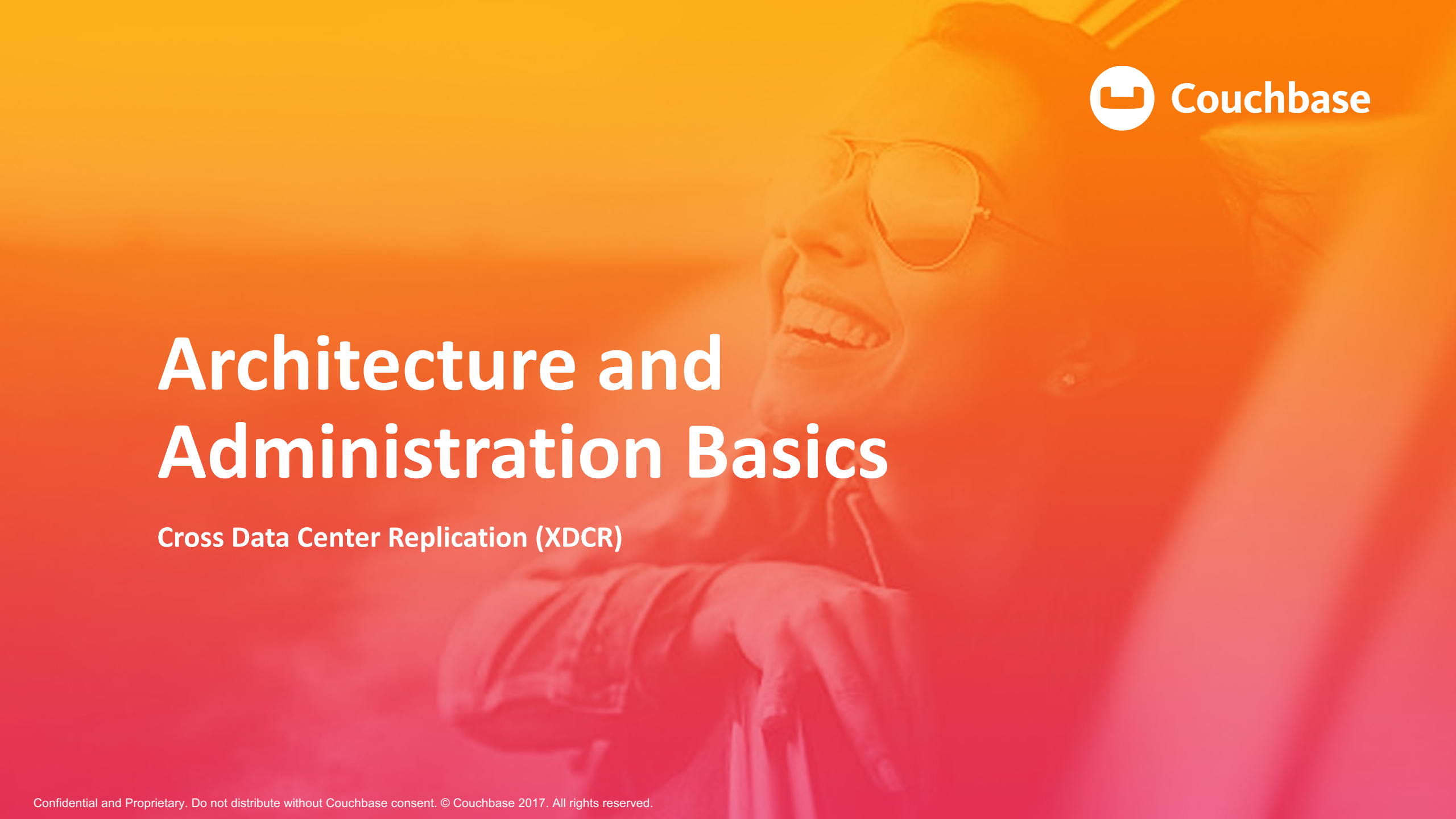


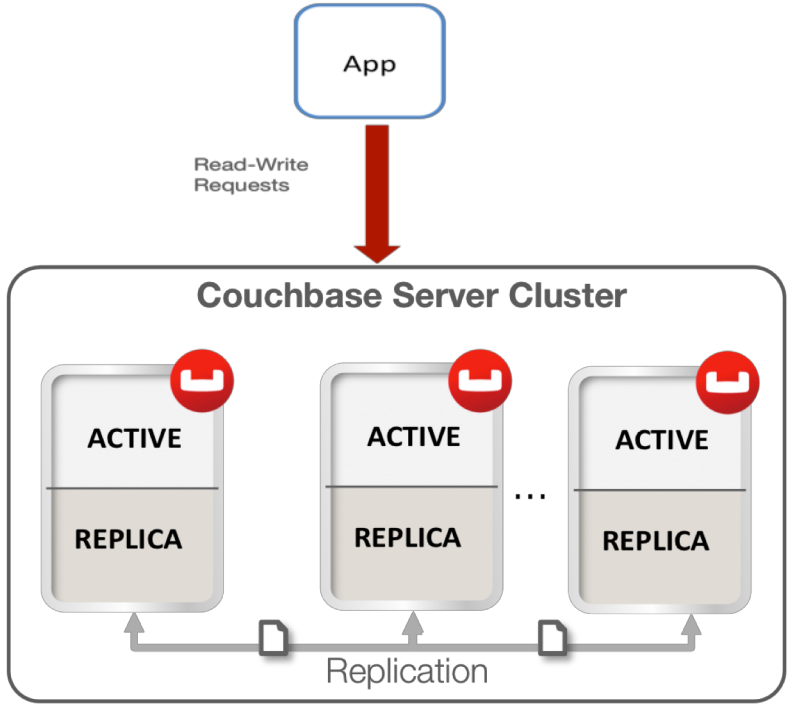
Couchbase

Architecture and Administration Basics

