



# Verifying MySQL Replication Integrity

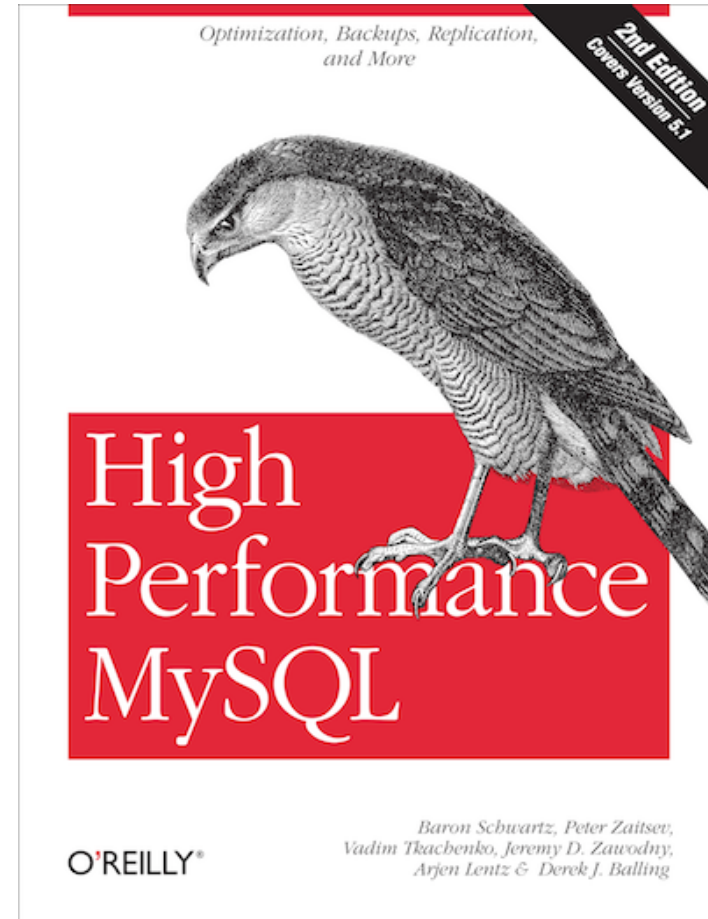
Baron Schwartz  
January 2012

# Agenda

---

- Introduction
- MySQL Replication Failure Modes
- How Data Drift Detection Works
- How To Use Percona Toolkit To Verify Data Integrity
- What's New In Version 2.0
- Resources

# About Me



# About Percona

- Software and services for MySQL users
- Software is 100% free and opensource:
  - Percona Server with XtraDB
  - Percona XtraDB Cluster
  - Percona XtraBackup
  - Percona Toolkit
  - A variety of other tools
- Services:
  - Consulting
  - Support
  - Training
  - Development

