MySQL Cluster (NDB)

An Introduction and Overview

Matthew Boehm

April 15 2015

PERCONA LIVE

Outline

- What is it?
- Who uses it?
- Why use it?
- How do you...?

Who - About Me

- Matthew (not Matt) Boehm
- Percona since November 2012 / PayPal Before
- Percona MySQL Architect
- Percona North American Trainer
- National Advanced-Level West Coast Swing

MySQL Cluster

WHAT IS IT?

What is it?

- MySQL Cluster is an open source, ACID-compliant, transactional database, designed to deliver real-time in-memory performance and 99.999% availability.

Break it Down

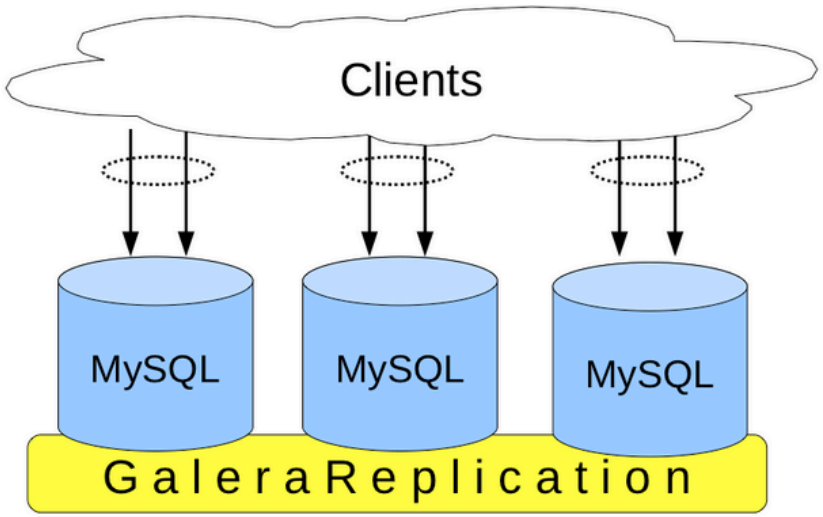
MySQL Cluster is an open source, ACID-compliant, transactional database, designed to deliver real-time in-memory performance and 99.999% availability.

What is it really?

- Just Another Engine
- Many Parts

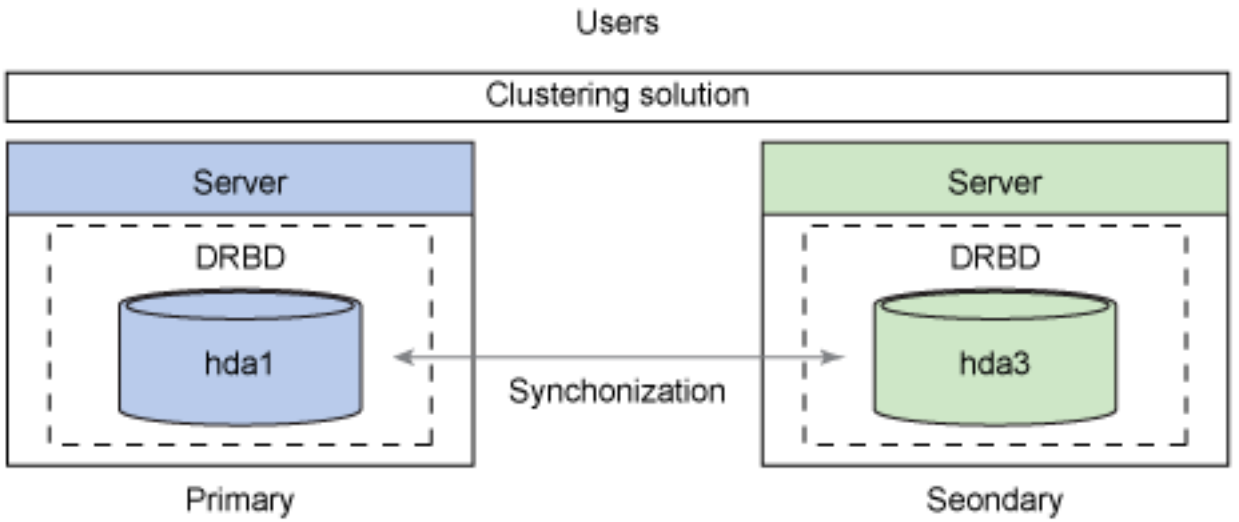
What is it NOT?