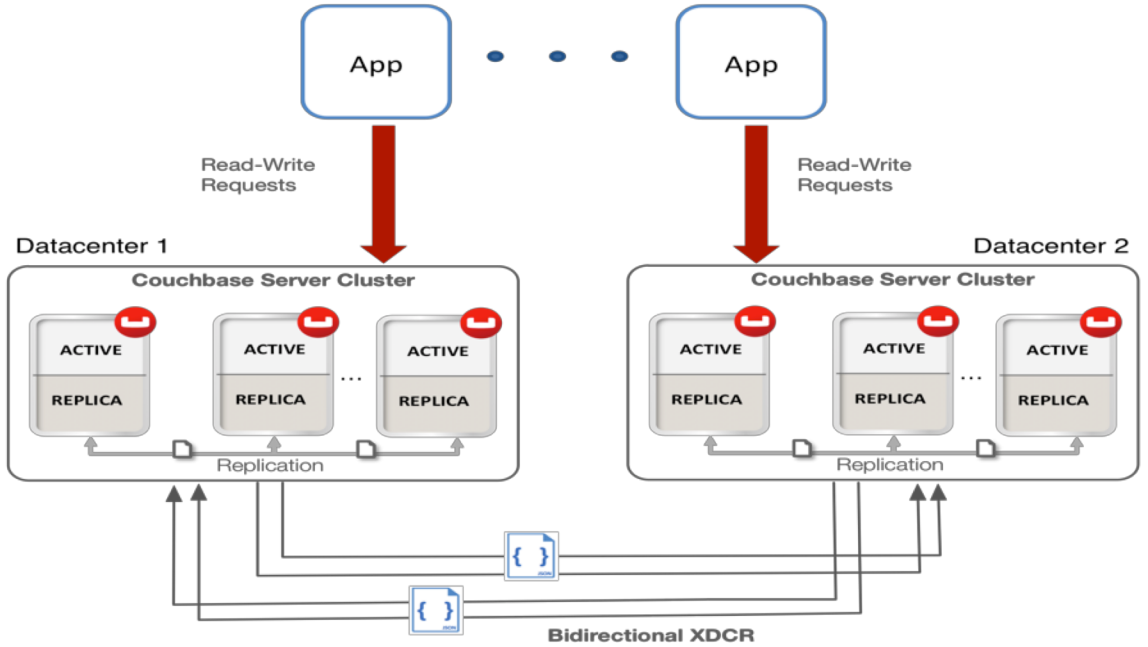
Cross Data Center Replication (XDCR)

1 | Introduction

Intra-Cluster vs. Inter-Cluster



VS.

