Rethinking Message Brokers on RDMA and NVM

2020 ACM SIGMOD SRC | June 14 - June 19, 2020 | Portland, OR, USA

Hendrik Makait

Contribution

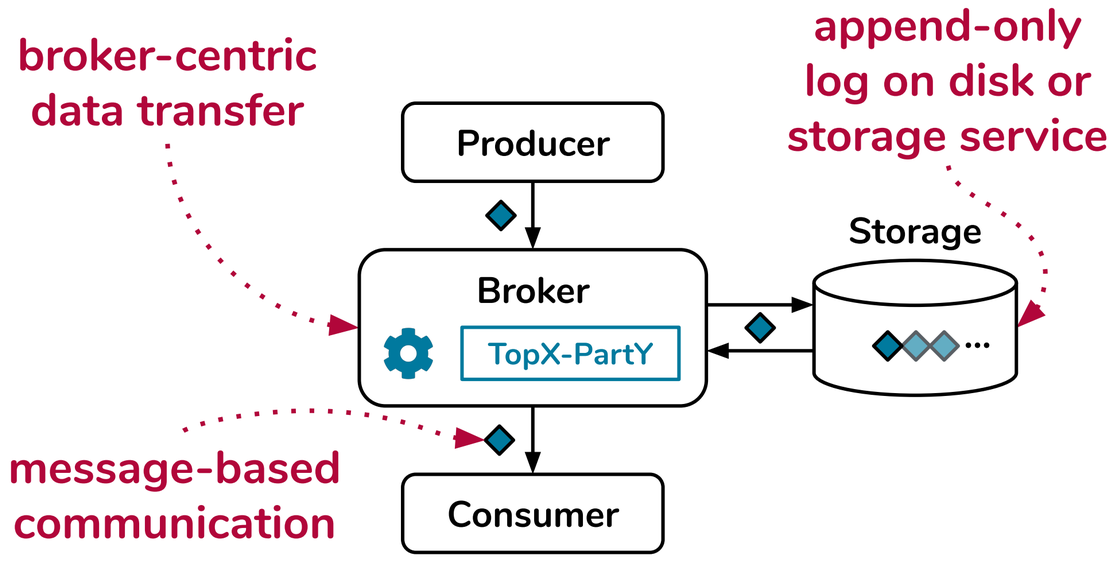
We propose an approach to improve the throughput of message brokers by utilizing RDMA-based data transfer to byte-addressable and non-volatile memory.

Our main contributions are:

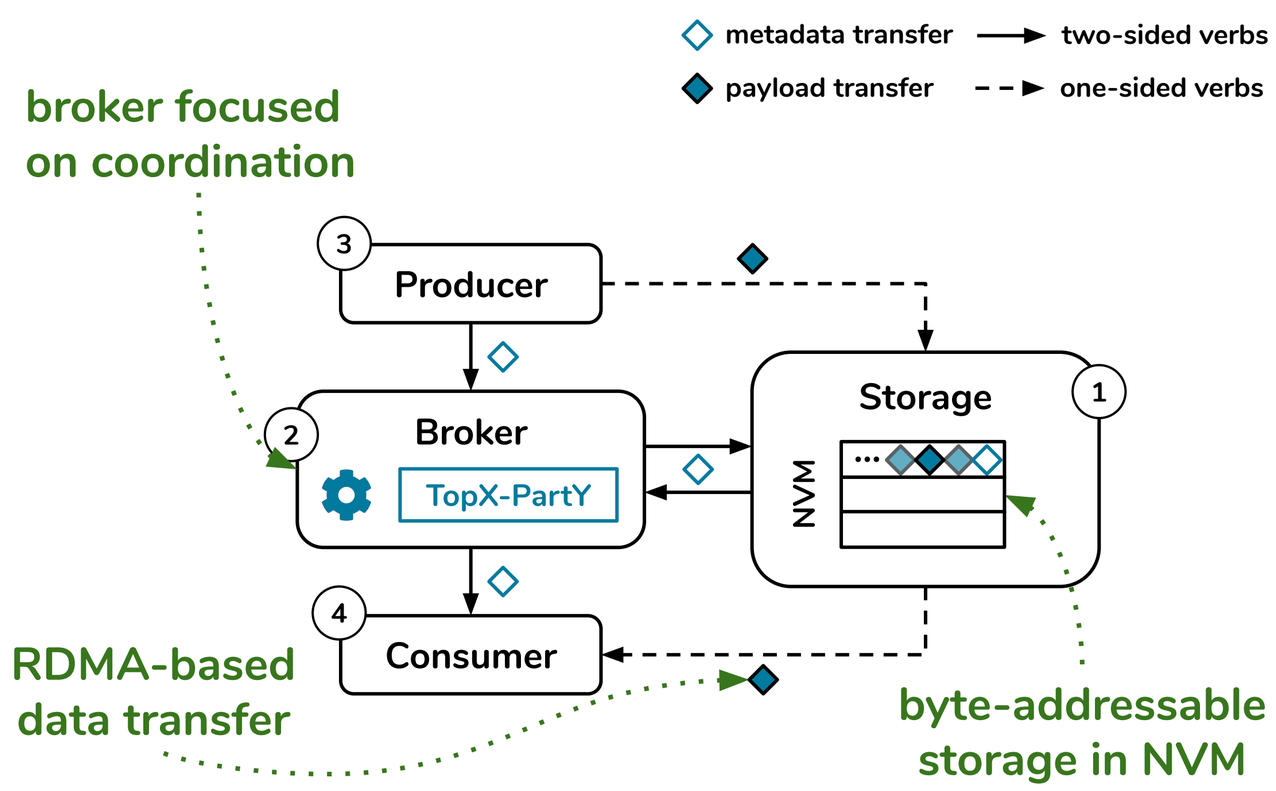
A broker architecture and protocol that ...

