



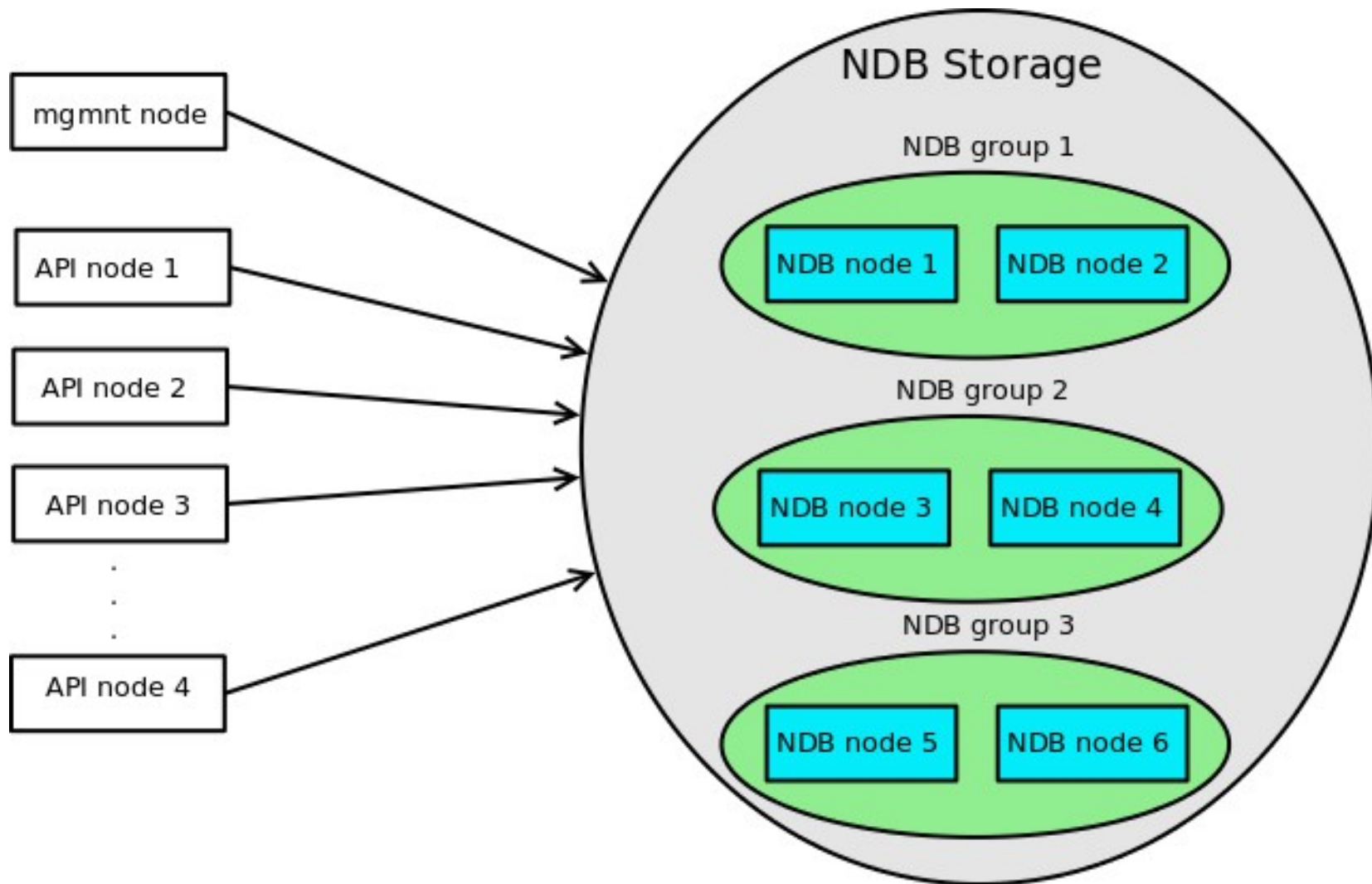
Introduction to MySQL NDB Cluster

Yves Trudeau Ph. D.
Percona Live DC/January 2012

Agenda

- What is NDB Cluster?
- How MySQL uses NDB Cluster
- Good use cases
- Bad use cases
- Example of tuning

What is NDB cluster?