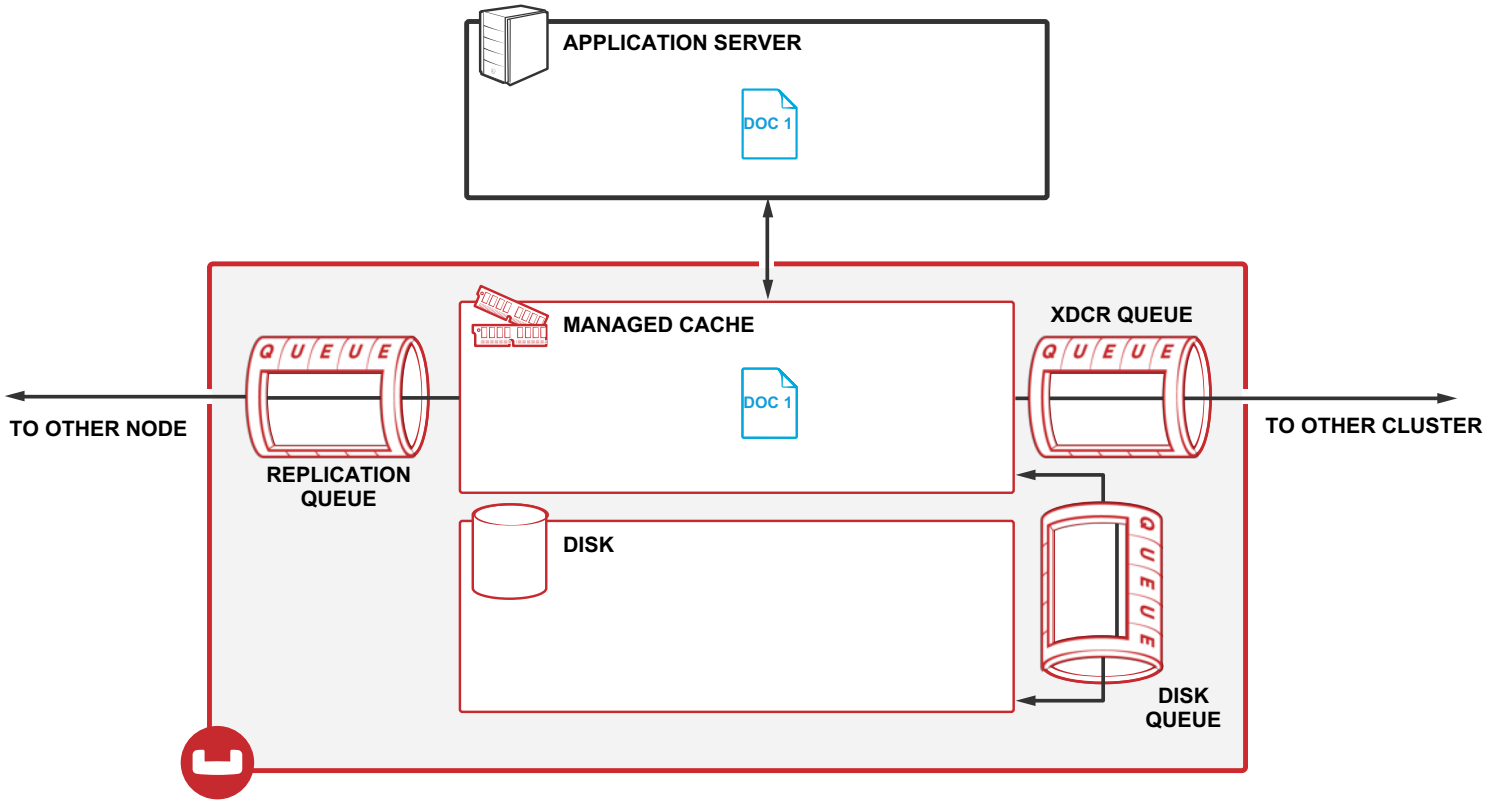


Purposes

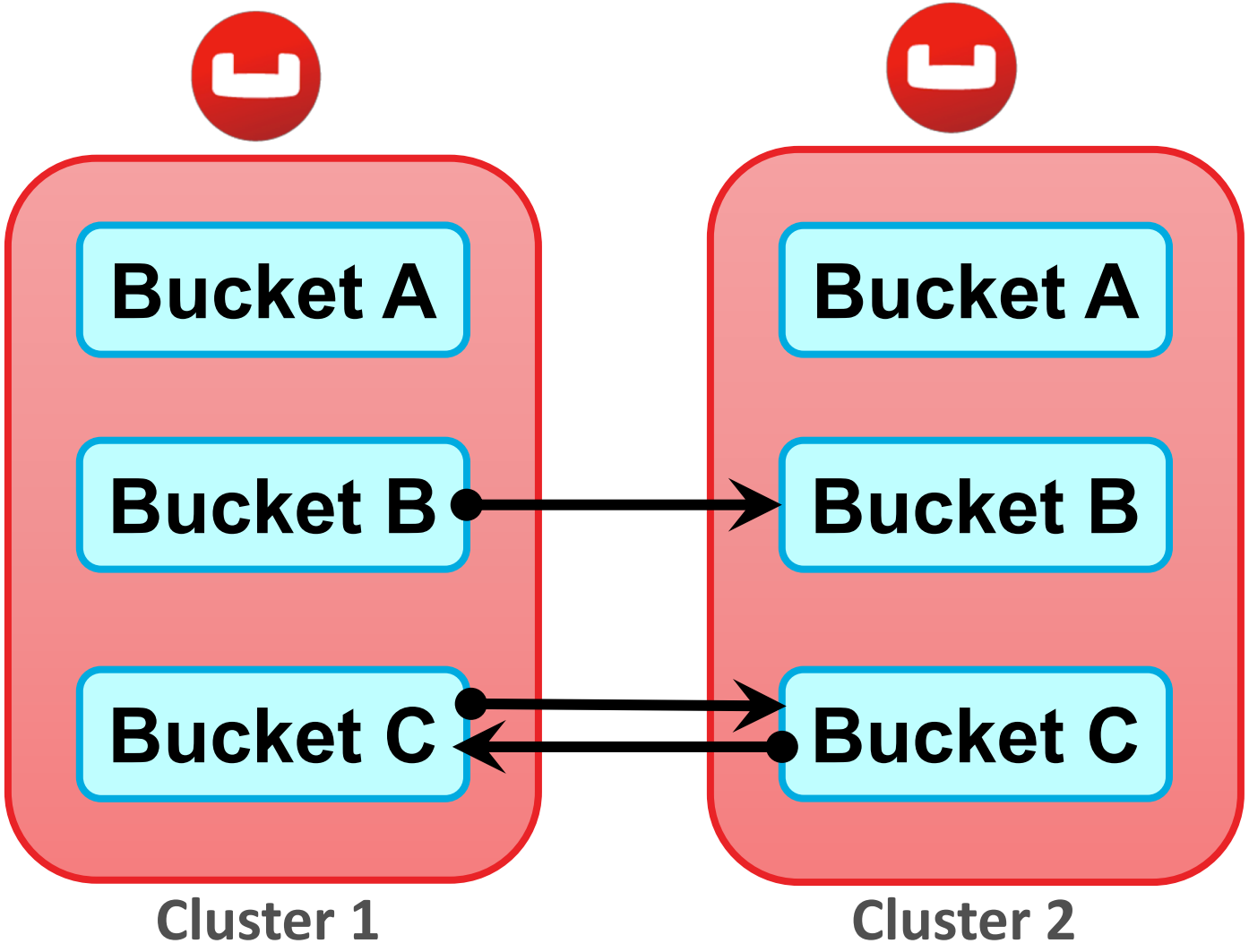
- Deliver high performing, async. data replication
- Provide disaster recovery and high availability across data centers
- Support data locality
- For load separation
- Support various topologies and replication schemes, including filtering
- Easy setup of development and test environments

