ORACLE
for Research

TECH TALK:

Oracle Cloud Foundations for Researchers

Friday, October 2, 2020
10:300 AM US EDT

Rajib Ghosh
Global Solutions Architect
Oracle for Research

ORACLE
for Research

TECH TALK HOUSEKEEPING

- Today's webinar is being recorded. We will share the link to the recording with you via email after the event. The recording will also be made available to the Oracle for Research community.
- We invite your comments and questions, both about the tech topic being discussed and about the series more generally. Questions may be submitted using the Q&A box on your screen or you may ask questions directly using your microphone. When not asking a question, please mute your microphone.
- Questions may be asked during the presentation and we will also have a Q & A time at the end of the presentation when you can ask questions directly and engage in discussion.
- At Oracle for Research, we believe that research and innovation happen best when a diverse and thoughtful community is free to engage in respectful, compassionate, and open dialog. To that end, when asking a question or providing feedback, we ask that all participants be respectful, collaborative, and constructive.

Oracle for Research Tech Talk Series Will Cover:

Technology training

Reference architectures
Best practices
Tools and automation
Cost control

Collaboration

Q & A Live discussions
Tips and tricks library
Community forum
Meet Oracle experts

Product announcements

OFR technology updates
OCI product updates
Images and containers
Public data

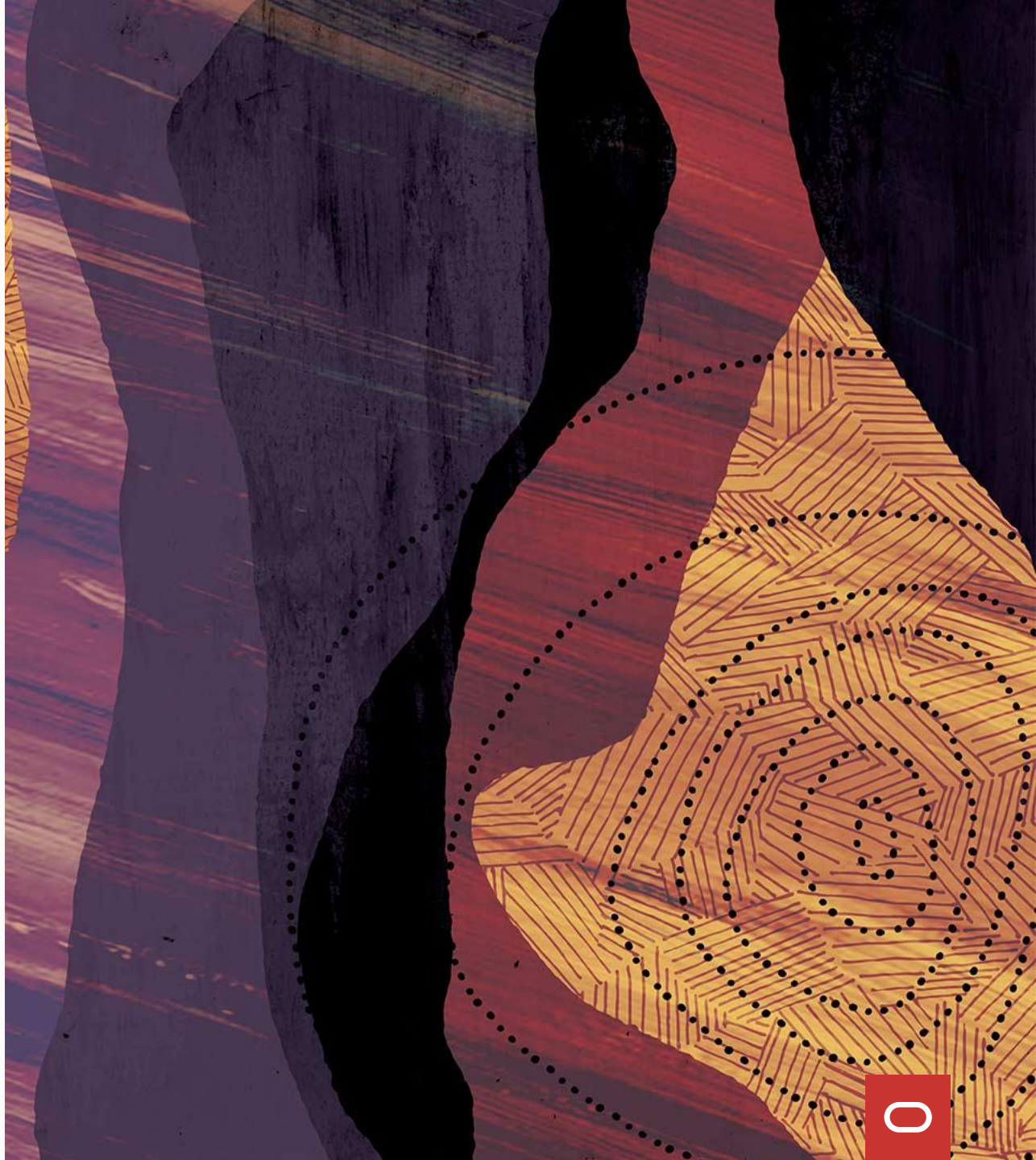
Researcher for researchers

Technology Innovations
Researcher publications
Benchmarks
Lessons learned

Oracle for Research Tech Talk Oracle Cloud Topics

Foundational	Oracle cloud – Getting you started and running Cloud instances and cloud storage options Migrating data and running computations
