Safe harbor and forward-looking statements

This presentation is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Statements in this presentation relating to Oracle's future plans, expectations, beliefs, intentions, and prospects are "forward-looking statements" and are subject to material risks and uncertainties. A detailed discussion of these factors and other risks that affect our business is contained in Oracle's Securities and Exchange Commission (SEC) filings, including our most recent reports on Form 10-K and Form 10-Q under the heading "Risk Factors." These filings are available on the SEC's website or on Oracle's website at http://www.oracle.com/investor. All information in this presentation is current as of January 2024 and Oracle undertakes no duty to update any statement in light of new information or future events.

Innovation at Oracle Cloud for HPC, Grid Computing, and Beyond

Attila Narin
Vice President of Cloud Engineering
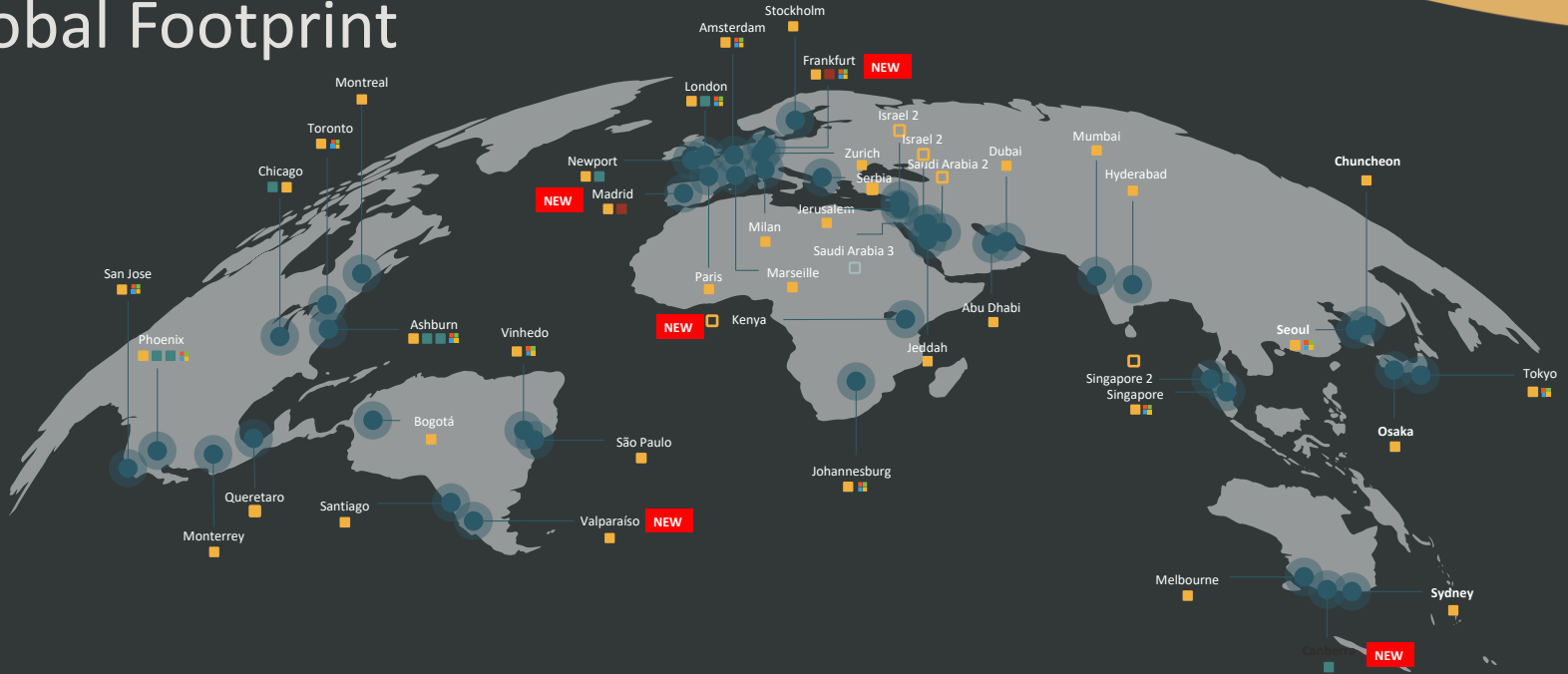Oracle

