



# Benchmarking Kafka Write Performance

2019 Update

John Hammink

Chicago Area Kafka Enthusiasts

[jammink@aiven.io](mailto:jammink@aiven.io)

twitter: @rijksband

# Agenda

1. Introduction
2. 2019 Aiven Kafka Benchmark setup
3. Aiven Kafka Business-4 benchmark results
  - msgs/sec; MB/Sec; monthly throughput cost  
(price per mb/sec/month)
4. Aiven Business-8 benchmark results
5. Aiven Kafka Premium-6x-8 benchmark results
6. Wrapping up
7. Q & A



# I am



@Rijksband

- Developer Advocate, Evangelist @ Aiven, a cloud company
- Previously: Skype, Mozilla, ScyllaDB, F-Secure, Treasure Data, Aloomo, others
- Kafka user since 2015



# Aiven is...

- #1 independent database service provider in all major clouds
- Based in Helsinki, Finland and Boston, MA
- PostgreSQL, MySQL, ZooKeeper, ElasticSearch, Cassandra, Redis, Grafana and others
- 8 products; 6 clouds; on nearly 80 regions around the world; virtual and bare metal instances



# Kafka Performance Factors

**Compute instances / nodes / servers set many constraints, but also**

- # of topics
- Partition count
- Replication factor
- Message size
- Spread of partitions and replicas vis a vis location in availability zones
- ...among other things



# Test Objective

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

Simulate customer usage:

- Typical Kafka client settings
- Over the network access



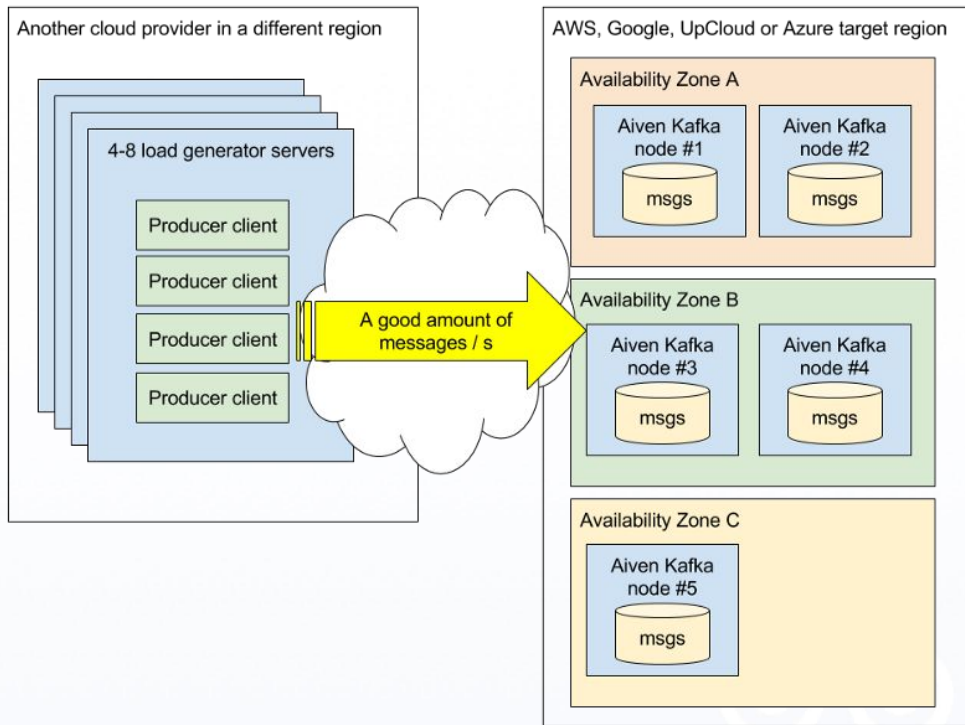
# Benchmark Setup

- `rdkafka_performance` from `librdkafka` client
- Single target topic
- Partition count set to either 3 or 6
- Message size 512 bytes
- Batch size set to 10k, compression disabled, SSL enabled
- Number of workers increased until saturation point was reached
- **Kafka version 2.1 running with Java 8**



