Apache Kafka overview

- Core
  - Pub/sub
- Connect
  - Integration
- Streams
  - Processing

Kafka adoption in enterprises

6 of the top 10 travel companies

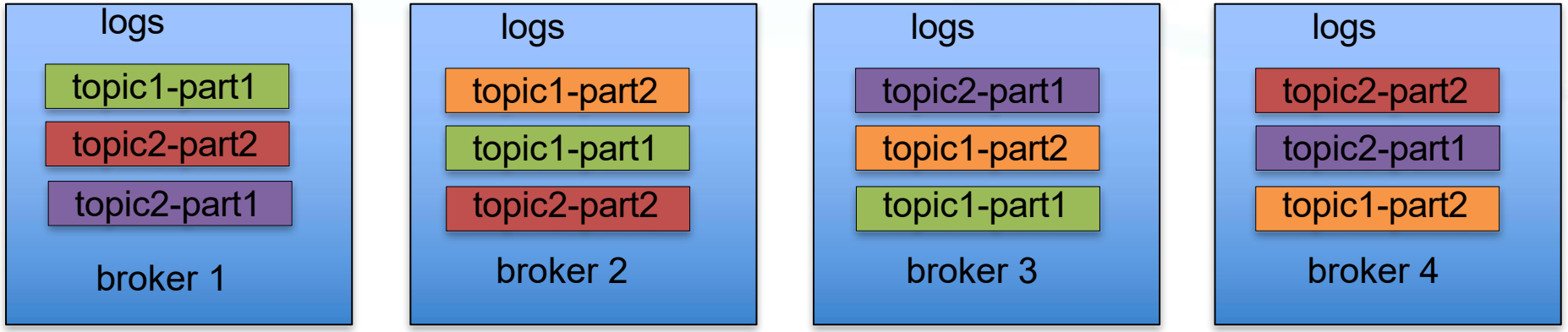
7 of the top 10 global banks

8 of the top 10 insurance companies

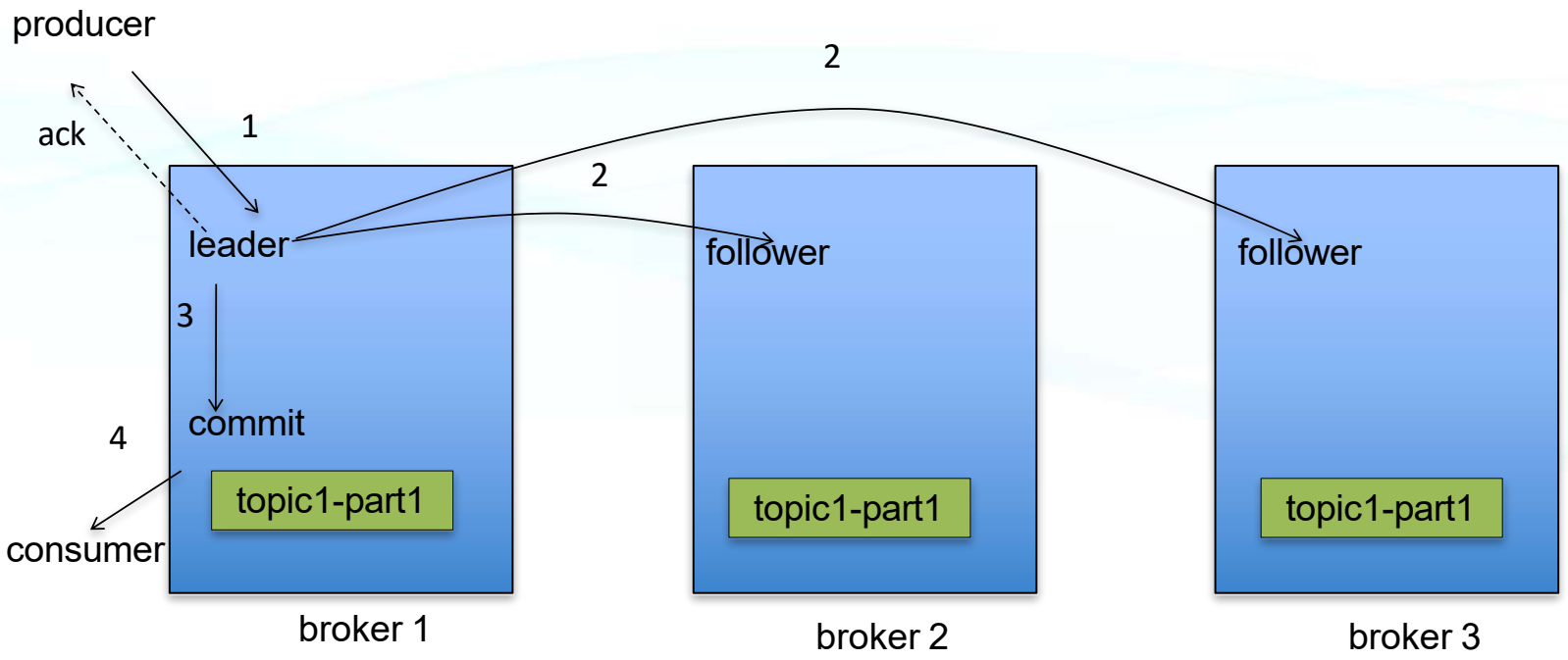
9 of the top 10 telecom companies

Kafka Replication

- Configurable replication factor
- Tolerating f – 1 failures with f replicas
- Automated failover