Oracle Cloud Infrastructure Global Footprint

Phoenix

Commercial

October 2016

Oracle Cloud Infrastructure Global Footprint

Today

Montreal
Toronto
Chicago
San Jose
Phoenix
Ashburn
Vinhedo
Bogotá
Queretaro
Santiago
Monterrey
Valparaíso **NEW**
São Paulo

Newport **NEW**
Madrid
London
Amsterdam
Stockholm
Frankfurt **NEW**
Paris **NEW**
Milan
Marseille
Zurich
Serbia
Jerusalem
Kenya
Israel 2
Israel 2
Saudi Arabia 2
Saudi Arabia 3
Dubai
Abu Dhabi
Jeddah
Johannesburg

Mumbai
Hyderabad
Chuncheon
Seoul
Tokyo
Osaka
Singapore 2
Singapore
Melbourne
Sydney
NEW

🟧	Commercial
⬜	Commercial Planned
🟥	Sovereign
🟩	Government
🟦🟥🟩🟨	Microsoft Interconnect Azure

OCI Gen2 Advantages

A 2nd Generation Cloud built for HPC, Grid Computing, and AI workloads

Off-Box Virtualization

Non-Blocking Networks

Remote Direct Memory Access (RDMA)

More Local NVMe Storage

Impacts are magnified with high-performance workloads

OCI Supercluster - run faster and more cost effectively

RDMA cluster networking

Nonblocking networks

More Local NVMe Storage

Highest performance, lowest cost GPU cluster technology in the world

Latency: ~2µs
Bandwidth:
- NVIDIA H100: 3.2Tbps
- NVIDIA A100: 1.6Tbps
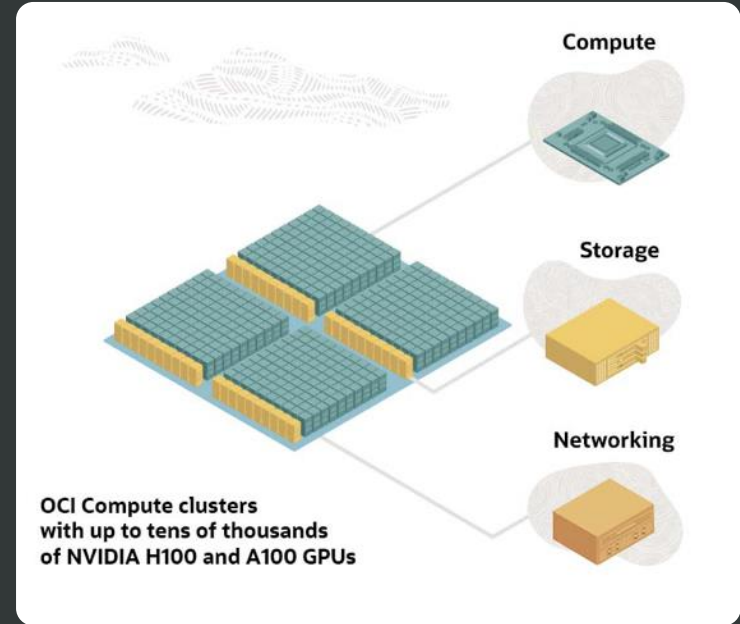Cluster size:
- Tens of thousands of NVIDIA H100 or A100 GPUs

Provides the largest cache for checkpointing
- H100: 61.4TB/node
- A100: 27.2 TB/node

Compute

Storage

Networking

OCI Compute clusters
with up to tens of thousands
of NVIDIA H100 and A100 GPUs

NOT STAC BENCHMARKS

Posted December 10, 2023

STAC Report: Oracle Cloud Infrastructure with NVIDIA SXM4 GPUs under STAC-A2 (derivatives risk)

New records in cloud price performance

NVDA231026

Compared to all publicly reported solutions to date, this solution demonstrated the best price-performance among all cloud solutions (48,404 options / USD) over 1 hour, 3 days, and one year of continuous use[1]

Compared to the most recently audited, single-server cloud-based solution (INTC221006b), this solution:

- Demonstrated 3.2x, 3.2x and 2.0x price-performance advantages (options / USD) for 1 hour, 3 days and 1 year of continuous use respectively[1]
- Had 14.6x the throughput (options / second)[2]
- Demonstrated 6.8x / 7.8x the speed in cold / warm runs of the baseline Greeks benchmarks[3]
- Demonstrated 9.0x / 4.3x the speed in cold / warm runs of the large Greeks benchmarks[4]

Stack under test (SUT): CUDA 12.2 with 8 x NVIDIA A100 SXM4 40 GiB GPUs in an OCI BM.GPU4.8 instance.