Architecture	Reference architecture patterns for researchers New features updates and recommended practices Performance benchmarks and data
Tools and automation	Tool selection, version and guidance Image repositories, Terraform and interfaces
HPC and cluster	High performance computing, workload classification, parallelization Cluster setup, utilization and monitoring
Machine and Deep Learning	Model selection guidance
Researcher guidance	Functional and data guidance and curation support Industry models with research computing

Agenda for today



Agenda (Foundational)

Common Researcher Issues	1. Where to start? 2. What shapes, images and storages do I need? 3. How to manage my credit allocation effectively?	1. Any automation to terminate idle instances? 2. How do I scale or burst my workload? 3. Can I use my on-campus identity to login?
OCI Architectures for Researchers	1. Standard researcher architecture 2. Shape and storage selection guidelines 3. Scaling and cloud bursting architectures	
Automation and usage control	Resource stacks and Terraform (demo) Tooling with OCI CLI Instance scaling (demo)	
Credit control mechanisms	Cost analysis and cost reports Setting budgets and alerts	
Github and links	Overview of the repositories Collaboration with Oracle and Researchers	
Q & A	Q & A What works best and researcher wish list?	

Oracle cloud image sources for Researchers

Platform Images

Operating system images supported by the Oracle cloud platform

- **Linux**
  - Oracle Linux 7.x /8.x
  - Centos 6.x / 7.x
  - Ubuntu 16.x/18.x/20.x
- **Windows**
  - Windows 2012 R2 DC
  - Windows 2012 R2 STD
  - Windows 2016 R2 DC
  - Windows 2016 R2 STD
  - Windows 2019 DC
  - Windows 2019 STD

Oracle Images

Oracle pre-built images tested and optimized for Oracle cloud and are available for use today.

