

รู้จักกับ MySQL Cluster



กิตติรักษ์ ม่วงมิ่งสุข

(Kittirak Moungmingsuk)

kittirak@clusterkit.co.th

May 19, 2012 @ossfestival #11

`whoami`

- A part of team at Cluster Kit Co.,Ltd. Since 2007.
- Adjacent Lecturer at Rajamangala University of Technology Thanyaburi (RMUTT)
- About Cluster Kit
 - We have the technological know-how and experience necessary in the high performance computing
 - Our Service including Implementation & Training

Cluster Kit: Achievement

- ThaiGrid (Tera Cluster)
 - 800 Cores, Linux Cluster
 - 133 Cores, Win Cluster
- Sila Cluster @Ramkhamhaeng U. 286 Cores
- BIOTEC (Eclipse Cluster) 704 Cores
- Virgin Radio Thailand
 - 7 nodes, Web Cluster
- Geo-Informatics and Space Technology Development Agency (GISTDA)
 - 10 nodes, Web Cluster
- HAI (HAI Cluster I, II) 480 Cores



Out of Scope, Related AEC.

รายได้ประชากรเฉลี่ยต่อคนต่อปีของประเทศสมาชิกอาเซียน

ประเทศ	รายได้เฉลี่ยต่อคนต่อปี (ดอลลาร์สหรัฐ)
สิงคโปร์	50,714
บรูไน	36,521
มาเลเซีย	8,617
ไทย	5,281
อินโดนีเซีย	3,469
ฟิลิปปินส์	2,255
เวียดนาม	1,362
สปป.ลาว	1,204
กัมพูชา	912
พม่า	804

ประเทศที่มีรายได้สูง : $\geq 12,276$ ดอลลาร์สหรัฐ

ประเทศที่มีรายได้ปานกลาง-สูง : 3,976-12,275 ดอลลาร์สหรัฐ

ประเทศที่มีรายได้ปานกลาง-ต่ำ : 1,006-3,975 ดอลลาร์สหรัฐ

ประเทศที่มีรายได้ต่ำ : $\leq 1,005$ ดอลลาร์สหรัฐ

หมายเหตุ : การแบ่งกลุ่มรายได้จำแนกตามเกณฑ์ของธนาคารโลก (World Bank)

ที่มา : World Bank, November 2011

Fundamental Information Technology Engineer Examination(FE)

Country	Status	2005 Autumn	2006 Spring	2006 Autumn	2007 Spring	2007 Autumn	2008 Spring	2008 Autumn	2009 Spring	2009 Autumn	2010 Spring	2010 Autumn	2011 Spring	2011 Autumn
Vietnam	Applicants	608	547	642	496	761	463	434	376	615	445	486	N/A	N/A
	Examinees	523	472	511	440	556	301	385	334	448	390	461	N/A	N/A
	Successors	71	57	85	99	115	56	0	73	76	86	145	N/A	N/A
	Pass Rate	13.58	12.08	16.63	22.50	20.32	18.60	0	21.86	16.96	22.1	31.50	N/A	N/A
Philippine	Applicants	498	456	723	535	614	613	819	586	598	366	371	N/A	N/A
	Examinees	432	391	666	484	546	534	747	555	571	349	325	N/A	N/A
	Successors	65	29	88	130	71	30	0	70	64	58	62	N/A	N/A
	Pass Rate	15.05	7.42	13.21	22.86	13.00	5.62	0	12.61	11.21	16.6	19.10	N/A	N/A
Myanmar	Applicants	77	110	80	-	156	235	162	88	104	33	87	N/A	N/A
	Examinees	74	93	78	-	130	196	123	65	91	30	76	N/A	N/A
	Successors	4	4	3	-	10	6	0	10	6	5	0	N/A	N/A
	Pass Rate	5.41	4.30	3.85	-	7.69	3.06	0	15.38	6.59	16.7	0	N/A	N/A
Malaysia	Applicants	181	189	135	131	65	71	47	70	62	79	3	N/A	N/A
	Examinees	141	162	0	120	49	0	45	31	42	32	3	N/A	N/A
	Successors	1	9	0	12	6	0	0	1	0	-	0	N/A	N/A
	Pass Rate	0.71	5.56	0	10.00	12.24	0	0	3.23	0	-	0	N/A	N/A
Thailand	Applicants	-	1,238	704	233	486	117	594	521	397	135	124	85	N/A
	Examinees	-	1,163	660	220	467	98	514	408	368	122	116	76	162
	Successors	-	79	62	24	48	21	40	39	29	5	4	9	17
	Pass Rate	-	6.79	9.39	10.91	10.28	21.65	7.78	9.56	7.88	4.1	3.40	11.84	10.49
Mongolia	Applicants	-	-	165	217	50	54	20	20	36	11	32	N/A	N/A
	Examinees	-	-	123	140	49	53	19	19	36	10	28	N/A	N/A
	Successors	-	-	6	3	8	7	0	3	2	1	5	N/A	N/A
	Pass Rate	-	-	4.88	2.14	16.33	13.21	0	15.79	5.56	10.0	17.90	N/A	N/A

