



PERCONA
LIVE EUROPE
AMSTERDAM

Replication Smackdown

Ronald Bradford
<http://ronaldbradford.com>

2016.10

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability

AGENDA

Replication

- What, Why, How, MySQL Types

Production Systems

- Requirements, Needs, MTBF

A New Mindset

- Availability, Classification, Pipeline

Which is the best replication approach to use?

Answer: It depends

Which is the best replication approach to use?

Answer: It depends



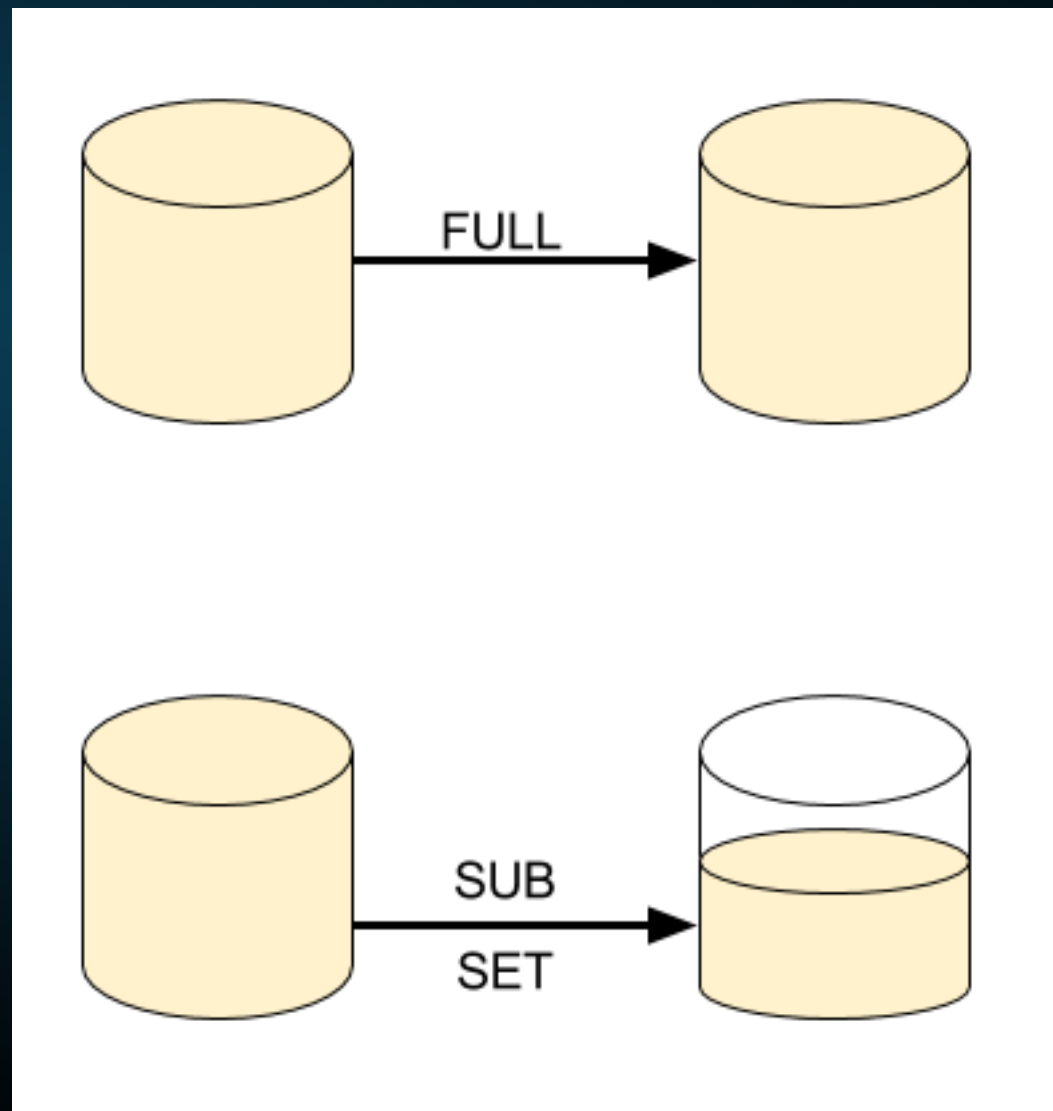
What is replication?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



WHAT?



WHAT?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

WHAT?

- Copy of data
 - Full or partial

WHAT?