# 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
```

Performance test code: <https://github.com/aiven/aiven-benchmark/>

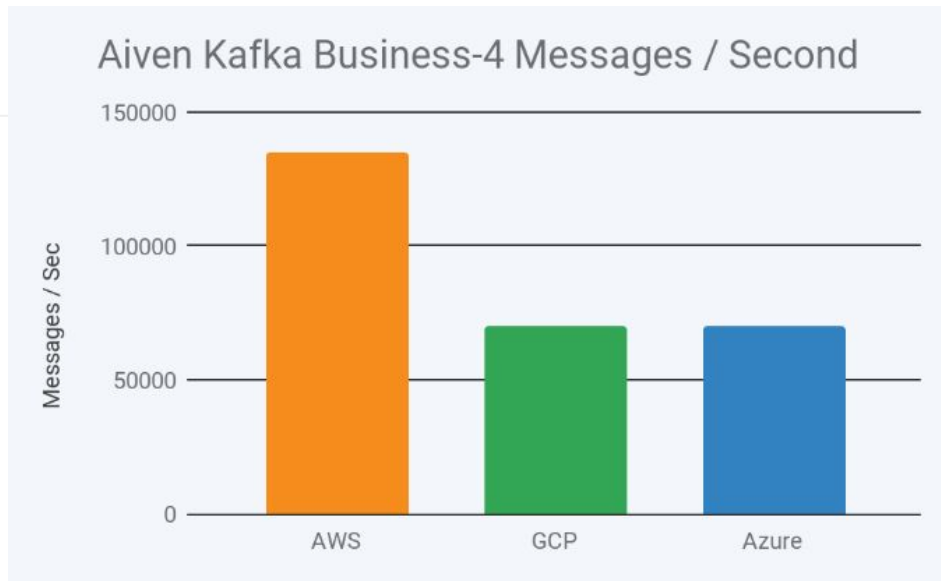


# Business-4 throughput in messages/second

3 Broker cluster; 1-2 CPU; 4 GB  
RAM/Instance

Google Cloud	70k
MS Azure	70k
AWS	135k

Messages/second

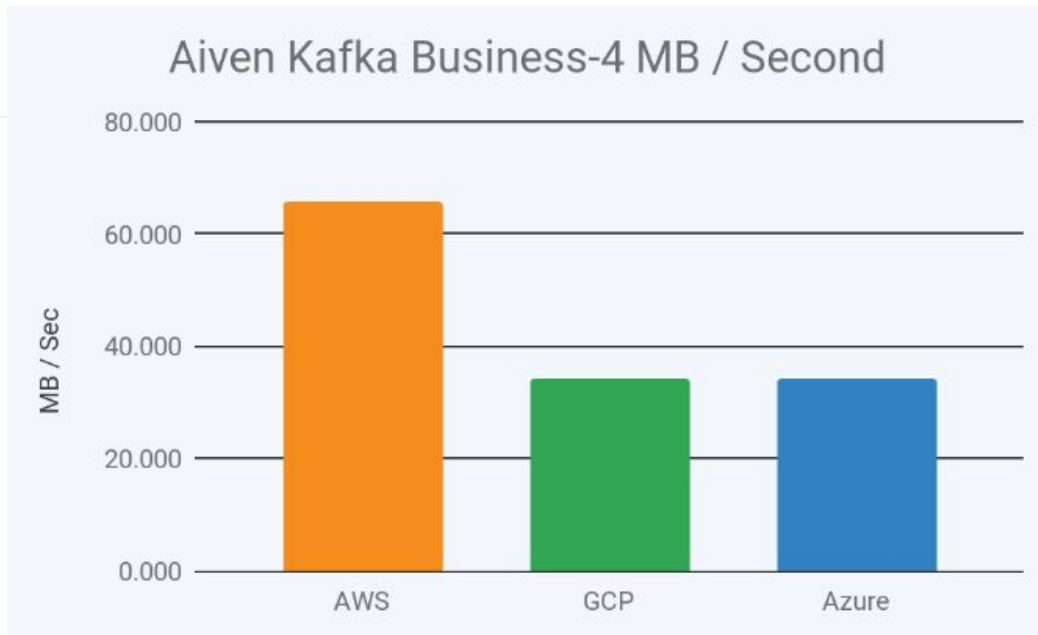


# Business-4 throughput in MB/second

3 Broker cluster; 1-2 CPU; 4 GB  
RAM/Instance

Google Cloud	>35
MS Azure	>35
AWS	65

MB/second