- **Data Science, AI & GPU**
  - AI (All in one) Data science image
  - Julia AI HPC GPU Image
  - NVIDIA GPU cloud machine image
- **HPC**
  - OL7 HPC Cluster Networking
  - OL7 BeeGFS parallel filesystem
  - OL7 DBRD,Corosync & Pacemaker
  - NVIDIA GPU cloud machine image
- **Research**
  - Genome analysis toolkit
  - Folding@Home GPU
- **Infrastructure**
  - OL Storage Appliance
  - OL8.1 ARM
  - Oracle secure global desktop

Marketplace & github Images

Oracle and third-party images available through Oracle marketplace and github today

- **Data Science, AI & GPU**
  - AI (All in one) Data science image
  - AI/ML Runbook
  - GPU Workshop
  - Driverless AI
- **HPC**
  - OL7 HPC Cluster with Slurm
  - Lustre and GlusterFS Quickstart
  - OCI HPC Filesystem
  - BeeGFS-beeond RDMA
  - PBS Professional
  - Spectrum scale
- **Molecular dynamics**
  - NAMD Runbook
  - GROMACS Runbook
- **Infrastructure**
  - OCI learning library
  - OCI Terraform base
  - Terraform Jupyter GPU build

Custom Images

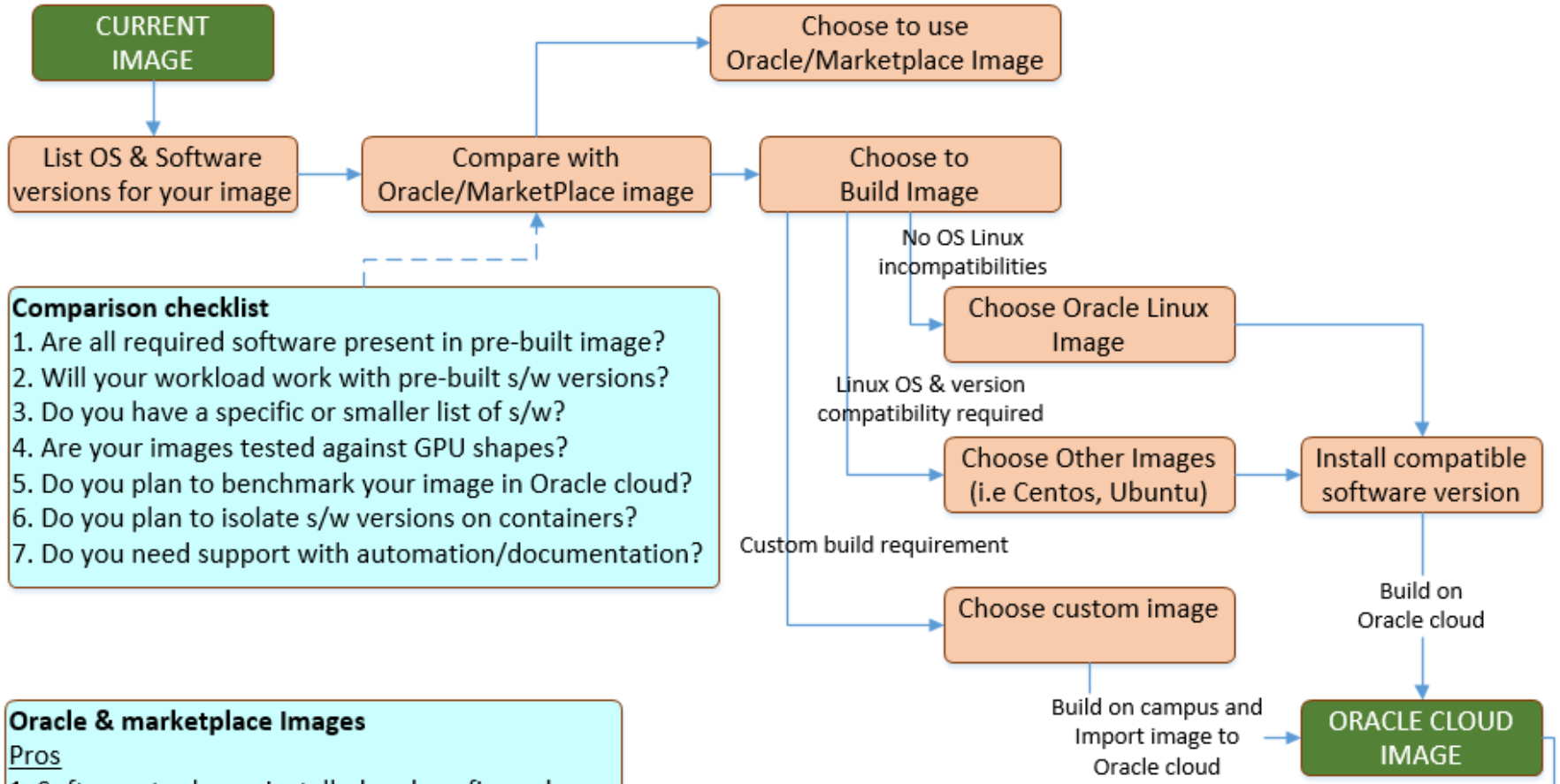
Any custom images (VM) created by you in campus or cloud

