

ORACLE

Oracle for Research

Tech Talk: High Performance Computing for Drug Discovery

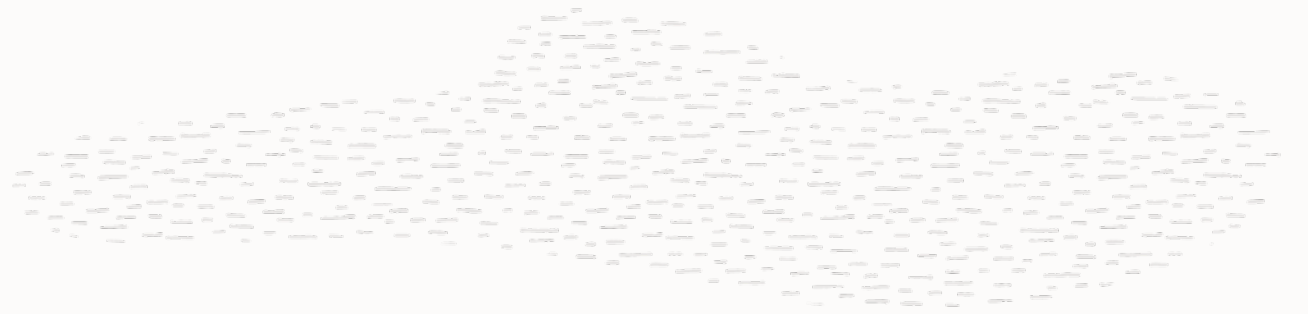
Dr. Harinda Rajapaksh, Peter Winn.

April 21, 2021



Agenda

1. Welcome to Oracle for Research
2. A case study in dedicated HPC
3. How did we do it:
 - An introduction to drug discovery
 - Challenges and solutions
 - Oracle HPC Infrastructure
 - How to use OCI Infrastructure for high-throughput drug discovery
 - Useful tips



Digital Research: Asking big questions & sharing big answers



**Commodity Super
Computing**



**Automation of
Tools and Platforms**



**Open
Collaboration**

Grant Program



Free Credits

Dedicated value of cloud credits to use on what you want, when you want it.



Cloud Expertise

Technical Support and functional supports to make your workloads faster than ever.



Marketing

Collaborative marketing and signal boosting of research and white papers.



Your IP

You remain in control, we're here to help.



Case Study

University of New South Wales (UNSW): Drug Screening on dedicated HPC

University of New South Wales – Rapid Drug Discovery



Drug Discovery requires the simulation of hundreds of millions of drug compounds



Growing libraries (200m+) can take years to scan utilizing public compute allocations



Purchases of in-house infrastructure is costly and difficult to justify

Dedicated HPC in OCI could change the game in drug discovery providing answers at predictable low operational costs.

Warp Speed Pipeline: Outcomes

1 year utilizing ad-hoc allocations on shared compute infrastructure.

~**10,000** compounds screened.

