

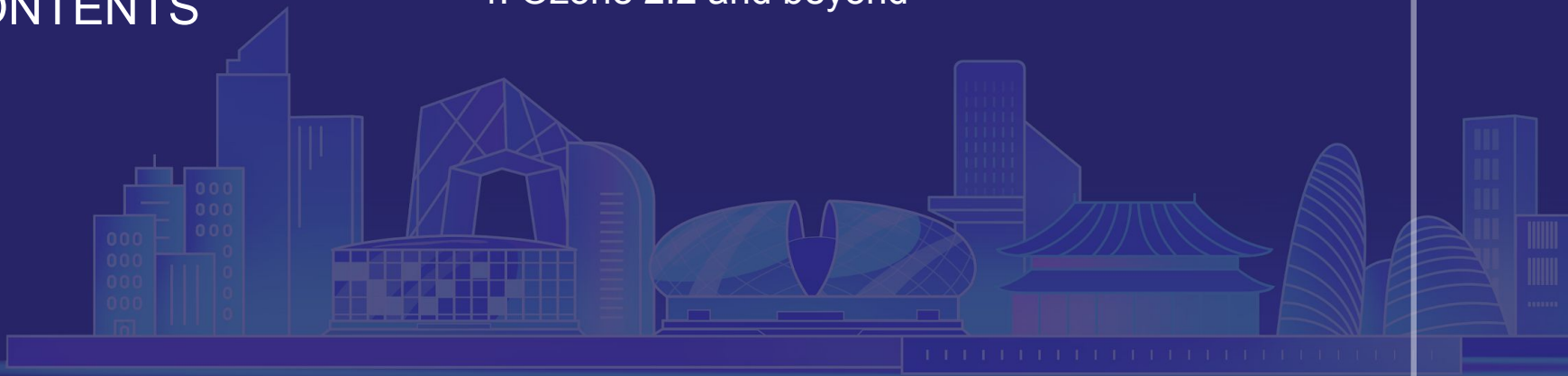
# Apache Ozone 2.0

Wei-Chiu Chuang



## CONTENTS

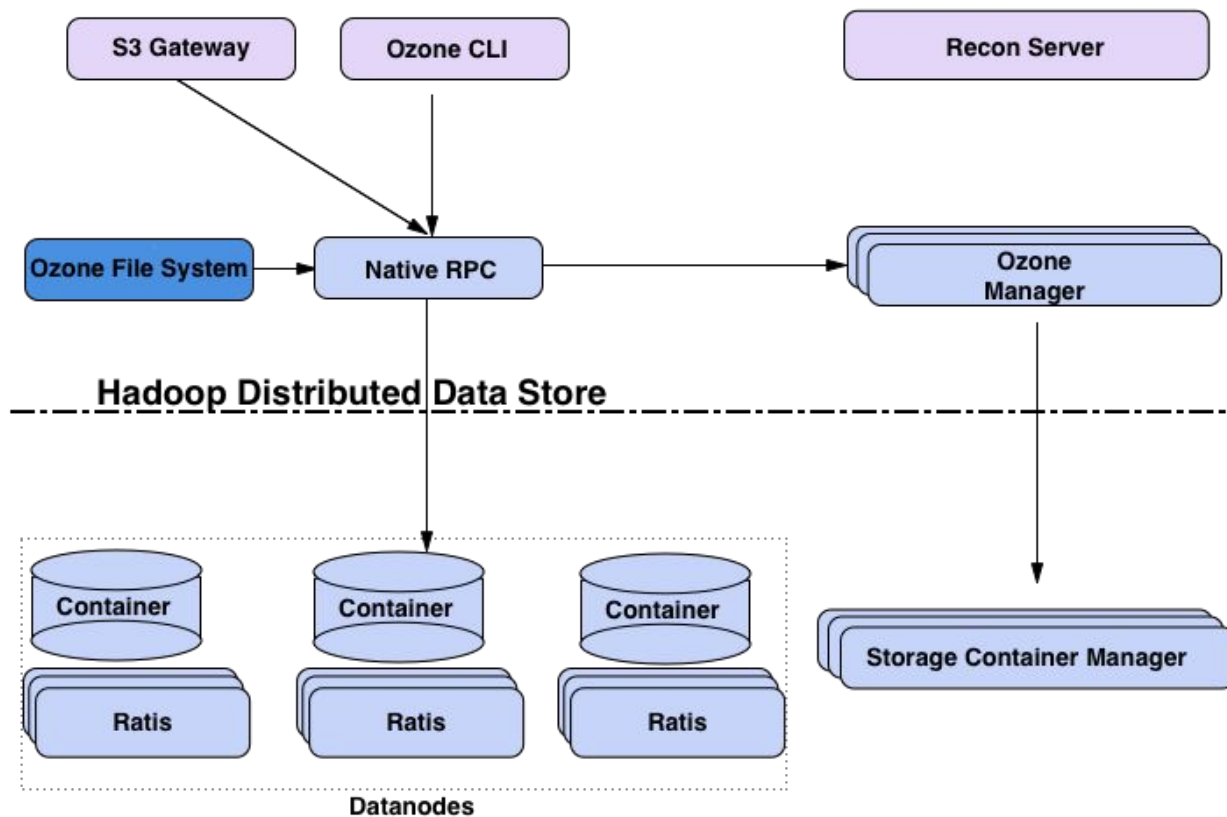
1. Ozone State of Union
2. Ozone 2.0
3. Ozone 2.1
4. Ozone 2.2 and beyond



