

# PRESENTATION TO **DISA TEM**

(OCIENT)<sup>TM</sup>



# AGENDA

**1**

**About Ocient**

**2**

**Technology  
Overview**

**3**

**Our Fit to DISA's Data  
Strategy**

**4**

**Sample Use Cases**

**5**

**Questions**

**6**

**Contact Info/Next  
Steps**

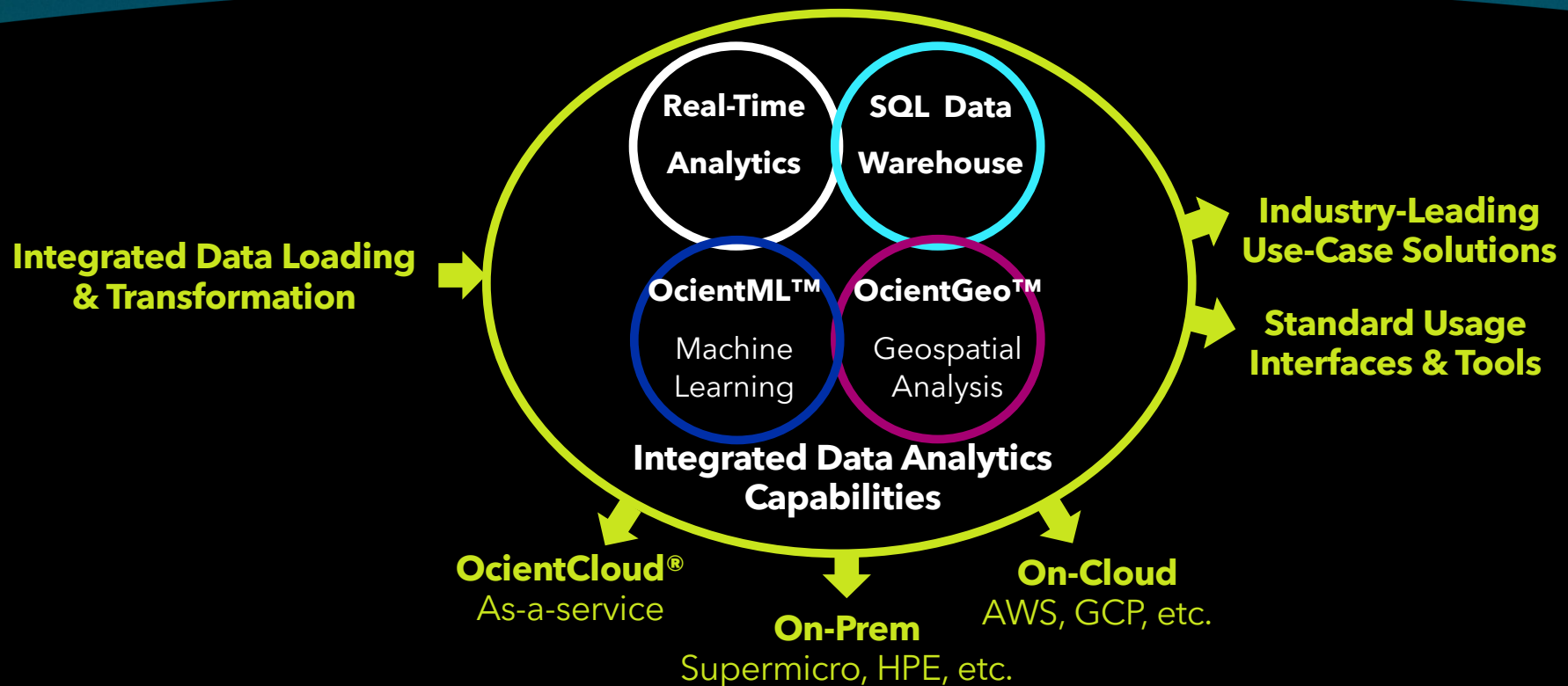
(OCIENT)™

1

# ABOUT OCIENT



# UNIQUE DATA ANALYSIS CAPABILITIES



*Same software, same capabilities; three deployment models*

# BACKGROUND

**2004-  
2015**

**Led** industry transformation of data storage



**2016**

**Founded** by request of the world's largest data analyzing enterprises

**2016-  
2023**

**Over \$130M** of financing

**OCA VENTURES**

**GREYCROFT** **iqt** IN-Q-TEL

**PSP** **V CAPITAL**  
PARTNERS

**Pritzker Group** **VALOR**  
VENTURE CAPITAL EQUITY PARTNERS

**Levy**  
Family Partners

**Riverwalk Capital, LLC**

**GREAT POINT**  
CAPITAL LLC

**BUOYANT**  
VENTURES

**WOLF CAPITAL**  
MANAGEMENT LLC

**2021**

**Product Launch after 300+ R&D person years**

Over 70 issued patents;  
Over 300 documented inventions

**2023**

Team of 160+ employees with 100+ engineers



