Customer Driven Database Design

ORACLE OPEN WORLD

October 1–5, 2017
SAN FRANCISCO, CA

2 Trillion Rows of DNS Query Volume

Charlie Baker
Sr. Director Product Management
Oracle Dyn GBU
October 4, 2017

charles.baker@oracle.com

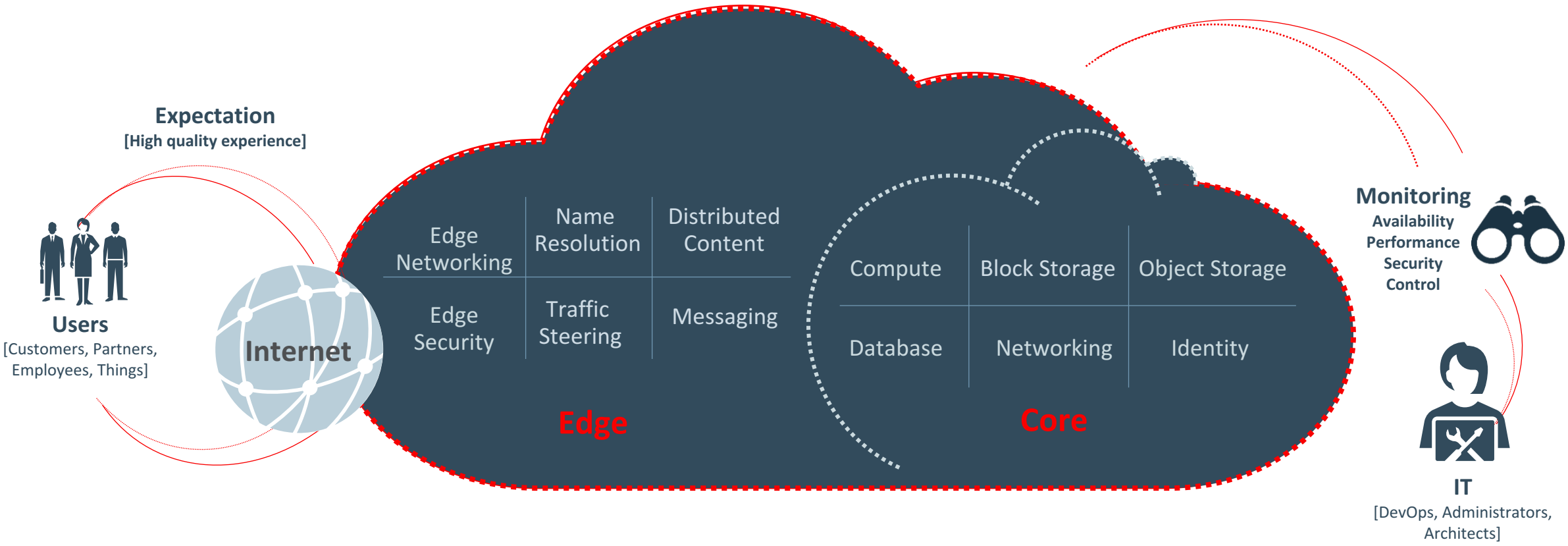
ORACLE

Quick Agenda

1. Who is Oracle Dyn
2. The Problem
3. Solutions Evaluated and Benchmarks
4. Deployed Solution Architecture
5. Benefits and Next Steps

Dyn Strengthens Oracle Cloud Infrastructure

Mission | Create the World's Most Capable Enterprise Cloud Infrastructure from Edge to Core