# What is Ozone

Distributed object store  
“data lake”, “data lakehouse”

# What is Ozone



## Why Ozone

Petabyte scale *per node*

100PB+ cluster scale, tens of billion keys

Apache license

Optimized for analytics workloads

## Why Ozone 2.0

- Ozone 1.x was more about getting feature parity with HDFS.
- With 2.0 and beyond, we are innovating beyond HDFS.

## How did we get here?

2014/10 - Ozone concept proposed ([HDFS-7240](#))

2018/4 - HDDS Subproject founded under Hadoop

2019/10 - Apache Ozone TLP founded

2020/09 - Apache Ozone 1.0

2025/04 - Apache Ozone 2.0

2025/09 - Apache Ozone 2.1 (in progress!)

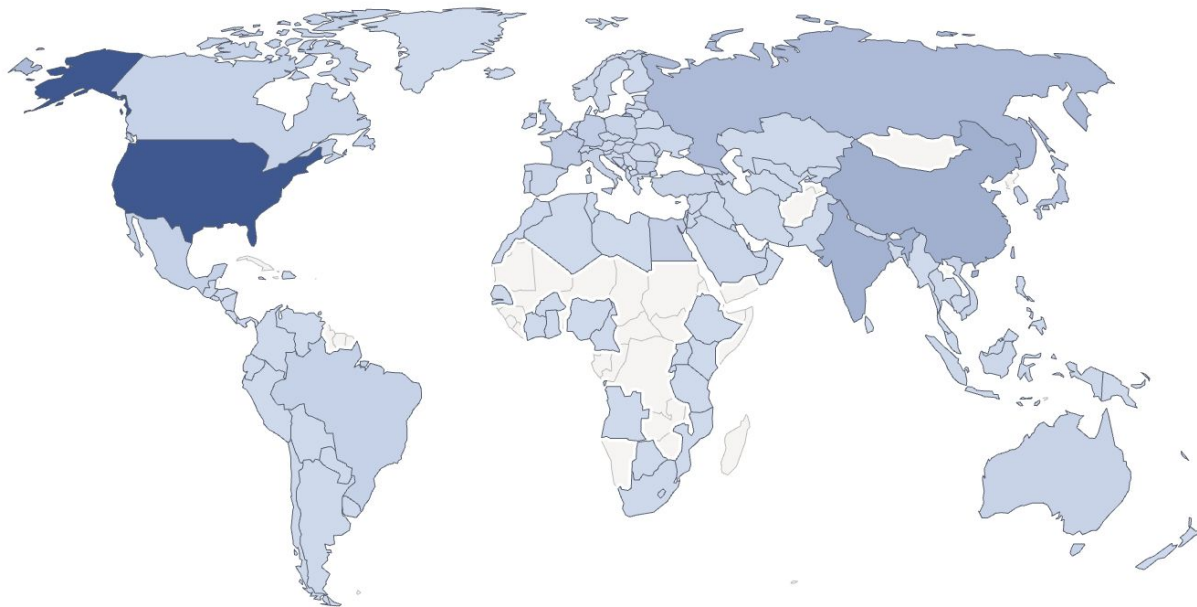
## What have we learned?