- Percona XtraDB Cluster



What is it NOT?

- DRBD / Pacemaker / Clusters FS



MySQL Cluster

WHO USES IT?

Who Uses MySQL Cluster

- Telecommunications
- Payment Gateways (CCs / ACH)
- Trading Systems (Store-And-Forward Message Systems)
- Online Games
- Session Management / eCommerce

Ultimate High Availability

- Traditional HA
  - Replication / Master->Slave[s]
  - MHA / HA-Proxy / Load Balancer
  - Manual Failover
- MySQL Cluster as HA
  - Shared Nothing

MySQL Cluster

WHY USE IT?

Why MySQL Cluster - Pros

- All In-Memory*
- High Availability
- Automatic Scaling / Synchronous Replication
- Concurrent / Parallel Writes
- Online Upgrades

Why MySQL Cluster - Cons

- All In-Memory* - 6.5TB practical limit
- No MVCC / READ-COMMITTED
- No row compression
- Many moving parts
- May require schema/query changes

MySQL Cluster

HOW DO YOU USE IT?

Moving Parts

- The 3 Components
  - MGM Nodes
  - NDBD Nodes
  - API / MySQL Nodes
- Internals
  - Fragments
  - Replicas / Node Groups

Set Up MySQL Cluster

NDB_MGMD

Management Nodes

Management Nodes

- Define the cluster configuration
- Minimal Resources
- Only really necessary to make changes
- Redundant

Cluster Config

- shell> mkdir /var/lib/mysql-cluster/
- shell> cd /var/lib/mysql-cluster/
- shell> $MY_FAV_EDITOR config.ini

config.ini


```
[ndb_mgmd default]
DataDir = /var/lib/mysql-cluster/

[ndb_mgmd]
NodeId = 1
Hostname = mgm01
```

config.ini


```
[ndbd default]
NoOfReplicas = 2
LockPagesInMainMemory = 1
DataDir = /var/lib/mysql-cluster/
DataMemory = 5G
IndexMemory = 1G
ODirect = 1
```

config.ini


```
[ndbd]
NodeId = 11
Hostname = ndbd11

[ndbd]
NodeId = 12
Hostname = ndbd12
```

```
[ndbd]
NodeId = 13
Hostname = ndbd13

[ndbd]
NodeId = 14
Hostname = ndbd14
```

config.ini


```
[mysqld]
[mysqld]
[mysqld]
[mysqld]
[api]
Hostname = mysql01
```

Light the Ignition


```
shell> ndb_mgmd \
  --config-dir=/var/lib/mysql-cluster \
  --config-file=/var/lib/mysql-cluster/config.ini
```

Set Up MySQL Cluster

NDBD NODES

Start ‘yer Engines


```
shell> mkdir /var/lib/mysql-cluster/
shell> ndbmtd -c mysql02[,mgm01]
```

Set Up MySQL Cluster

MYSQL / API NODES

Get me to the SQL!


```
shell> $EDITOR /etc/my.cnf

[mysqld]
ndbcluster
ndb-connectstring = mysql02[,mgm02]
ndb-cluster-connection-pool = 4

shell> $START_MYSQLD
```