- **Characterized by**
  - Researcher created
  - Any operating system
  - Researcher software
  - Sharable across tenancy
  - Import as-is to OCI
  - Download to campus
  - Share with other clouds
- **Licensing support**
  - Open source
  - BYOL

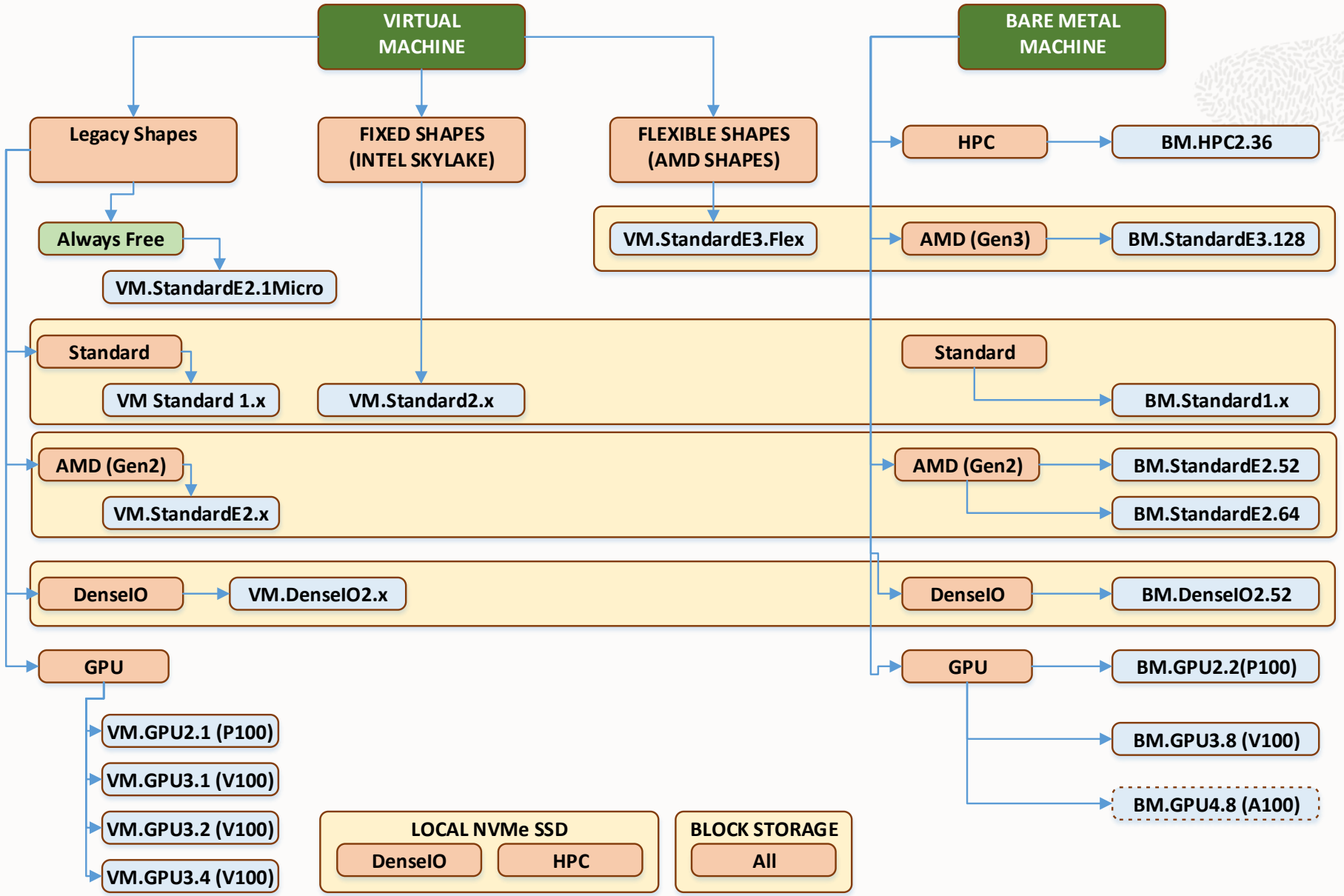
Object store images

Any OCI images (mostly in development) sharable via custom object store links

CHOOSING A PRE-BUILT IMAGE vs BUILDING YOUR OWN IMAGE

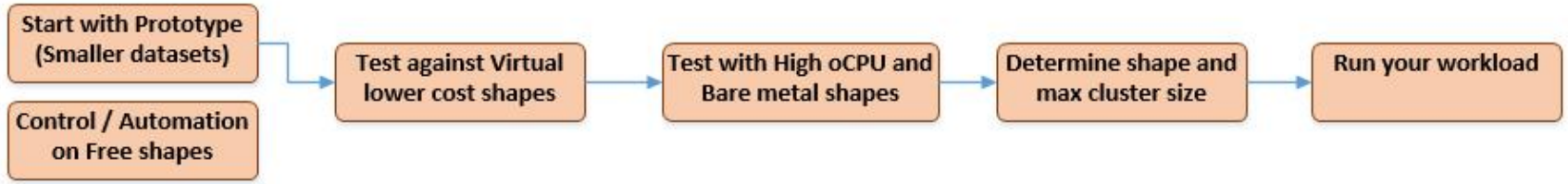


ORACLE CLOUD MACHINE SHAPES

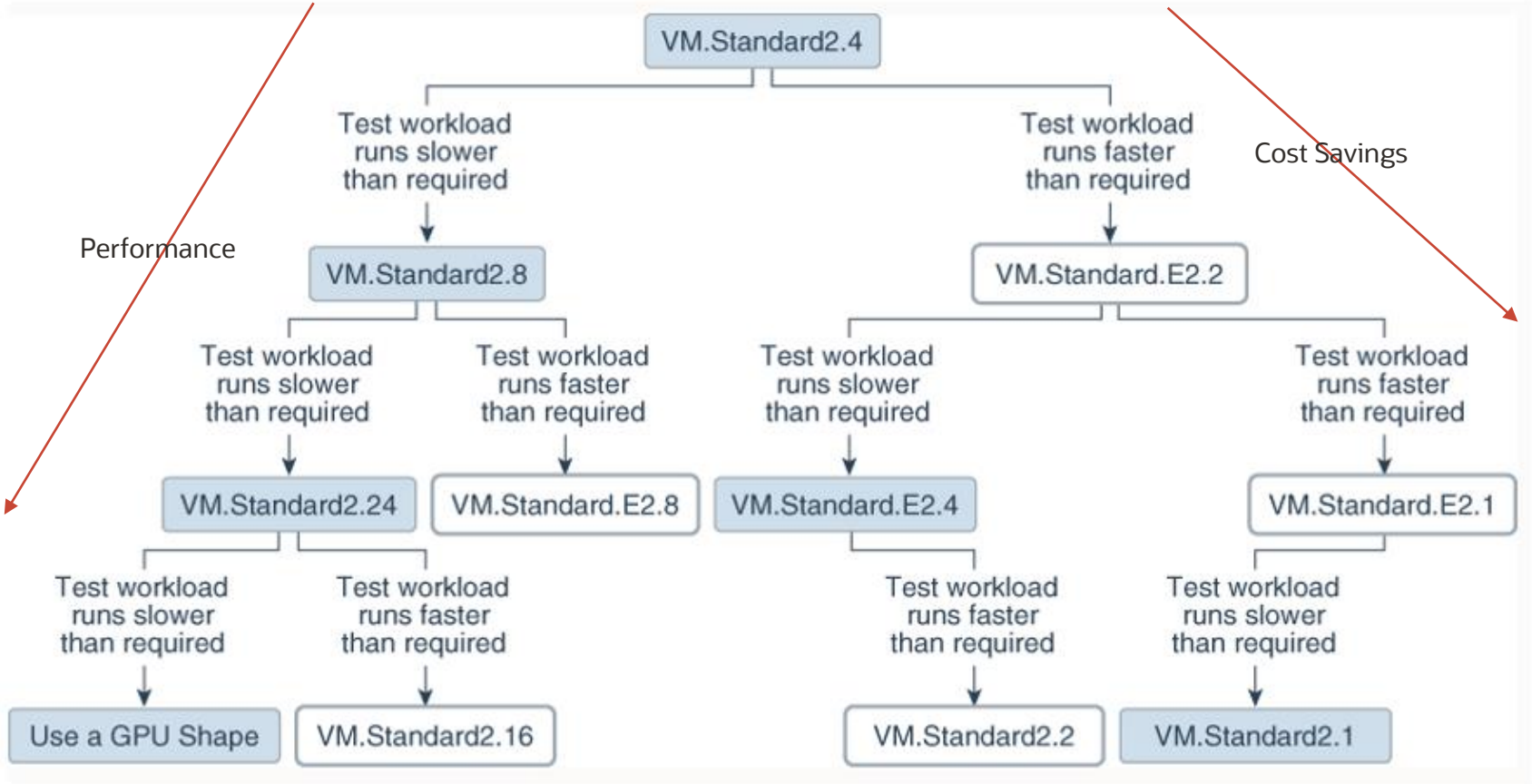


Instance Type	Shape series	Shape	Purpose