[1]STAC-A2.β2.HPORTFOLIO.PRICE_PERF.[BURST | PERIODIC | CONTINUOUS]
[2]STAC-A2.β2.HPORTFOLIO.SPEED
[3]STAC-A2.β2.GREEKS.TIME.[COLD | WARM]
[4]STAC-A2.β2.GREEKS.10-100K.1260.TIME.[COLD | WARM]
[5]STAC-A2.β2.GREEKS.[MAX_ASSETS | MAX_PATHS]

STAC Report: Oracle Cloud Infrastructure 8-node Intel cluster under STAC-A2 (derivatives risk)

Numerous new records for non-accelerated cloud solutions

INTC231113

Compared to all publicly reported, non-accelerated cloud solutions to date, this solution set numerous records:

- The best price-performance (19,788 options / USD) over 1 hour and 3 days of continuous use[1]
- The highest throughput (119.2 options / second)[2]
- The fastest warm time (24.3 ms) in the baseline Greeks benchmark[3]
- The fastest cold (2.98 s) and warm (1.96 s) times in the large Greeks benchmarks[4]
- The most correlated assets (145) and Monte Carlo paths (7,000,000) simulated in 10 minutes[5]

Compared to the previously most price-performant, non-accelerated cloud solution (SUT ID INTC221006b), this solution demonstrated a 1.3x price-performance advantage (options / USD) for 1 hour and 3 days of continuous use[1]

Stack under test: STAC-A2 Pack for Intel® oneAPI (Rev N) with 8 x OCI BM.Optimized.3.36 bare-metal servers, each server providing 2 x Intel® Xeon® Gold 6354 (Ice Lake) CPUs and 512 GiB of DRAM, networked via RDMA over 100 Gbps Ethernet.

[1]STAC-A2.β2.HPORTFOLIO.PRICE_PERF.[BURST | PERIODIC]
[2]STAC-A2.β2.HPORTFOLIO.SPEED
[3]STAC-A2.β2.GREEKS.TIME.WARM
[4]STAC-A2.β2.GREEKS.10-100K.1260.TIME.[COLD | WARM]
[5]STAC-A2.β2.GREEKS.[MAX_ASSETS | MAX_PATHS]

OCI for Financial Services

Tier1 Multinational Investment Bank

Monte Carlo Simulation with significantly higher performance based on 2x network throughput
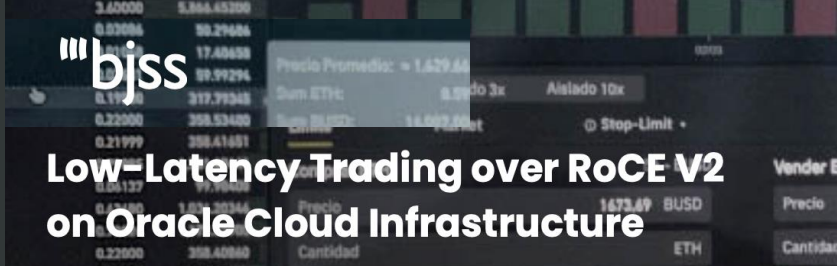

ADVANCED LABORATORY ECONOMICS AND FINANCE s.r.l.

powers complex HPC financial simulations 30% faster with OCI

NRI
Nomura Research Institute Selects OCI Dedicated Region to Support Hundreds of Millions of Financial Trades a Day

Low-Latency Trading over RoCE V2 on Oracle Cloud Infrastructure

HPC workload boosted by 5x and deterministic latencies at the μs level at very high message volumes

HSBC | bjss | NVIDIA | alef | Deutsche Bank | BBVA | NRI

ORACLE

Thank you!

Attila Narin
attila.narin@oracle.com