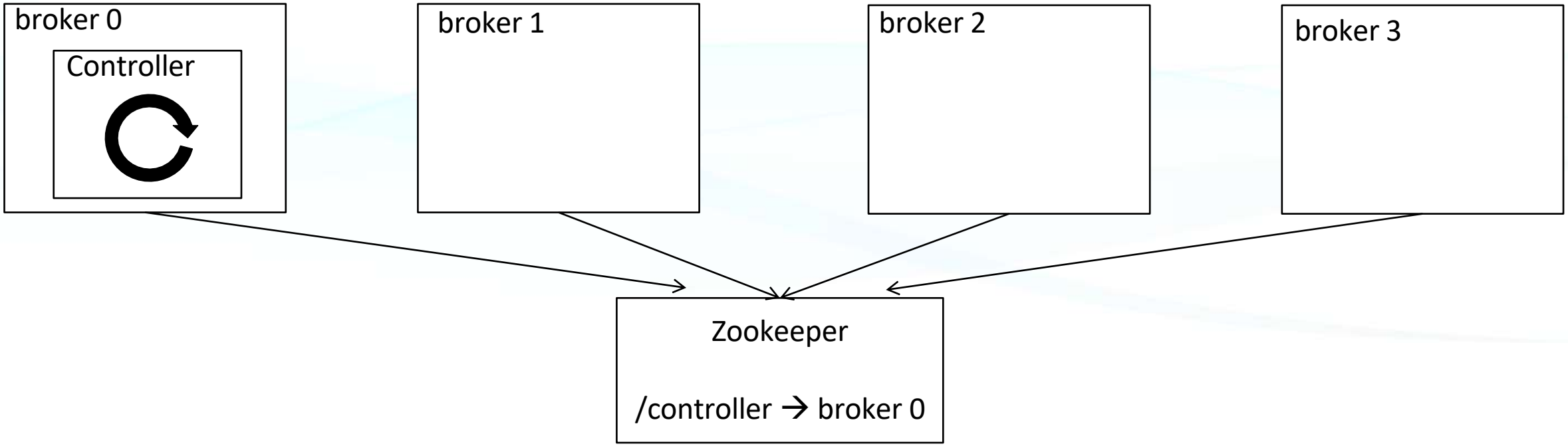
High Level Data Flow in Replication



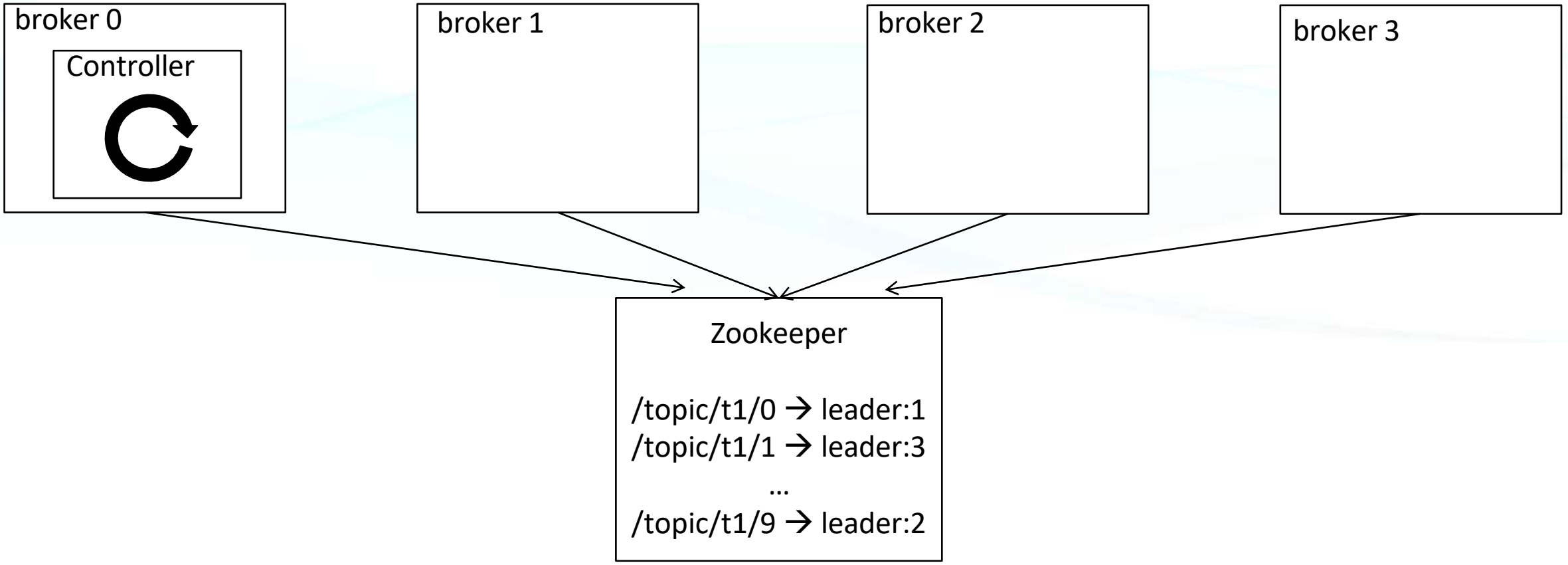
What's controller

- One broker in a cluster acts as controller
- Monitor the liveness of brokers
- Elect new leaders on broker failure
- Communicate new leaders to brokers

