Billions of Records in MySQL

Lessons Learned

Lukas Vileikis, BreachDirectory

lukasvileikis.com

BreachDirectory.com

YouTube.com/@DatabaseDive

Queries

Query #1: <u>0.01s</u>

```sql
1    SELECT * FROM table WHERE MATCH(column) AGAINST("demo");
```

Query #2: <u>Timeout</u>

```sql
1    SELECT * FROM table WHERE MATCH(column) AGAINST('"demo@demo.com"' IN BOOLEAN MODE);
```

BUG#104263

lukasvileikis.com
BreachDirectory.com

whoami

- Author, Database & Security Guy
- *Hacking MySQL: Breaking, Optimizing, and Securing MySQL for Your Use Case*
- I've built BreachDirectory.com
- Blogs = {BreachDirectory, Arctype, DbVisualizer, Redgate…}
- @DatabaseDive (YouTube)
- Bugs in MySQL (BUG#104263)

lukasvileikis.com
BreachDirectory.com

Agenda

- Performance Basics

- MySQL Engines

- Optimizing Schema and Data Types

- Optimizing for Big Data & Security

- Backups

- Advanced MySQL Features

- Tools to Improve Performance

lukasvileikis.com
BreachDirectory.com

YouTube.com/@DatabaseDive

Performance Basics

- Configure your servers! (my.cnf)

- Only search for what you need

- **Use InnoDB**

- Follow performance best practices

lukasvileikis.com
BreachDirectory.com

MySQL Storage Engines

- **InnoDB / XtraDB**
- **MyISAM**
- MEMORY
- CSV
- ARCHIVE
- BLACKHOLE
- ...

Optimizing the Schema

- "Back to the Future" principle – no need to innovate
- Simple data types, indexes, normalization, partitioning.
- Disable STRICT_TRANS_TABLES & STRICT_ALL_TABLES for automated truncation
- Avoid TEXT
- VARCHAR – length in bytes + [1|2] bytes

lukasvileikis.com
BreachDirectory.com

Optimizing for Big Data

- LOAD DATA INFILE [CHARACTER SET x|FIELDS TERMINATED BY|LINES TERMINATED BY|ESCAPED BY|…]
- utf8mb4_unicode_ci

```
PARTITION BY RANGE COLUMNS(a,b,c)
(
PARTITION p0 VALUES LESS THAN ('a', 'a', '0'),
PARTITION p1 VALUES LESS THAN ('b', 'b', '1'),
...
PARTITION p27 VALUES LESS THAN ('z', 'z', ''),
PARTITION p28 VALUES LESS THAN (MAXVALUE, MAXVALUE, MAXVALUE)
);
```

lukasvileikis.com
BreachDirectory.com

Optimizing for Big Data

- ALTER TABLE functionality
- DEFAULT == UPDATE (sometimes)
- utf8mb4_unicode_ci
- SET SESSION alter_algorithm='INSTANT';
  ALTER TABLE demo ADD COLUMN …; **(MariaDB >=10.3.2)**

lukasvileikis.com
BreachDirectory.com

Backups

- `mysqldump` won't work
- Physical? Logical?
- `SELECT * INTO OUTFILE` -> `LOAD DATA INFILE`

lukasvileikis.com
BreachDirectory.com

Advanced Features

- DEFAULT == UPDATE (LOAD DATA INFILE / sometimes)
- UNIQUE INDEX -> `sort -u data.txt > uniq.txt`
- `INSERT -> INSERT IGNORE`
- Deleting data according to time periods:
  `DELETE FROM ... FOR PORTION OF time_period`
  `FROM '2024-09-10' to '2024-09-15' [WHERE ..]`
- `TRUNCATE`
- InnoDB == Killing of Queries

lukasvileikis.com
BreachDirectory.com

Tools

- **You**
- `mariadb-secure-installation`
- Slow query log
- MySQL Enterprise Edition
- DbVisualizer, Arctype, Redgate Monitor, ClusterControl...

lukasvileikis.com
BreachDirectory.com

Interesting Stuff

- Schrödinger's Paradox – data file missing, but .frm file exists

lukasvileikis.com
BreachDirectory.com

Thank You! ☺

lukas@lukasvileikis.com

lukasvileikis.com/feedback

BreachDirectory.com

YouTube.com/@DatabaseDive

lukasvileikis.com
BreachDirectory.com