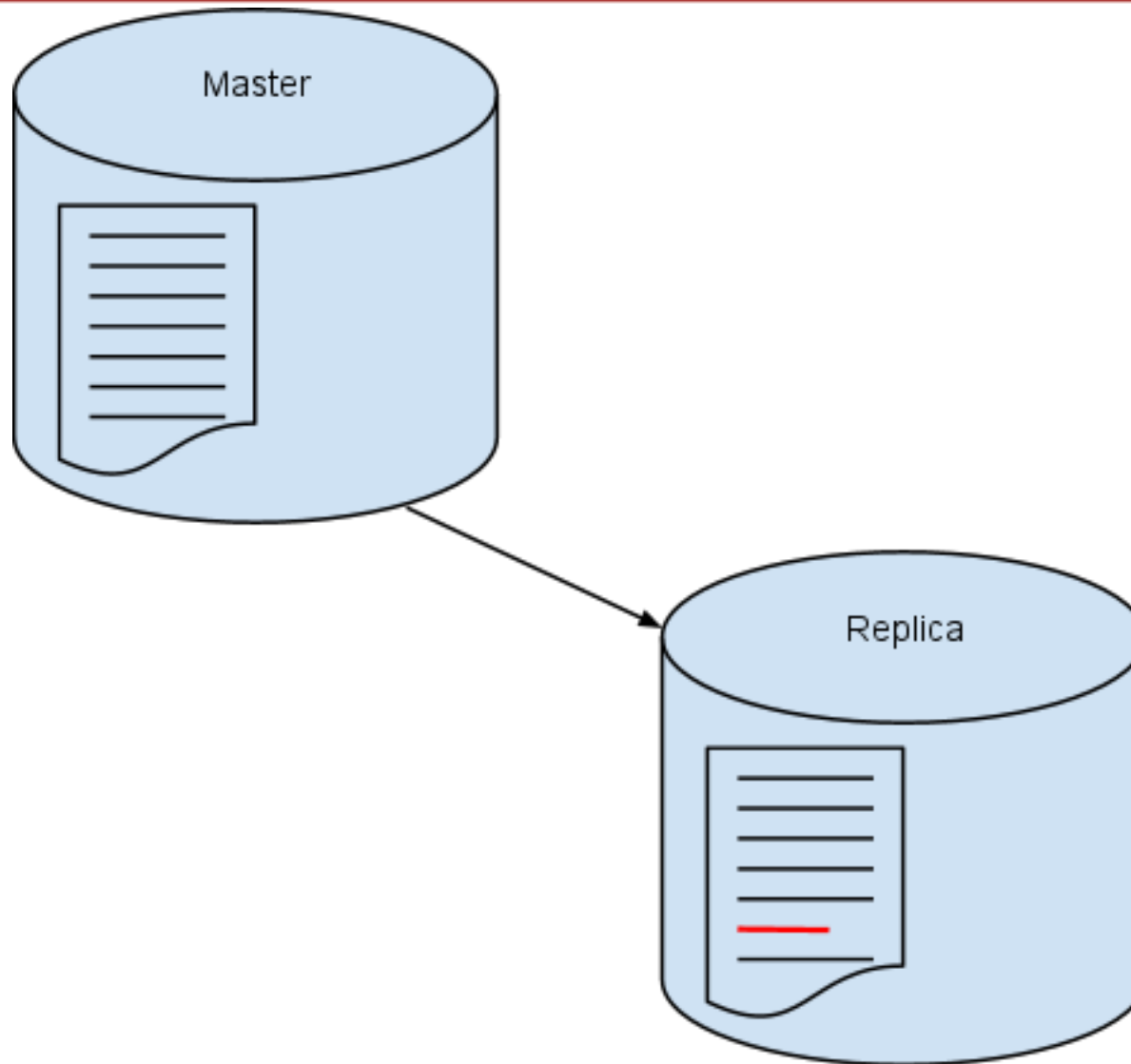
# MySQL Replication Failure Modes

- Complete Failure
  - Replication stops
  - Typical cause: duplicate key violation
- Replication Delay
  - Replicas can't keep up with the master
  - Application reads stale data
- Data Drift
  - Can be invisible
  - Can ultimately cause complete failure
  - Are you taking backups from your replica?

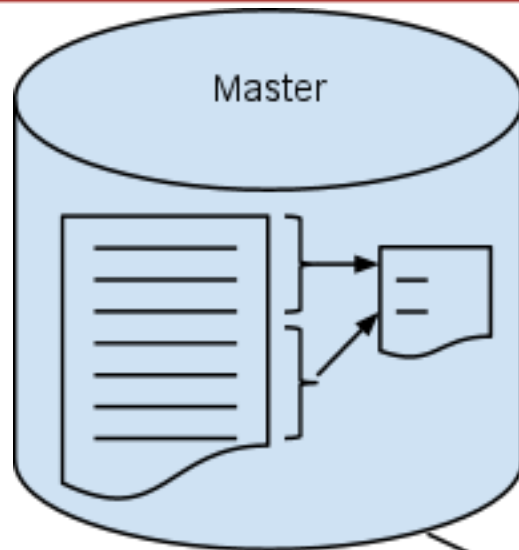
---

If You Use Replication, Then You  
Need To Guard Against Data Drift.