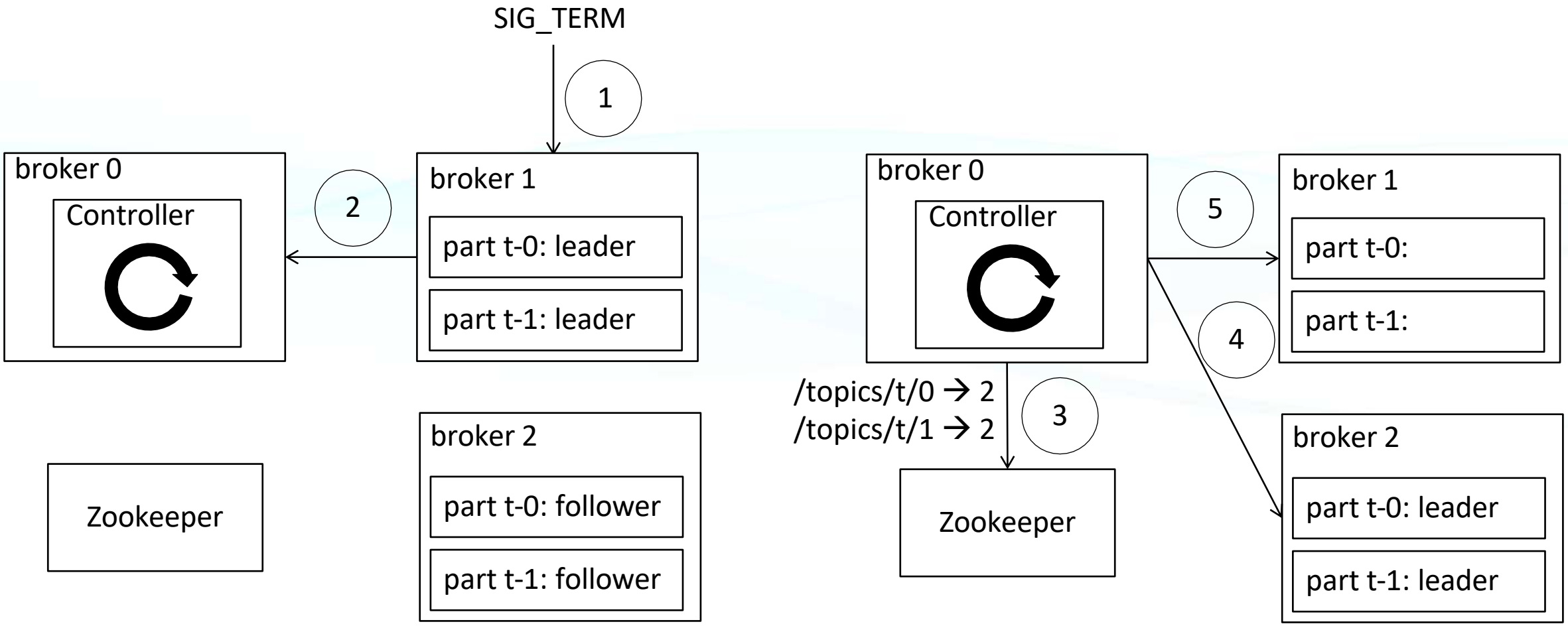
Controller election



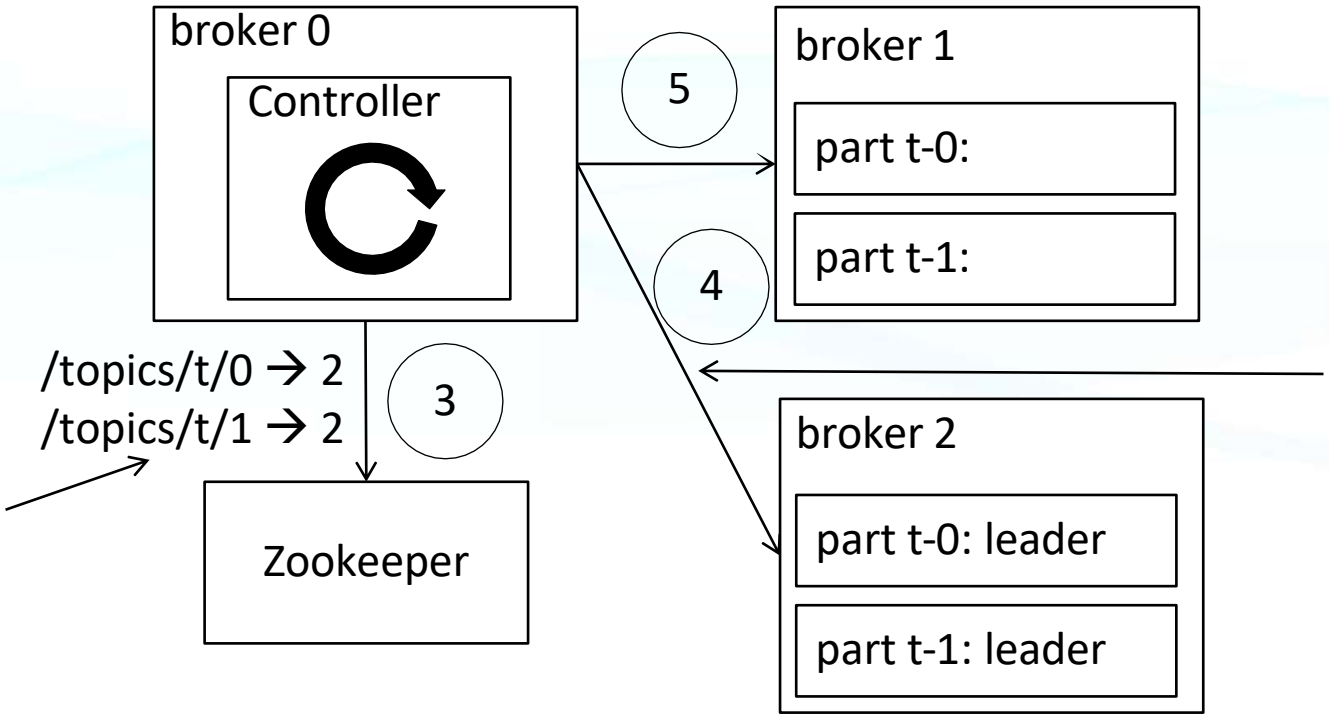
Partition state: stored in ZK, cached in controller