2 | How it works

Asynchronous, Memory to Memory

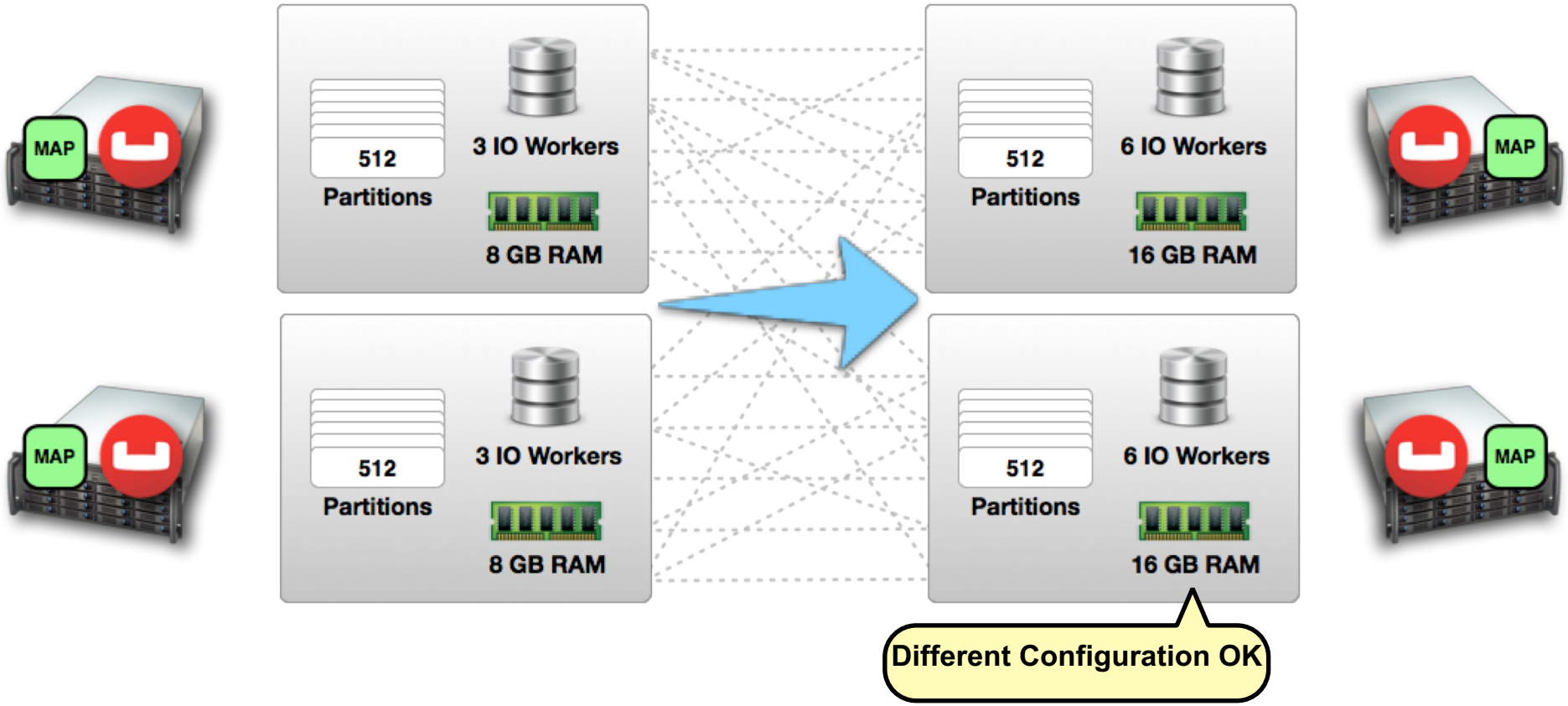


From Bucket to Bucket



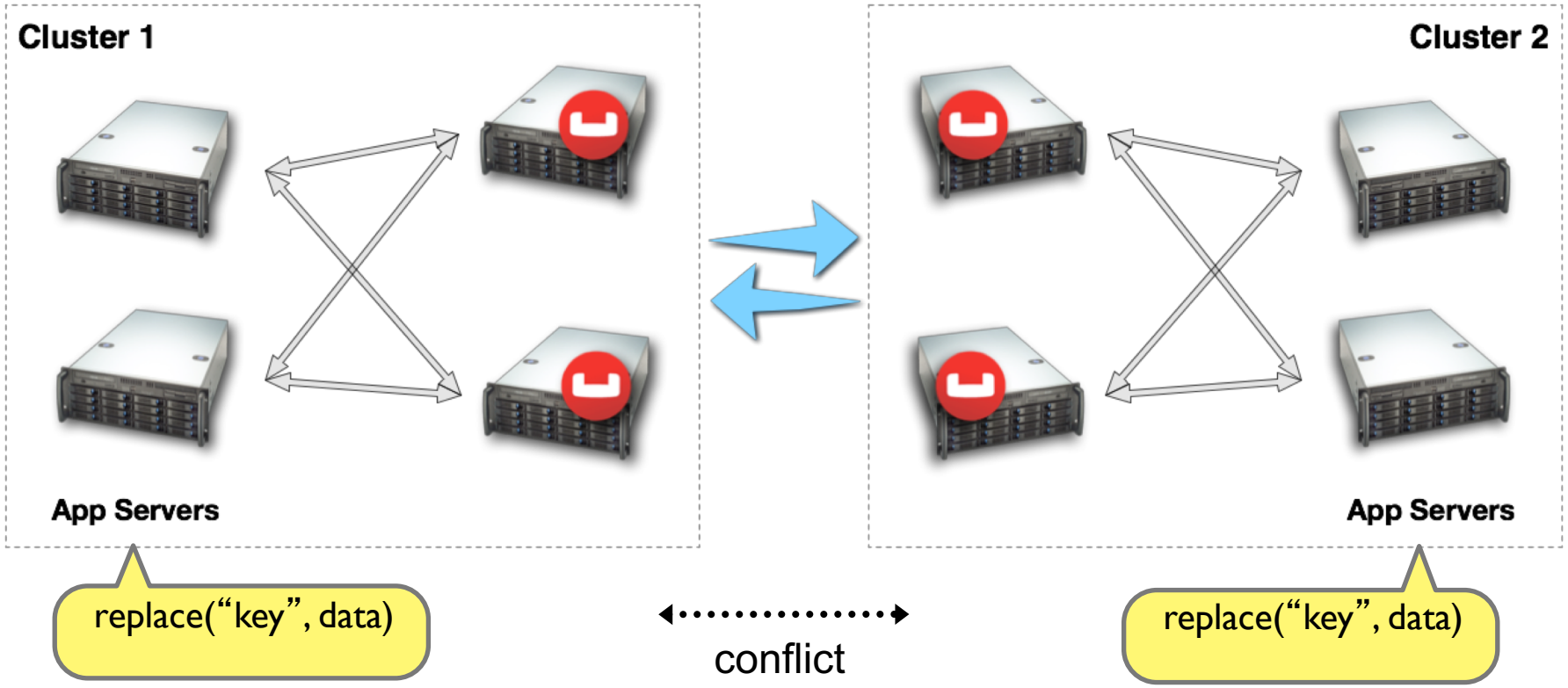
Cluster-aware

- Follows the cluster map
- Source and destinations can have different number of servers
- Takes topology update into account if E.G. a node of the destination cluster goes down