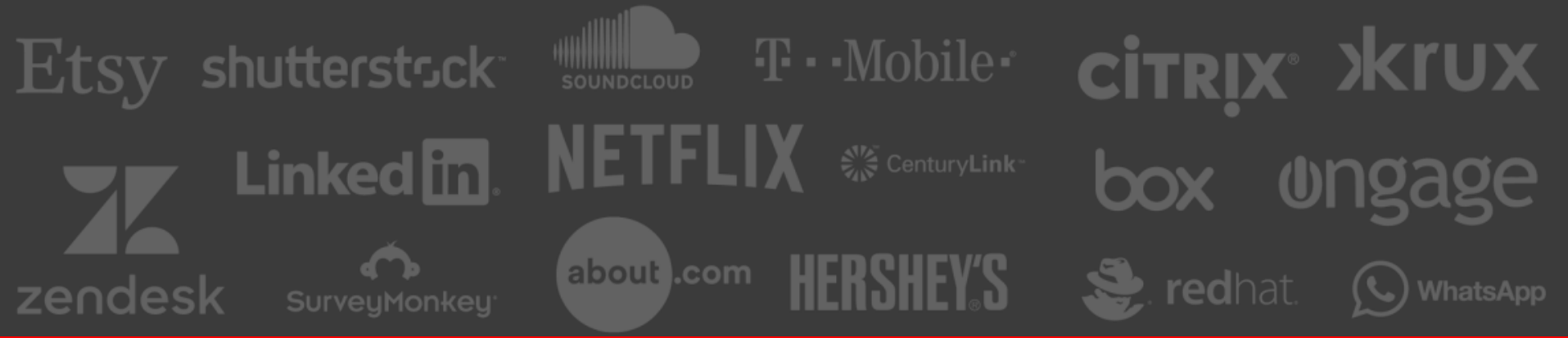


The Next Generation Cloud

Trusted by Over 3,500 Customers, Including Some of the Most Preeminent Digital Brands

ORACLE + Dyn

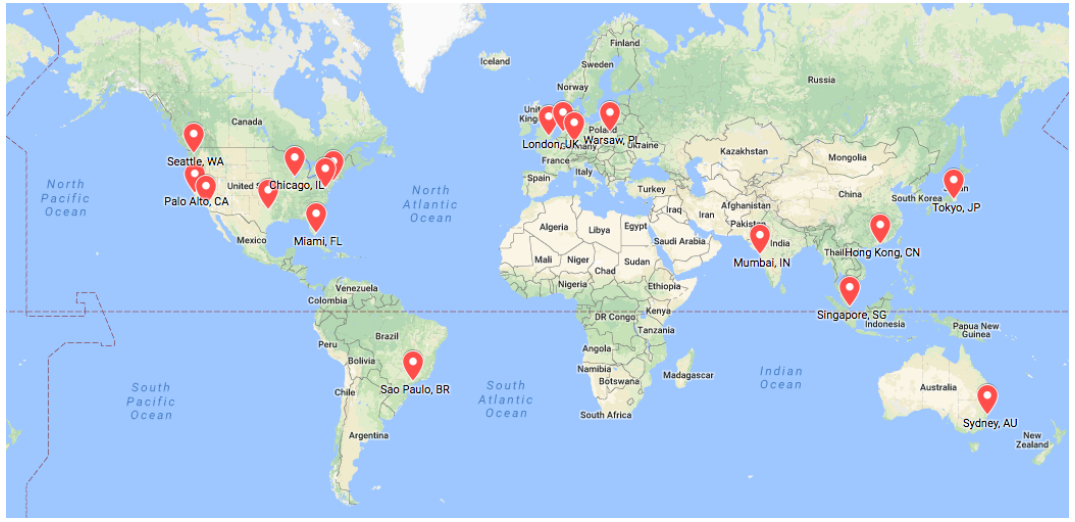
Over 40 Billion DNS Queries Per Day
Over 214 Billion data points - powering analytics and steering decisions



Visit dyn.com/OOW17 for more information and downloads.
Visit our product experts in the Oracle Cloud Infrastructure expo

The Problem – High-Scale Billing, Reporting & Analytics

- Globally Distributed Name servers answering DNS Queries
  - Fully attributed with ASN, GEO-ID, Country Code
- Data Volume Generated:
  - Estimated at 2 Trillion Rows for 90 days of query data
  - Raw data of 1.6 TB/day working out to over 144 TB across 90 days
- Query Latency Measured across Multiple Systems:
  - UI
    - Splash page queries < 5s
    - Detailed investigation queries <20s
  - API
    - Single customer queries <20s
    - Multiple customer queries (billing/metering) <1min