Show me the money

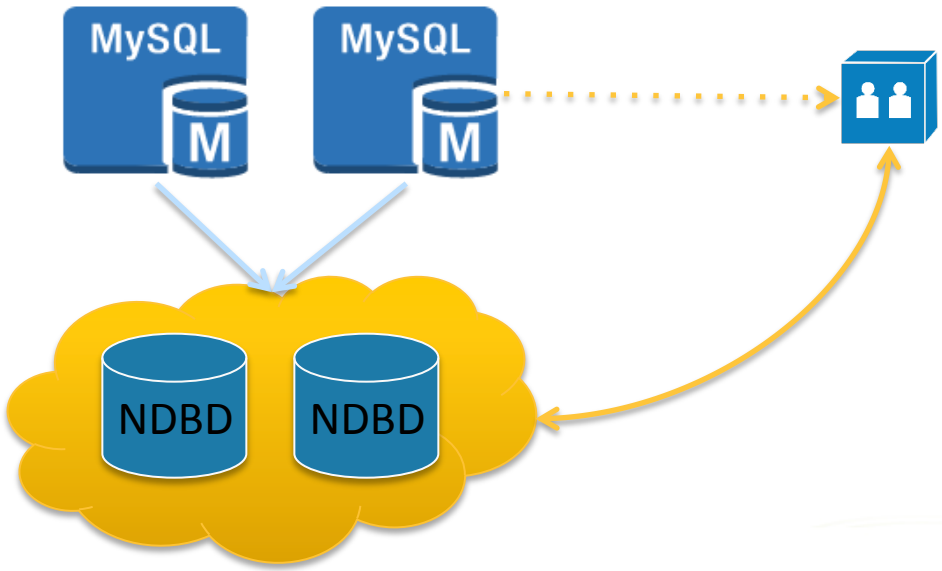

```
mysql> SHOW ENGINES\G
...
*************************** 12. row ***************************
Engine: NDBCLUSTER
Support: YES
Comment: Clustered, fault-tolerant, memory-based tables
*************************** 13. row ***************************
Engine: NDB
Support: YES
Comment: Alias for NDBCLUSTER
```

Got Tables?

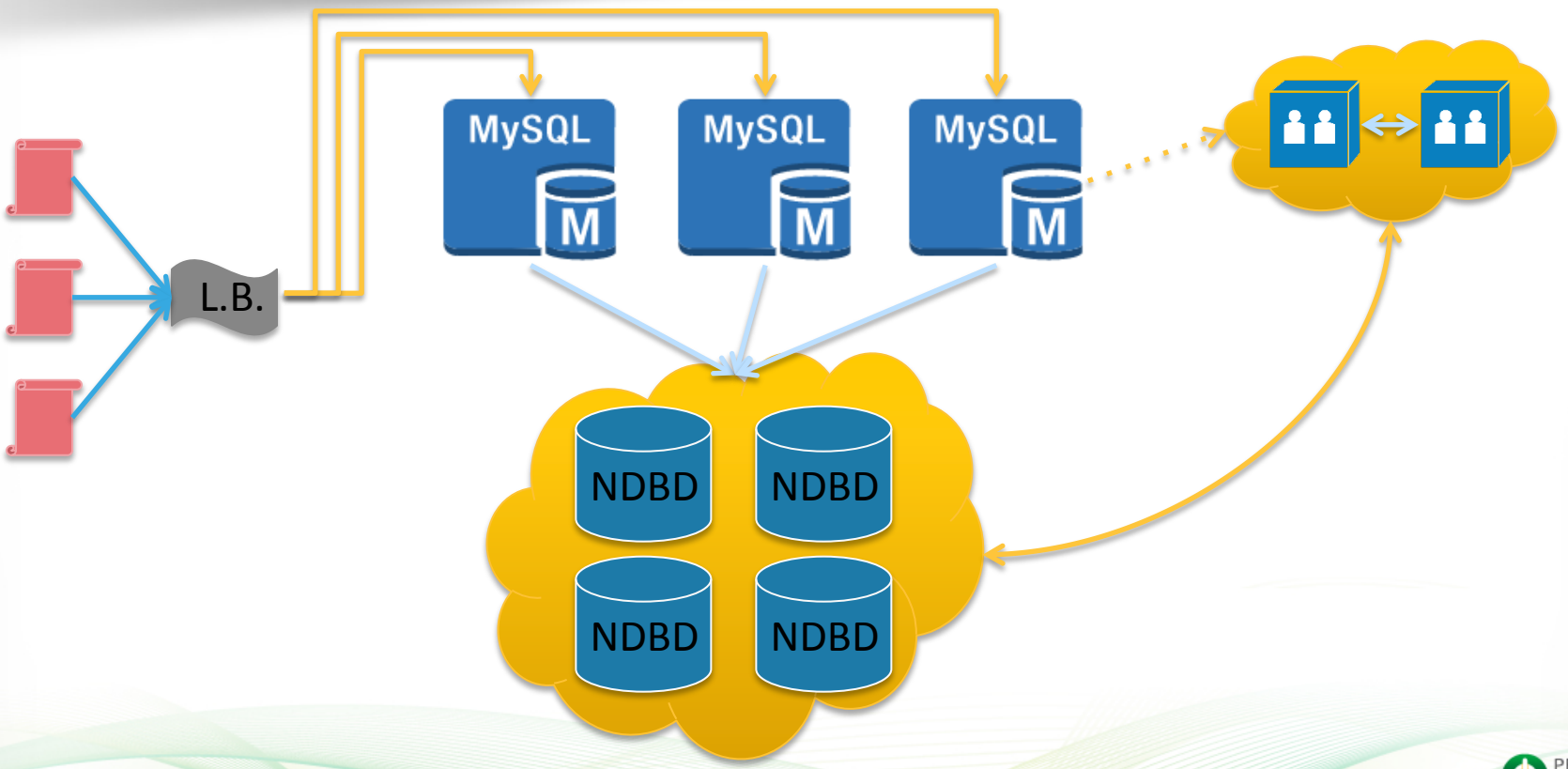

```
mysql> CREATE TABLE ctest (i INT) ENGINE=NDBCLUSTER;
Query OK, 0 rows affected (0.09 sec)

mysql> SHOW CREATE TABLE ctest \G
***********1. row **********
Table: ctest
Create Table: CREATE TABLE `ctest` (
  `i` int(11) default NULL
) ENGINE=ndbcluster DEFAULT CHARSET=utf8
1 row in set (0.00 sec)
```