ORACLE
for Research



2 Million top compounds screened on OCI in **2 Hours** for a running cost of ~\$**100**

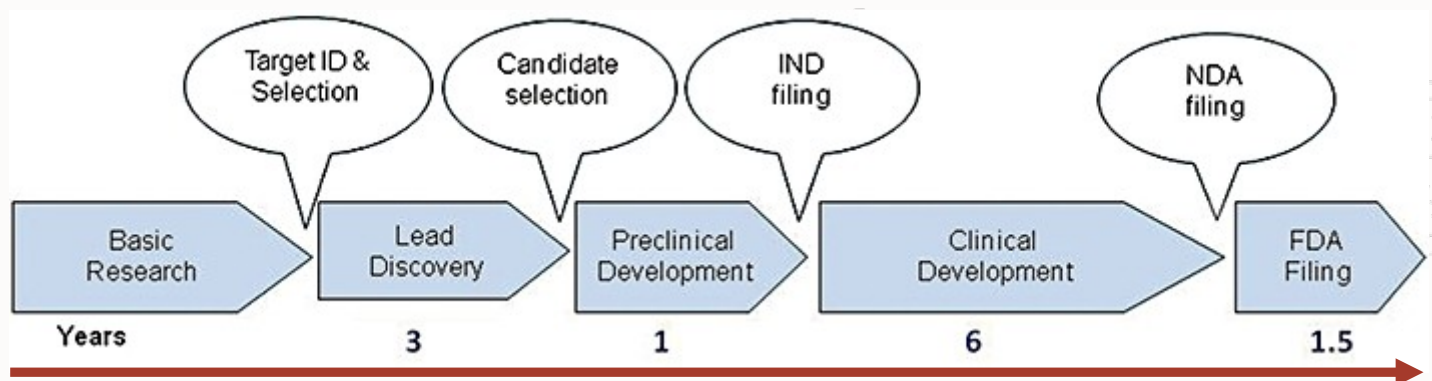
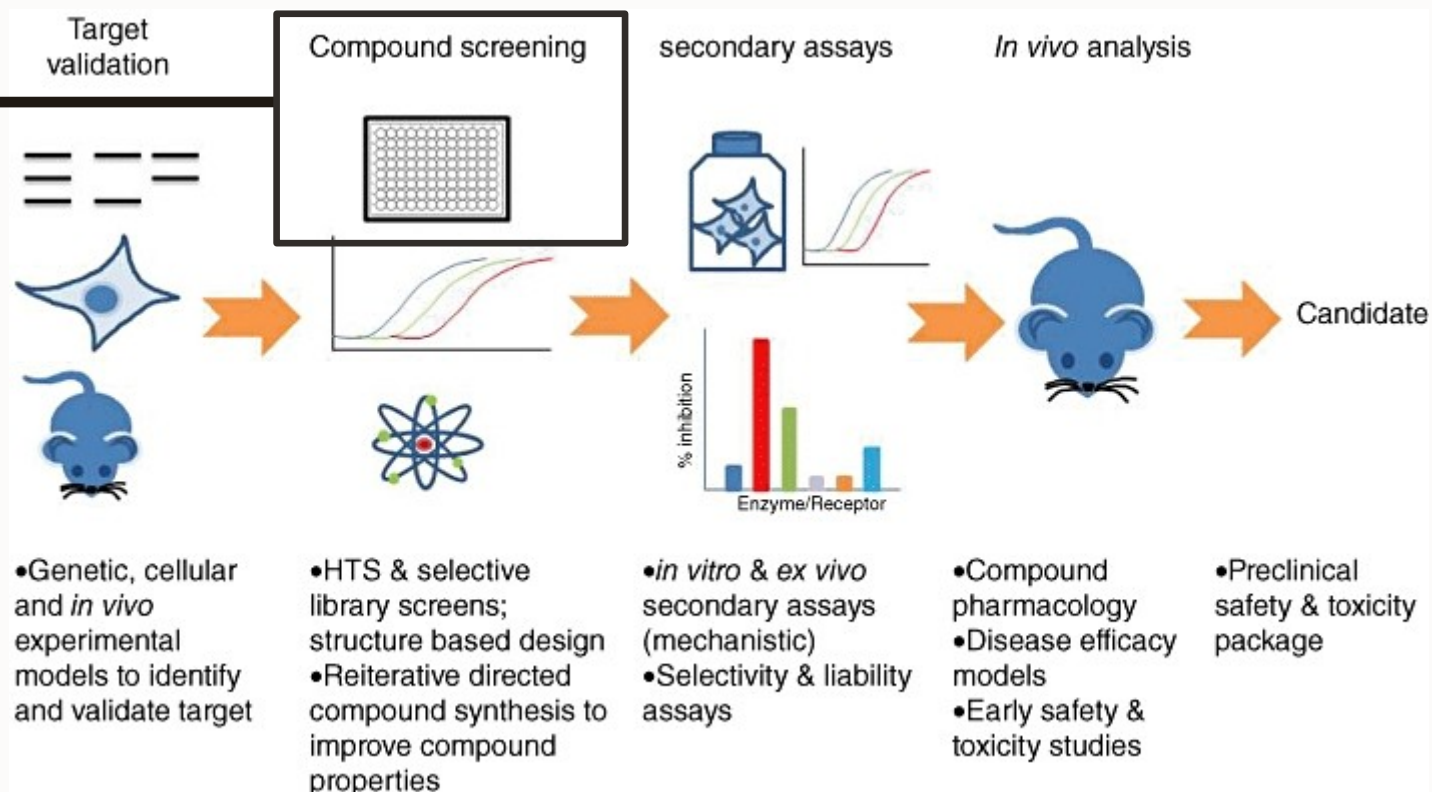


How did we do it

Deep dive on approach and tips and tricks

Drug discovery

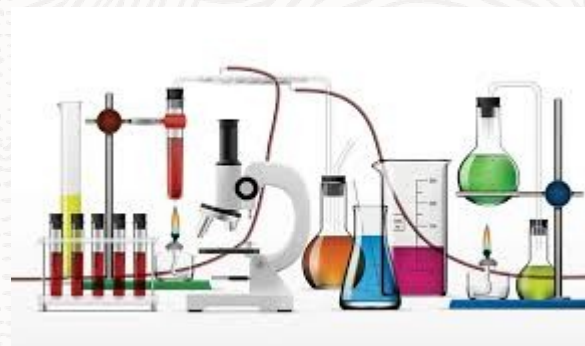
Billions of compounds





Drugs screening methods

Conventional methods (Wet
lab screening)

