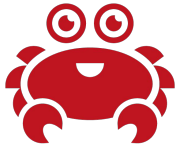


aiven

Apache Kafka Cloud Performance

Heikki Nousiainen
@hnousiainen

Speaker

- **Heikki Nousiainen, @hnousiainen**
- **CTO, co-founder @ Aiven**
- **Kafka user since 2014 (rel 0.8)**

Aiven

#1 independent database as a service provider in all major clouds

Based in Helsinki, Finland and Boston, MA

PostgreSQL, Kafka, ZooKeeper, Elasticsearch, Redis, InfluxDB and Grafana in 67 regions around the world

Cassandra and ScyllaDB to be launched in 2018 Q1

Kafka Performance

Kafka Performance Promise

Kafka is a **high-performance**

…streaming platform

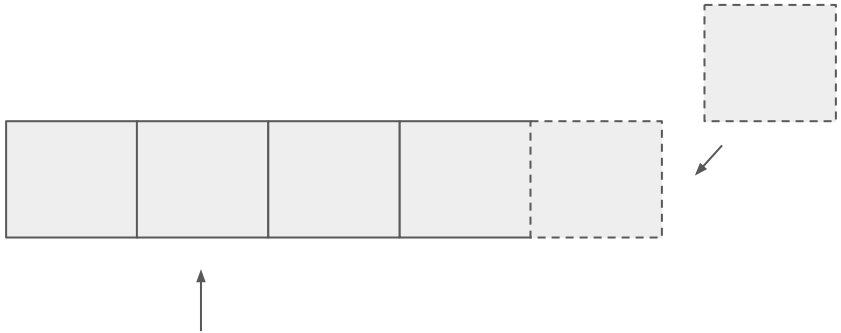
…distributed message queue

…distributed log

Kafka Performance Promise

Kafka is an append-only log.

State / cursor is held by consumer.

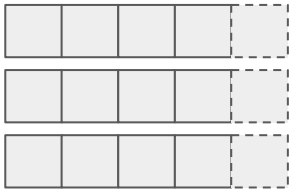


Kafka Performance Promise

Kafka data is organized in **topics**

Topics are divided into one or more **partitions**

Partitions are independent from each other



Scaling by adjusting number of partitions

Kafka Performance Factors

1. Partition layout
2. Replication levels
3. Producer acknowledgement levels (1 vs. all)
4. Consumer ratios

* Network Bandwidth, CPU, Disk IO

Benchmarks

Test Objective

Figure out the raw write performance of Aiven Kafka plan tiers in supported clouds.

Simulate customer usage:

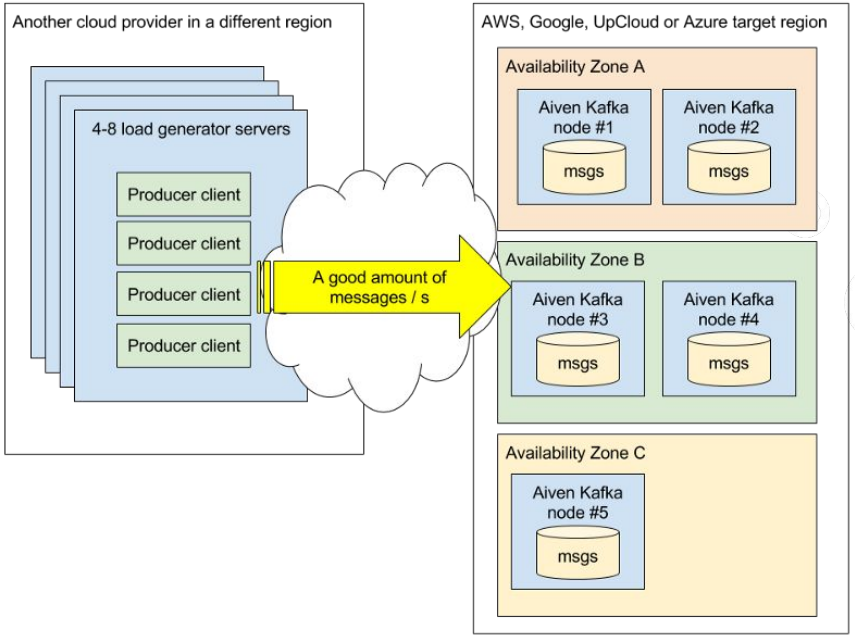
1. Typical Kafka client settings
2. Over the network access

Test Setup

- **Rdkafka_performance** from librdkafka client
- Single target topic with perfect distribution of partitions
- 1 replica
- Message size 512 bytes
- Batch size 1953, compression disabled, SSL enabled
- Number of workers increased until saturation point was reached

Test Setup

```
metadata.broker.list=kafka-tgt.aiven.io
security.protocol=ssl
ssl.key.location=client.key
ssl.certificate.location=client.crt
ssl.ca.location=ca.crt
request.timeout.ms=60000
```

