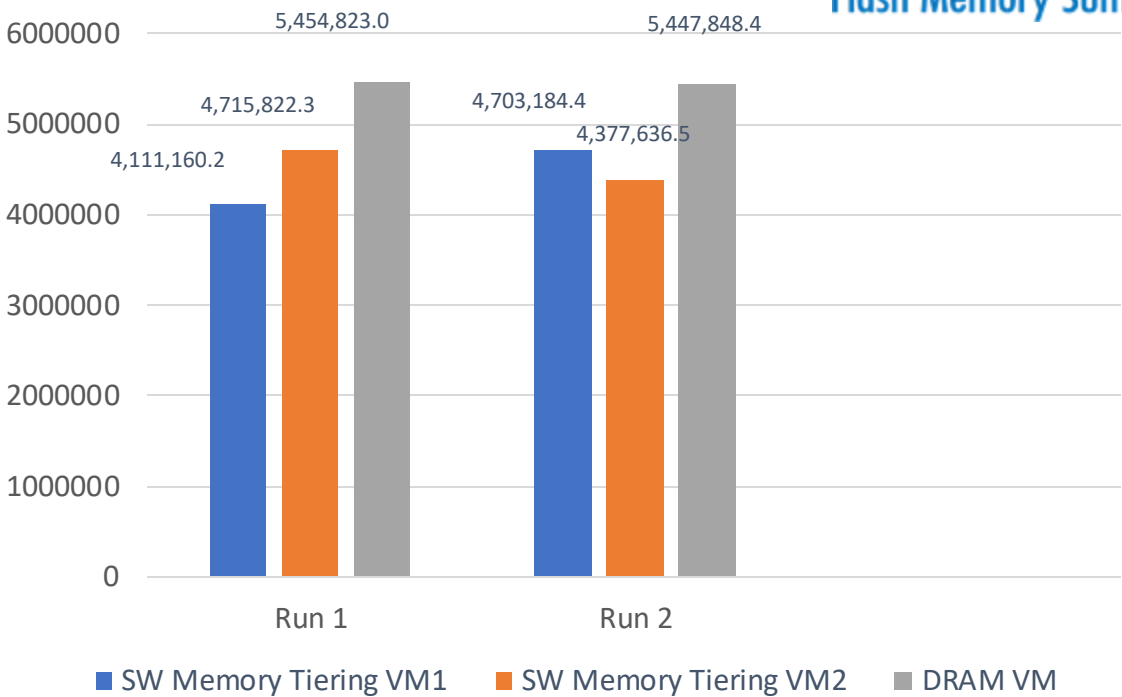


Executes (SQL) per second



- Load Generator chosen as SLOB 2.5.4.0
  - UPDATE\_PCT=0 - READ only test - performance comparison between SW Memory Tiering v/s DRAM Mode
  - RUN\_TIME=1200 secs (20mins)
- Test Results
  - Executes(SQL) / second
    - Run 1
      - Aggregate SW Tier VM1+VM2 = 69,841/sec
      - DRAM Mode VM - 41,917.1/sec
    - Run 2
      - Aggregate SW Tier VM1+VM2 = 69,811.9/sec
      - DRAM Mode VM - 41,880.9/sec

Logical Reads (blocks) per second



- Test Results
  - Logical Reads (blocks) per second
    - Run 1
      - Aggregate SW Mem Tier VM1 + VM2 = 8,826,982.5/sec
      - DRAM Mode VM - 5,454,823.0/sec
    - Run 2
      - Aggregate SW Mem Tier VM1 + VM2 = 9,080,820.9/sec
      - DRAM Mode VM - 5,447,848.4/sec

# SW Memory Tiering & DRAM Mode – OS & Oracle Metrics - Summary



Run	Metric	SW Mem Tier VM1	SW Mem Tier VM2	SW Mem Tier VM Aggregate	SW Mem Tier VM Average	DRAM VM	Difference (%)
Run 1	Executes(SQL) / second	33,603.80	36,237.20	69,841.00	34920.5	41,917.10	16.69
Run 2	Executes(SQL) / second	36,165.70	33,646.20	69,811.90	34,905.95	41,880.90	16.65
Run 1	Logical Reads (blocks) per second	4,111,160.20	4,715,822.30	8,826,982.50	4,413,491.25	5,454,823.00	19.09
Run 2	Logical Reads (blocks) per second	4,703,184.40	4,377,636.50	9,080,820.90	4,540,410.45	5,447,848.40	16.66

Run 1

SW Mem Tier VM1

CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.16	12.50	75.7	2.3	0.0	21.2

SW Mem Tier VM2

CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.15	413.04	86.8	2.5	0.0	9.8

DRAM Mem Tier VM1

CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.09	141.57	88.0	2.5	0.0	8.8

Run 2

SW Mem Tier VM1

CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.25	312.22	86.6	2.6	0.0	10.0

SW Mem Tier VM2

CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.19	60.59	80.8	2.3	0.0	16.1

DRAM Mem Tier VM1

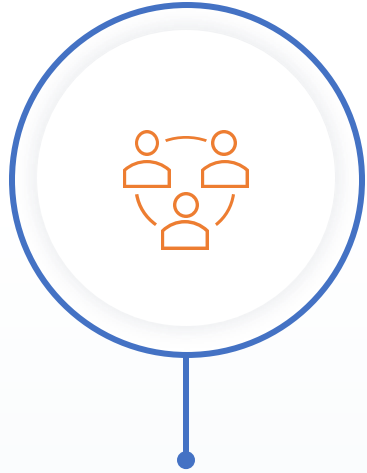
CPUs	Cores	Sockets	Load Average Begin	Load Average End	%User	%System	%WIO	%Idle
12	12	12	0.22	234.71	87.9	2.5	0.0	8.8