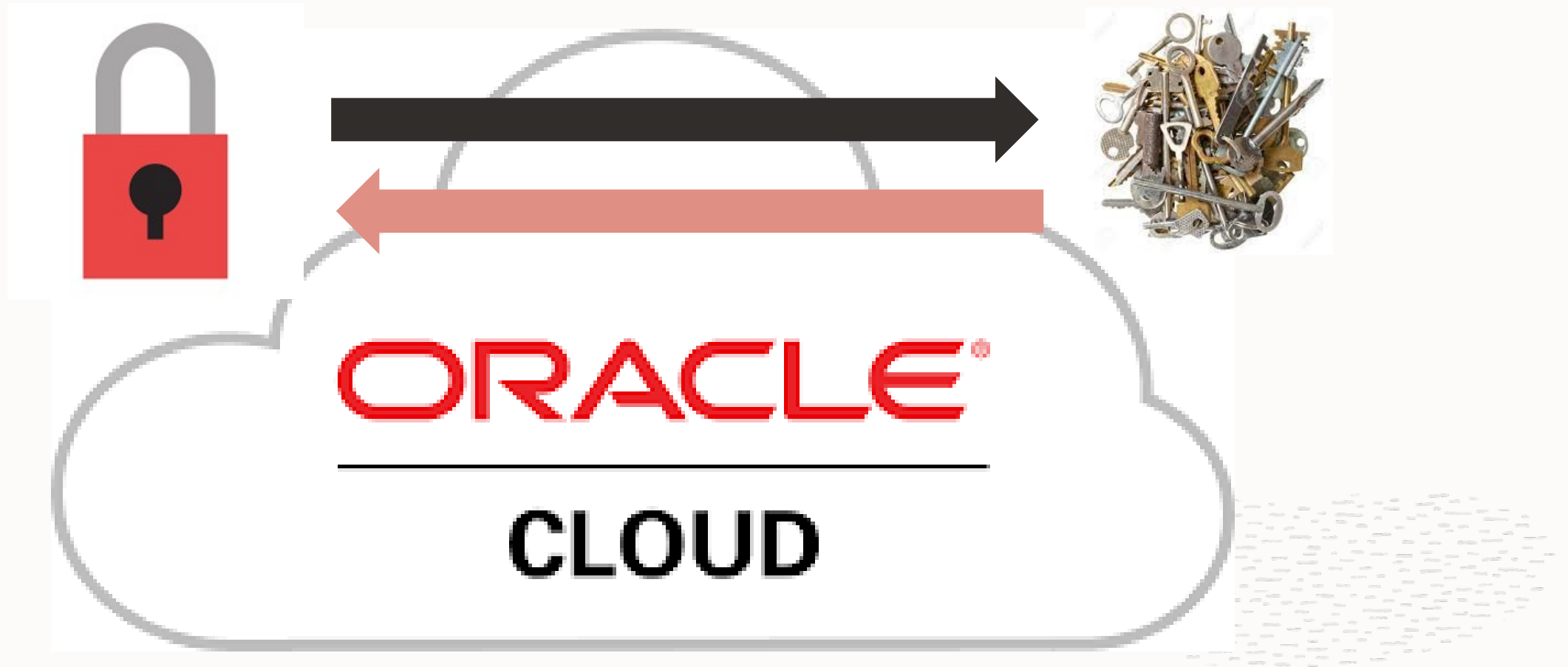


Computational methods
(Molecular docking)

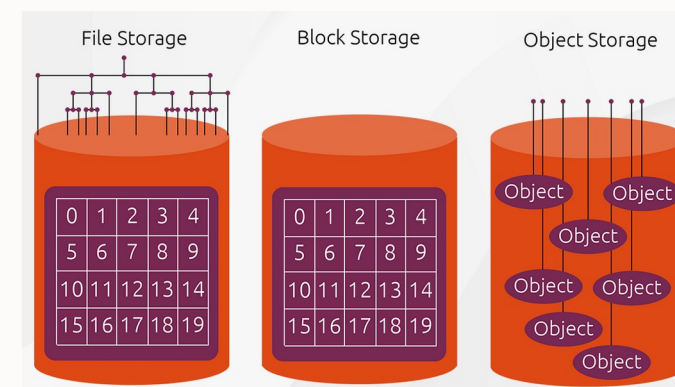
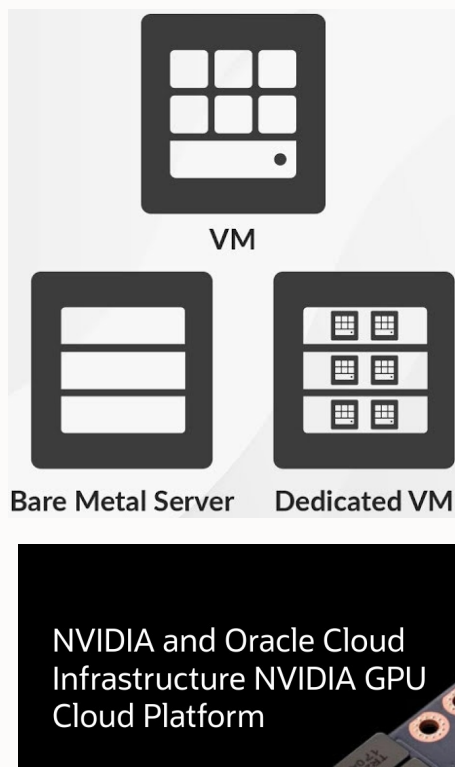
What is molecular docking?



What is molecular docking?