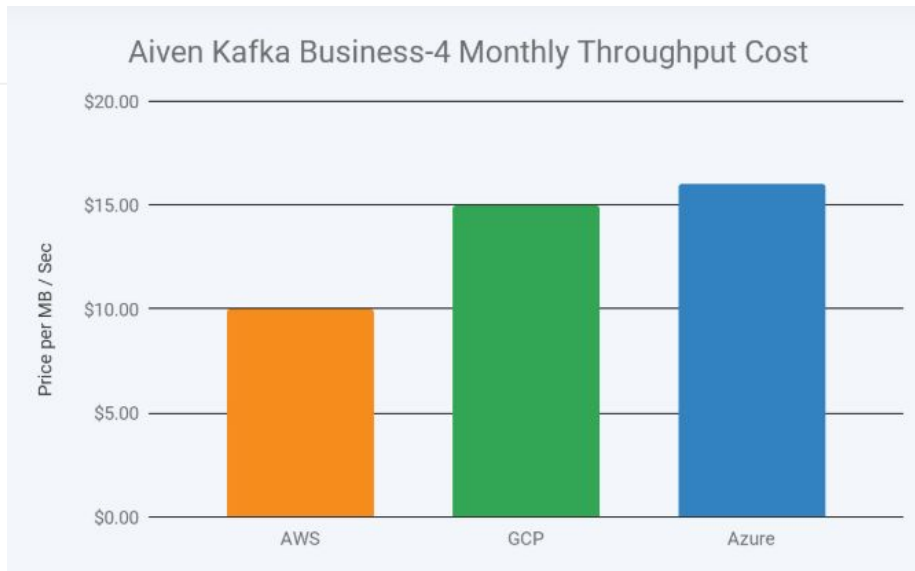


# Business-4 Monthly throughput cost

3 Broker cluster; 1-2 CPU; 4 GB  
RAM/Instance

Google Cloud	\$15
MS Azure	\$16
AWS	\$10

Price per MB/sec/Month

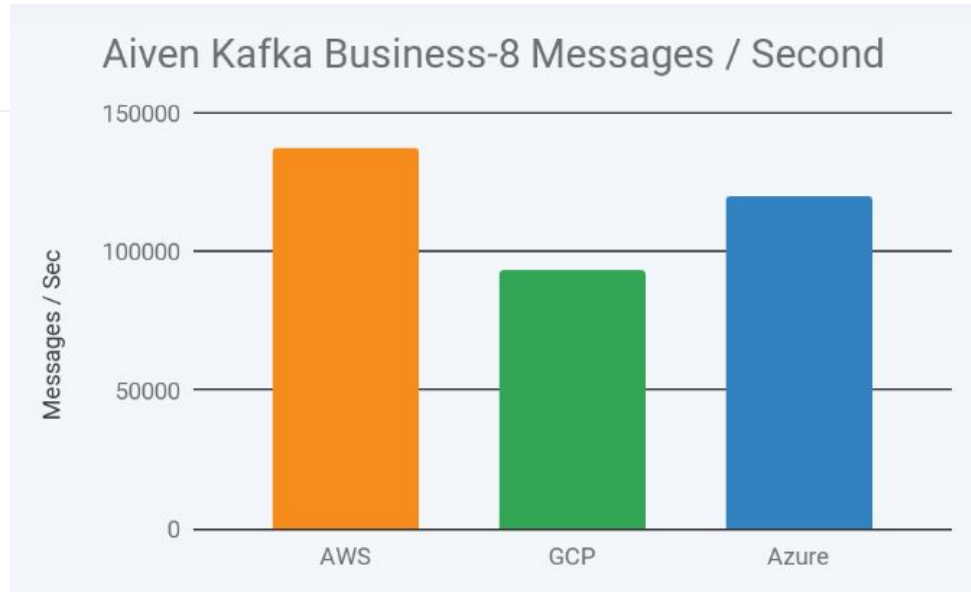


# Business-8 Throughput in messages/sec

3 Broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	95k
MS Azure	120k
AWS	137k

Messages/second

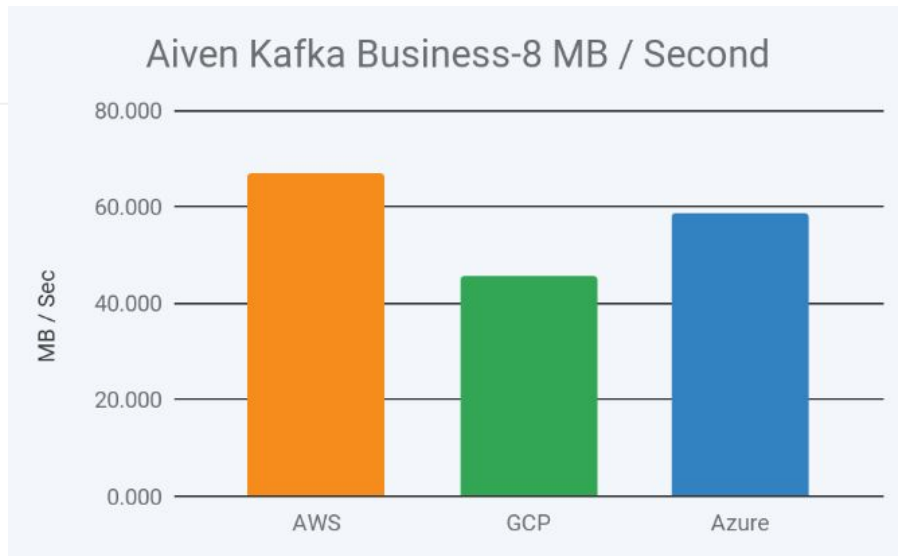


# Business-8 throughput in MB/second

3 Broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	46
MS Azure	59