# OCIENT DELIVERS SOLUTIONS

## OPTIMIZED BY USE CASE



### Telecommunications and Networks:

CDR Search & Analysis

*Video On-Demand Optimization*

ICR & IPDR Search & Analysis



### Climate & Environmental Intelligence:

*Weather, Climate and Environmental Analysis, Alerting, Modelling & Prediction*

Resource and Supply Chain Planning

*Insurance, Agricultural, Real Estate, Business and Financial Decision Making*



### Intelligence and Security:

Lawful Intercept

*Threat Hunting with Investigative Analysis*

Space & Satellites

*Financial Security*



### AdTech:

*Campaign Forecasting*

*Audience Analytics*

*Trade Area & Traffic Pattern Analysis*



### Data Fabrication:

*Data Assimilation & Publishing*

*LLM Data Preparation and Enhancement*

(OCIENT)<sup>TM</sup>

ORDERS OF MAGNITUDE FASTER  
THAN LEGACY SOLUTIONS, MORE  
COST EFFECTIVE THAN CLOUD  
SYSTEMS - AT *HYPERSCALE*

Complex and Continuous  
Hyperscale advanced  
analytics

- ✓ SQL Database Analysis
- ✓ Geospatial Data
- ✓ Machine Learning

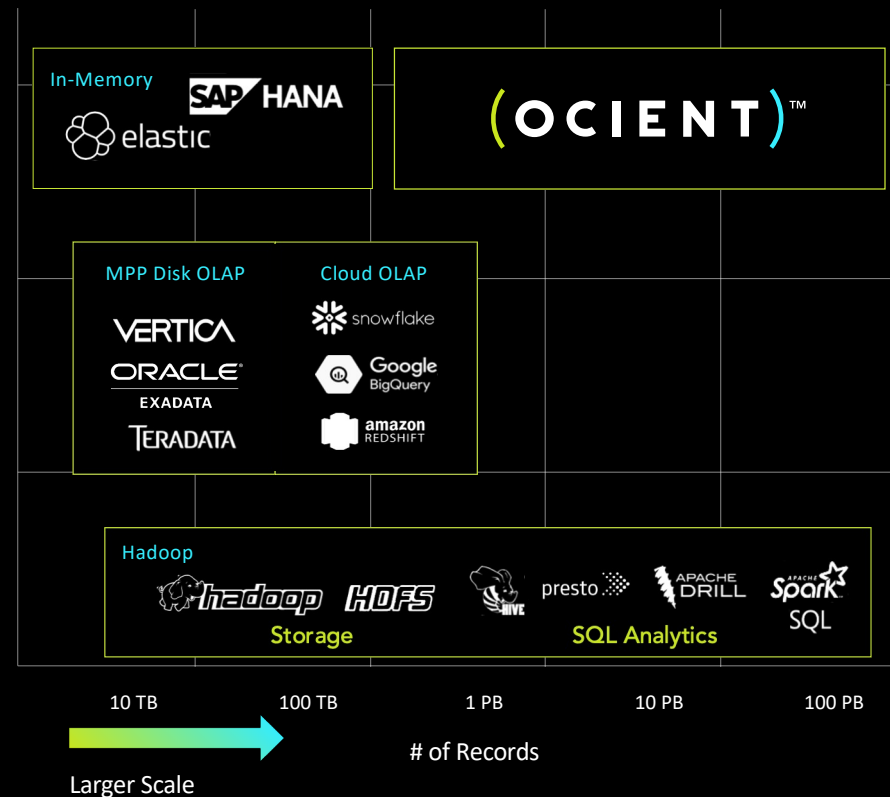
Faster Speed of Full  
Table Read and  
Aggregation