Controlled shutdown



Issues with controlled shutdown (pre 1.1)



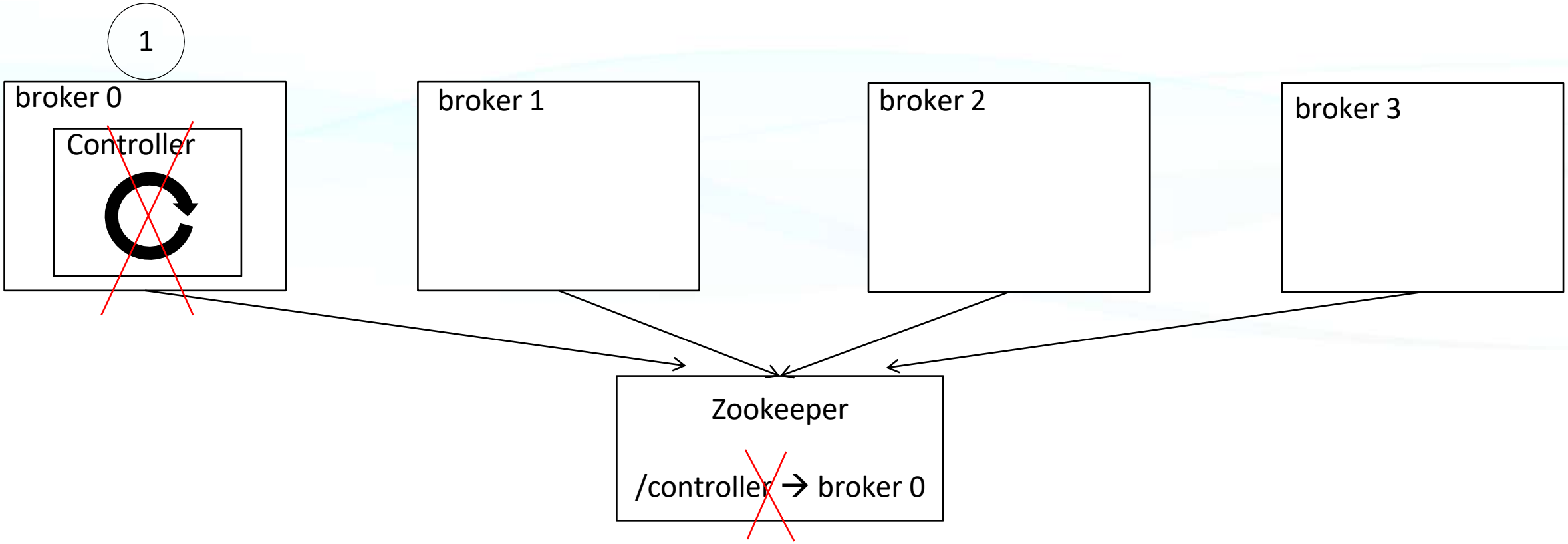
Writes to ZK are serial

Impact: longer shutdown time

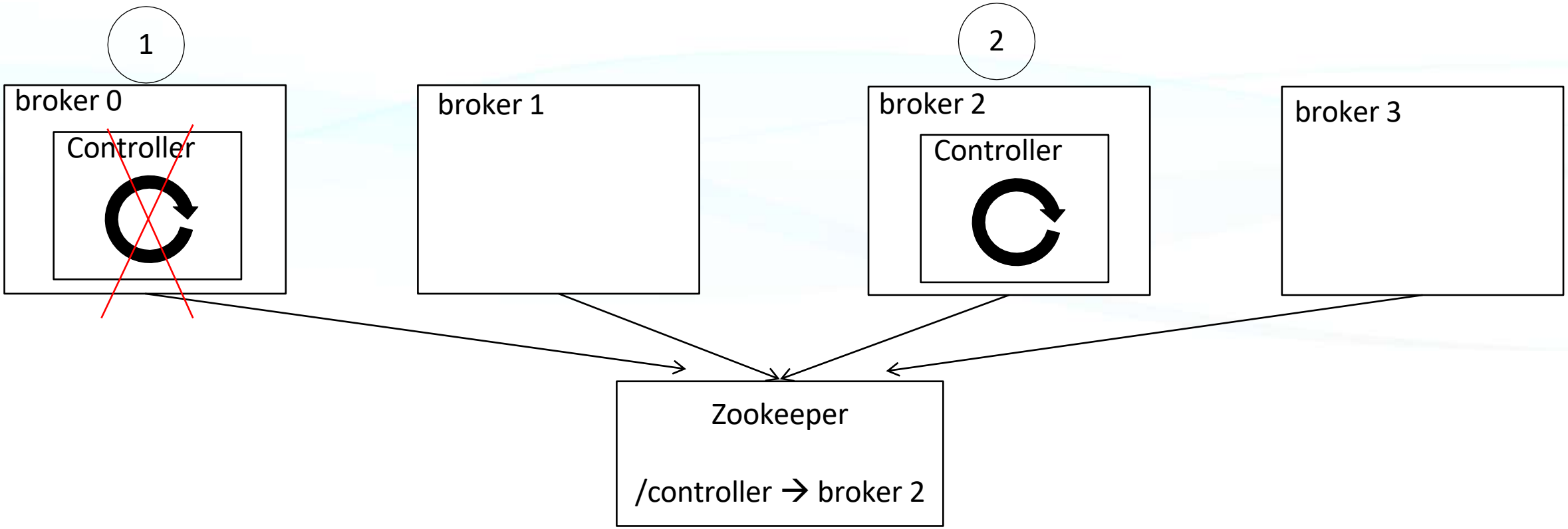
Communication of new leaders not batched