Oracle Infrastructure for Drug Discovery

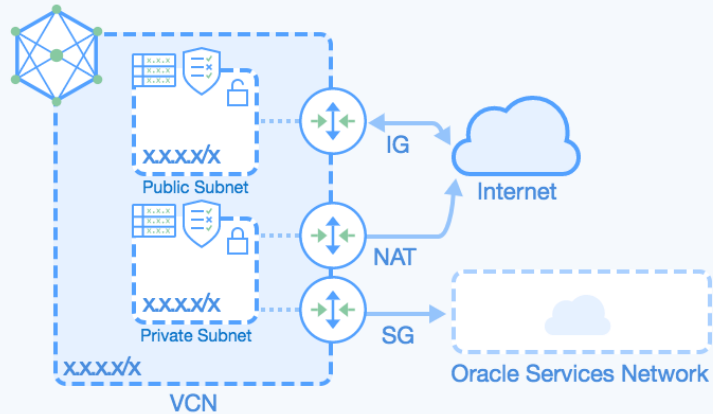


Storage

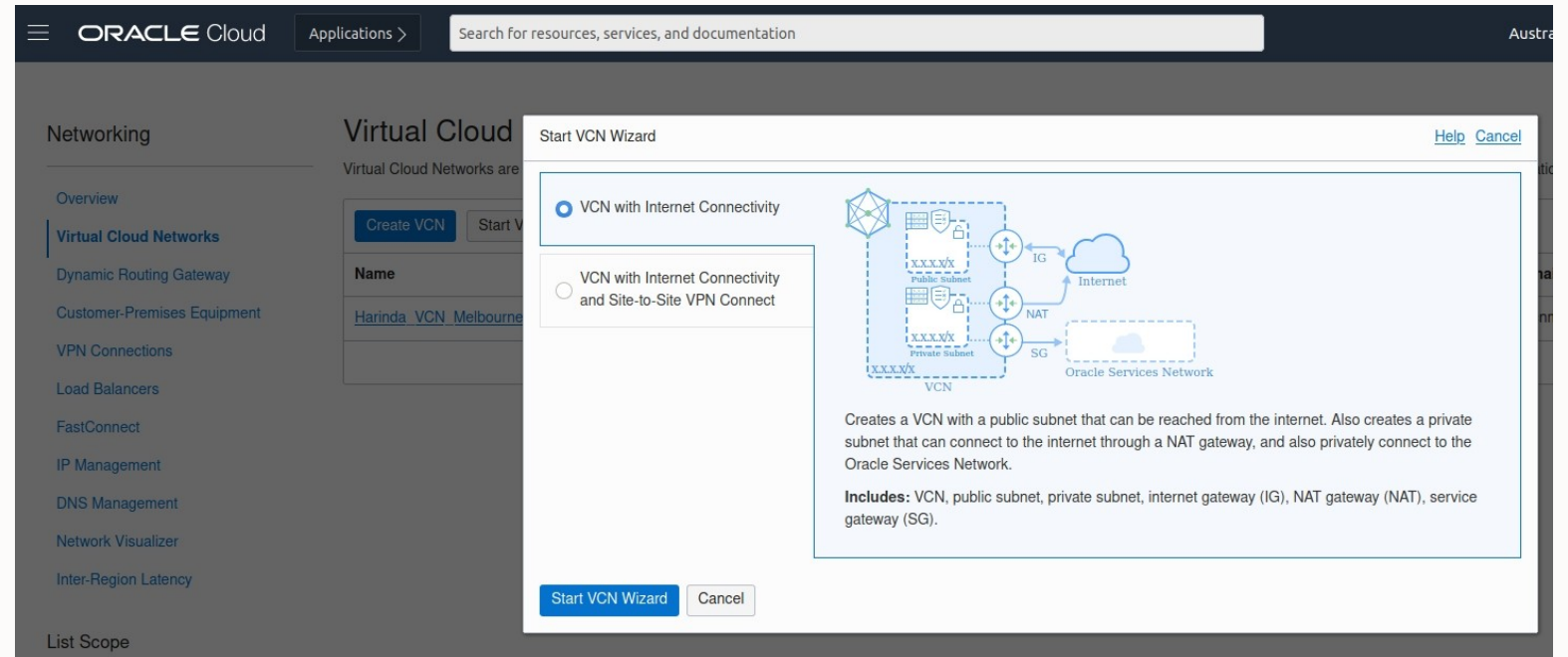
Computing resources

How to use OCI Infrastructure for high-throughput drug discovery

Create the network



VCN with Internet Connectivity



Help @ -> <https://docs.oracle.com/en-us/iaas/Content/Network/Tasks/quickstartnetworking.htm>