SECONDS

MINUTES

HOURS



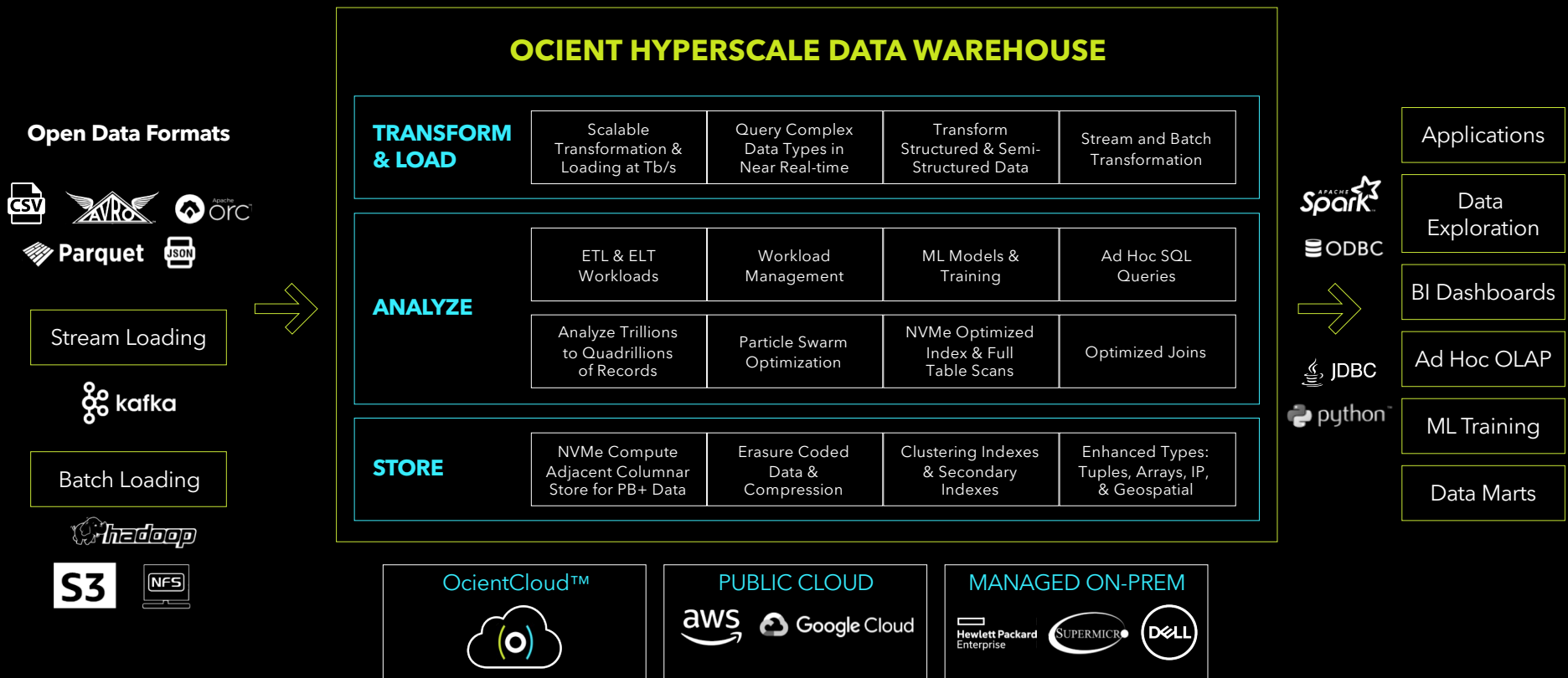
(OCIENT)<sup>™</sup>

2

# TECHNOLOGY OVERVIEW



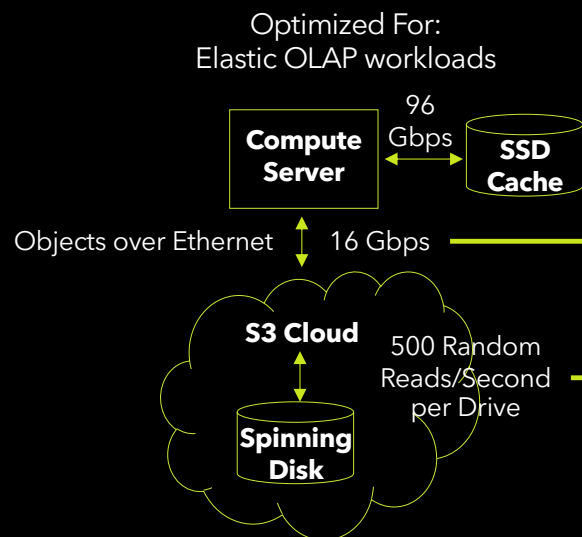
# COMPLETE SOLUTION STACK



# ROSA (REMOTE OBJECT) VS. CASA (STORAGE-ADJACENT)

## PERFORMANCE BENCHMARK ON 7 TRILLION NORMALIZED ROWS AT SAME COST

### Remote