# Oracle Cloud Infrastructure

ASCENDANCE

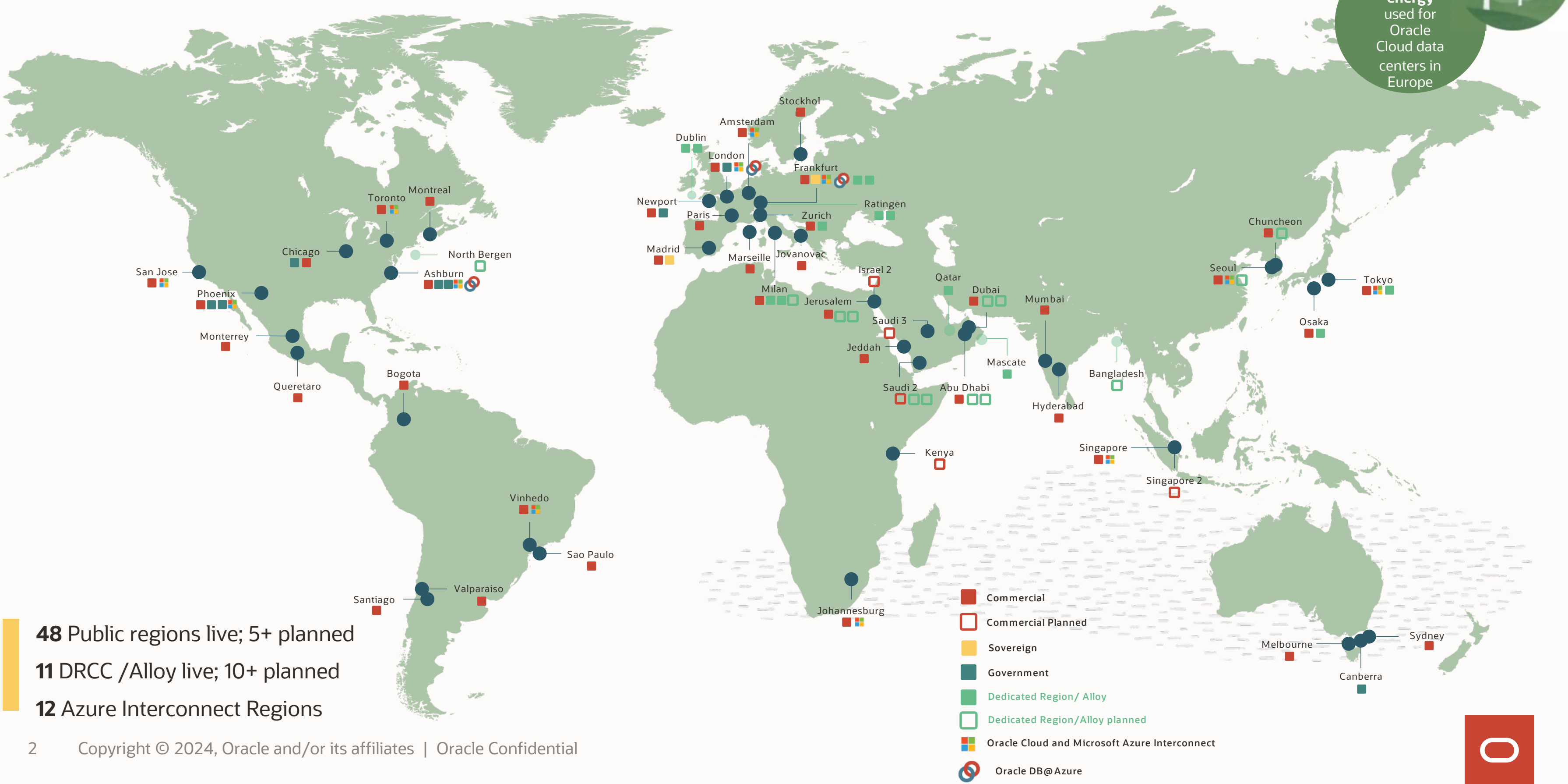
FLIGHT TECHNOLOGIES

Jorge Quintero – Ascendance Software Leader
Romuald Josien – Head of AI OCI

May 2024

# Oracle's global footprint positions it as a competitive player in the hyperscaler market

100% **renewable energy** used for Oracle Cloud data centers in Europe

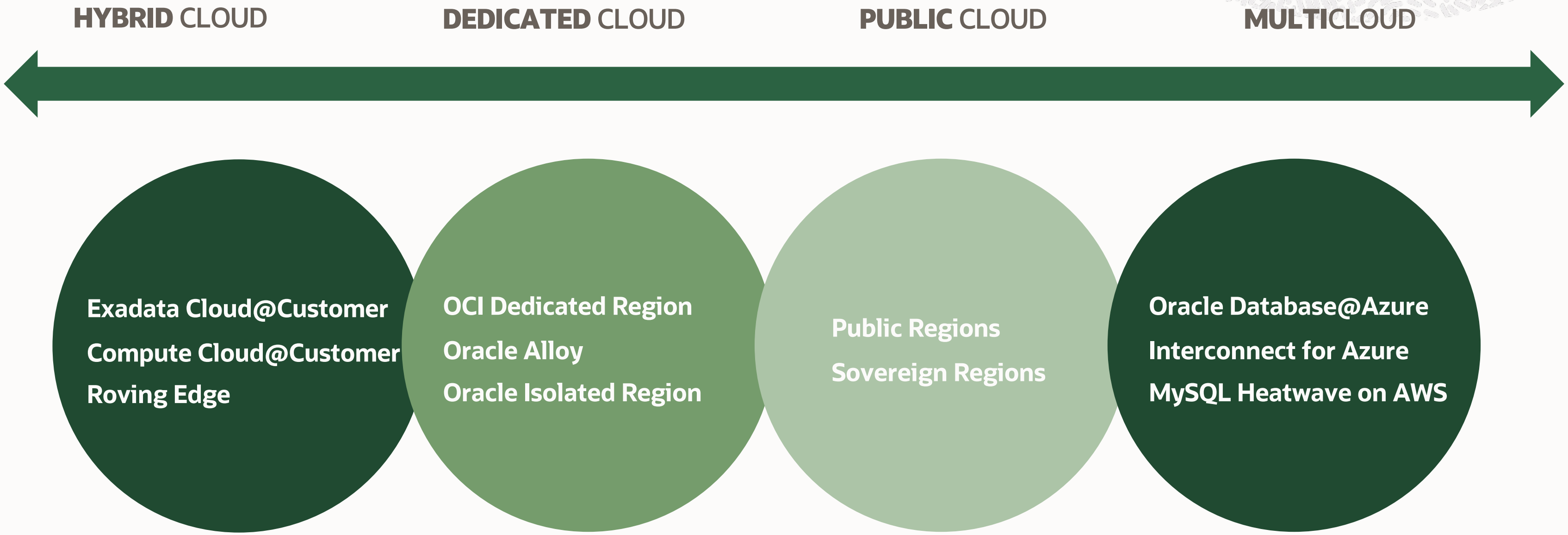


**48** Public regions live; 5+ planned

**11** DRCC /Alloy live; 10+ planned

**12** Azure Interconnect Regions

# With its distributed cloud strategy, Oracle is the only hyperscaler to offer its cloud services, including AI, in the deployment model a customer selects



**HYBRID** CLOUD

- Exadata Cloud@Customer
- Compute Cloud@Customer
- Roving Edge

**DEDICATED** CLOUD

- OCI Dedicated Region