Impact: client timeout

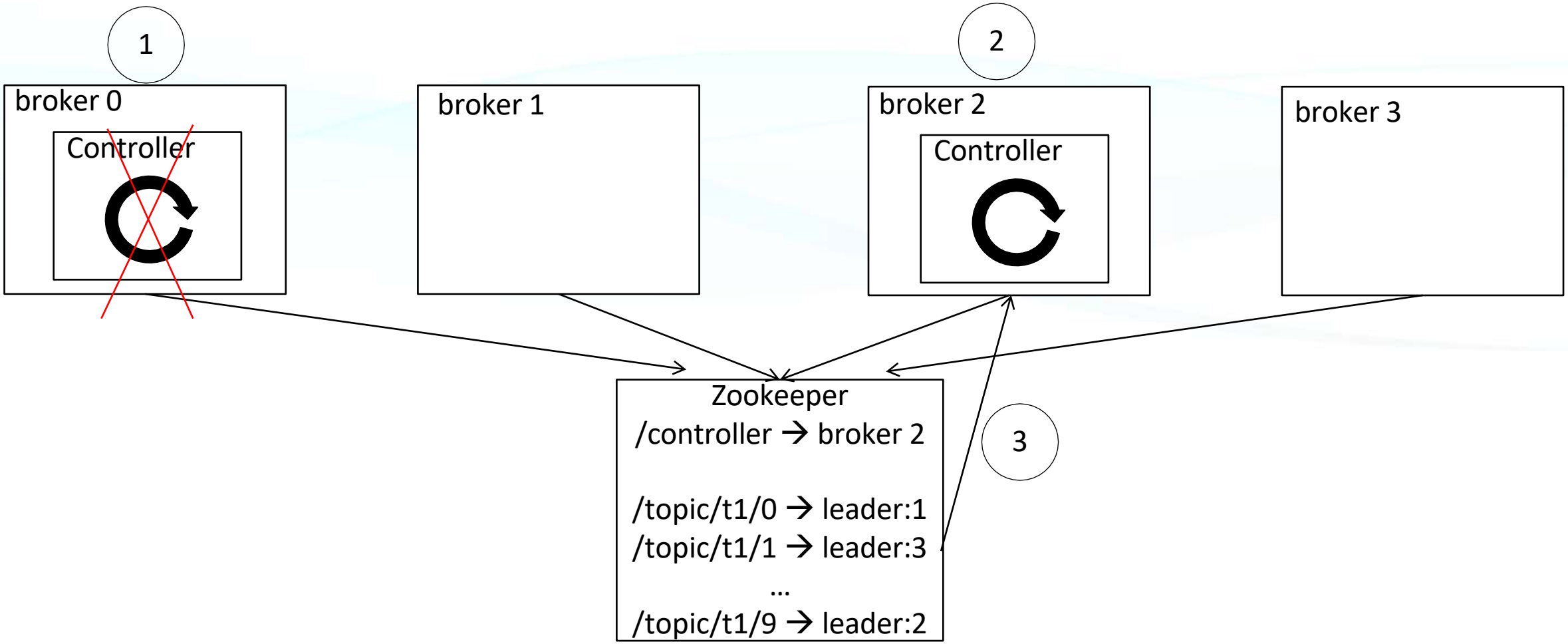
Controller failover



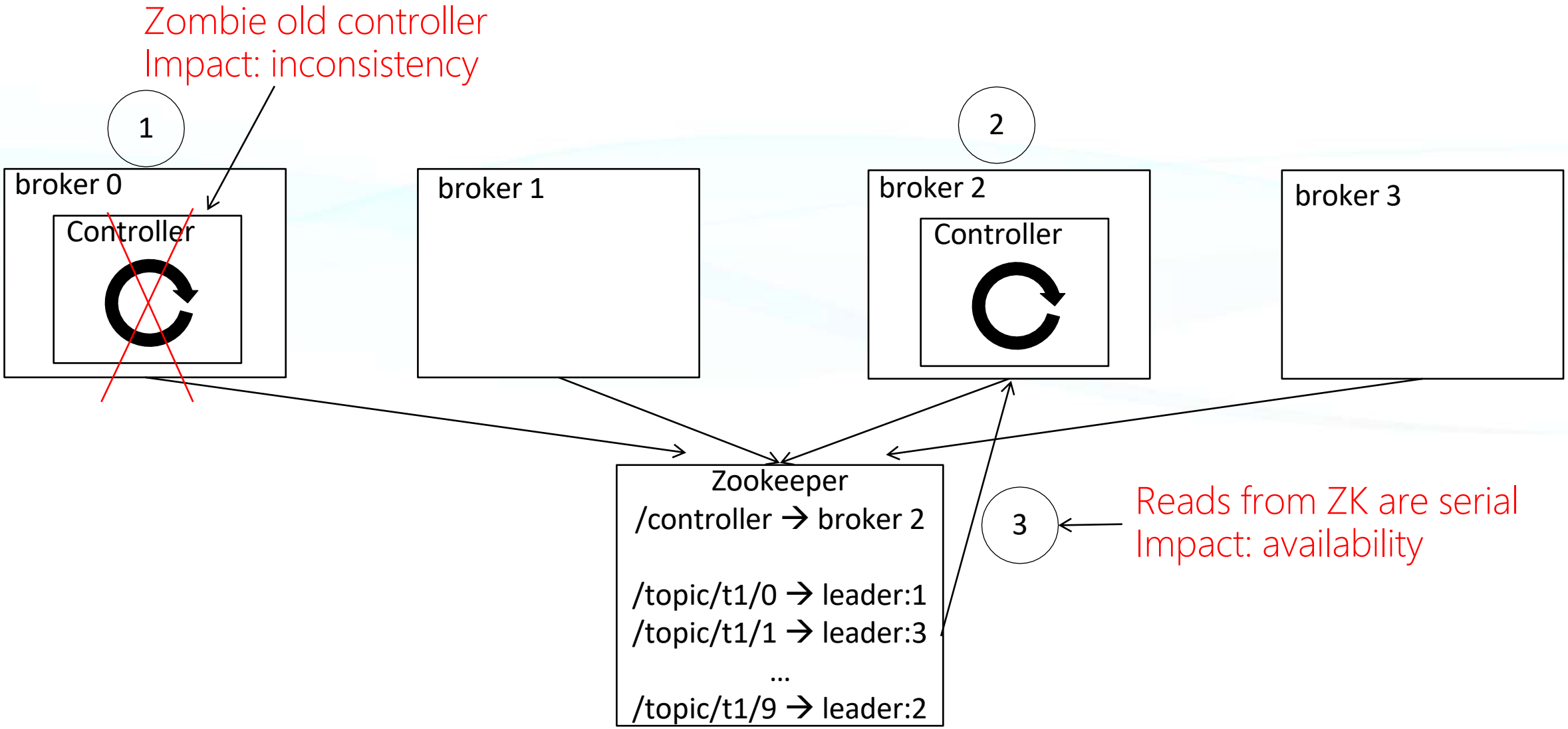