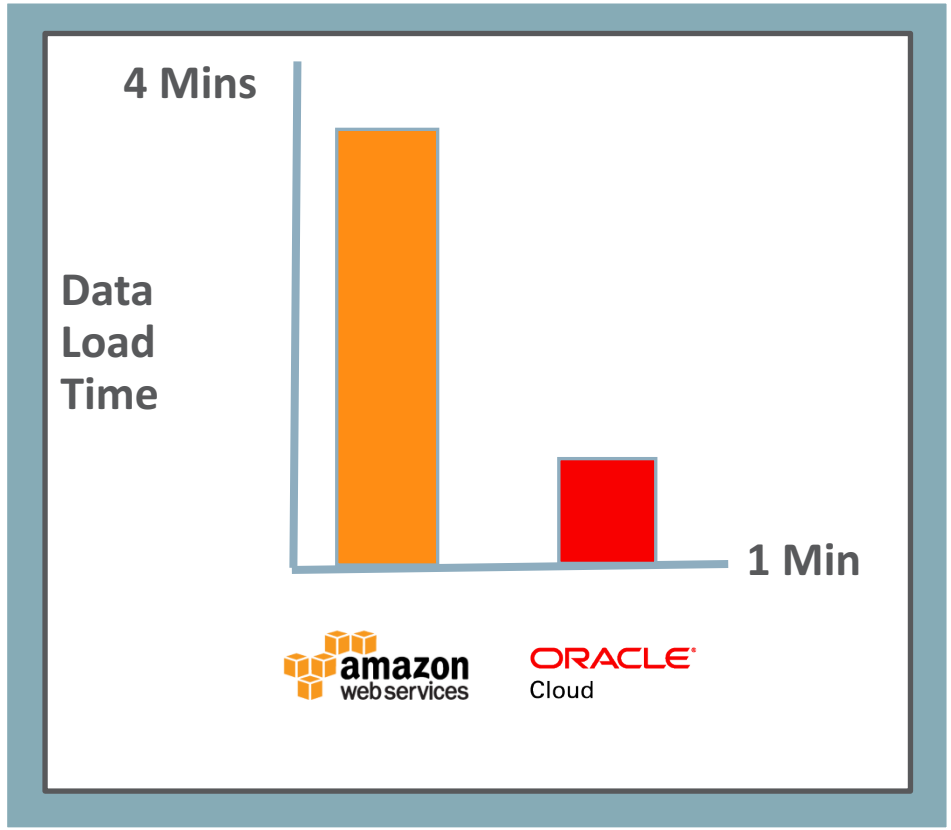


Solutions Evaluated

Amazon Redshift Amazon EMR		Complex system design Data loads could not keep up with incoming data volume Queries were unacceptably slow
druid		Could ingest and index at scale Required large shared storage Difficulties running multiple GROUP BY queries
Apache CASSANDRA		Limited expertise at Dyn Unable to support complex queries, forcing application to implement complex query processing logic
ORACLE Cloud - **Single Database**		Good Performance and Ingest speed Complex query support Performance and Storage limited to Single Database
ORACLE Cloud - **Sharded Database**		Superior Ingest speeds Scale queries by distributing across as many servers as needed (up to 1000) Near limitless TB of storage by adding more Shards (up to 1000)

Evaluation Results

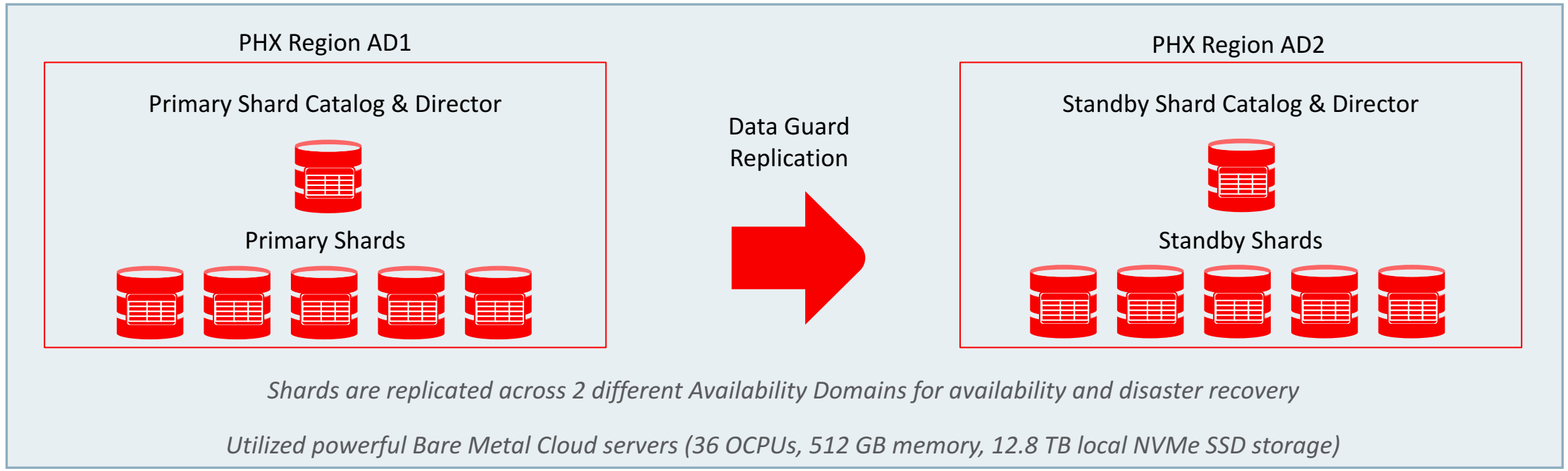
- Loading from S3 to Redshift
  - 3 to 4 minutes
- Loading from Oracle Object Storage to Shards
  - 500k rows / second
  - Full load in 1 minute
  - 4 minutes of headroom for queries