Virtual	Always Free	VMStandardE2.1Micro	Automation control, gateway,configurations
	Standard	VMStandard1.1~1.16	Low workload testing / Image builds / installs
	AMD (Gen 2)	VMStandardE2.1~2.8	Prototype workload testing
	DenseIO	VMDenseIO2.x (NVMe)	Heavy IO workload testing
	GPU (P100)	VM.GPU2.1	AI / ML or other GPU prototype testing
	GPU (V100)	VM.GPU3.1~3.4	Tensor core AI / DL workloads
	Intel Skylake (Fixed)	VM.Standard2.1~2.24	Workloads to save on credits
	AMD Rome (Flex)	VM.StandardE3.Flex	Benchmarking / price-performance
Bare metal	HPC	BM.HPC2.36 (NVMe)	CPU+high throughput for HPC workloads
	AMD (Gen 3)	BM.StandardE3.128	High CPU/throughput workloads
	Standard	BM.Standard1.36/B1.44	Low CPU/RAM utilization at lowest BM cost
	AMD (Gen 2)	BM.StandardE2.52	Best price-performance for BM workloads
	AMD (Gen 3)	BM.StandardE2.64	Best Gen3 price-performance for BM workloads
	DenseIO	BM.DenseIO2.52 (NVMe)	Best price performance for IO intensive workloads
	GPU (P100)	BM.GPU2.2	Benchmarking pascal based GPU workloads
	GPU (V100)	BM.GPU3.8	Best price performant for large GPU workloads
	GPU (A100)	BM.GPU4.8	Fastest GPU – large DL applications (pre-GA)