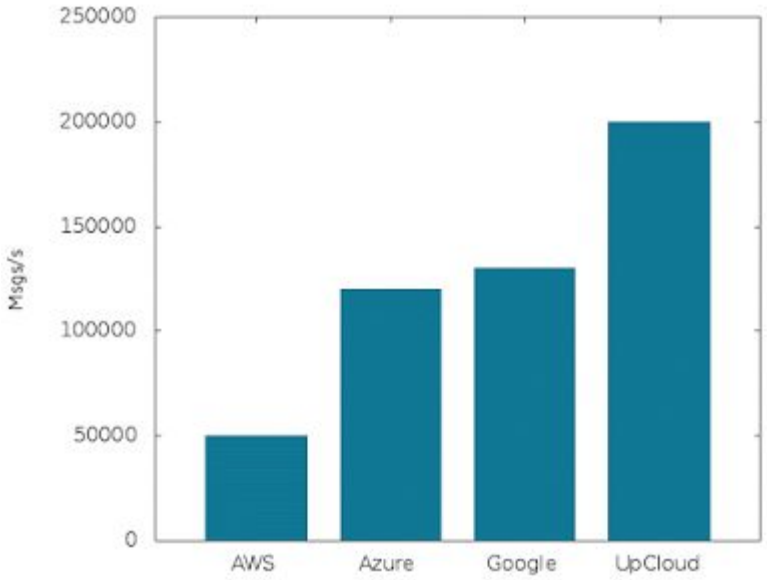



```
Rdkafka_performance -P -s 512 -t target-topic -X file=producer.props
```

Tested Aiven Kafka Plans

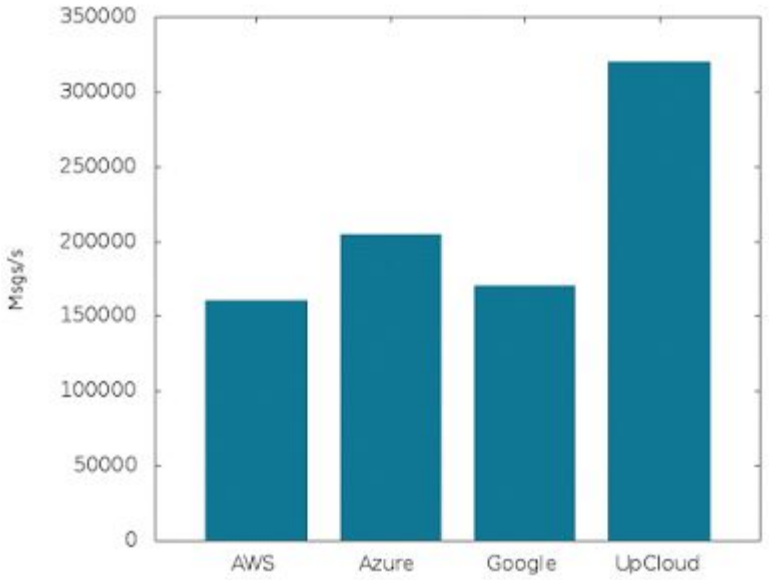
- Business-4, 3-node cluster with 1-CPU, 4GB memory
- Business-8, 3-node cluster with 2-CPU, 8GB memory
- Premium-8-5x, 5-node cluster with 2-CPU, 8GB memory

Results, Business-4



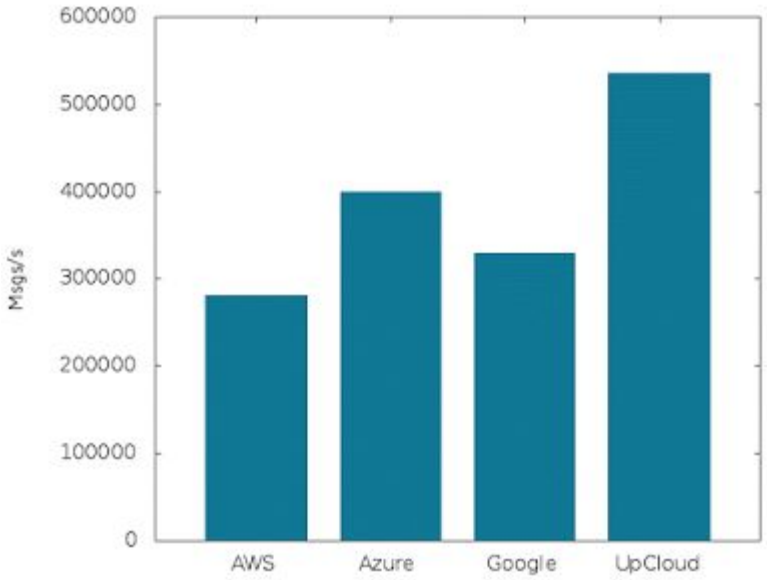
UpCloud	200,000 msgs/s
Google Cloud	130,000 msgs/s
Microsoft Azure	120,000 msgs/s
Amazon Web Services	50,000 msgs/s

Results, Business-8



UpCloud	320,000 msgs/s
Microsoft Azure	205,000 msgs/s
Google Cloud	170,000 msgs/s
Amazon Web Services	160,000 msgs/s

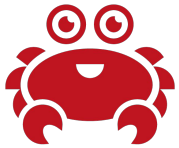
Results, Premium-8-5x



UpCloud	535,000 msgs/s
Microsoft Azure	400,000 msgs/s
Google Cloud	330,000 msgs/s
Amazon Web Services	280,000 msgs/s

Conclusions

- Even fairly modest specs deliver impressive throughput.
- Kafka scales with the resources in use.
- Kafka scales horizontal with the cluster growth.



aiven

Thank you!

@hnousiainen

@aiven_io