- ... **scale up single partition throughput** for producers and consumers to utilize modern IB networks.
- ... **ensure message ordering and delivery** by writing to byte-addressable locations in non-volatile memory.
- ... **reduce CPU overhead** using one-sided RDMA verbs.

Current Message Brokers



Architecture



① **Storage Node** provides one large persistent memory region split into segments.
② **Broker Node** allocates individual segments and coordinates access to them.
③ **Producer** writes to non-volatile memory locations provided by the broker.
④ **Consumer** reads fully-written data directly from storage node.

Protocol Design Space

Producer

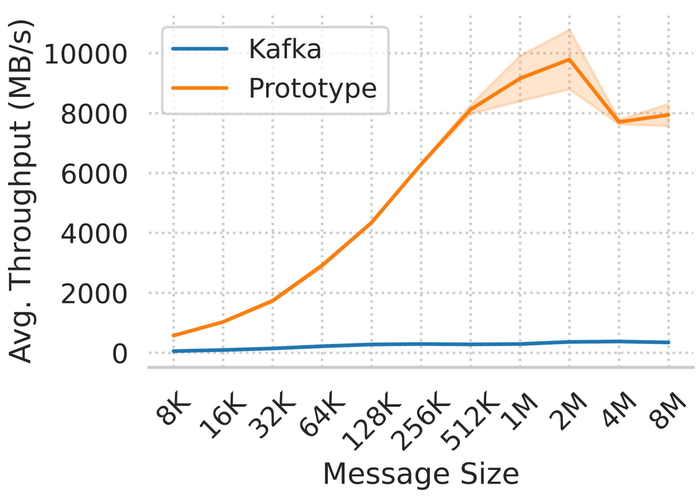
- Exclusive access to partition allows caching segment metadata and reduce control flow.
- Shared access to partition requires broker to manage write offsets.
- Committing multiple writes reduces control flow but increases publish latency.
- Sequential staging allows to interleave data transfers and commits while maintaining order.

Consumer

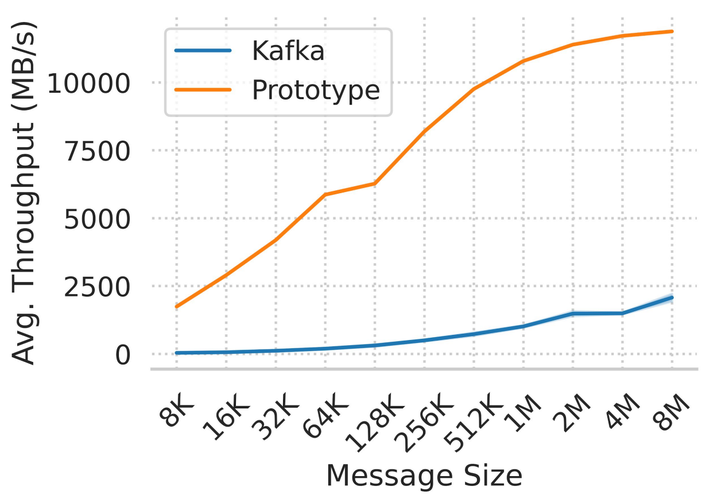
- Replication to multiple storage nodes allows to load balance data access.
- Cache segment metadata to manage sequential read offsets locally.
- Mix pull-based data access and push-based segment metadata updates.

Performance Evaluation

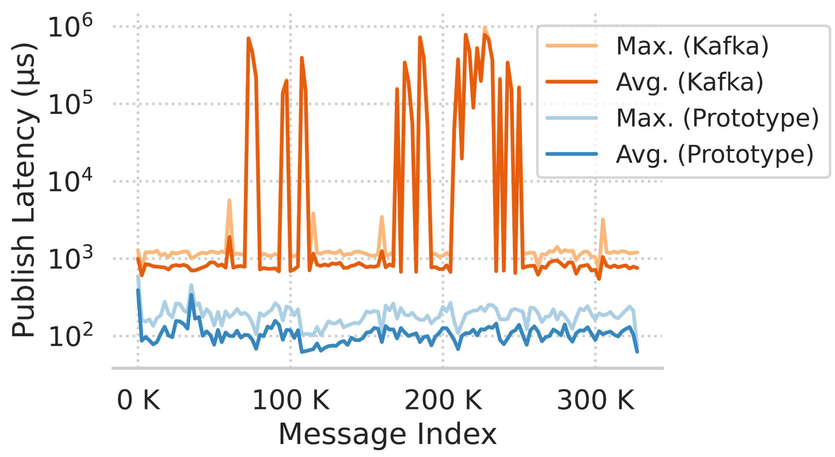
Producer Throughput



Consumer Throughput



Publish Latency (Window Size: 100 Msgs)



1 Producer | 1 Consumer | 1 Partition | No replication | IB EDR 4x (100Gbit/s) | Kafka storage in /dev/shm

HPI Hasso Plattner Institut

Data Engineering Systems Group
Hendrik Makait
Hasso Plattner Institute, Potsdam, Germany
E-Mail: hendrik.makait@guest.hpi.de | hendrik.makait@campus.tu-berlin.de

Technische Universität Berlin