Deployed Solution

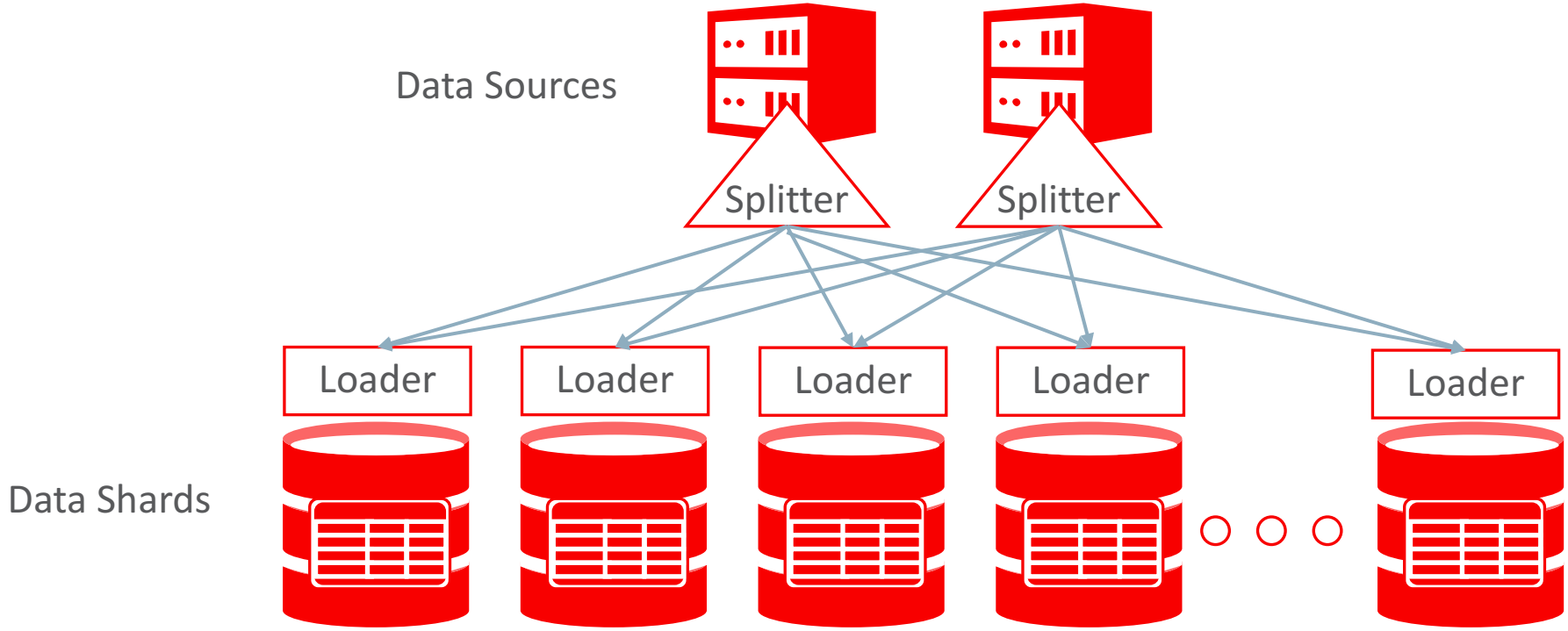
Oracle Sharded Database

- Proven linear scalability
- Ingest speeds scale with number of shards
- Constant query time even as we grew size of dataset
- Geo-distributed to be close to our customers