More %Idle in case of SW Memory Tiering VM's – Better CPU Utilization

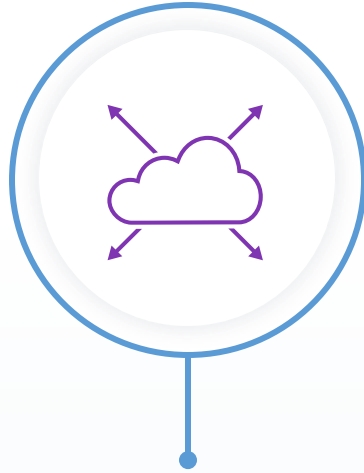
We were able to run '2 SW Memory Tiering' VM's on SMT Server on 1 NUMA node as compared to '1 DRAM VM' on DRAM Server on 1 NUMA node, even with the current VM size and DRAM capacity constraints , with increased workload performance with lower TCO as compared to DRAM only mode - VMware 'Software Memory Tiering' capability



# Summary



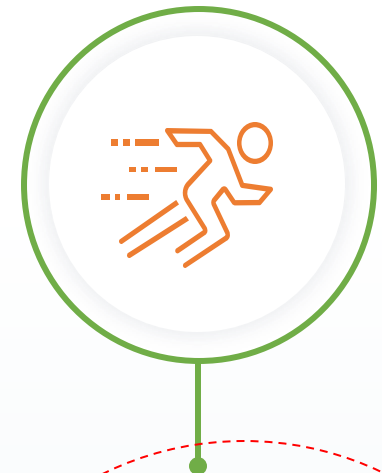
VMware has been collaborating with CXL partners to bring tiering to address cost, scale, and density challenges



Business transformation is leading to larger datasets and real-time analytics that requires more performance and larger memory capacity



VMware performance results show software tiering is ready for the next phase of the big memory evolution. Applications benefit from being in memory. Mission critical applications like Oracle also benefit from such innovations



Software Memory Tiering will bring scale without adding any operational complexity. Software tiering is also ready for future technologies like CXL that can bring pooling and disaggregation

# Oracle on VMware Collateral – One Stop Shop



All Oracle on vSphere white papers including Oracle on VMware Hybrid Multi-Clouds (vSphere / vSAN / vVols / VMware Clouds) Best practices, Deployment guides, Workload characterization guide can all be found in the URL below

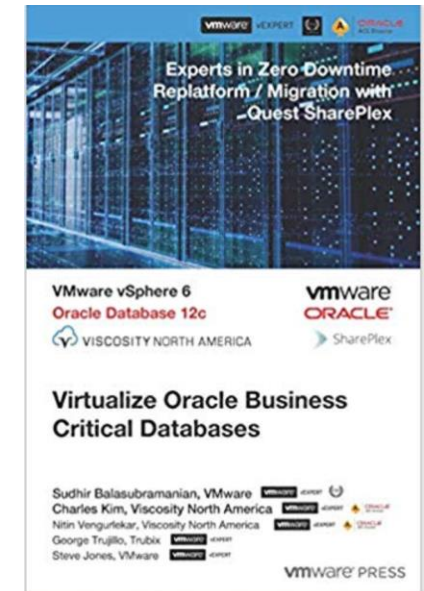
Oracle on VMware Collateral – One Stop Shop

<https://via.vmw.com/Oracle-on-VMware>

# Sudhir Balasubramanian



- **27 + years Oracle hands on experience** - Principal Oracle DBA / Architect, Oracle RAC/Data Guard Expert, Experienced in EMC SAN Technologies
- Currently - VMware [2012-] Senior Staff Solution Architect & Global Oracle Practice Lead
- Principal Oracle DBA/Oracle Architect (1995 – 2011) [ Tata Consultancy Services (TCS), Sony Electronics, Newgen Results (Aspen) ,Teletech Corp, SAIC, Active Network, Sempra Energy Holdings]
- VMware VCA – Cloud ,VMware vBCA Specialist
- VMware vExpert & vExpert Application Modernization - <https://vexpert.vmware.com/directory/1038>
- Member of the Office of the Chief Technical Ambassador VMware (Alumni)
- Oracle ACE - <https://apex.oracle.com/pls/apex/f?p=19297:3>
- Leading Author of “Virtualizing Oracle Business Critical Databases on VMware SDDC”
- Recognized Speaker@ Oracle Open World, IOUG, VMworld, VMware Partner Exchange, EMC World, EMC Oracle Summit and Webinars
- Industry recognized expert in Oracle Virtualization technologies
- Blogs
  - <http://vracdba.com/> | <https://community.oracle.com/blogs/sudhirb>
  - <https://blogs.vmware.com/apps/author/sudhirbalasubramanian/>
- Twitter : @vracdba [ <https://twitter.com/vRacDbA> ]
- LinkedIn : <https://www.linkedin.com/in/sudhirbalasubramanian/>



Arvind Jagannath

Sr. Product Line Manager for vSphere at VMware



Arvind Jagannath works in Product Management at VMware. With over 25 years of experience in the industry working on memory, networking, storage, embedded, and kernel development, he currently leads infrastructure and core platform enablement for vSphere, working across the VMware ecosystem of server/CPU, IO, and storage partners. Arvind most recently drove platform product management at Cohesity and NetApp. Arvind holds an MBA from the University of Chicago, Booth school of Business and a Bachelors in Computer Science and Engineering from India.

# Thank you