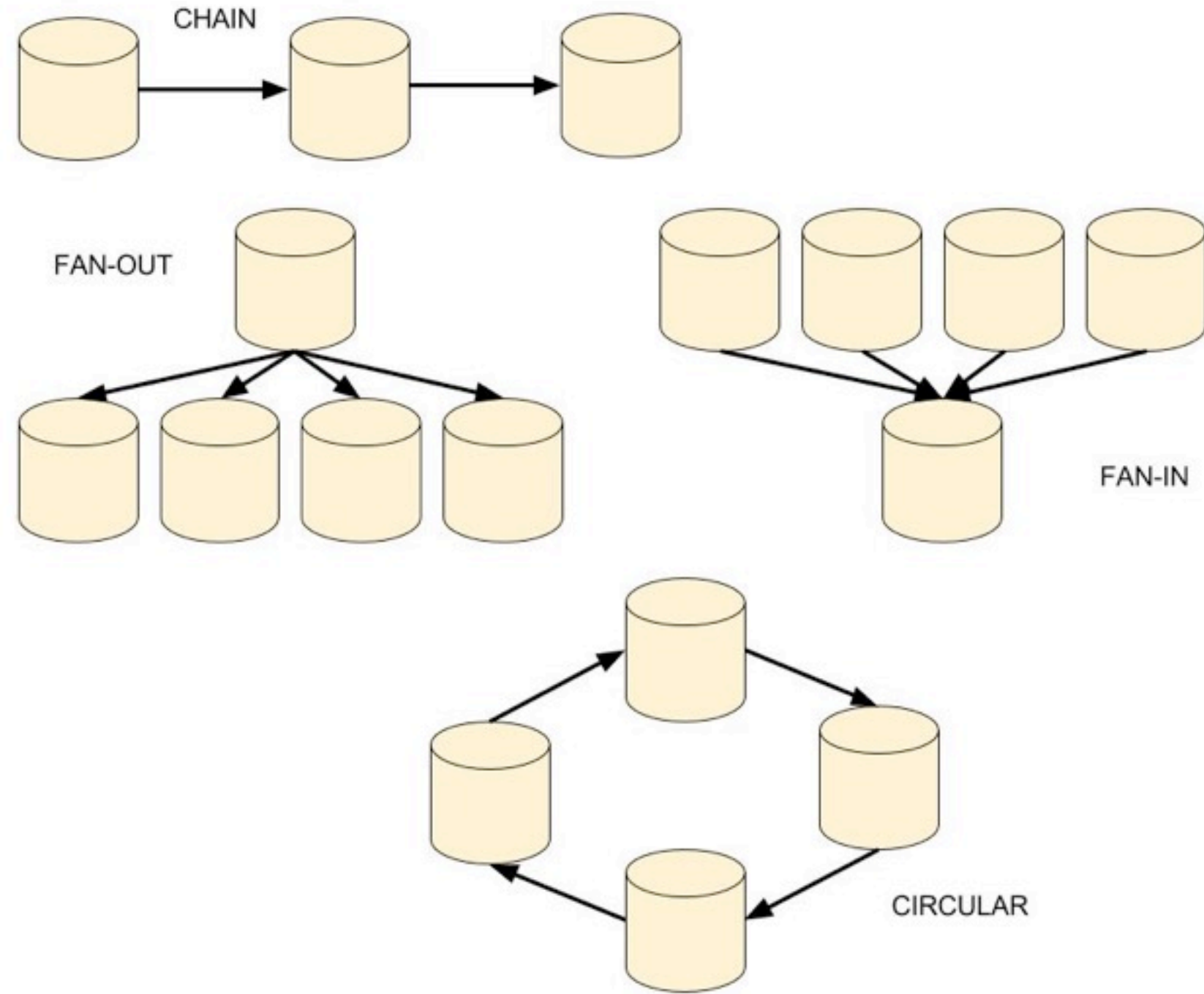
- Copy of data
 - Full or partial
- Translation of data
 - e.g. Oracle to MySQL

WHAT?