A GENERIC WORKFLOW TO DECIDE OPTIMAL SHAPES



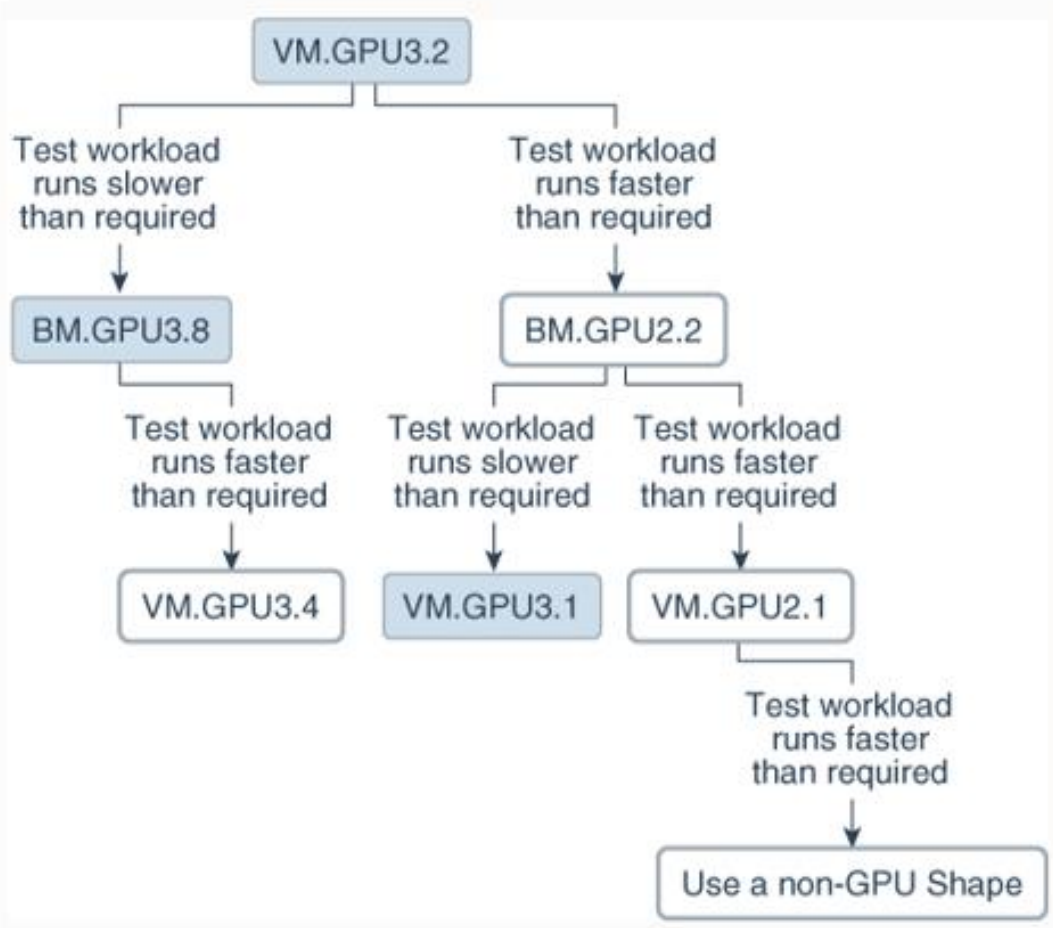
SHAPE TESTING HIERARCHY FOR NON-GPU USAGE USE-CASE



NOTE : This should be used as a guideline and may differ based on researcher data and computation scenario

Courtesy : Oracle architecture center

SHAPE TESTING HIERARCHY FOR GPU USAGE USE CASE



NOTE : This should be used as a guideline and may differ based on researcher data and computation scenario