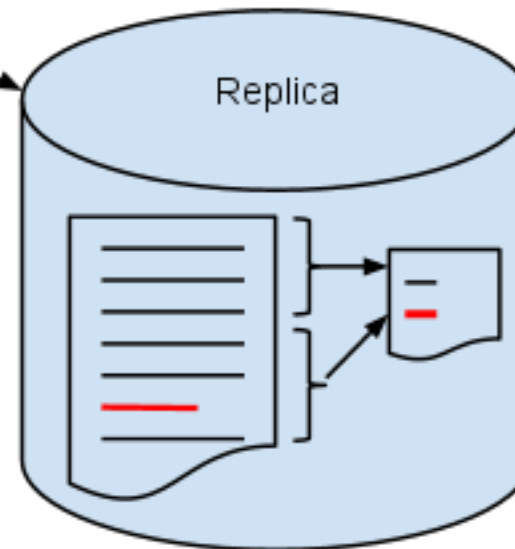
# Data Drift Detection



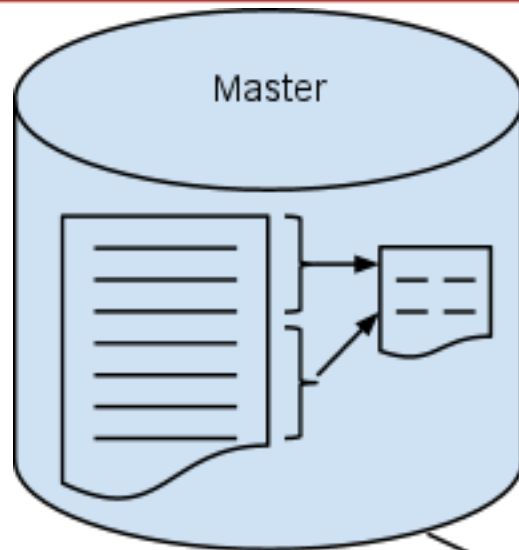
# Data Drift Detection



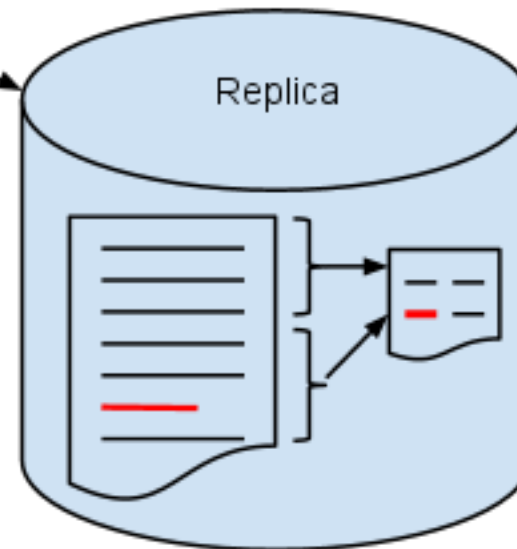
Execute `INSERT... SELECT`  
on ranges of rows to generate  
checksums of the data