AWS	67

MB/second

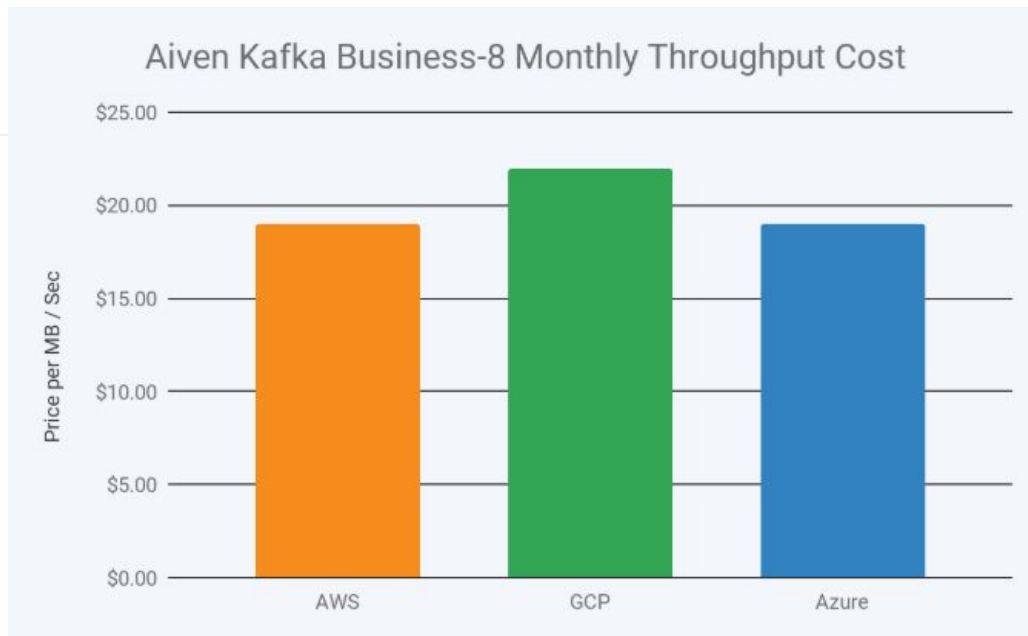


# Business-8 Monthly throughput cost

3 Broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	\$22
MS Azure	\$19
AWS	\$19

Price per MB/sec/Month

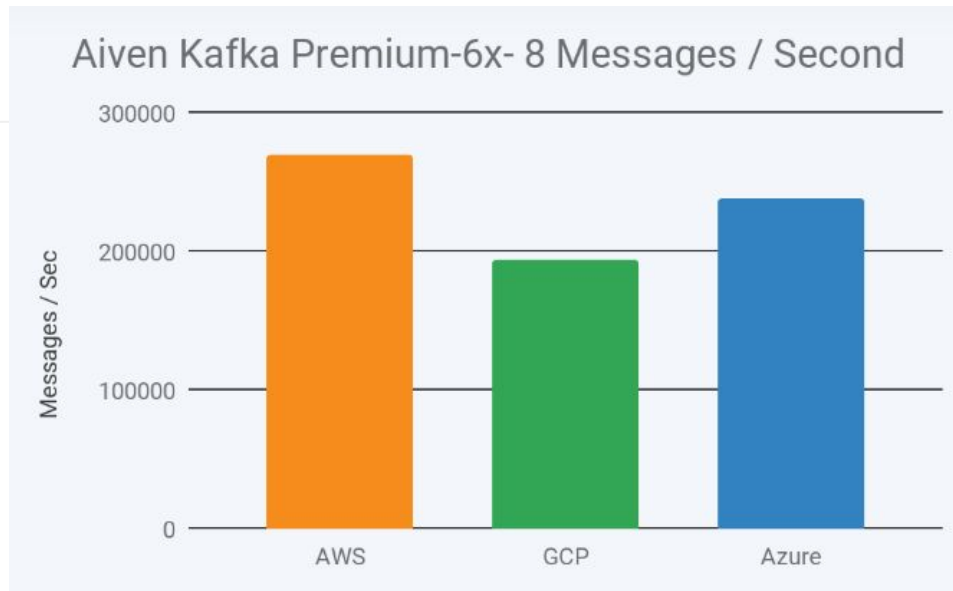


# Premium-6x-8 throughput in messages/sec

Six broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	167k
MS Azure	238k
AWS	270k

Messages/second

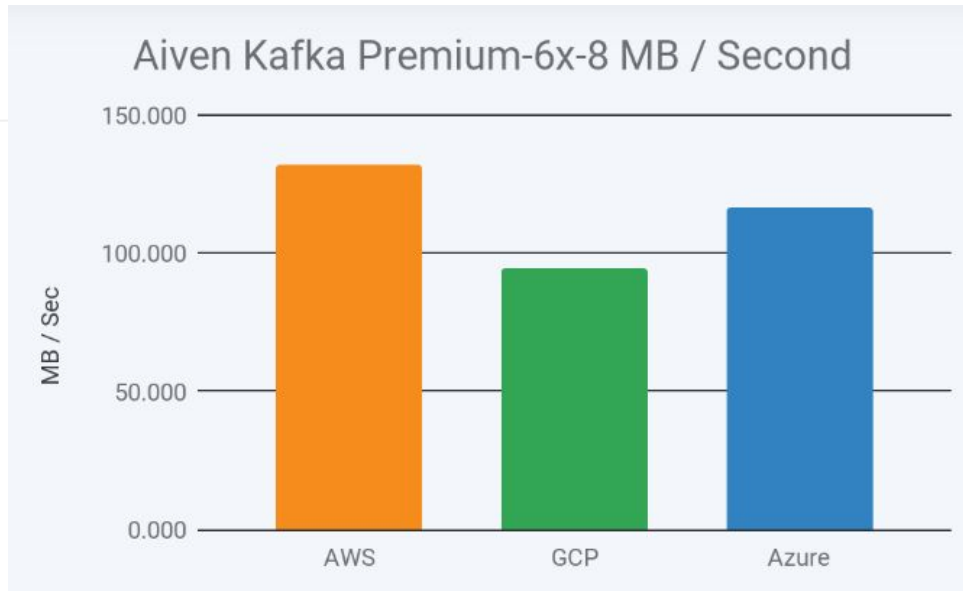


# Premium-6x-8 throughput in MB/sec

Six broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	82