Controller failover



Controller failover

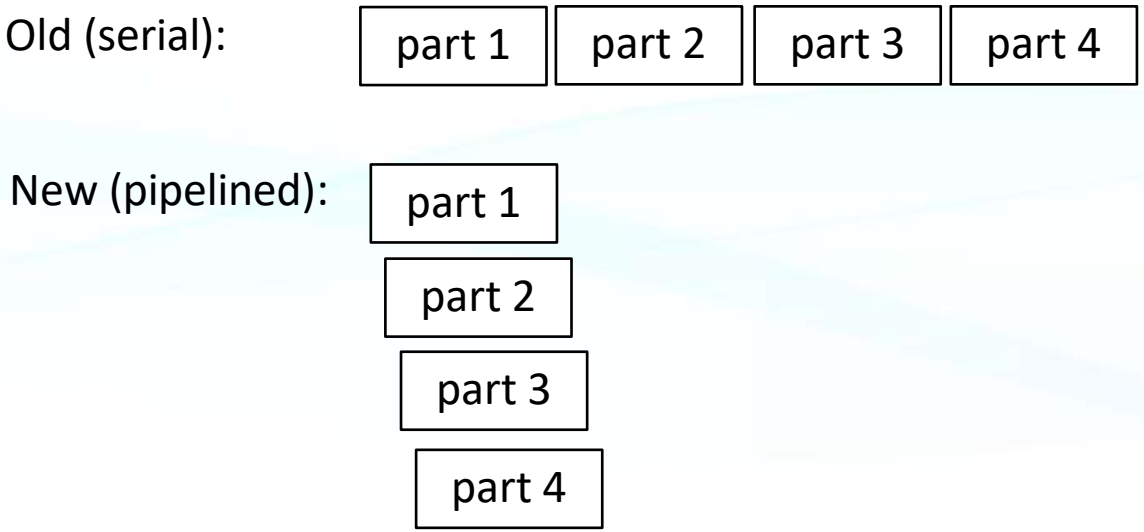


Issues with controller failover (pre 1.1)



