Percona Backup for MongoDB

Solution for consistent backups of MongoDB sharded cluster

PERCONA
Backup for MongoDB

PERCONA

The Main Problem

Consistency

If there are write operations during the dump operation, the dump will not reflect a single moment in time.

Changes made to the database during the backup process can affect the output of the backup.

Source:
https://docs.mongodb.com/manual/reference/program/mongodump/

Single-Shard Backups

Use "mongodump --oplog" option

Creates a oplog file as part of the mongodump output. Which contains oplog entries that occur during the mongodump operation. This file provides an consistent snapshot of the state of a mongod instance.

Source:
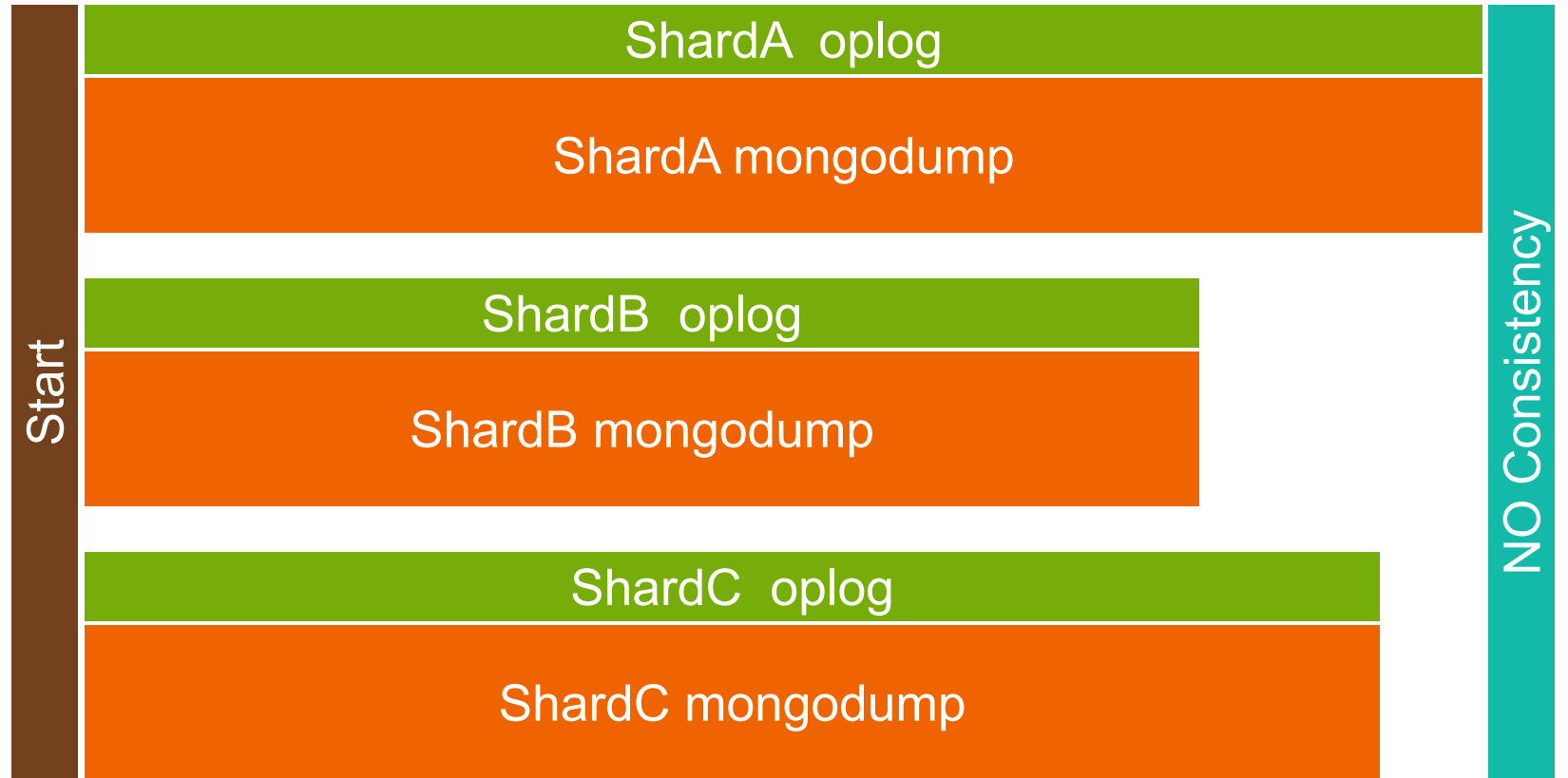https://docs.mongodb.com/manual/reference/program/mongodump/

Single-Shard Backups

Start

oplog

mongodump

Consistency at the end

PERCONA

Multi-Shard Backups

Backups created with mongodump do not maintain the atomicity guarantees of transactions across shards.

