Taking Omid to the Clouds:
Fast, Scalable Transactions for Real-Time Cloud Analytics

Ohad Shacham
Yahoo Research
ohads@verizonmedia.com

Yonatan Gottesman
Yahoo Research
yonatang@verizonmedia.com

Aran Bergman
Technion
aranb@campus.technion.ac.il

Edward Bortnikov
Yahoo Research
ebortnik@verizonmedia.com

Eshcar Hillel
Yahoo Research
eshcar@verizonmedia.com

Idit Keidar
Technion and Yahoo Research
idish@ee.technion.ac.il

ABSTRACT

We describe how we evolve Omid, a transaction processing system for Apache HBase, to power Apache Phoenix, a cloud-grade real-time SQL analytics engine.

Omid was originally designed for data processing pipelines at Yahoo, which are, by and large, throughput-oriented monolithic NoSQL applications. Providing a platform to support converged real-time transaction processing and analytics applications – dubbed *translytics* – introduces new functional and performance requirements. For example, SQL support is key for developer productivity, multi-tenancy is essential for cloud deployment, and latency is cardinal for just-in-time data ingestion and analytics insights.

We discuss our efforts to adapt Omid to these new domains, as part of the process of integrating it into Phoenix as the transaction processing backend. A central piece of our work is latency reduction in Omid's protocol, which also improves scalability. Under light load, the new protocol's latency is 4x to 5x smaller than the legacy Omid's, whereas under increased loads it is an order of magnitude faster. We further describe a *fast path* protocol for single-key transactions, which enables processing them almost as fast as native HBase operations.

1 BIBLIOGRAPHIC REFERENCE

2 KEY ASPECTS OF THE SOLUTION

- Optimizing Omid for low latency.
- Redesigning Omid to eliminate its principal bottleneck.
- Novel fast-path API for single-key transactions.
- Integrating Omid as the transactional layer in Phoenix SQL engine.

3 WHY A HIGHLIGHT PRESENTATION?

- Real-world solution: used in production in a popular platform, as part of an Apache incubator project.
- Pushes the state-of-the-practice in transaction processing systems.
- Addresses challenges that arise in modern use cases.