- Oracle Alloy
- Oracle Isolated Region

**PUBLIC** CLOUD

- Public Regions
- Sovereign Regions

**MULTI**CLOUD

- Oracle Database@Azure
- Interconnect for Azure
- MySQL Heatwave on AWS

All built on the same foundations

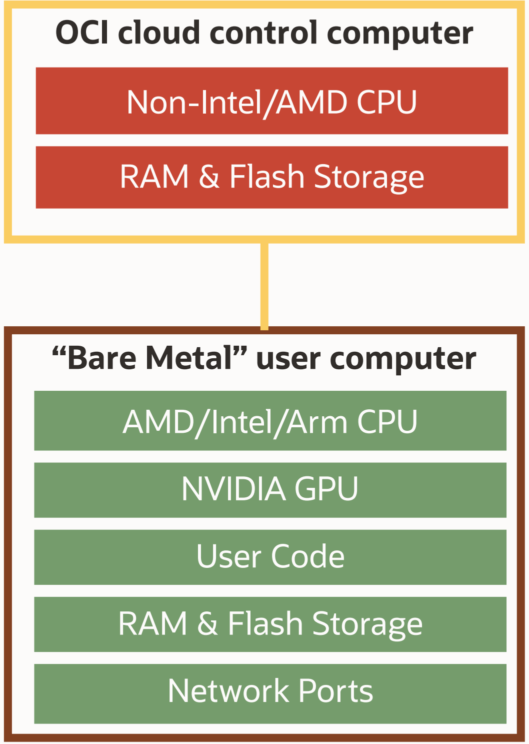
ORACLE

NVIDIA®

# OCI GPU Shape

# BARE METAL

# Our unique Bare Metal offering is designed for the performance needed for AI workloads



- OCI has off-box virtualization throughout the fleet
- These “cloud control computers” run OCI’s control plane, offloading OCI’s use of resources
- Off-box virtualization enables:
  - High performance bare metal compute instances
  - Greatly reduced performance overhead in virtual machines and containers
  - Greater isolation from other OCI customers for better security and more consistent performance