Object Storage Architectures: ROSA

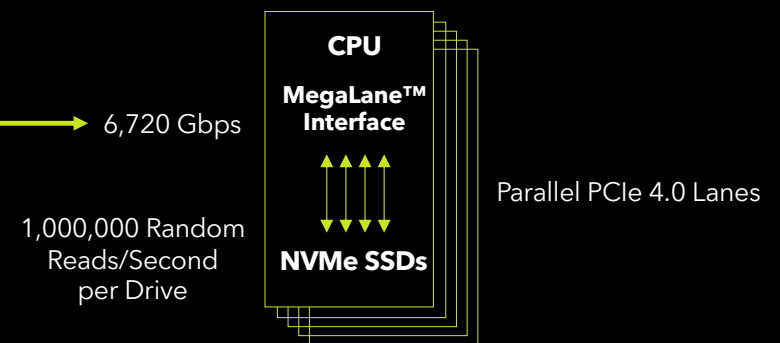


#### Drawbacks:

- Slow and expensive at hyperscale

### Ocient's Compute Adjacent Storage Architecture:™ CASA

Optimized For:  
Hyperscale OLAP workloads and AdHoc queries



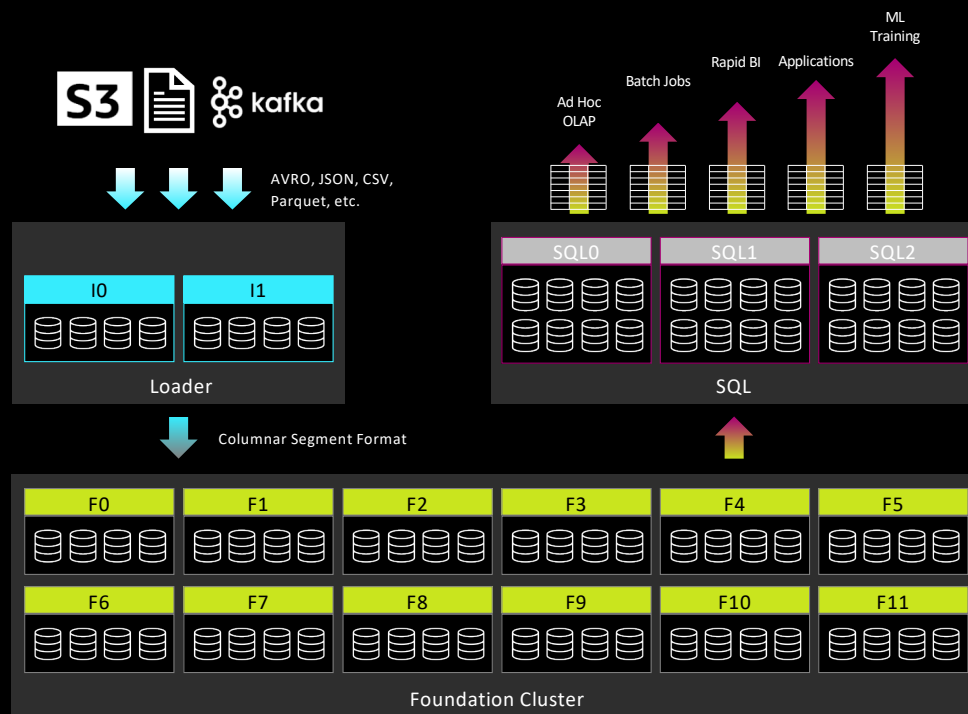
#### Advantages:

- Interactive-performance Hyperscale Data Analysis at up to 80% lower cost

**Queries  
10-1,000X faster  
than Snowflake, AWS  
Redshift, Databricks, etc.**

# EXAMPLE SYSTEM

## NOTIONAL SYSTEM ARCHITECTURE



|                    |                                                                                                                                                                                          |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Queuing / Kafka    | Receive, protect, queue                                                                                                                                                                  |
| Load and Transform | Accepting multiple, simultaneous streams of data; parse and transform from semi-structured to relational schema, protect, deduplicate and index - made available to query within seconds |
| Foundation         | Erasur storage protection and processing of data                                                                                                                                         |
| SQL                | Provides query results to JDBC / ODBC connections as well as performs admin / metadata duties                                                                                            |

| Node             | CoresSQL | Memory | Storage               | Qty |
|------------------|----------|--------|-----------------------|-----|
| Queuing          | 32       | 192GB  | 12 x 10TB HDD         | 2   |
| Loader           | 56       | 1TB    | 8 x 2TB NVMe          | 2   |
| Foundation       | 56       | 1TB    | 12 x 15.3TB NVMe      | 12  |
| SQL              | 56       | 1TB    | 4 x 2TB NVMe          | 3   |
| Total Core Count |          |        | 896                   |     |
| Raw Disk Storage |          |        | 2,200TB               |     |
| Protection       |          |        | 8 Data + 2 Parity + 2 |     |

System design is dependent on complexity of transformation, sampling rate, scale, number of mixed workloads, concurrency etc. Workshops will be used to determine an appropriate BOM and build the financial and technical business case based on solution requirements both now and in the future.