Source:
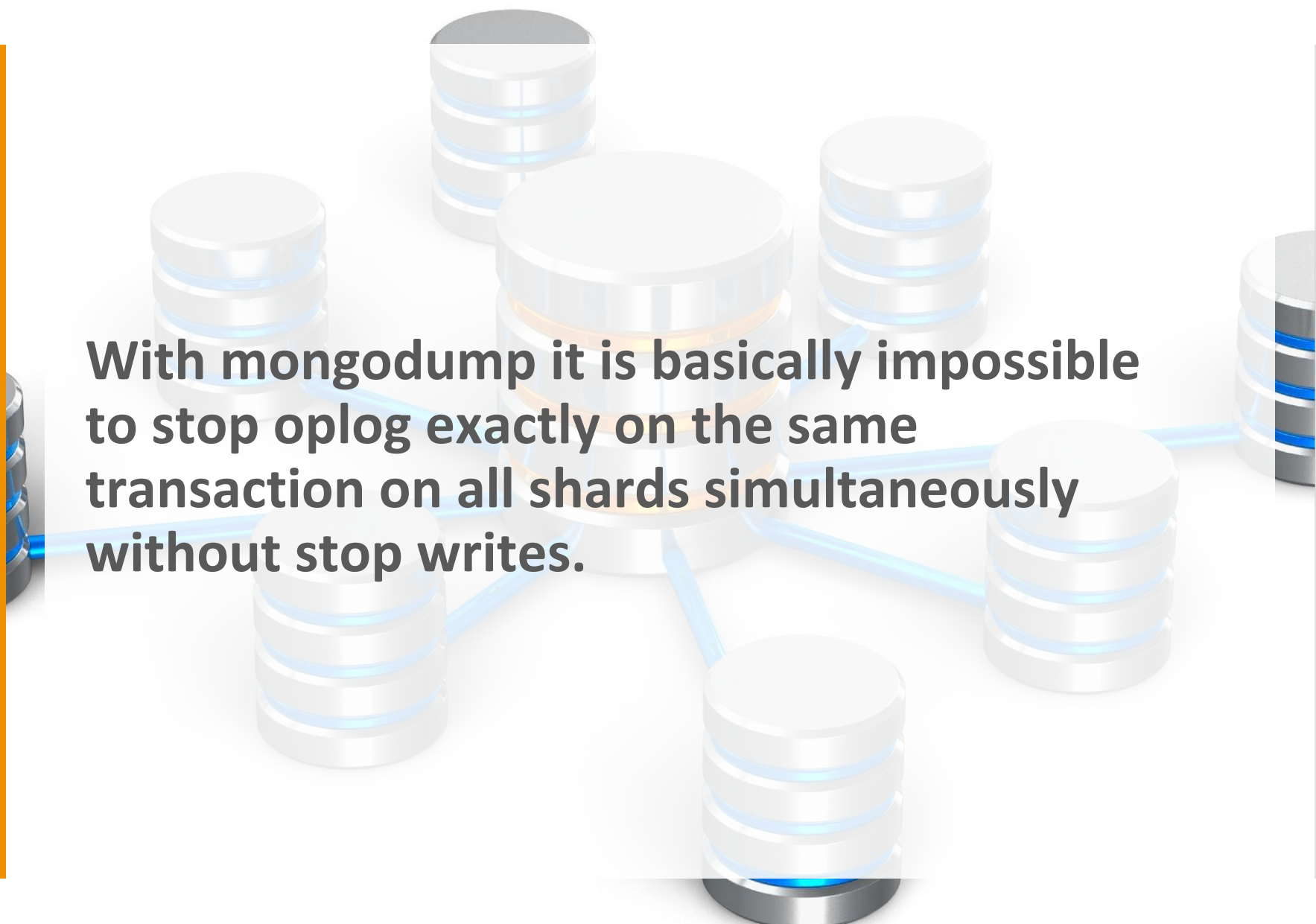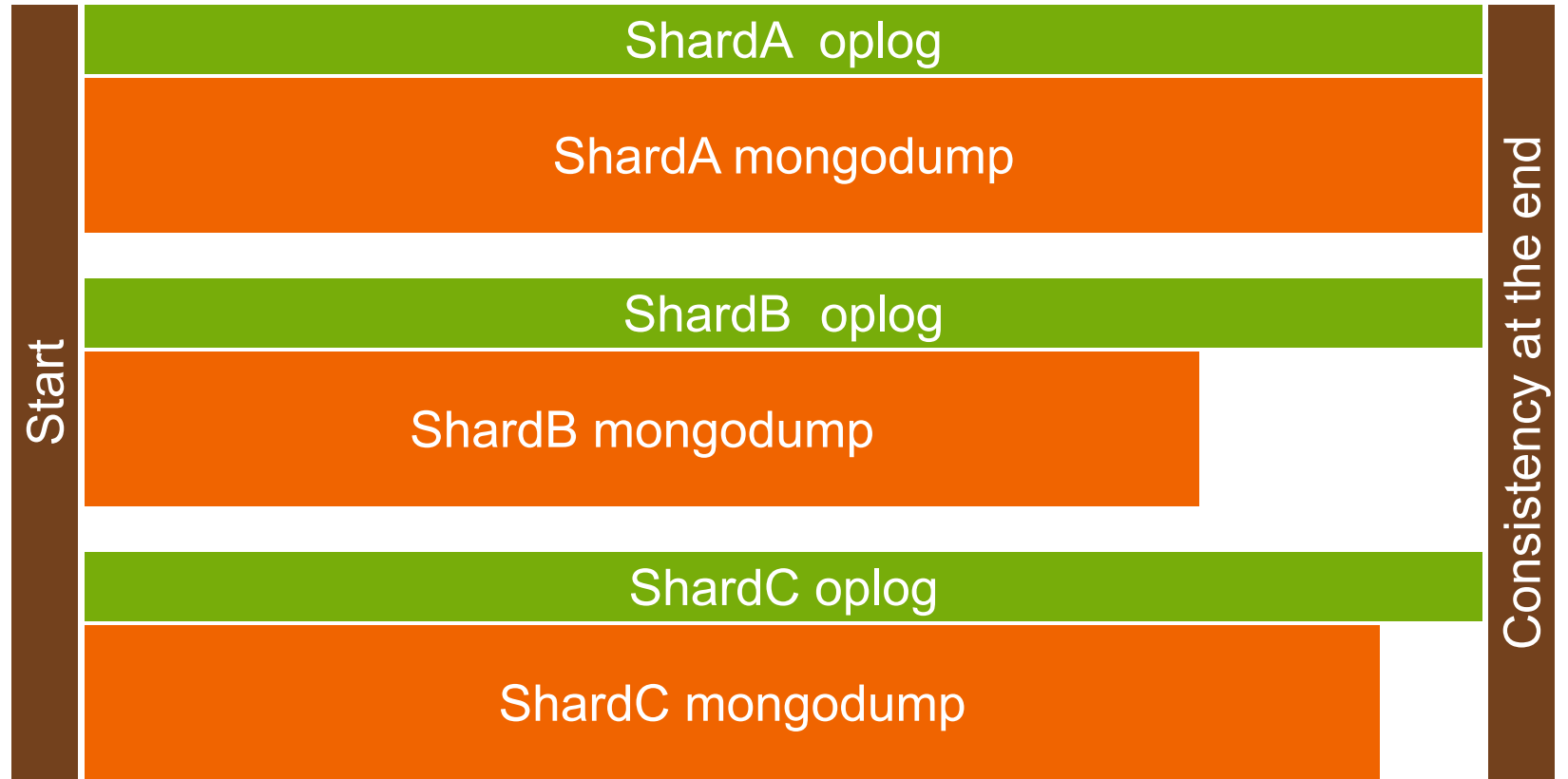https://docs.mongodb.com/manual/reference/program/mongodump/