What is NDB cluster?

- A high performance, highly available storage system
- Three types of node
 - Data nodes
 - Management nodes
 - API nodes
- Original use case
 - Storage for telco switches

What is NDB cluster?

- Highly-Available
 - Can reach 99.999% if well deployed
- Data distributed by hash of the primary key or chosen key
 - Sharding!
 - Automatic and transparent

What is NDB cluster?

- By default 2 copies of the data are kept
 - 1 copy per node in a node group
 - Each node master for half
- In memory data
 - async persistence
 - Can reach one millions transactions/s
 - Designed for high concurrency
 - Impacts of network latency

What is NDB cluster?

- Disk based data possible, lower performance
 - Columns with keys on them in memory
 - Disk latency lowers performance
 - Good for large payloads
- Fully transactional
 - read-committed only
 - XA between nodes involved

What is NDB cluster?

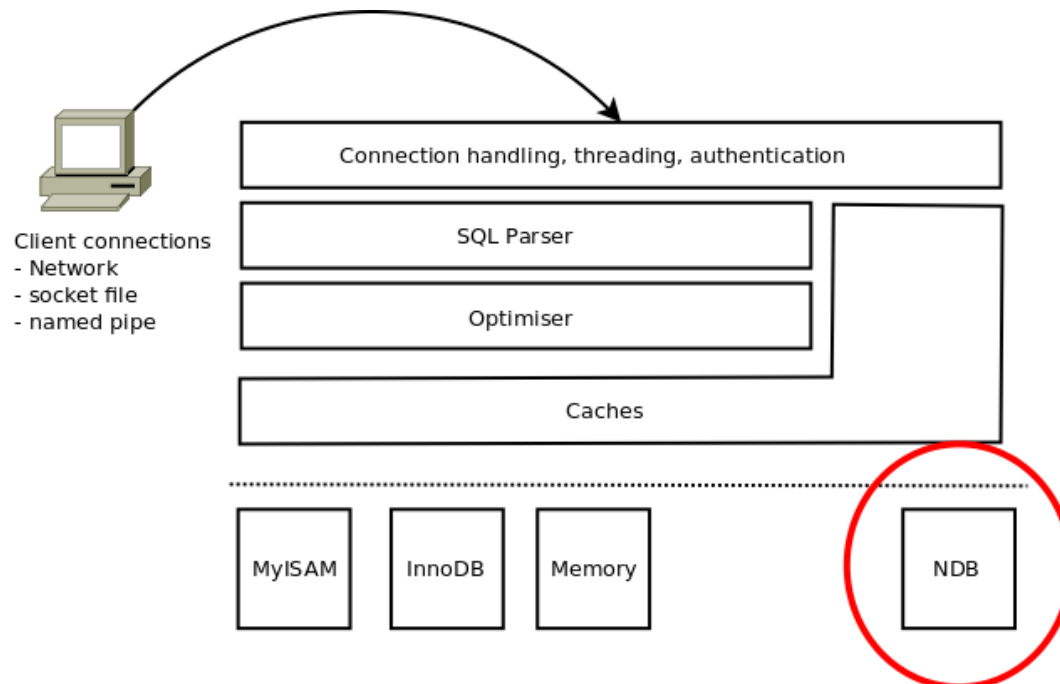
- NoSQL type of storage
 - A bit like memcached but richer API
 - Persistence
 - similar access time than memcached
- Low level C++ API for data manipulation
 - NDB “alone” is not SQL

What is NDB cluster?

- Consistent distributed backups
 - Data save to disk by all data nodes
 - Background operation, non-blocking
- Online maintenance
 - Configuration changes
 - Backups
 - Upgrades
 - add node group (rehashing)

What about MySQL?

- NDB API is not SQL
- NDB API pretty close to storage engine API
- The idea is to use MySQL as a proxy



MySQL and NDB?

- Brings a scalable SE to MySQL
 - Add node for processing
 - Add node for more storage
- Brings SQL features to NDB
 - Custom partitioning
 - Access rights
 - Replication
 - Joins...

Issue with MySQL and NDB?