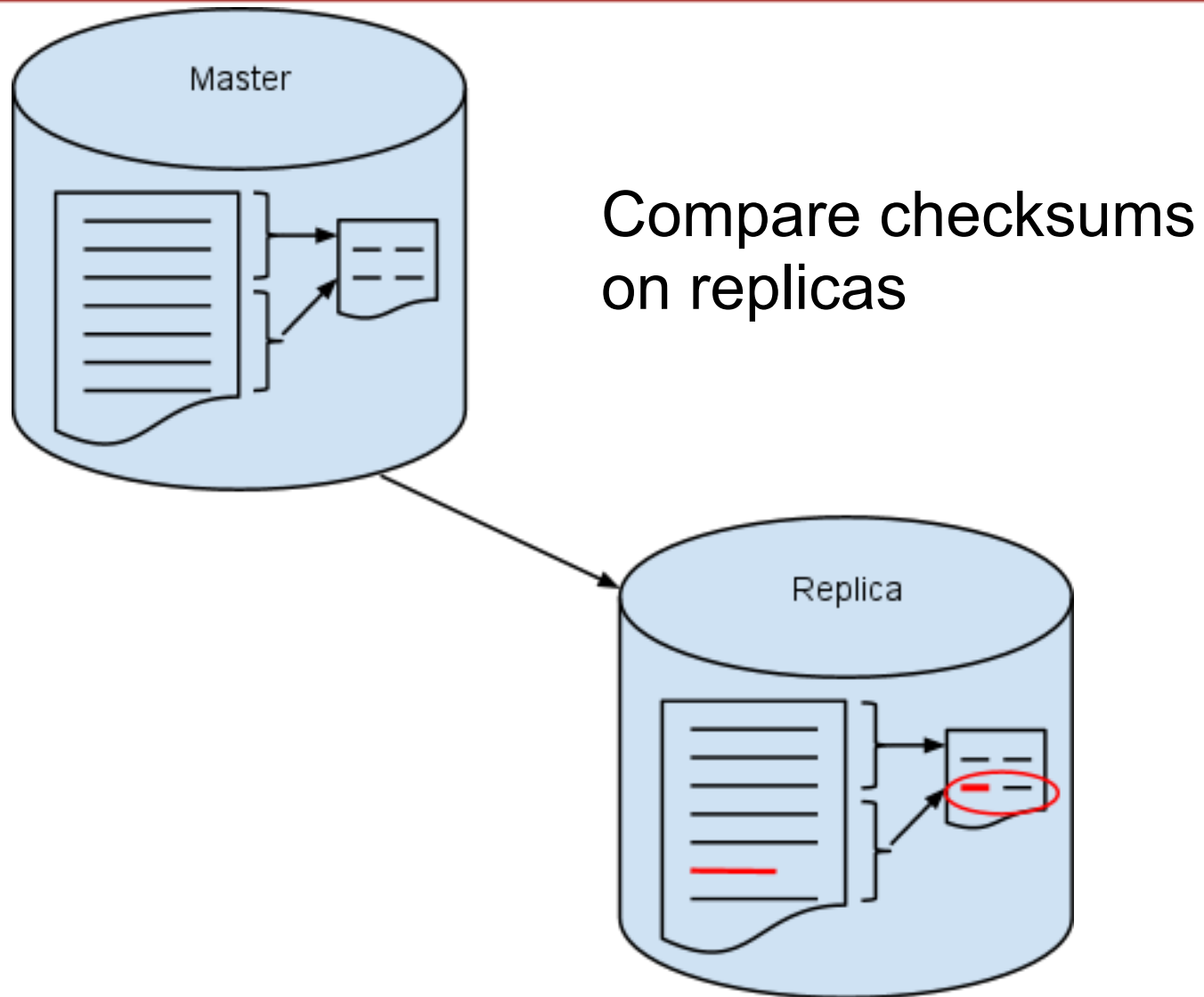
# Data Drift Detection



Insert the master's  
checksum as a literal  
value



# Data Drift Detection



# Data Drift <> Replication Delay

- This technique compares logical data
- Replication is asynchronous and always has some delay
- Logical drift detection is immune to delay

# Using Percona Toolkit 2.0

- Percona Toolkit's tool: pt-table-checksum
- How to use:
  - Download
  - Install
  - Execute

```
[baron@localhost ~]$ ./pt-table-checksum --databases sakila
      TS ERRORS  DIFFS      ROWS  CHUNKS  SKIPPED    TIME TABLE
01-18T11:02:22    0     0       200     1      0    0.037 sakila.actor
01-18T11:02:22    0     0       603     1      0    0.021 sakila.address
01-18T11:02:22    0     0        16     1      0    0.029 sakila.category
01-18T11:02:22    0     0       600     1      0    0.026 sakila.city
01-18T11:02:22    0     0       109     1      0    0.020 sakila.country
01-18T11:02:22    0     0       599     1      0    0.037 sakila.customer
01-18T11:02:22    0     0      1000     1      0    0.040 sakila.film
01-18T11:02:22    0     0     5462     1      0    0.029 sakila.film_actor
01-18T11:02:22    0     0      1000     1      0    0.025 sakila.film_category
01-18T11:02:22    0     0      1000     1      0    0.019 sakila.film_text
01-18T11:02:22    0     0     4581     1      0    0.047 sakila.inventory
01-18T11:02:22    0     0         6     1      0    0.023 sakila.language
01-18T11:02:22    0     0    16049     1      0    0.082 sakila.payment
01-18T11:02:23    0     0    16044     1      0    0.077 sakila.rental
01-18T11:02:23    0     0         2     1      0    0.018 sakila.staff
01-18T11:02:23    0     0         2     1      0    0.018 sakila.store
[baron@localhost ~]$
```

# How It Works

---

- Connect to a server
- Discover and connect to replicas
- Run safety checks
- Discover databases and tables
- Paginate ("nibble") through each table
- Report the results