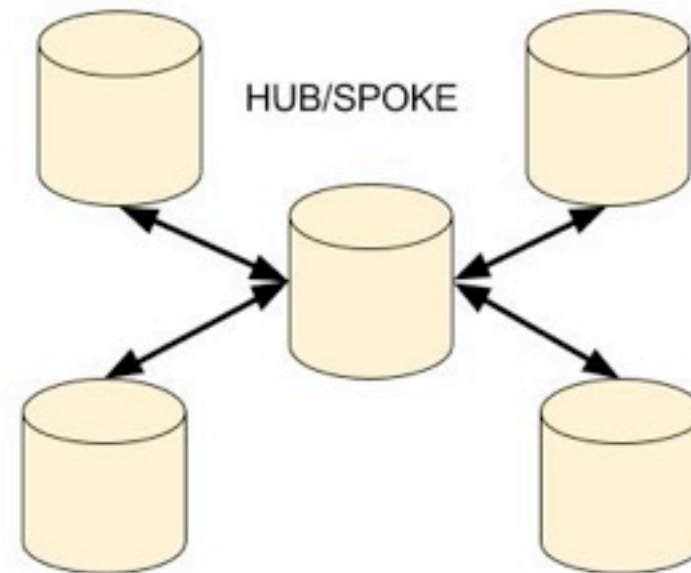
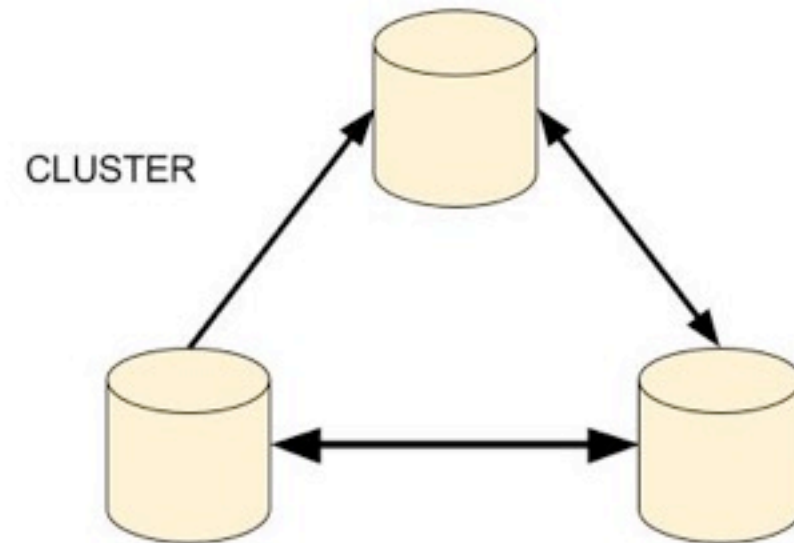
- Copy of data
 - Full or partial
- Translation of data
 - e.g. Oracle to MySQL
- Transformation of data
 - e.g. MySQL to DW
 - e.g. MySQL to Hadoop



ONE WAY FORM



TWO WAY FORM



Why use replication?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

WHY?

- Redundancy (copies)
- Availability (distribution)
 - Failover
- Scalability (read/write)
- Performance (optimizations)
- Backups (locking, load)
- Consolidation



What are production requirements?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

PROD REQS

- Acceptable latency
- Acceptable throughput
- Performance under load
- High availability
- Failover
- Disaster recovery
- Security
- Backups
- Load testing
- Monitoring
- Alerting
- Sizing

How does replication help achieve production requirements?

What does MySQL offer for replication?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

TYPES

- [Traditional] MySQL Replication
 - Asynchronous / Semi-synchronous
- MySQL Cluster
- Galera (MySQL/Percona/MariaDB)
- MySQL Group Replication (RC) / InnoDB Cluster (TBD)
- Amazon RDS MAZ & Aurora
- Others (e.g. Google Cloud SQL, Clustrix, DRBD)



What are barriers to usage?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

TRADITIONAL

async - since 3.23

- lag
- drift
- consistency
- throughput

semi-sync - since 5.5

- lag
- drift
- consistency
- throughput

MYSQL CLUSTER

- Additional installation complexity
 - Data/SQL/Admin nodes
 - Different admin interface
 - Different backup strategy
 - LAN based
- Same SQL syntax
 - Limited large join options



- Syntax limitations
- Feature limitation (e.g. MEMORY table), primary key
- OS Limitations (Linux Only)
- Hot data spots /Large transactions in multi-master write
 - Timeouts
- Schema upgrades
- LAN v WAN

GROUP REPLICATION

- Syntax Limitations, Feature Limitation
- Hot Spots/Large transactions in multi-master write
- Supported on Linux, Windows, Solaris, FreeBSD, OSX

Requirements

- MySQL 5.7, GTID, binlog_format=ROW
- Other configuration settings
- RC only

INNODB CLUSTER

- Only in labs
- Based on group replication limitations
- Helps solve the routing problem
- Simplified orchestration in JS



What are improvements to replication?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

MYSQL VERSIONS

- 5.6
 - crash save slaves, GTID, group commit, multi-threaded
- 5.7
 - semi-sync improvements, multi-threaded V2, multi-source, XA support, group replication



- MySQL
- Percona
- MariaDB
- Custom
- Others

Does not improve async/semi sync replication?
Improves [faster] failover

<https://mariadb.com/kb/en/mariadb/gtid/>

<https://www.facebook.com/notes/mysql-at-facebook/lessons-from-deploying-mysql-gtid-at-scale/10152252699590933/>

<http://code.openark.org/blog/mysql/refactoring-replication-topology-with-pseudo-gtid>

MULTI-THREADED

- parallel schema applier (5.6)
- parallel query applier (5.7)

Does improve async/semi sync replication?
Improves performance (i.e. lag)
Does not eliminate lag

How do we apply replication to our systems?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

What are your business needs?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



What are your business needs? What are [ideal] business needs?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



What are your business needs?
What are [ideal] business needs?
What are [acceptable] business needs?