Our End Result



Practical End Result



MySQL Cluster

FRAGMENTS, REPLICAS, NODE GROUPS - OH MY!

Divide and Conquer

- Fragment (Shard) Each Table
  - Built-In Partitioning
  - Equal to Number of Nodes
  - MD5 on Primary Key

Divide and Conquer


```
ndb_desc -c mysql02 -d test ctest -n -p
...
-- Per partition info --
Partition   Row count   Commit count ...  Nodes
0           2           2                 11,12
1           2           2                 12,11
```

Schema Changes


```
CREATE TABLE Orders (
  orderId INT(11) UNSIGNED AUTO_INCREMENT,
  PRIMARY KEY (orderId)
);
CREATE TABLE Orders_LineItems (
  lineItemId INT(11) UNSIGNED AUTO_INCREMENT,
  orderId INT(11) UNSIGNED,
  PRIMARY KEY (lineItemId, orderId)
);
```

Related data will reside on different nodes; increasing inter-node communications

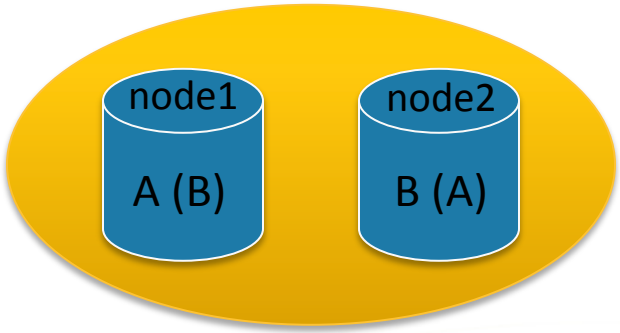
Schema Changes


```
CREATE TABLE Orders (
  orderId INT(11) UNSIGNED AUTO_INCREMENT,
  PRIMARY KEY (orderId)
);
CREATE TABLE Orders_LineItems (
  lineItemId INT(11) UNSIGNED AUTO_INCREMENT,
  orderId INT(11) UNSIGNED,
  PRIMARY KEY (orderId, lineItemId)
);
```