How to use OCI Infrastructure for high-throughput drug discovery

Launch the computing resources

Compute

Instances

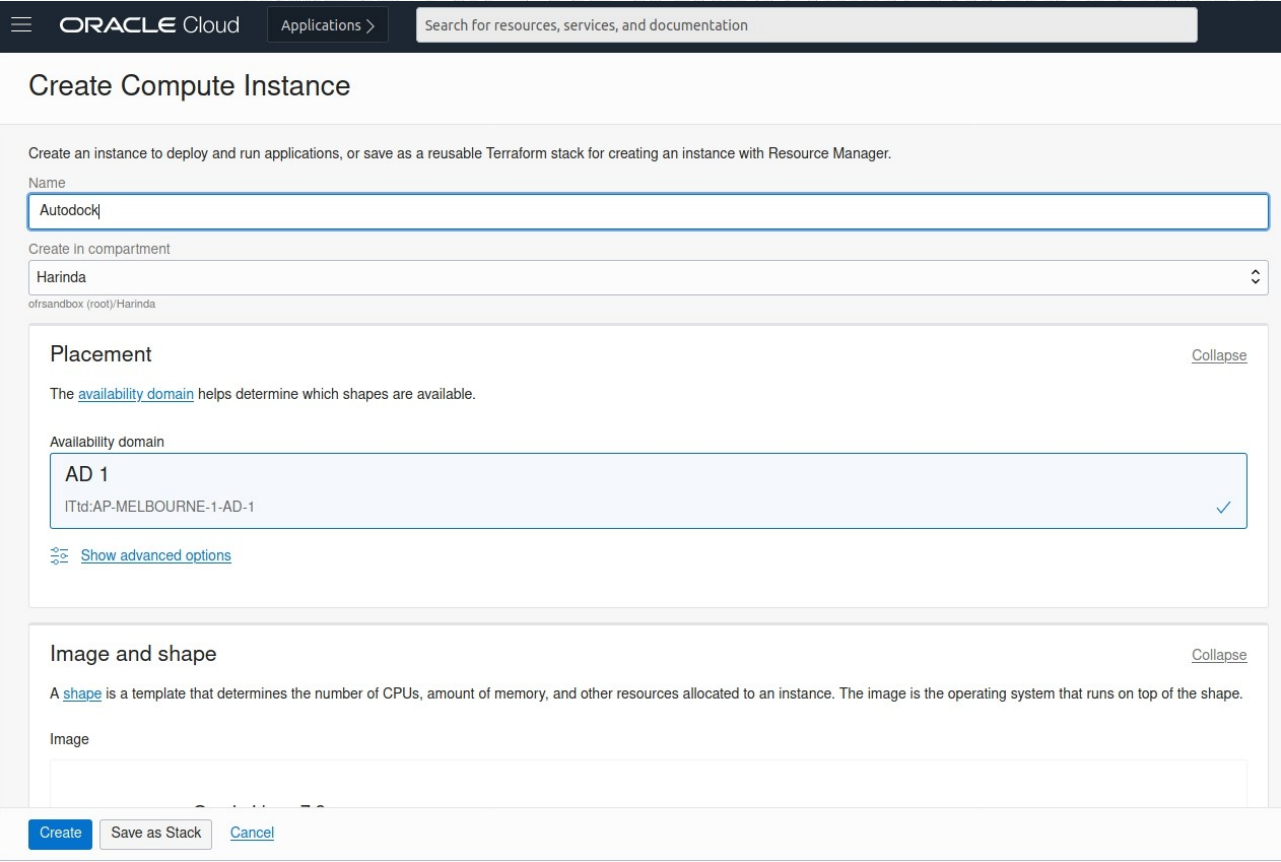
Dedicated Virtual Machine Hosts

Instance Configurations

Instances in  Compartment

The [Compute service](#) helps you provision VMs and bare metal instances to run applications. Launching an instance determines its operating system and other software.

Create Instance



ORACLE Cloud Applications > Search for resources, services, and documentation

Create Compute Instance

Create an instance to deploy and run applications, or save as a reusable Terraform stack for creating an instance with Resource Manager.

Name
Autodock

Create in compartment
Harinda
ofrsandbox (root)/Harinda

Placement Collapse

The [availability domain](#) helps determine which shapes are available.

Availability domain
AD 1
ITId:AP-MELBOURNE-1-AD-1 ✓

[Show advanced options](#)

Image and shape Collapse

A [shape](#) is a template that determines the number of CPUs, amount of memory, and other resources allocated to an instance. The image is the operating system that runs on top of the shape.

Image

Create Save as Stack Cancel

Help @ -> <https://docs.oracle.com/en-us/iaas/Content/Compute/Concepts/computestartthere.htm>

How to use OCI Infrastructure for high-throughput drug discovery

Create the storage

ORACLE Cloud Applications > Search for resources, services, and documentation

File Storage

File Systems in Compartment

File Storage provides durable, scalable, secure, enterprise-grade network file :

Create File System Delete

<input type="checkbox"/>	Name	State
--------------------------	------	-------

Mount Targets

List Scope

Create File System

[Help](#)

File Storage provides durable, scalable, and secure file systems.

File System Information

[Edit Details](#)

Name: FileSystem-20210420-0345-24
Availability Domain: ITId:AP-MELBOURNE-1-AD-1
Compartment: ofrsandbox (root)/Harinda
Encryption Key: Oracle-managed key

Export Information

[Edit Details](#)

Exports control which file systems are available to a given mount target. Create a new export to make your file system available through the selected mount target. [Learn more.](#)

Export Path: /FileSystem-20210420-0345-24
Use Secure Export Options: Disabled ⓘ

Mount Target Information

[Edit Details](#)

Mount targets are NFS endpoints used to access your file systems. The following mount target will be created and associated with your new file system. [Learn more.](#)