- Distributed object store is *really* hard to develop
- Ecosystem shift
  - Hadoop-centric → Kubernetes, S3 centric
- New applications
  - Flink, Trino, Iceberg
- Hardware trend: Super dense storage nodes, NVMe
- Day 2 operations seems boring, but they are essential
- Supercycle: AI



# Where are Ozone users

35,138 visits



  Countries

Worldwide

Visits

- USA
- China
- India
- Russia
- Taiwan
- Japan
- France
- UK
- Hong Kong
- Germany

...

COMMUNITY  
THE ASF CONFERENCE  
**CODE**

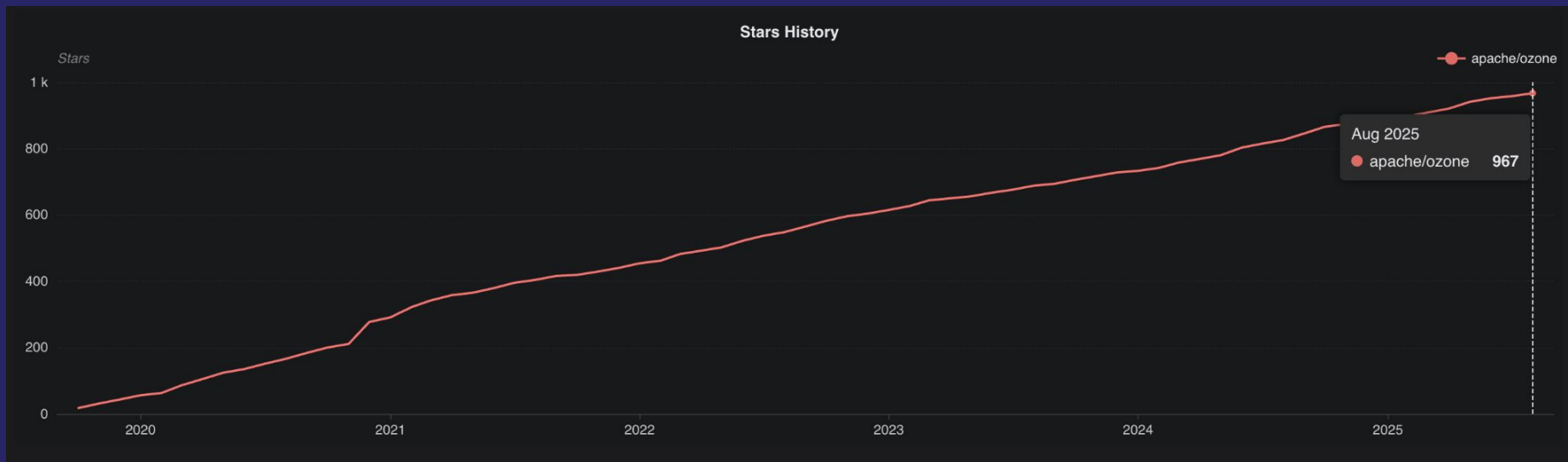
# Community growth

1000th github stars in September 2025!



Starred

1k



<https://ossinsight.io/analyze/apache/ozone#overview>

COMMUNITY  
THE ASF CONFERENCE  
CODE

# Community

- Preferred Networks
- Shopee
- DiDi <https://ozone.apache.org/blogs/>
- Qihoo
-

## Ozone 2.0

Apache Ozone 2.0.0 adds 1708 new features, improvements and bug fixes on top of Ozone 1.4.