ORACLE
Cloud Infrastructure

# RoCEv2

+

## Oracle fine tuning

# OCI Supercluster - Train faster and more cost effectively

**RDMA cluster networking**

**Nonblocking networks**

Highest performance, lowest cost GPU cluster technology in the world

Latency: ~2µs
Bandwidth:

- NVIDIA H100: 3.2Tbps
- NVIDIA A100: 1.6Tbps

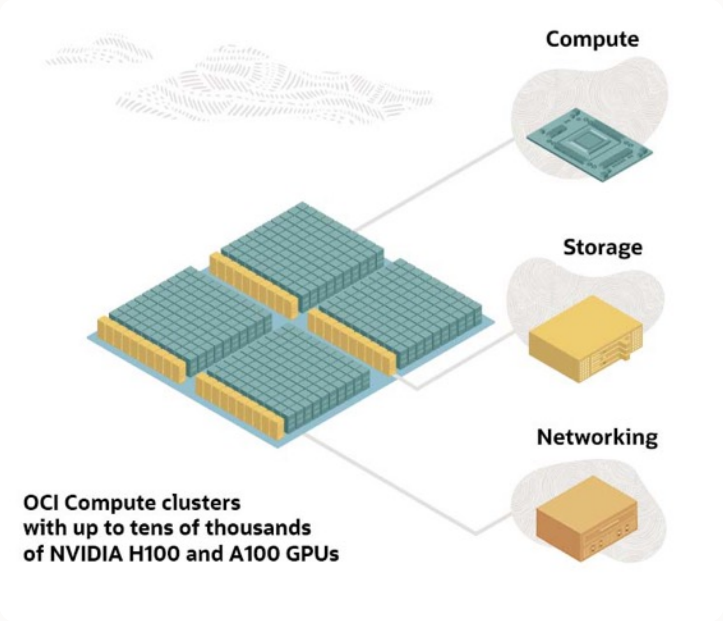
Cluster size:

- Tens of thousands of NVIDIA H100 or A100 GPUs

**More Local NVMe Storage**

Provides the largest cache for checkpointing

- H100: 61.4TB/node
- A100: 27.2 TB/node



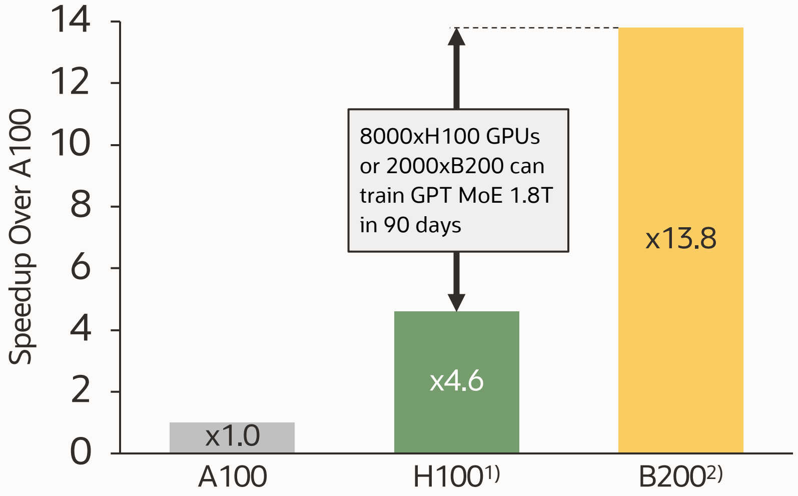
OCI Compute clusters with up to tens of thousands of NVIDIA H100 and A100 GPUs