Conflict resolution

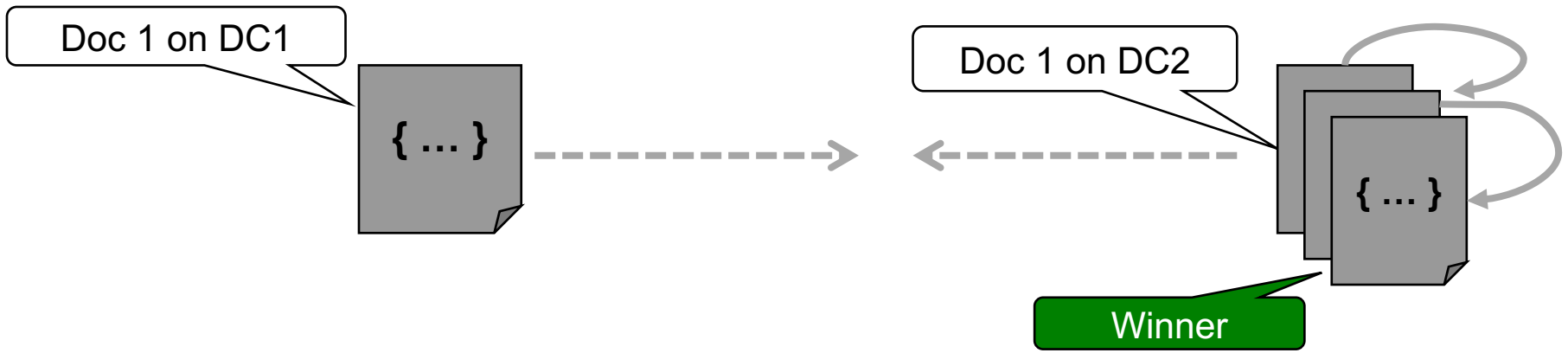
What happens when you write the same key in multiple clusters?



Conflict resolution

- XDCR is eventually consistent; checks document metadata to resolve conflicts:
  1. Numerical sequence (incremented on each mutation)
  2. CAS value
  3. Expiration (TTL) value

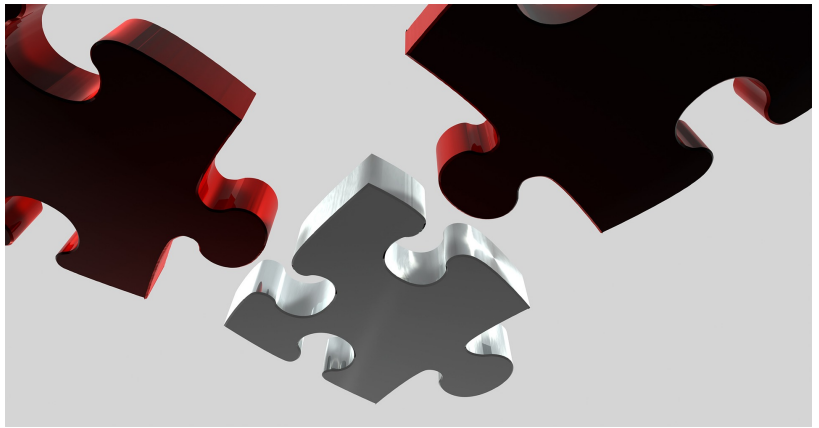
→All clusters will pick the same “winner”



Other Properties

- Automatic Resume
  - Push based replication
  - Compares revisions before transfer
  - Source tracks what destination last received via checkpoints
  - Resume from last checkpoint