Notable features:

- [HDDS-7593](#) Supporting HSync and lease recovery
- [HDDS-7852](#) SCM Decommissioning Support
- [HDDS-8829](#) Symmetric Keys for Delegation Tokens
- [HDDS-10656](#) Atomic Key Overwrite and Key Replacement

# Ozone 2.0

## Other noteworthy changes:

- [HDDS-6263](#) ARM64 support
- [HDDS-8246](#) Java 11/17/21 Support and Testing
- [HDDS-11577](#) AWS SDK v2 client support



## Observability:

- [HDDS-9307](#) Ozone performance and operational dashboards
- [HDDS-11153](#) [UI] Recon UI Improvements
- [HDDS-11825](#) Support interactive mode for ozone CLI

## Developer:

- [HDDS-6729](#) JUnit 4 to 5 upgrade
- [HDDS-8113](#) Dropped Hadoop 2.7 ~ 2.9 support.
- [HDDS-10986](#) Publish SBOM artifacts
- [HDDS-11617](#) Hadoop dependency was updated to 3.4.1

# Overview

Auto Refresh ☒ | Refreshed at 12:45 PM  | DB Synced at 12:45 PM 

Health

Datanodes

Containers

✓ Healthy

✓ Healthy

	Available	Actions
Datanodes	1/1	<a href="#">View More</a>
Containers	0/0	<a href="#">View More</a>

Cluster Capacity

51%

29.8 GB / 58.4 GB

Usage	Size
Ozone Used	4 KB
Non Ozone Used	29.8 GB
Remaining	28.5 GB
Container Pre-allocated	0 B

Volumes

[View More](#)

2

Buckets

[View More](#)

1

Keys

0

Pipelines

[View More](#)

1

Deleted Containers

0

Open Keys Summary

[View Insights](#)

Total Replicated Data	N/A
Total Unreplicated Data	N/A
Open Keys	N/A

Delete Pending Keys Summary

[View Insights](#)

Total Replicated Data	N/A
Total Unreplicated Data	N/A
Delete Pending Keys	N/A

## Future plan

### Innovate, faster

- Minor release every 6 months.

### More community building

- LinkedIn page <https://www.linkedin.com/company/apache-ozone>
- X page <https://x.com/ApacheOzone>
- The Apache Blog
- More frequent meetups

### More user-focused

- Release notes, user documentation

# Ozone 2.1

## Stability, Day 2 operations

### New Features

- [HDDS-10685 Short circuit read support in Ozone](#)
- [HDDS-10239 Storage Container Reconciliation](#)
- [HDDS-5713 Add a disk balancer for datanode](#)
- [HDDS-11523 Support listener OM](#)
- [Replace gofys with mountpoint-s3](#)

### Improvements

- [HDDS-12940 Ozone Snapshot Phase 3: Scale up Snapshot Operations](#)
- [HDDS-8387 Container and volume scanners phase II](#)
- [HDDS-12564 Handling disk issues in Datanodes - Phase II](#)
- [HDDS-12716 Ozone S3 gateway Phase 4](#)

### Packaging

DEB/RPM packages

ARM builds



## Ozone 2.2 and beyond

We've been super focused on HDFS feature parity, getting users onboard, and didn't realize the world has evolved.

- A single shared object store across the entire enterprise
- More comprehensive S3 API support
- Modernize the platform
- Iceberg
- *AI*

## Ozone ❤️ AI

\* Not roadmap, but items I think would make sense

- AI assisted code development
- MCP Server
- RAG with Ozone <https://www.youtube.com/watch?v=fVGV3oYlyqw>
- GPUDirect, RDMA

# Ozone 2.2 and beyond

## Innovation

- [HDDS-13513 Event notification support in Ozone](#)
- [HDDS-8342 S3 Lifecycle Configurations - Object Expiration](#)
- [HDDS-11233 Ozone Storage Policy Support](#)
- [HDDS-13323 STS - temporary, limited-privilege credentials service](#)
- S3 Vectors
- [HDDS-12289 Support Trino](#)

## Performance

- [HDDS-8691 Ozone Streaming Write Pipeline Phase 2](#)
- [HDDS-11897 Migrating Ozone Manager replication from post Ratis execution to Pre Ratis execution.](#)
- Replace Hadoop RPC



Thanks