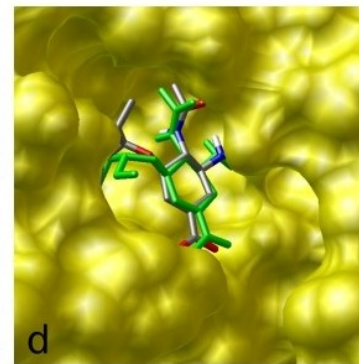
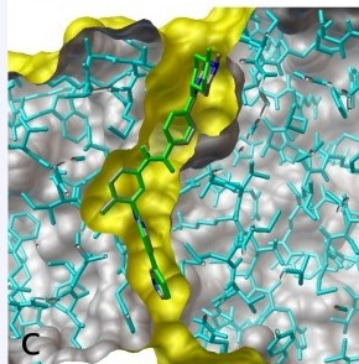
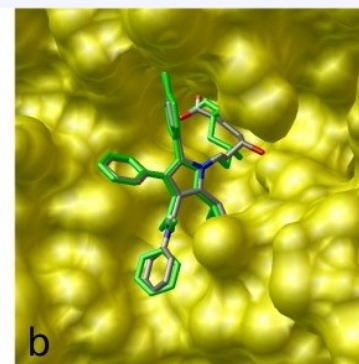
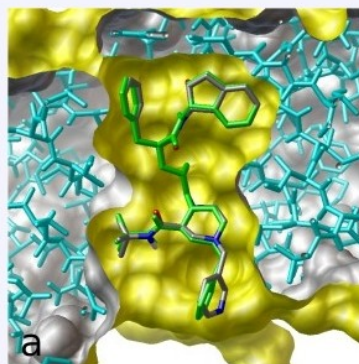
New Mount Target Name: MountTarget-20210420-0345-24
Compartment: ofrsandbox (root)/Harinda
Virtual Cloud Network: Harinda_VCN_Melbourne
Subnet: Public Subnet-Harinda_VCN_Melbourne

Create Cancel

Help @ -> <https://docs.oracle.com/en-us/iaas/Content/File/Tasks/creatingfilesystems.htm>



Drugs screening Software



AutoDock Vina

AutoDock Vina is an open-source program for doing [molecular docking](#). It was designed and implemented by [Dr. Oleg Trott](#) in the Molecular Graphics Lab at The Scripps Research Institute.

The image on the left illustrates the results of flexible docking (green) superimposed on the crystal structures of (a) indinavir, (b) atorvastatin, (c) imatinib, and (d) oseltamivir bound to their respective targets.

Publication

If you used AutoDock Vina in your work, please cite:

[O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *Journal of Computational Chemistry* 31 \(2010\) 455-461](#)



Drugs screening on your Desktop or Laptop



```
#####
# If you used AutoDock Vina in your work, please cite:      #
#                                                            #
# O. Trott, A. J. Olson,                                     #
# AutoDock Vina: improving the speed and accuracy of docking #
# with a new scoring function, efficient optimization and    #
# multithreading, Journal of Computational Chemistry 31 (2010) #
# 455-461                                                    #
#                                                            #
# DOI 10.1002/jcc.21334                                     #
#                                                            #
# Please see http://vina.scripps.edu for more information. #
#####

WARNING: The search space volume > 27000 Angstrom^3 (See FAQ)
Output will be ZINC000001351167_out.pdbqt
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 914682656
Performing search ...
0%  10  20  30  40  50  60  70  80  90 100%
|----|----|----|----|----|----|----|----|----|----|
*****
done.
Refining results ... done.

mode |  affinity | dist from best mode
     | (kcal/mol) | rmsd l.b. | rmsd u.b.
-----+-----+-----+-----
  1   |    -6.5   |    0.000   |    0.000
  2   |    -6.4   |    2.366   |    2.884
  3   |    -5.6   |    3.101   |    4.031
  4   |    -5.4   |    3.755   |    6.733
  5   |    -4.3   |    4.462   |    7.591
  6   |    -3.9   |    5.772   |    7.075
  7   |    -3.6   |    5.697   |    8.262
Writing output ... done.
```





Drugs screening on OCI

Single node screening

Multi-node screening

Drugs screening on OCI

Single node screening

Multi-node screening

ORACLE Cloud Marketplace

Oracle Cloud Home

ORACLE
High Performance
Computing

HPC Cluster

High Performance Computing - RDMA cluster network

Oracle Cloud Infrastructure | High Performance Computing

Improving the performance

Filtering compounds before docking

Analyze the receptor

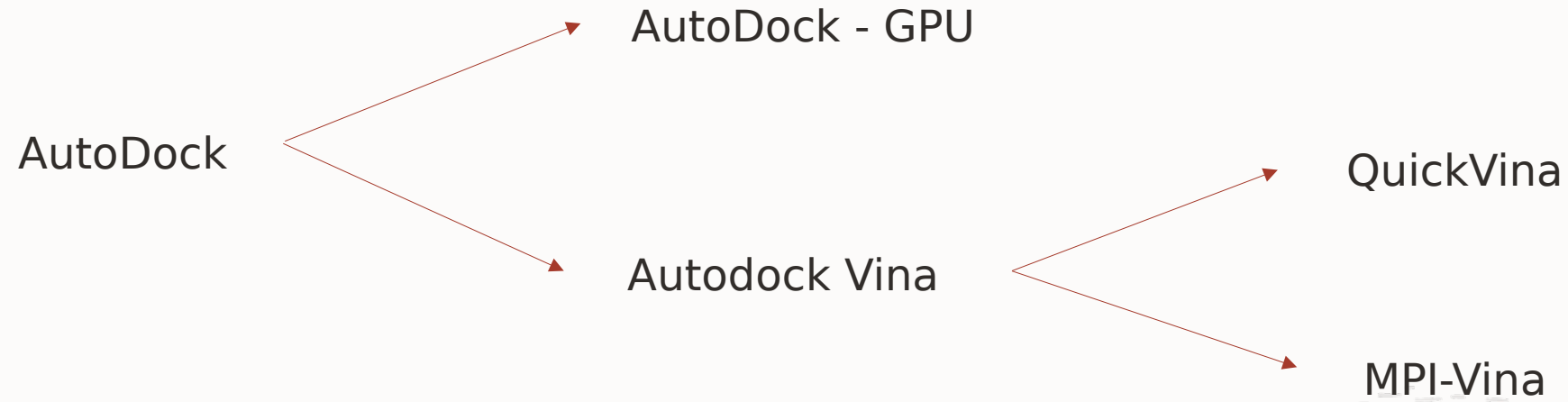


1. Molecular weight
2. xlogP
3. Net Charge
4. Number of rotatable bonds
5. Polar surface area
6. Hydrogen donors
7. Hydrogen acceptors
8. Polar desolvation
9. Apolar desolvation