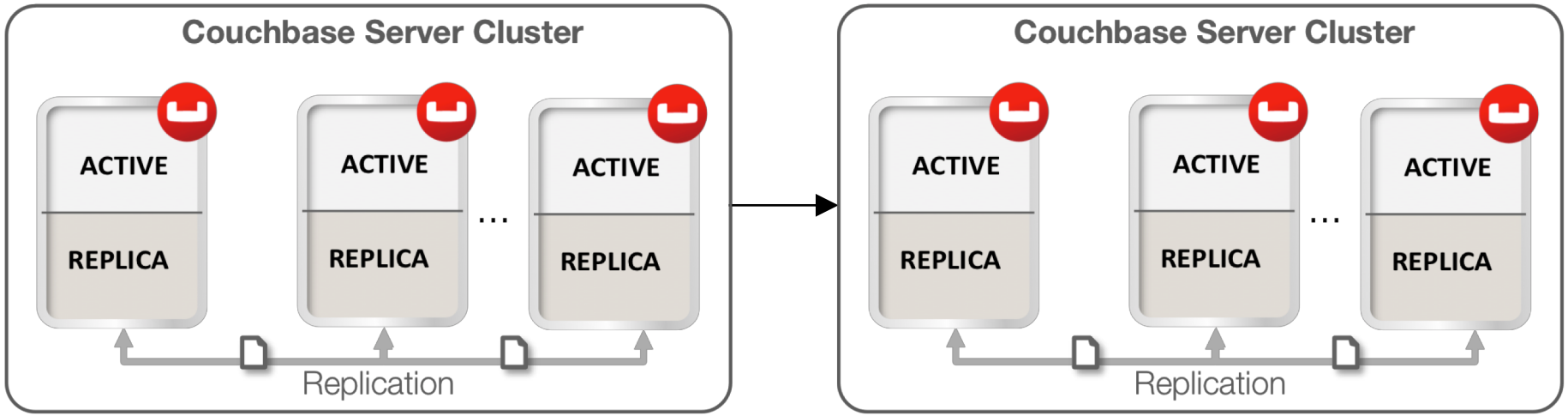
- Transfer Protocols
  - V1: HTTP
  - V2: XMEM (uses Memcached protocol)
  - Encrypted transfer (SSL)



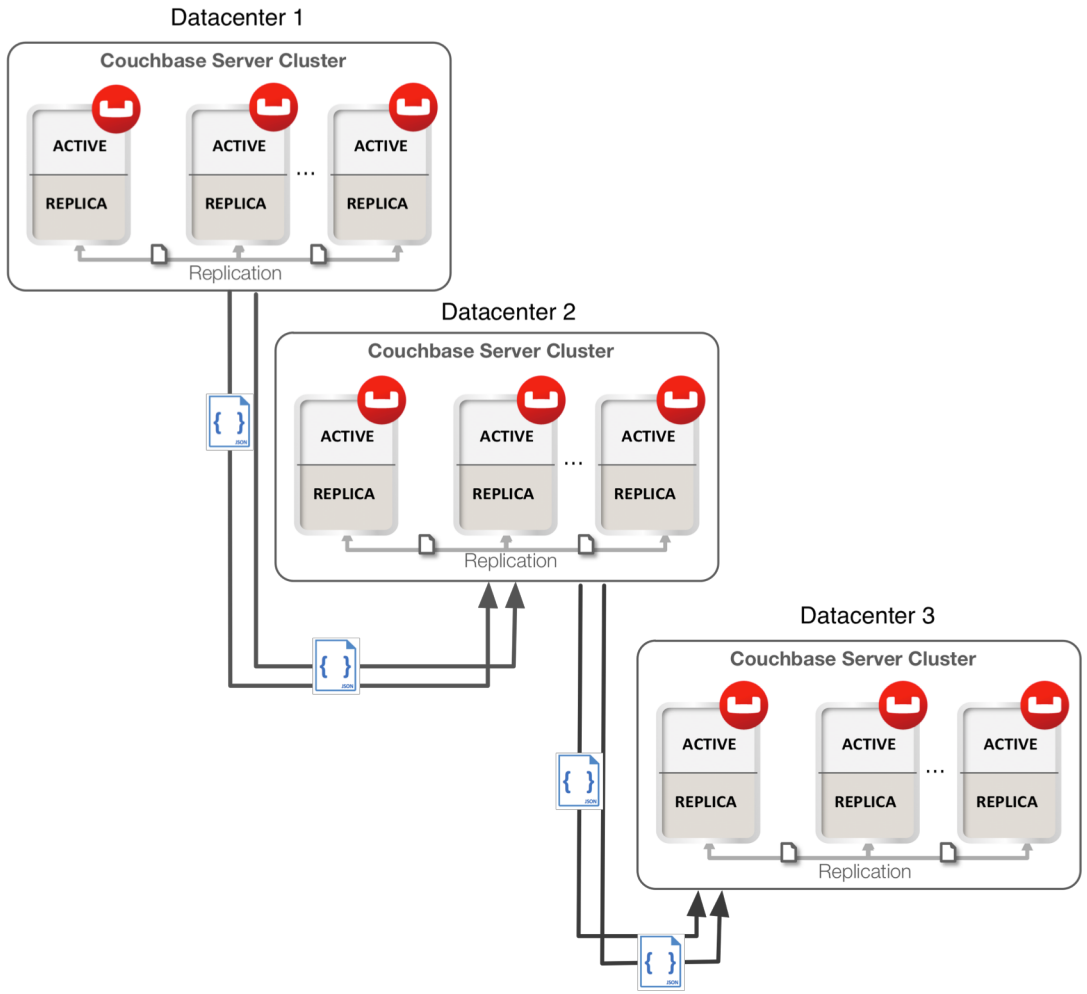
3 | Topologies & Use Cases

Uni-Directional

- Hot spare / Disaster recovery
- Development/testing copies
- Heavy reporting (since 4.0 via MDS)
- Integrate to Elasticsearch
- Integrate to custom consumer