MS Azure	116
AWS	132

MB/second



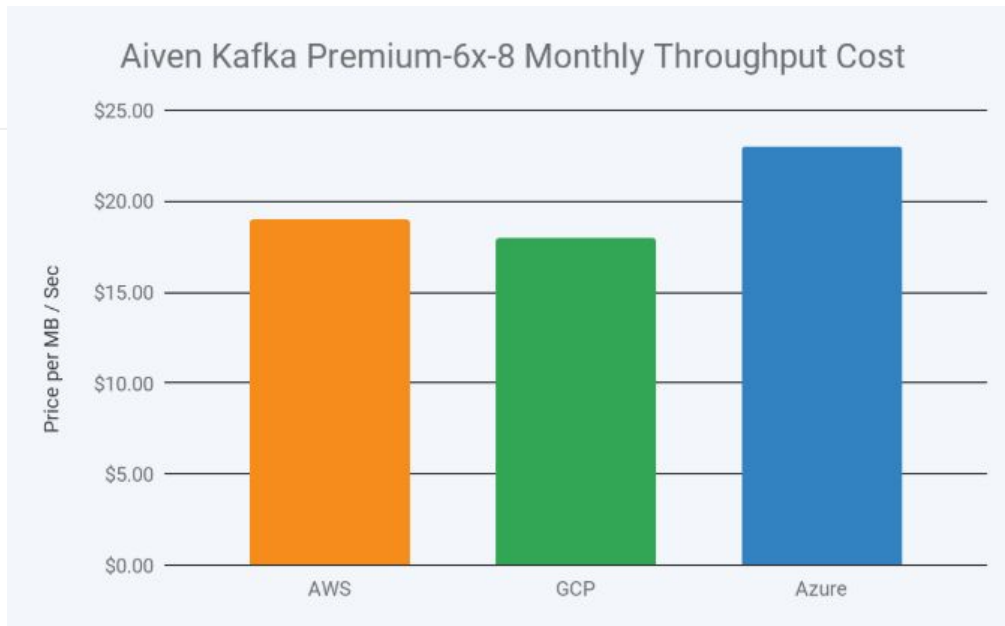


# Premium-6x-8 monthly throughput cost

Six broker cluster; 2-4 CPUs; 8 GB  
RAM/Instance

Google Cloud	\$18
MS Azure	\$18
AWS	\$23

Price per MB/sec/Month



# Summary

- Performance varies greatly between different clouds and instances
- Monthly throughput cost should not be considered in isolation when comparing plans.
- Larger instance sizes don't always offer better performance
  - However, additional nodes improve performance nearly linearly
- Test your workload on multiple instances before deciding what to use
- The most popular managed services may not offer all the options out there
- For a more robust test, we'll be addressing read/write tests in the near future

# Questions?

Try out all the clouds and plans on <https://aiven.io>

[jammink@aiven.io](mailto:jammink@aiven.io); @rijksband

Performance test code:

<https://github.com/aiven/aiven-benchmark/>





# Thanks!



<https://aiven.io>



@aiven\_io