(OCIENT)<sup>™</sup>

3

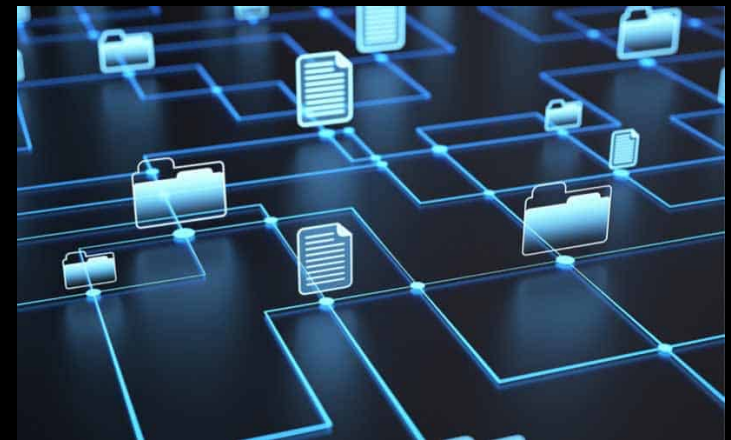
# OCIENT'S FIT TO DISA'S DATA STRATEGY



# LINE OF EFFORT #1

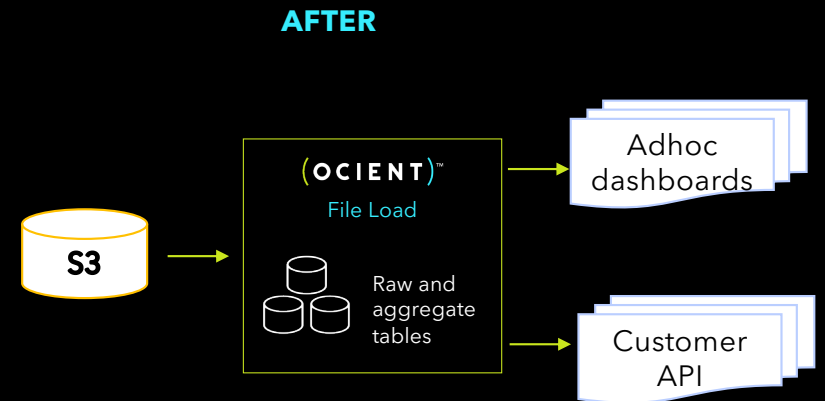
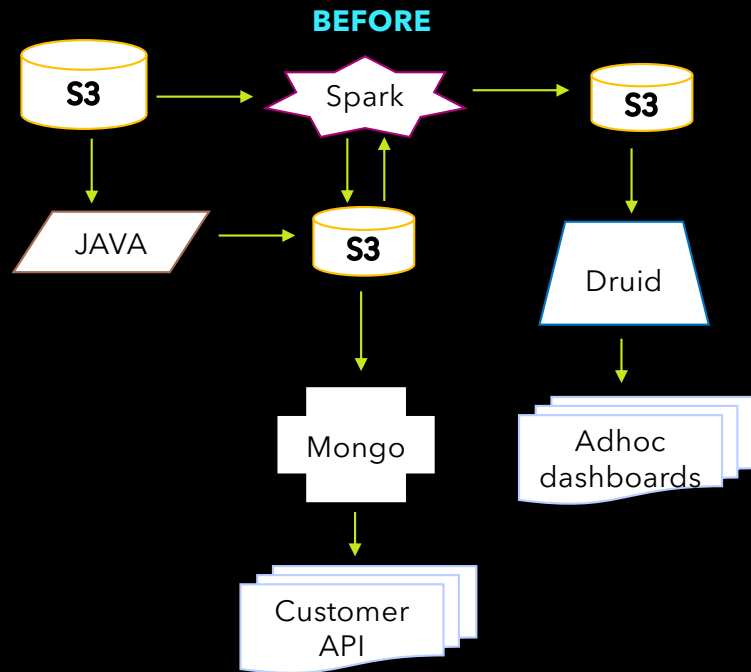
## DATA ARCHITECTURE

- DISA wishes to support data assets through their entire life cycle
- Data life cycle needs to be understood in context of the volume, richness, and time-criticality of said data
- Larger datasets will be more complex to support, may need special tools to extract full value
- Some datasets lose value as they age
- **The right architecture can help with all of this!**



# ETL IS HARD

## OCIENT MAKES IT EASIER



...

# WORKLOAD MANAGEMENT ALLOWS SEAMLESS RESOURCE SHARING

## ENABLE COLLABORATION WITHOUT IMPACTING MISSION CRITICAL ACTIVITIES

CASE  
STUDY

| TYPE                   | CORES                                                                                                                           |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| scheduling_priority    | The priority given to queries when executing. This is a proxy for how much CPU time a query receives relative to other queries. |
| max_concurrent_queries | Maximum number of concurrent queries running for the service class.                                                             |
| max_elapsed_time       | The max time a query can run in seconds.                                                                                        |
| max_temp_disk_usage    | Max percentage of temp disk that a query can use when spilling intermediate result sets to disk.                                |
| max_rows_returned      | Maximum rows that may be returned for a query.                                                                                  |
| result set caching     | Size and duration of query results to ensure sub-second results                                                                 |

### High Priority Service Class



Campaign  
Forecast +  
Price Curve  
API



Campaign  
Exec  
Service



Inventory  
Analysis  
API



PMP Deal  
Troubleshooting  
API

### Medium Priority Service Class



Bid Shading  
+ Freq  
Capping  
Analysis



ML Model  
Training &  
Verification



Nightly  
Inventory  
Analysis

### Lower Priority Service Class



Bidding Ops  
Analytics

AdHoc Queries  
Realtime Monitoring  
Fraud Detection  
Delivery Analysis



Partner Data  
Extracts

# LINE OF EFFORT #2

## ADVANCED ANALYTICS

- DISA wishes to harness the predictive power of data
- Accurate predictions depend on access to full-resolution data
- Increasingly, this will mean hyperscale operations on exotic data types
  - Trillions of records
  - PBs of data
- Oh, and make it all ML-ready
- **Ocient delivers interactive-time results on these workloads**