# Using Percona Toolkit 2.0

```
[baron@localhost ~]$ ./pt-table-checksum --databases sakila
```

	TS	ERRORS	DIFFS	ROWS	CHUNKS	SKIPPED	TIME	TABLE
01-18T11:02:22		0	0	200	1	0	0.037	sakila.actor
01-18T11:02:22		0	0	603	1	0	0.021	sakila.address
01-18T11:02:22		0	0	16	1	0	0.029	sakila.category
01-18T11:02:22		0	0	600	1	0	0.026	sakila.city
01-18T11:02:22		0	0	109	1	0	0.020	sakila.country
01-18T11:02:22		0	0	599	1	0	0.037	sakila.customer
01-18T11:02:22		0	0	1000	1	0	0.040	sakila.film
01-18T11:02:22		0	0	5462	1	0	0.029	sakila.film_actor
01-18T11:02:22		0	0	1000	1	0	0.025	sakila.film_category
01-18T11:02:22		0	0	1000	1	0	0.019	sakila.film_text
01-18T11:02:22		0	0	4581	1	0	0.047	sakila.inventory
01-18T11:02:22		0	0	6	1	0	0.023	sakila.language
01-18T11:02:22		0	0	16049	1	0	0.082	sakila.payment
01-18T11:02:23		0	0	16044	1	0	0.077	sakila.rental
01-18T11:02:23		0	0	2	1	0	0.018	sakila.staff
01-18T11:02:23		0	0	2	1	0	0.018	sakila.store

```
[baron@localhost ~]$ █
```

# What's New In Version 2.0?

---

- Simpler
- Self-tuning / zero-config
- Safe and unintrusive
- Resilient

# Simpler

---

- Does only one thing
- Many behaviors combined or eliminated
- Fewer command-line options
- All of the chunking/nibbling complexity removed

# Self-Tuning

---

- The remaining command-line options are sensible
- Most things self-adjust
- You can run it with no options at all

# Safe and Unintrusive

- Self-throttles to prevent overloading
  - Auto-adjusts to changing server load
- Self-configures to be the victim
  - e.g. sets `innodb_lock_wait_time = 1`
- Won't impact replicas
  - Waits if they begin to lag
- Runs extensive safety checks to prevent problems
  - Stops if replication stops
  - Won't run if there are replication filters
  - Checks table size on replicas when needed

# Resilient

---

- Most things restart, resume, retry, reconnect
- This is very important for Getting Things Done

# Checksumming in Chunks

- Checksums  $N$  rows at a time
  - Self-adjusts chunk size to achieve a target time
- Works on any table with an index
  - Tables without indexes will be skipped
  - Oversized chunks will be skipped
- Chunks are smartly paginated without overhead

# Key Command-Line Options

---

- --chunk-time
- --replicate
- --resume
- --max-load
- --quiet
- --retries

# Tweaking Replica Detection

- Automatic detection doesn't always work
- You can override it
  - Use `--recursion-method = DSN`
  - This specifies a table full of DSNs

# What If You Find Differences?

- If pt-table-checksum finds differences, you can use pt-table-sync to resolve them.
- Be sure to read the documentation, because *it does change data*

# Future Releases

- We now have two dedicated developers
  - Daniel Nichter
  - Brian Fraser
- Releases about once a month
  - Coming soon: new pt-stalk and pt-diskstats
  - Thereafter, pt-summary, pt-mysql-summary, pt-table-sync, pt-online-schema-change
- Percona Toolkit is included in Percona Support contracts

# Resources

- Docs, downloads, PDF manual:
  - <http://www.percona.com/software/percona-toolkit/>
- Forum:
  - <http://forum.percona.com/>
- Mailing list:
  - <https://groups.google.com/group/percona-discussion/>
- Training courses worldwide:
  - <http://www.percona.com/training>
- High Performance MySQL
  - 3rd Edition available in April <http://t.co/OWG817iz>
- Online MySQL Configuration Wizard
  - <http://tools.percona.com/>

# More Resources

---

- Espen Braekken's webinar on preventing downtime
  - January 25th, <http://www.percona.com/webinars/>
- MySQL User's Conference April 10-12
  - <http://www.percona.com/live/>
  - Breakout session schedule announced next week

# Kill the Roots, not the Leaves

---

- Asynchronous replication is inherently fallible
- Percona XtraDB Cluster should avoid many failure cases
- More information:
  - <http://www.percona.com/software/percona-xtradb-cluster/>



**baron@percona.com**  
**@xaprb**