Courtesy : Oracle architecture center

Oracle cloud storages

Storage type	Features	Recommended usage
On Campus storage	Good for on-campus data processing requirements Data could be distributed (laptops) or centralized Quick data retrieval for on-campus computations	Store data for on-campus computational purpose Store data if storage is available and it is already paid for
Object storage archive	Unlimited data storage in Oracle cloud Low cost ($0.0026/GB/Month) by consumption Better if cost/GB is lower than on-campus storage	Not so good for frequent large data retrievals Not readily available for computational purpose
Object storage standard	Unlimited data storage with faster access than OS archive Costs (0.0255/GB/Month) by consumption Good for frequently accessed data across cloud tenancies Secured and encrypted data at rest and in transit (https)	Store data backups for quick downloads in cloud / campus Store large data volumes at relatively lower cost Store data that is infrequently processed
Block volume	Most common storage for compute/databases Mountable across multiple instances within an AD Cost (0.0255/GB/month) – based on total volume in GB Supports parallel filesystem & best price performance	Leverage to store computational data for most loads Extend storage / instance as needed Fill up allocated capacity to save on costs
File system storage	NFSv3 unlimited file system storage mountable across AD Higher cost (0.3/GB/month) on consumption Good for file sharing across tenancies and OS Performs linearly or better with higher data set size	Not good for cross AD data transfer in computational cycle Good for moving large volumes of data quickly across AD Use sparingly
Local NVMe	Highest IOPS and throughput & good for IO intensive loads Higher cost (built into compute) Non-persistent data Part of DenseIO and HPC shapes	Use for IO intensive parallel CPU workloads

Useful links

Getting Started

Key concepts and terminology
Signing in to console, Sign-in options and changing your password
Setting up your tenancy
Tutorial – Launching your first Linux instance
Tutorial – Launching your first windows instance
Object storage and Pre-authenticated access
Image import and export
File storage system concepts
OCI Hands on labs
New features and navigation updates
Oracle cloud Free tier and FAQ
Custom key generation with puttygen or ssh-keygen
Frequently asked Questions
Getting help and contacting support

Identity federation

Federated identity for single sign-on
Migrating from on-premise to Oracle identity service

Databases and moving data

Oracle Autonomous databases and Tools
MySQL and NoSQL Services
Key concepts and terminology
Migrating databases to cloud
Loading data to autonomous with OCI Functions
Single-click move to autonomous

Data science and AI/ML

Oracle Data science platform and Tutorials
Genome analysis toolkit
Julia AI/HPC GPU Image
NVIDIA images and NVIDIA GPU image
Building a ML sandbox on Oracle cloud
Setting up an open-source ML and AI Environment
Machine learning autonomously

High performance computing (HPC)

Oracle HPC Cluster and Oracle HPC File system
NVIDIA GPU Cloud machine image
Oracle Linux 7 Cluster Networking Image
Oracle marketplace slurm image (HPC + Slurm combo)
Oracle cloud slurm image
Github OCI-HPC
Enabling HPC Cluster networking
Deploy High performance computing on Oracle cloud Infrastructure
Deploy scalable and distributed file system using Lustre
Deploy BEEGFS parallel file system
UoB Cluster in the cloud
Cluster in the cloud - github
Molecular dynamics NAMD runbook and GROMACS runbooks

Usage, billing and credit control

Oracle cloud storage costs
Resource billing for stopped instances
Oracle cloud universal credit PaaS and IaaS service descriptions

ORACLE
for Research

TECH TALK:

Oracle Cloud Foundations for Researchers

Questions, Answers & Discussion

ORACLE
for Research

TECH TALK:

Oracle Cloud Foundations for Researchers

Questions? Comments? Feedback?

Contact us!

Website: oracle.com/oracle-for-research/
Twitter: @OracleResearch
Email: OracleForResearchTech_ww@oracle.com

Next Tech Talk: October 23, 2020, 10:30AM EDT