- Joins
 - MySQL still runs joins in nested loops
 - One network hop per loop per table
 - Will change with BKA
 - NDB 7.2.2, MySQL can push joins
- Range statistics
 - Statistics on key range hardcoded to 10

When is MySQL with NDB good?

- Application with many small transactions coming over multiple threads/connections
- When many of queries are by PK
- When custom partitioning can colocate data on the same data node
- HA is required

Example good use cases?

- Telco switches
 - High rate, small transactions, HA
- Gaming
 - High rate, small transactions
- Financial logger
 - HA, high capacity
- Queuing

When is MySQL with NDB bad?

- Applications with long running transactions
- Applications with highly normalized schema
- Applications relying on huge joins
- Applications with low concurrency
- Applications with TB of data

Example bad explain?

id	select_type	table	type	possible_keys	key	rows
1	SIMPLE	a	range	FK_1,FK_2, coll1	coll1	146096
1	SIMPLE	b	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	c	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	d	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	e	eq_ref	PRIMARY,idx_1	PRIMARY	1
1	SIMPLE	f	eq_ref	PRIMARY,idx_1	PRIMARY	1
1	SIMPLE	g	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	h	eq_ref	PRIMARY,idx_2	PRIMARY	1
1	SIMPLE	i	eq_ref	PRIMARY,ID_IDX	PRIMARY	1
1	SIMPLE	j	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	k	ref	b_fk	b_fk	1
1	SIMPLE	l	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	m	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	n	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	o	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	p	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	q	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	r	eq_ref	PRIMARY,idx_6	PRIMARY	1
1	SIMPLE	s	eq_ref	PRIMARY	PRIMARY	1
1	SIMPLE	t	eq_ref	PRIMARY,idx_1	PRIMARY	1
1	SIMPLE	u	eq_ref	PRIMARY,idx_1	PRIMARY	1
1	SIMPLE	v	ALL	NULL	NULL	2
1	SIMPLE	w	ref	FK_2,idx_1	FK_2	2
1	SIMPLE	x	ref	idx_tr	idx_tr	18
1	SIMPLE	y	eq_ref	PRIMARY	PRIMARY	1

Why is the query slow?

- Base table has 146046 rows to process, this is the outer loop
- More than 146k x 24 network hops
 - That's 3.5 M hops
 - If one hop takes 0.4 ms, it is 23 min.
 - Add the processing to this...
- BKA will help queries like this
 - Dimension tables queried by IN clauses

Tuning for NDB

- Configuration, all memory pre-allocated, many settings
- Add hints to queries because of range stats
- Custom partitioning
- Denormalization or local tables for dimensions
- Rewrite of joins, joining in code with IN clauses

Percona Live DC Sponsors



MySQL Conference & Expo 2012

Presented by Percona Live

The Hyatt Regency Santa Clara & Santa Clara Convention Center

April 10th-12th, 2012

Tutorials & Initial Breakout Sessions Announced

<http://www.percona.com/about-us/pressreleases/percona-live-mysql-conference-and-expo-heads-to-santa-clara-california-april-10-12-2012/>

Featured Speakers

Mark Callaghan (Facebook), Jeremy Zawodny (Craigslist), Marten Mickos (Eucalyptus Systems)

Sarah Novotny (Blue Gecko), Peter Zaitsev (Percona), Baron Schwartz (Percona)

Learn More at www.percona.com/live/mysql-conference-2012/

Want More MySQL Training?

Percona Training in Washington DC Next Week

January 16th-19th, 2012

MySQL Workshops

MySQL Developers Training - Monday

MySQL DBA Training - Tuesday

MySQL InnoDB / XtraDB - Wednesday

MySQL Operations – Thursday

Use Discount code DCPL30 to save 30%

Visit <http://bit.ly/dc-training>

Upcoming Percona MySQL Training

Washington, DC – January 16, 2012

Vancouver, Canada - February 6, 2012

Frankfurt, Germany - February 13, 2012

Irvine, CA – February 14, 2012

Denver, CO - February 20, 2012

San Francisco, CA - March 12, 2012

Austin, TX - March 19, 2012

Visit <http://percona.com/training>



Yves Trudeau Ph. D.
yves@percona.com

We're Hiring!
www.percona.com/about-us/careers/



PERCONA
LIVE

www.percona.com/live