OBJECTIVES

- Mean Time Between Failure (MTBF)
- Mean Time To Detect (MTTD)
- Mean Time To Recover (MTTR)
- Recovery Point Objective (RPO)
- Recovery Time Objective (RTO)

The new mindset in architecture

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

1. Availability

2. Classification

3. Pipeline



Data Availability

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



DATA AVAILABILITY

- Ability to write data
- Ability to read data
- Ability to [read|write] cached data
- Ability to operate with no data

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

What do I mean?

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



What do I mean?

What is your definition of downtime?



AVAILABILITY

- Not database availability
 - e.g. those maintenance windows
- Not data availability
 - e.g. Write/Read/Cache/None
- It is all about service availability
 - i.e. endpoints

Data Class

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



DATA CLASS

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

DATA CLASS

All data in a RDBMS schema (or even table) is not equal

- Users data (register, modify)
 - Change password
 - Last login date
- Add content
- Comment/Rate/Score

DATA CLASS

All data in a RDBMS schema (or even table) is not equal

- Users data (register, modify)
 - Change password
 - Last login date
- Add content
- Comment/Rate/Score
- Current Order
- Last Order
- Historical Orders
- Credit Card Details

DATA CLASS

- Some data needs to be more highly available than other data
- Some data access requires more responsiveness than others
- Some data has acceptable data loss
- Some data can be unavailable some of the time
- Some data visibility can vary between users
- All data should be secure, some more secure

DATA CLASS

- Some data needs to be more highly available than other data
- Some data access requires more responsiveness than others
- Some data has acceptable data loss
- Some data can be unavailable some of the time
- Some data visibility can vary between users
- All data should be secure, some more secure

Reclassification of data changes replication requirements

PROD SYSTEMS

Requirements

- Acceptable latency
- Acceptable throughput
- Performance under load
- High availability
- Failover
- Disaster recovery
- Security

Responsibilities

- MTBF
- MTTD
- MTTR
- RPO
- RTO



Example (Financial)

DATA CLASS

Single greatest feature loss

- Referential Integrity
 - A & C of ACID

Data Pipeline

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

DATA PIPELINE

- A single request does not produce one synchronous response
- Data is not stored in one RDBMS or product type
- Data locality for responsiveness
- Use product strengths for data manipulation

Rethinking how to choose MySQL Replication

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



PERCONA
LIVE EUROPE
AMSTERDAM

The next great social experience

UTF-FTV

you-twit-face-flix-talk-vr

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



OLD WAY

- Found in many “traditional” frameworks/OSS products
 - Users table in monolithic schema
 - Synchronous web requests for information
 - Polling for new/streaming information
 - Data for application is available or not available
 - Replication enables read scalability only

NEW WAY STORES

- Highly available synchronous store
- Columnar store
- Messaging (PUB/SUB)
- Graph
- Queue
- Search

USER PATH

- Microservices

- Login/Logout

- Register/Maintain

- Log actions (login good/bad, click, mouse movement)

- Friends

- Friends interactions

Availability: What type of data access is available



NEW WAY

- Graph Store
 - Friends
 - Customized for relevancy/strength algorithm
- Queue
 - Password changes, User profile changes
 - Lost password

NOTIFICATIONS

- User actions (success, failure) are published (PUB)
- (SUB) Subscriber logs information
- (SUB) Subscriber audits for unexpected behavior
- (SUB) Subscriber notifies friends user is online/offline

Pipeline of multiple asynchronous actions

RATE ME

- Rate a comment
 - Client only feedback
 - Action held on client
 - batched, overloaded, timed transmission
 - Supports rate/unrate (client side only)

Class: Optimized for payload

Conclusion

@RonaldBradford
#PerconaLive

EffectiveMySQL.com - Performance, Scalability, Site Reliability



CONCLUSION

- Do you choose a replication approach to match your [ideal] business needs and data store(s)

OR

- Do you architect a data infrastructure to meet your [ideal] business needs and target specific replication (aka availability) approaches where applicable



What does this have to do with replication?

Answer: Everything