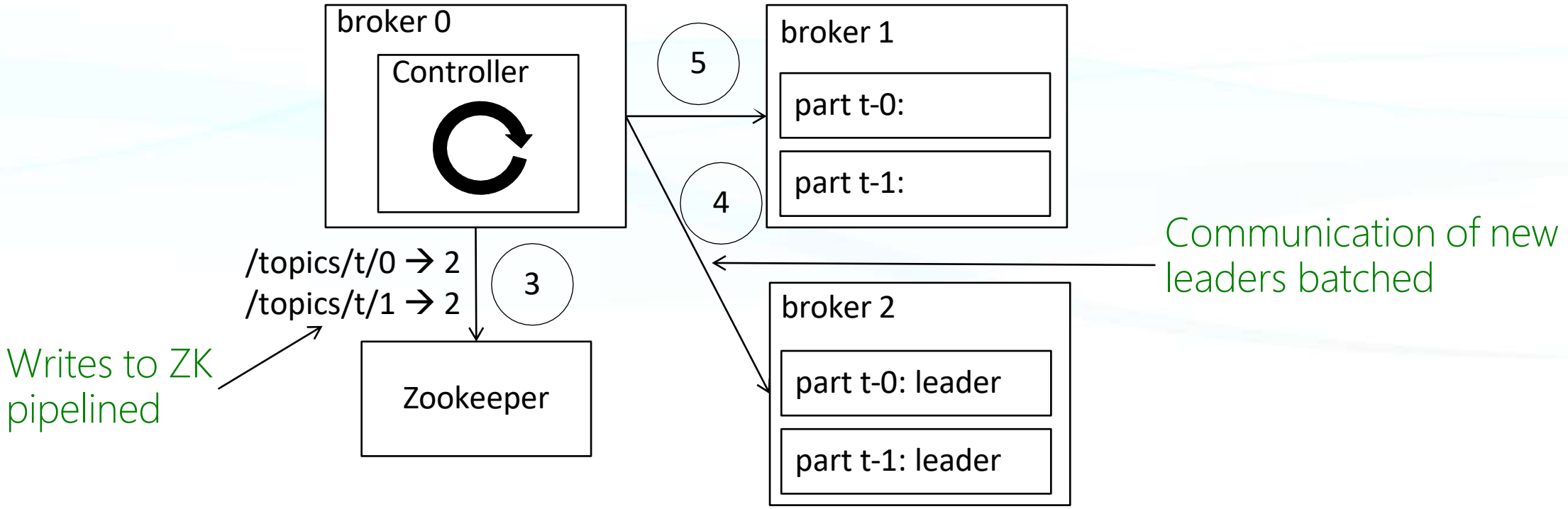
Performance improvements in 1.1

- Controller uses async ZK api for reads/writes

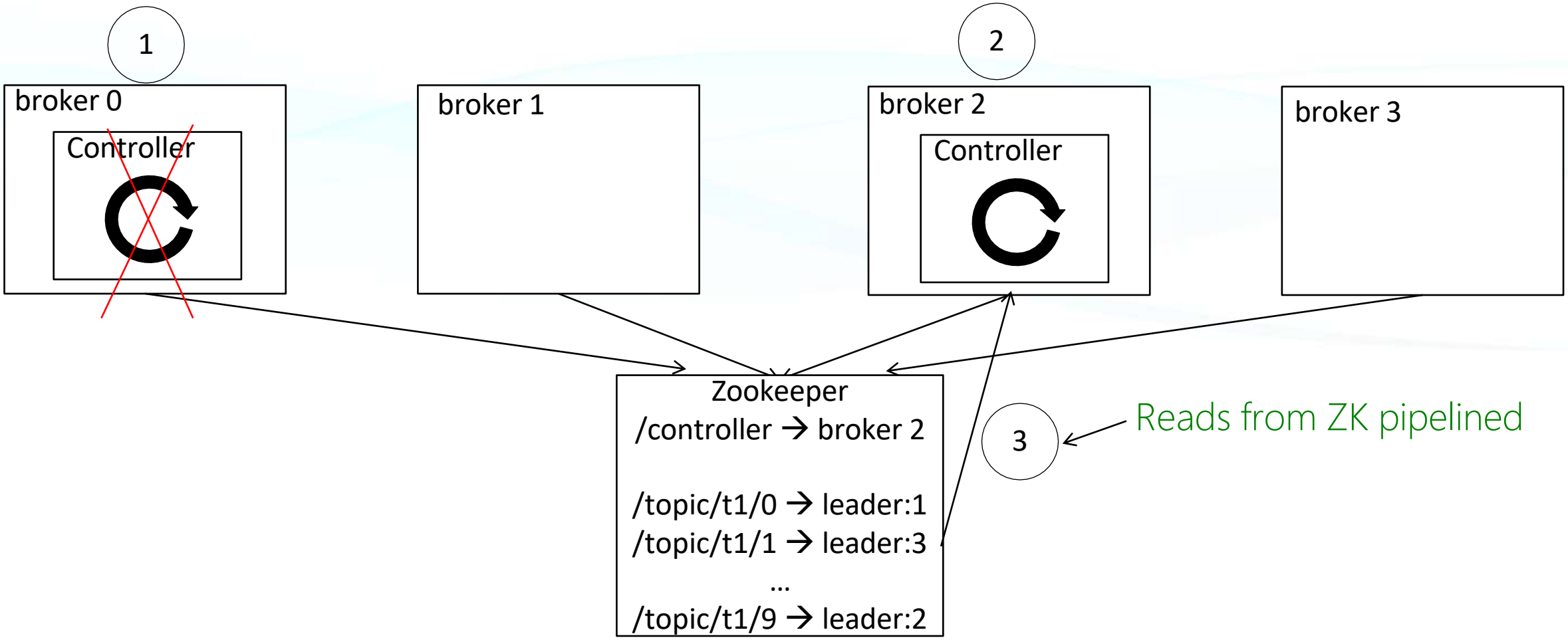


- Controller communicates new leaders to brokers in batches

Controlled shutdown (post 1.1)



Controller failover (post 1.1)



Results for controlled shutdown

- 5 ZK nodes and 5 brokers on different racks
- 25K topics, 1 partition, 2 replicas
- 10K partitions per broker

	Kafka 1.0.0	Kafka 1.1.0
Controlled shutdown time	6.5 minutes	3 seconds

Results for controller failover

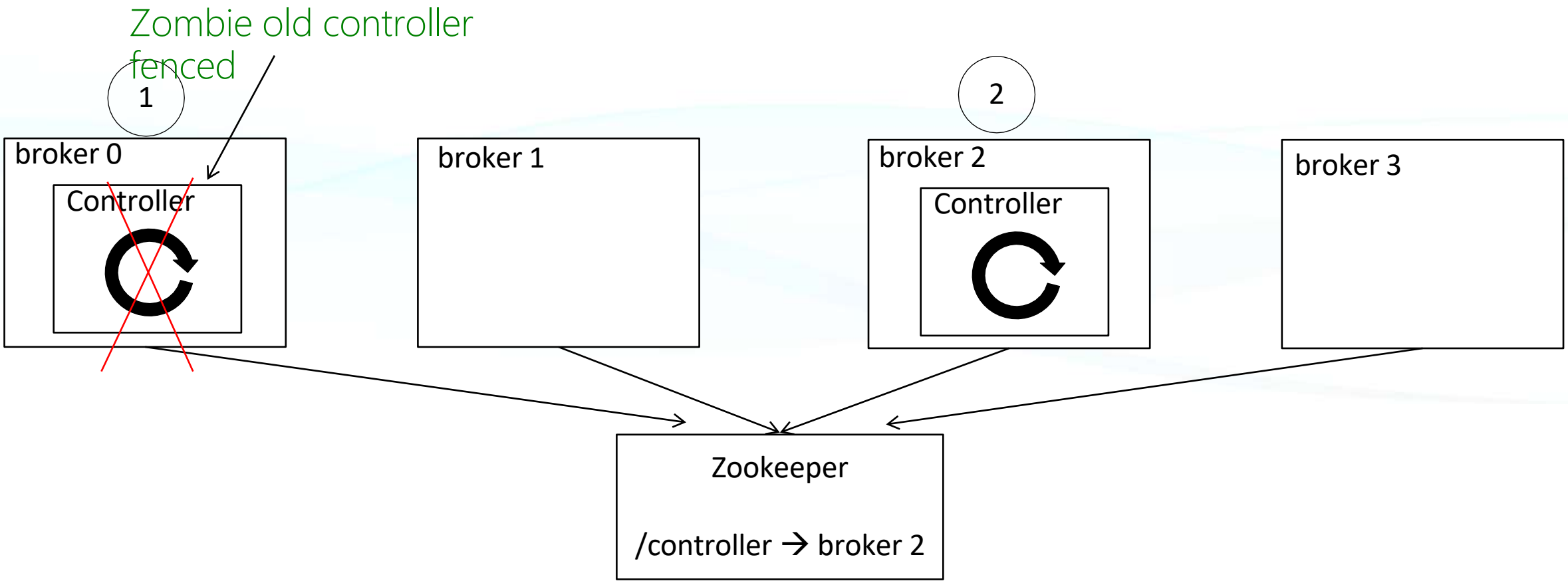
- 5 ZK nodes and 5 brokers on different racks
- 2K topics, 50 partitions, 1 replica
- Controller failover: reload100K partitions from ZK

	Kafka 1.0.0	Kafka 1.1.0
State reload time	28 seconds	14 seconds

Fencing zombie controller

- ZK session expiration
  - Better handling in the controller (1.1)
- Controller path deletion
  - Writes to ZK conditioned on controller epoch (to be in 2.1)

Controller failover (expected in 2.1)



Summary

- Significant performance improvement in controller in 1.1
  - Allow 10X more partitions in a Kafka cluster
- Better fencing of zombie controller in 1.1 and 2.1
- More details in KAFKA-5027