# MEASUREMENTS FROM PROOF OF TECHNOLOGY

## SHOCK AND AWE

| Metric                                 | Rate                                                                     | Notes                                                     |
|----------------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------|
| SQL node returning results             | 28Gbps x 2 (35% CPU)                                                     | Equivalent to ~16TBs an hour outbound                     |
| Queries per second                     | 344/sec (20k in 58 seconds)                                              | All 20k views reported in < 1 minute                      |
| Filter rate (20 LTS nodes)             | 1.4T rows/sec                                                            | Analyze 4.4 years of flights per second                   |
| LAT sustained rate                     | 2.05M rows/sec = 3.5Gbps (53% CPU)                                       | Load 2.5TBs per hour per loader                           |
| Concurrent loading and querying        | ~2.5% increase in query elapsed time loading 2.9M rows/sec vs 0 rows/sec | Compare to other vendors showing 20-70% query degradation |
| Multi-column sorted aggregate at scale | Return min, max, avg speed/altitude across 3.6B records in 0.5 seconds   | Profile worldwide flight points across 8 days in 1 second |
| Efficient row removal                  | Delete ~4 trillion rows = 8 seconds                                      | Wipe 1.6 years of flights in a second                     |

*Assume 10k aircraft in the air sending data every second*

# CUTTING-EDGE GEOSPATIAL ANALYTICS

> 120 UNIQUE FUNCTIONS



## • NATIVE GEOSPATIAL AND COMPLEX COLUMN TYPES

- Geospatial Types
  - Points
  - LineStrings
  - Polygons
- Complex Types
  - Arrays
  - Tuples
- Standard latitude/longitude WGS84 Ellipsoid model



## • NATIVE GEOSPATIAL FUNCTION SUPPORT

- 120+ Geospatial functions
- Based on Non-Euclidian Spherical Geometry
  - Leads to minimal distortion vs projections
- In-Database Machine Learning
- Unique Spatiotemporal functions answer when *and* where

- Enabling organizations to store and analyze geospatial data at hyperscale natively in Ociant
- Enabling analysis of streaming geospatial data that is queryable seconds after ingest
- Enable complex analysis with numerous native Geospatial functions
  - ST\_CONVEXHULL
  - ST\_PROJECT
  - ST\_CLUSTERDBSCAN
  - ST\_INTERSECTION
  - 100+ more

# MACHINE LEARNING IN PLACE

## ENABLING HYPERSCALE MACHINE LEARNING



- **TRAIN IN PLACE**

- Train ML models **without moving data**, enabling **larger training sets** and better **data security**.



- **ANALYZE AND ITERATE**

- Choose and customize data analysis using a large toolbox of **ML algorithms** in multiple categories:

|                     |                                                                                                                                                 |
|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| Regression          | Simple Linear, Multiple Linear, Polynomial, Linear Combination, Logistic, Nonlinear                                                             |
| Classification      | K-Means Clustering, K Nearest Neighbor, Naïve Bayes, Feedforward Neural Network, Decision Tree, Support Vector Machine, Gaussian Mixture Models |
| Feature Engineering | Principal Component Analysis, Linear Discriminant Analysis                                                                                      |



- **EVALUATE AT LOAD (FUTURE)**

- Evaluate trained ML models **at load time** as part of an ETL pipeline. **Classify, filter, and/or enrich** streaming or batch loaded data in **one step**.