# OCI against

# OBSOLESCENCE

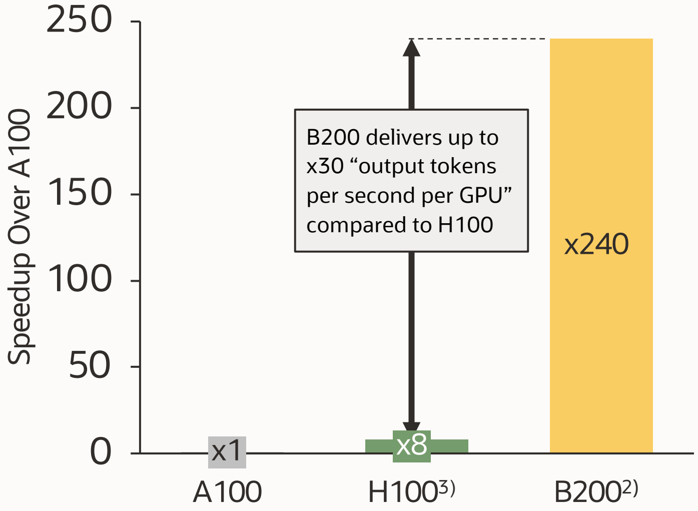
# With a lifespan up to 5 years, Clients deploying GPUs ’on-premises’ are missing out on next gen tech with significant performance enhancement …

## Deep Learning Training GPU Performance



8000xH100 GPUs or 2000xB200 can train GPT MoE 1.8T in 90 days

## Deep Learning Inference GPU Performance



B200 delivers up to x30 “output tokens per second per GPU” compared to H100

⬜ A100 🟩 H100 🟨 B200

1) GPT-J 6B, 2) GPT-MoE 1.8T 3) Llama2
Source: nVIDIA

# OCI Services Level Agreement

## OCI engages by default on solid SLA for its accelerated compute services, including Manageability and Performance[1]

| | |
|---|---|
| 100%> availability >99.99%<br>0,744h > interruption | SLA reached |
| 99.99%> availability >99%<br>0,744h < interruption < 7,44h | 10% Credit |
| 99%> availability >95%<br>7,44h < interruption < 37,2h | 25% Credit |
| 95%> availability<br>37,2h < interruption | 100% Credit |

[1] - Oracle PaaS and IaaS Public Cloud Services Pillar Document (PDF)

# Customers that trust OCI for AI/ML

Adept

Recently launched ACT-1, new large-scale Transformer model

mosaicML

Delivering Composer, a library for accelerating ML training by 7x

SoundHound

Speech recognition platform powers Mercedes and Pandora

cohere

Cohere was GCP's largest TPU customer, migrating to OCI

character.ai

#1 on the App Store, create your own AI characters

Reka

Exited stealth with $50M in funding, creating AI assistants

UNIVERSITY OF MICHIGAN

Improves AI text summaries for for academic journals

Twelve Labs

Building a best in class video search model

MIT

Creating the next frontier of AI research

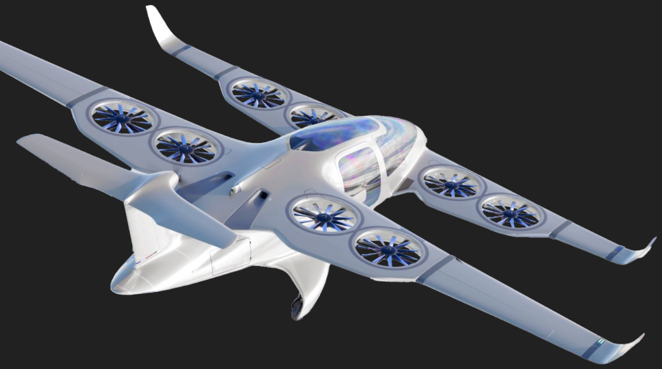
# ascendance

Where innovation takes flight.

# Decarbonizing air transport with hybrid electric propulsion technology and cleaner aircraft.

# Atea

The 1st aircraft powered by Sterna
› 555 pre-orders



**RANGE**

400 km

with full payload incl. 30min reserve

**PAYLOAD**[1]

450 kg

(for 400 KM) or 4 PAX + 1 operator

**SPEED**

+200KM/H

**SAFE**

BY REDUNDANCY

VS Helicopter up to

-75% noise

POWERED BY
Sterna

-80% carbon emissions

-50% Direct Ops. Costs

[1] Data displayed for manned configuration
The payload could be increased in an unmanned configuration

# Atea is the only aircraft matching the specific requirements of helicopters and regional markets

## Compared to helicopters

- Similar level of performance
- up to -50% of operating costs to increase profitability
- up to -75% noise for quieter flight
- up to -80% of CO2 emissions for more sustainable flights

## Compared to other battery eVTOLs

- x2 productivity thanks to hybrid
- Regional range consistent with RAM distances
- Infrastructure agnostic thanks to in-flight charging