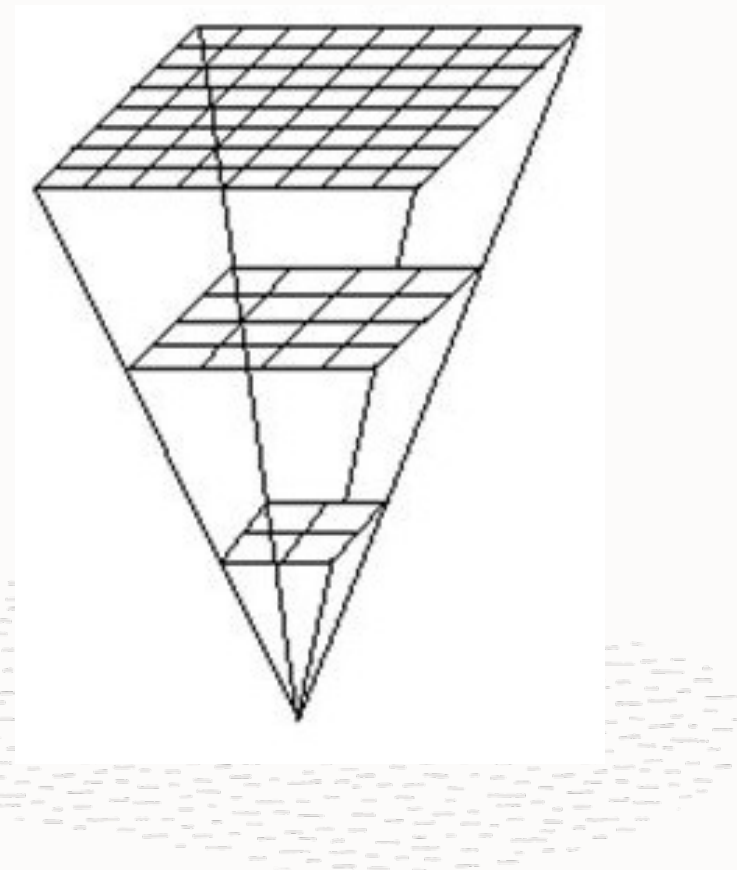
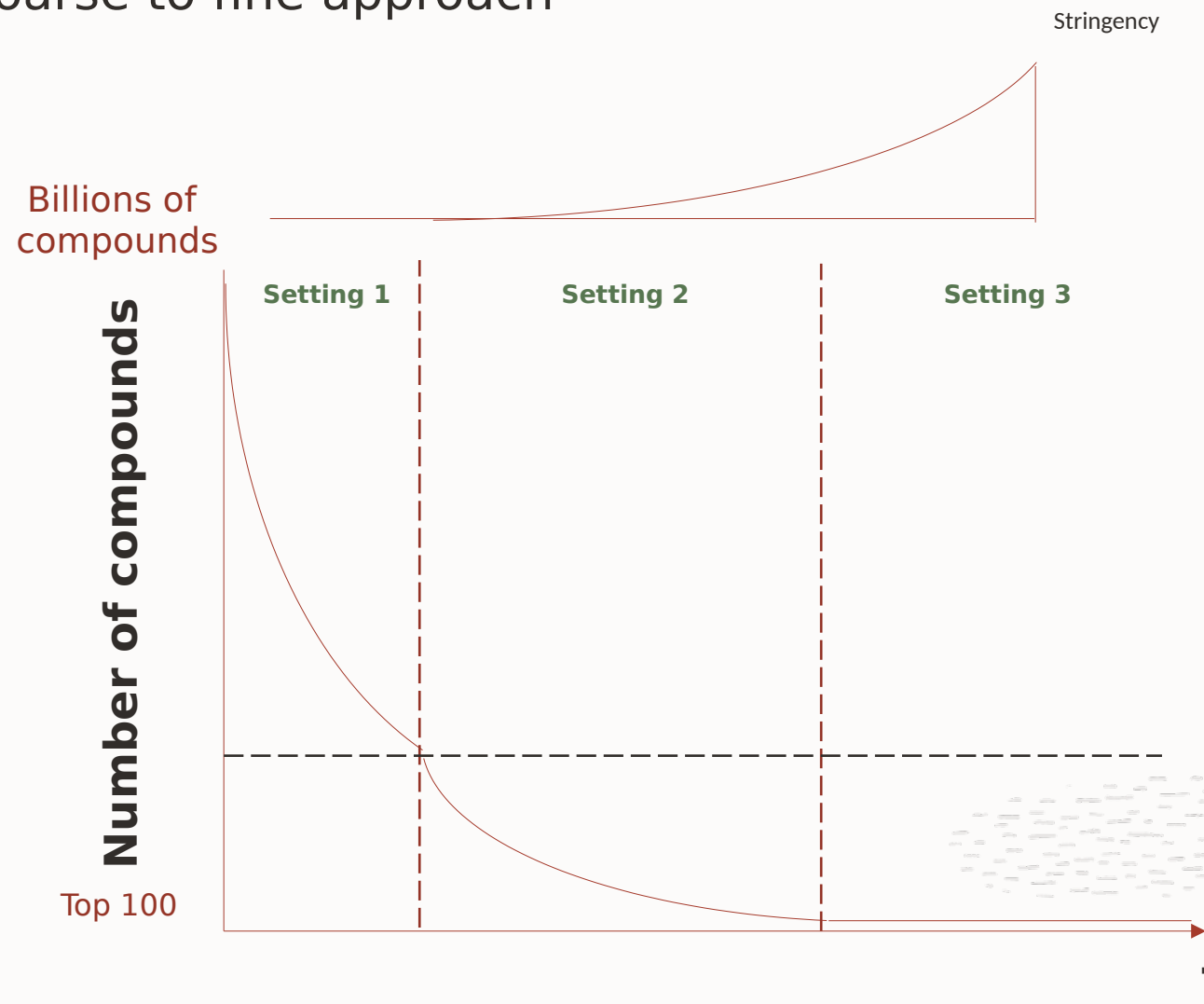
Improving the performance