Why?

With mongodump it is basically impossible to stop oplog exactly on the same transaction on all shards simultaneously without stop writes.

Multi-Shard Backups

Start

ShardA oplog

ShardA mongodump

ShardB oplog

ShardB mongodump

ShardC oplog

ShardC mongodump

Consistency at the end

PERCONA

Percona Backup for MongoDB

Solution for consistent backups of MongoDB sharded cluster

PERCONA

Agent

- Agent must be attached on localhost of each mongod instance (including secondaries and config server instances).

- Agent connects and works only with/via MongoDB. **No** need to handle additional security aspects: open ports, TLS/SSL, authorization etc.

- Agent detects if it is a good candidate to do the backup or restore operations and coordinates with the other instances to accomplish the requested actions.

PERCONA

CLI

- Administrators observe and control the backups or restores with a pbm CLI command that they can run from any host with the access to the MongoDB cluster.

- CLI connects and works only with/via MongoDB. **Not** needed to handle agent addresses, TLS/SSL, authorization etc.

PERCONA

How to start using it?

- Create user If authentication is enabled
- Run agent locally for each mongod instance
- Configure location of remote storage
- Run "pbm backup" command

PERCONA

Current state and Roadmap

PERCONA

Current state

- **MongoDB 3.6, 4.0, 4.2 versions are supported.**

- **Backup and restore via S3-protocol.**
  Officially supported engines: AWS S3, GCP GCS (interoperability), MinIO.

- **Backup and restore from locally mounted NAS/SAN.**

- **Take backup from Secondary or hidden nodes if possible.**

- **GZIP compression of oplogs and logical backups by default.**

© 2019 Percona

PERCONA

Roadmap

- **Recovery to any point-in-time (PITR)**

- **Add ability to delete old backups from CLI**

- **Centralize logs**

- **Add --progress and --wait modes for backups and restore commands**

- **Supporting cancelling a backup in progress**

- **Utilize physical backups for both backup and restore for Percona Server for MongoDB only**

PERCONA

We are Hiring!

Speak to me ;)

© 2019 Percona

PERCONA

Q & A

Mykola Marzhan
Director of Server Engineering

PERCONA