Chain

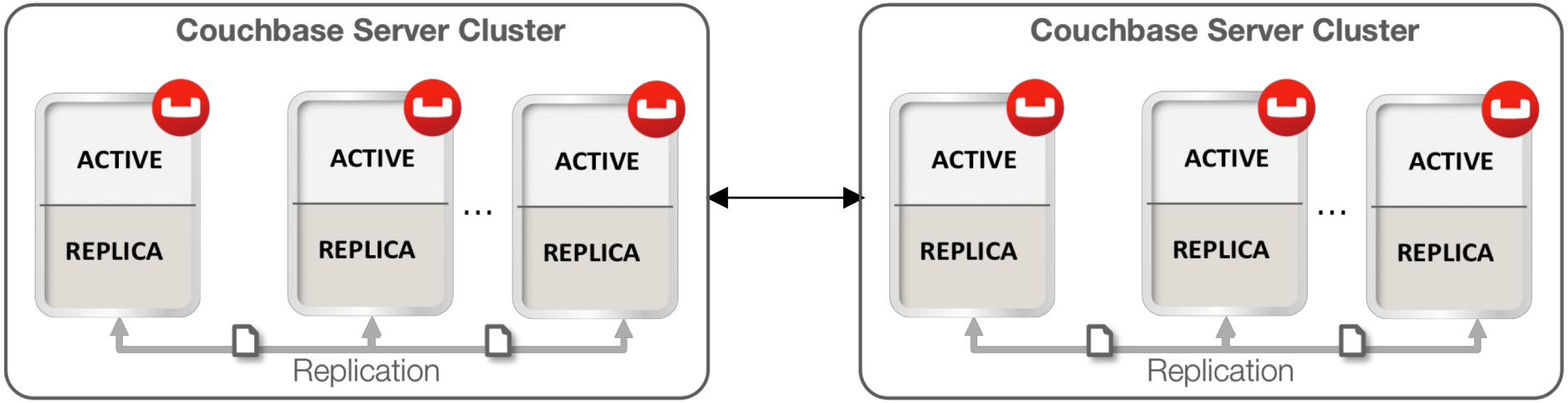

Datacenter 1
Couchbase Server Cluster
ACTIVE
REPLICA
ACTIVE
REPLICA
ACTIVE
REPLICA
Replication
Datacenter 2
Couchbase Server Cluster
ACTIVE
REPLICA
ACTIVE
REPLICA
ACTIVE
REPLICA
Replication
Datacenter 3
Couchbase Server Cluster
ACTIVE
REPLICA
ACTIVE
REPLICA
ACTIVE
REPLICA
Replication

Data Aggregation

(Filtered) Propagation

Bi-Directional (aka Active-Active)

- Multiple active masters
- Disaster Recovery
- Data locality



Caution

- **Avoid updating the same document in multiple clusters with bi-directional XDCR**
  - Be sure to understand the conflict resolution rules
- **Best Practices**
  - Data Center stickiness
    - Keep users/transactions isolated to a DC
    - Only redirect to another DC in case of major outage
  - Use separate key spaces (e.g. DC prefix) to avoid conflicts on individual documents. Example:
    - dc1::user:a9838-s92-s00
    - dc2::user:293ba-293-922

4 | Tuning Parameters

Advanced Settings

Parameter	Default	Description
Optimistic replication threshold	256	If the size of a document is higher than this threshold then XDCR will send a getMeta request (in batches) from the source cluster to the destination cluster in order to find out if the document needs to be sent over.
Source nozzles per node	2	Controls the parallelism
Target nozzles per node	2	Controls the parallelism
Checkpoint interval	1800	Time in seconds between checkpoints. This defines the amount of data which has to be resent in case of a communication failure.
Batch count	500	Controls the number of documents to be transferred in one batch.
Batch size (kB)	2048	Limits the size of a batch in KB.
Failure retry interval (s)	10	Time in seconds before XDCR retires to resume the replication after a failure.
Filter	None	The filter expression allows you to limit the data which will be sent over the wire by using a regular expression on the document key.

Thank you

Couchbase