Choice of the software / algorithm

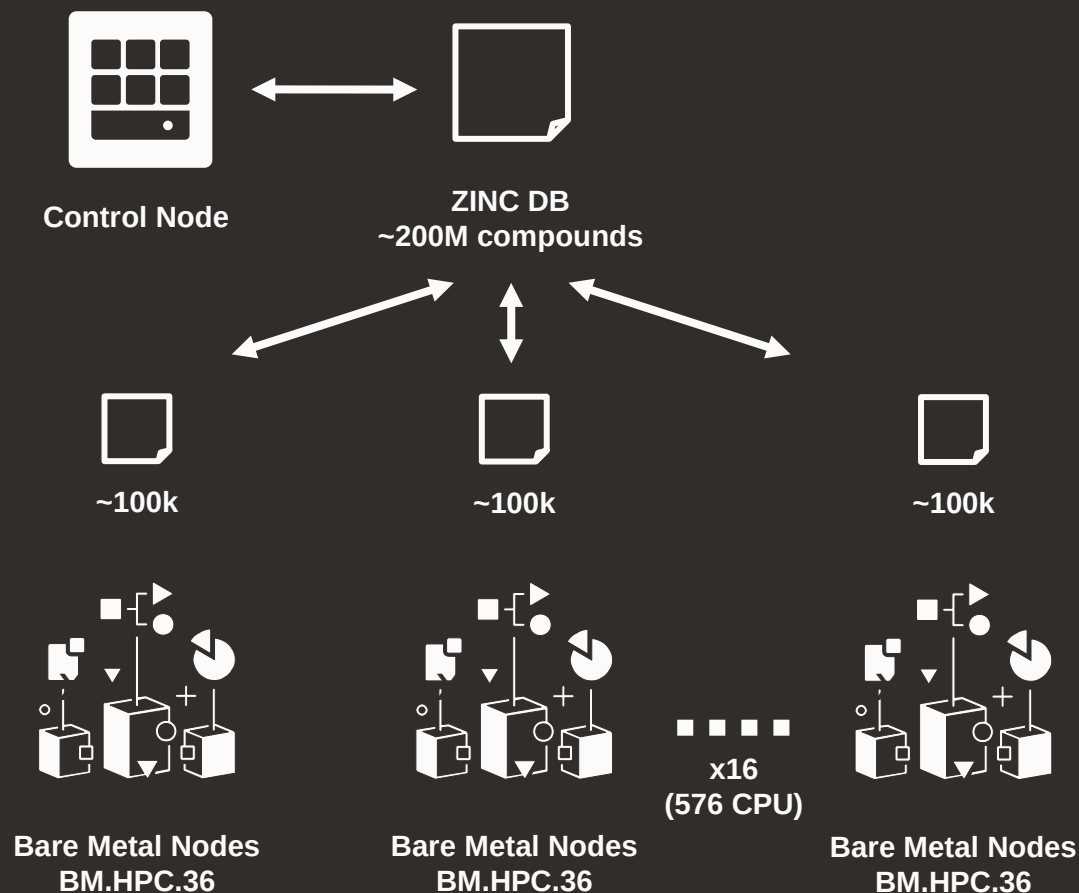


Improving the performance

Coarse to fine approach



Warp Speed Pipeline



- Drug database is filtered to prioritize top 2M compounds from >750M possible.
- Control nodes serve drug batches, optimizing for utilization, to dedicated HPC nodes.
- Autodock Vina codebase customized for Oracle HPC shapes to utilize in-memory disks and BM CPU architecture for high performance.
- Pipeline is embarrassingly parallelized to avoid network communication and scale linearly for maximum performance.
- HPC nodes are destroyed on completion with results saved to shared storage to manage costs.





Questions?

w: oracle.com/oracle-for-research

e: oracleforresearch_ww@oracle.com



Our mission is to help people
see data in new ways, discover
insights,
unlock endless possibilities.

Thank You