Related data now on same node

Node Groups

- Created Implicitly
- Sized by NoOfReplicas

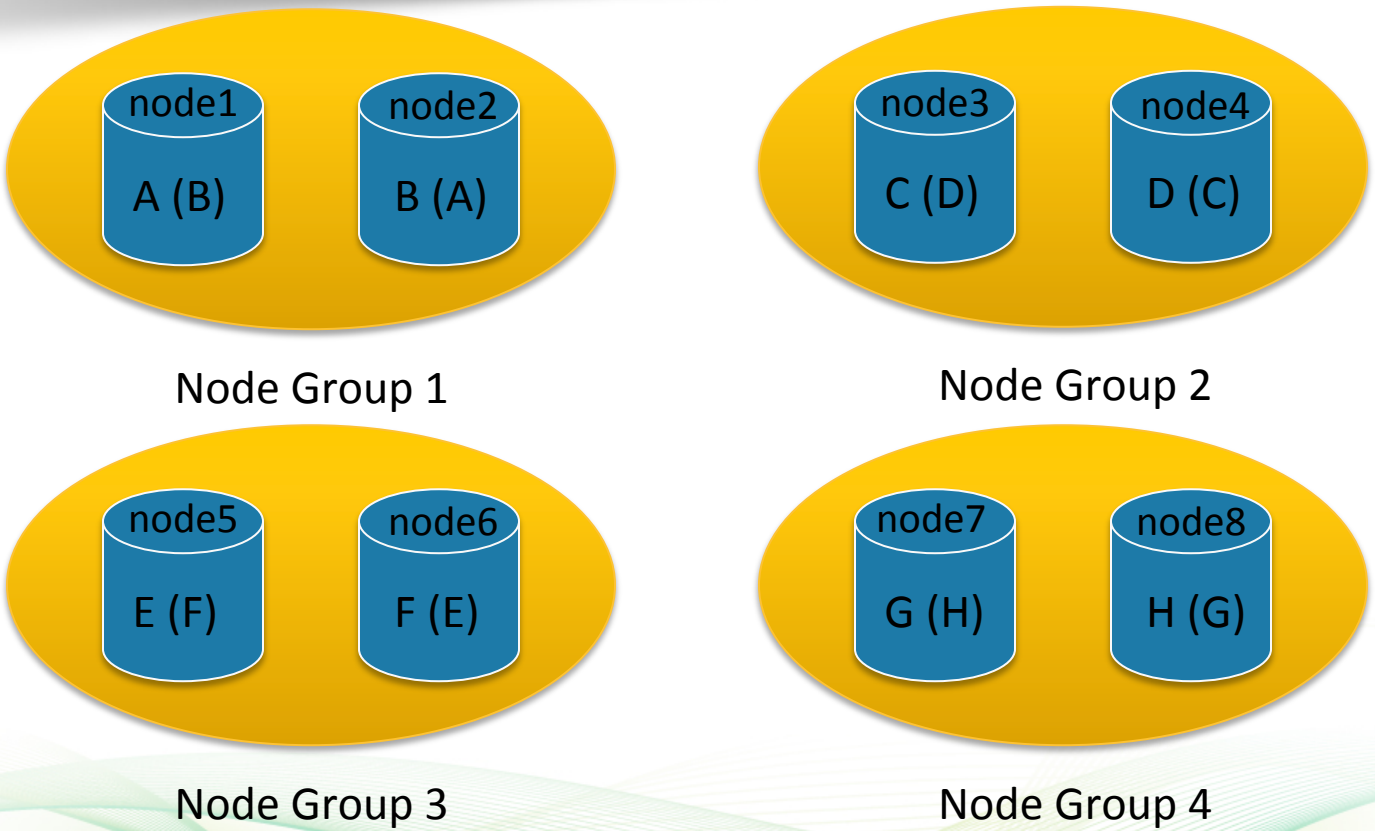


More Node Groups

Node Group 1

Node Group 2

Node Groups Galore



MySQL Cluster

QUESTIONS

Thank You!

Thank You For Attending!

http://www.percona.com/training