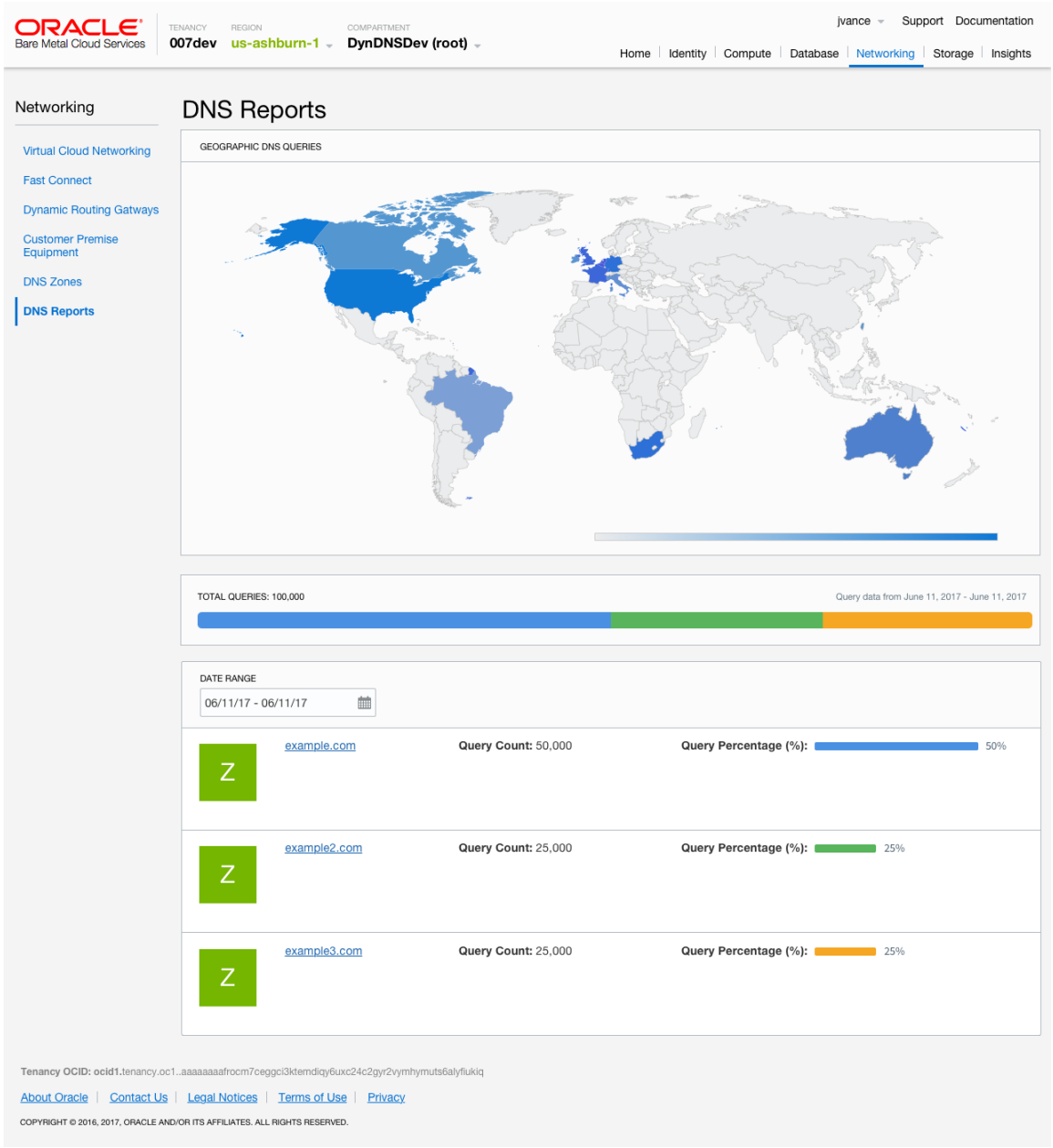
High Speed Data Load

- Utilized fully parallel direct-to-shard data loader
- Sharded architecture scales the CPU, flash, and network interfaces
- Can add shards as needed to accommodate higher ingest rates
- Architecture applies to IIoT, IoT and edge compute scenarios



Testing the User Interface

- Basic query time < 2 seconds
  - Total requests over period of time
- Complex query time <20 seconds
  - Requests by geography, operator, recursive over period of time



Wrap Up

- Oracle Sharding Allows Dyn to…
  - Deliver fast, real-time analytics to our customers, billing and analytics systems
  - Dynamically scale our computing infrastructure to handle ever-growing ingest rates, while keeping query times constant even though the data volume grows
  - Easily support geographic data distribution so that we can keep data closer to our customers
  - Collect more data per customer, so that we can make better traffic steering decisions and provide more value to customers
  - Create our data lake for analytics – including recursive, RUM, synthetic, remote access, CDN and WAF data sets
- Oracle Dyn DNS and Email services officially launch at OOW – leveraging the sharded DB