# Our roadmap

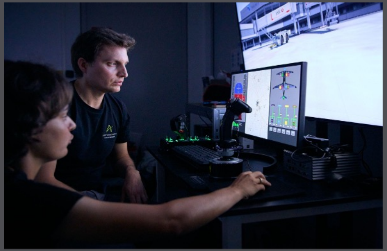
2022 → 2025 → 2026 → 2027+

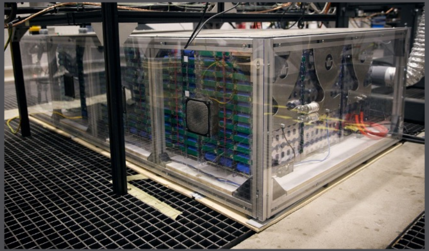
**2022**
Developed & Demonstrated Technologies

**2025**
First customers STERNA
First flight ATEA

**2026**
Production & industrialization

**2027+**
Entry into service 1st client aircraft powered by STERNA
ATEA Entry Into Service

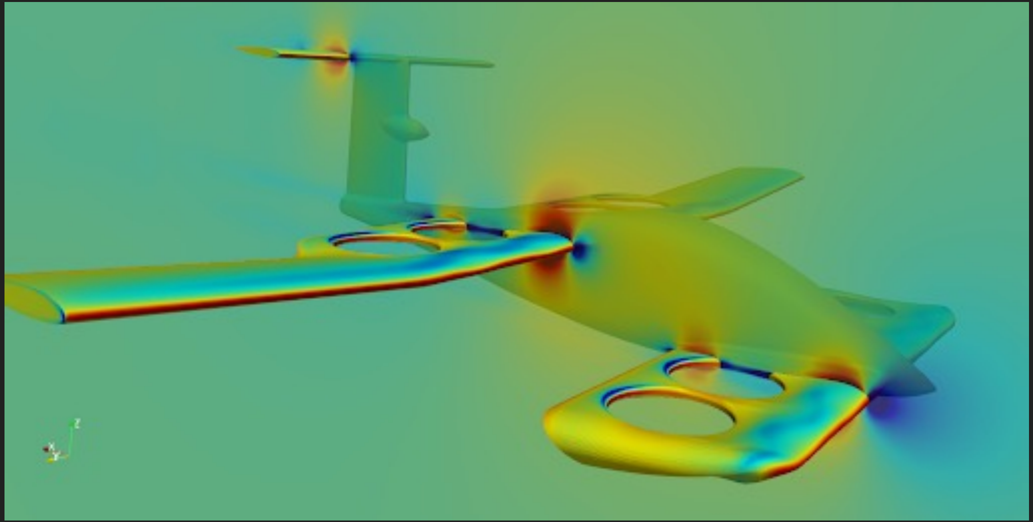




# Atea

Our aircraft empowers an atypical & innovative design.

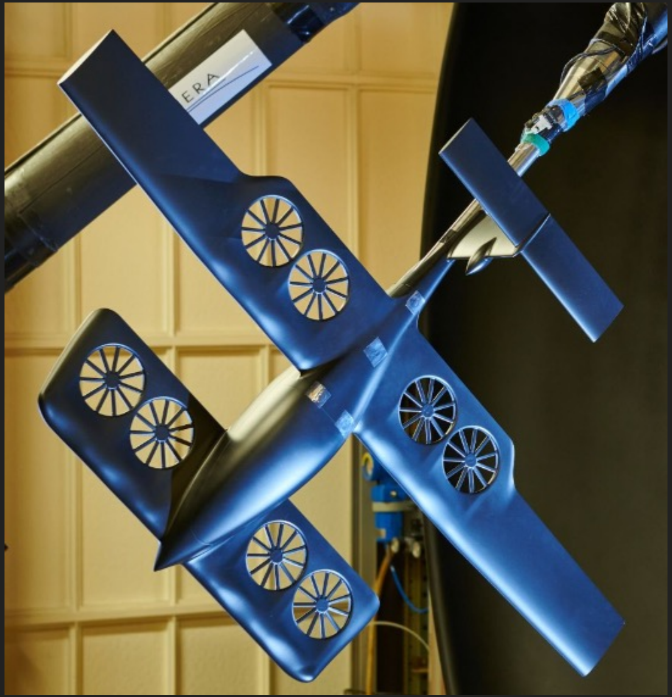
We need to master and validate all its aerodynamical aspects



# Atea

Our aircraft empowers an atypical & innovative design.

We need to accelerate convergence to a satisfying physical architecture



# Atea

Our aircraft empowers an atypical & innovative design.

We need to achieve all this on-time and within a reasonable budget.

# Atea

**In short, we believe that we need high-fidelity simulations from the beginning.**

We believe that our goals can be achieved through:

- An innovative CFD approach based on GPU architectures
- Performant solvers based on physicsML
- Scalable & compatible HPC infrastructure to master computing needs & budget