- **UNIFY OLAP SQL AND ML**

- ML as a first-class citizen in the database enables **full expressivity of SQL** for data analysis with ML and linear algebra mixed in

(OCIENT)<sup>TM</sup>

4

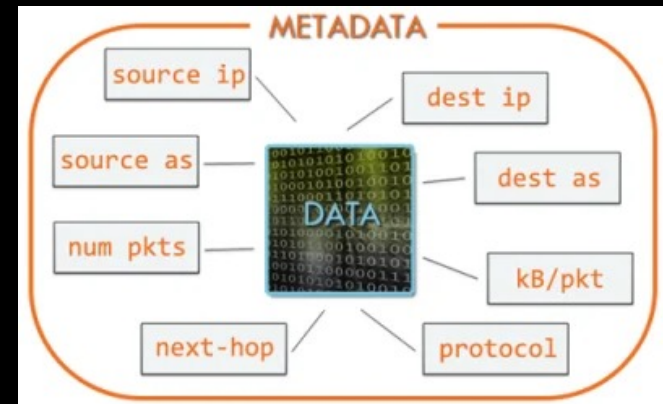
# SAMPLE USE CASES



# NETWORK TRAFFIC ANALYSIS

## SEARCH AND ANALYZE YEARS OF NETWORK METADATA

- Built-in support for IP address data types
- Support for Netflow/IPFIX records
- Existing integrations with top probe vendors such as Gigamon, SS8, NetQuest
- Scalable ingest to handle millions of records/sec
- Interactive-time query processing can handle comparisons across years of cyber data
- Built-in ML enables rapid model iteration and full-resolution training to enable accurate predictive analytics



# DIGITAL WARFARE IN THE 21<sup>ST</sup> CENTURY

- [illegible]

(OCIENT)<sup>™</sup>

5

**QUESTIONS?**



# DIFFERENTIATED CAPABILITIES

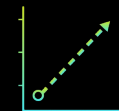
**OCIENT DELIVERS  
BEYOND ALTERNATIVES**

(OCIENT)<sup>™</sup>



## Breakthrough Performance

Transform & load data at terabits per second and run queries 10-100x faster



## Increased Scale

High core count systems that harness petabytes of data / trillions of records



## New Hyperscale Capabilities

ETL/ELT plus full ANSI SQL plus Geospatial plus AI/ML all at Hyperscale in Near Real Time



## Reduced Complexity

Support multiple workloads and concurrent users from a single platform



## Accelerated Deployment

Get to market faster by leveraging Ocient's turnkey solutions on-prem, cloud or in OcientCloud



## Lower Cost

Cut costs by 50% or more while upgrading to a modern hyperscale data solution

(OCIENT)<sup>™</sup>

6

**CONTACT INFO/  
NEXT STEPS**





# ADDITIONAL OCIENT INFORMATION

Acquisition Contracts and Hardware Options for On Prem deployments

## Government Approved Contract Vehicles

<https://www.carahsoft.com/ocient>

### **1. NASA SEWP V**

*NNG15SC03B/NNG15SC27B  
May 01, 2015- Apr 30, 2025*

### **2. ITES-SW2**

*W52P1J-20-D-0042  
Aug 31, 2020- Aug 30, 2025  
\*Additional Option Years Available*

## Hardware Options for On Prem Deployments

We are strategic partnerships with:

1. Ociant Appliance - SuperMicro
2. DELL
3. Hewlett Packard

Please be sure to consult with Ociant to build and confirm the appropriate configurations for the Servers with NVMe drives prior to acquisition.



## CONTACT OUR TEAM

**BILL MINARCHI / FEDERAL SALES  
LEADER**

[bminarchi@ocient.com](mailto:bminarchi@ocient.com)

(571) 528-3696

**JOHN TEMPLE / DOD SALES**

[jtemple@ocient.com](mailto:jtemple@ocient.com)

(703) 980 - 5428

**ADRIAN BURKE / FEDERAL  
SOLUTIONS ARCHITECT**

[aburke@ocient.com](mailto:aburke@ocient.com)

(817) 937-4879

**TONY IBÁÑEZ / FEDERAL  
SOLUTIONS ARCHITECT**

[tibanez@ocient.com](mailto:tibanez@ocient.com)

(443) 699-3700

(OCIENT)<sup>™</sup>

# THANK YOU