Agenda

- Quick review MySQL
- MySQL Cluster Concept
- What's MySQL Cluster Do?
 - Load Balance
 - High Availability
- NDB Storage Engine
- MySQL Cluster Features
- In-Memory Vs. Disk-Bases
- Monitoring Tools

MySQL Review

- Latest stable version is 5.5
- Written in C and C++.
- Works on many different platforms.
- Designed to be fully multi-threaded using kernel threads
- Support for large databases
 - 200,000 tables and about 5,000,000,000 rows.
- supports high-availability database clustering using the NDBCLUSTER storage

MySQL Development Roadmap

Subqueries	4.1
R-trees	4.1 (for the MyISAM storage engine)
Stored procedures and functions	5.0
Views	5.0
Cursors	5.0
XA transactions	5.0
Triggers	5.0 and 5.1
Event scheduler	5.1
Partitioning	5.1
Pluggable storage engine API	5.1
Plugin API	5.1
Row-based replication	5.1
Server log tables	5.1
Foreign keys	6.x (implemented in 3.23 for InnoDB)

Storage Engines

- MyISAM
- InnoDB
- NDBCLUSTER
- Memory
- Merge
- Archive
- Federated
- CVS
- Blackhole
- Example

MyISAM

- The default MySQL storage engine
- supported in all MySQL configurations
- default storage engine unless you have configured MySQL to use a different one by default.

InnoDB

- A transaction-safe (ACID compliant)
- Row-level locking
- Oracle-style consistent nonlocking reads increase multi-user concurrency and performance.
- Stores user data in clustered indexes to reduce I/O for common queries based on primary keys.
- Supports FOREIGN KEY referential integrity constraints.

NDBCLUSTER

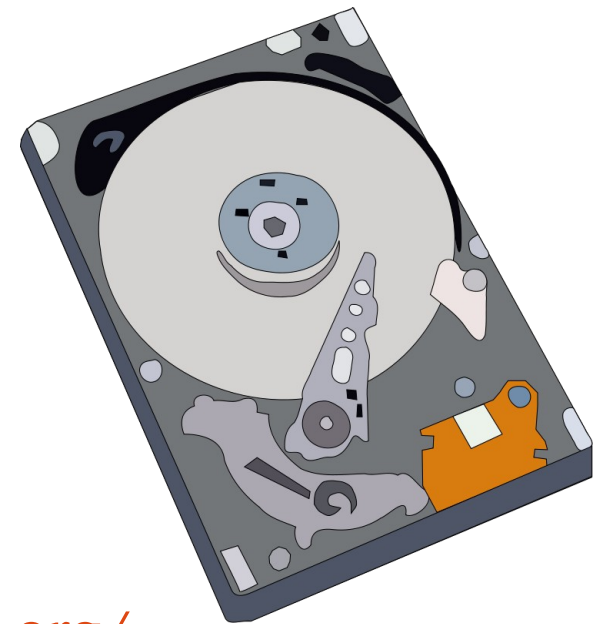
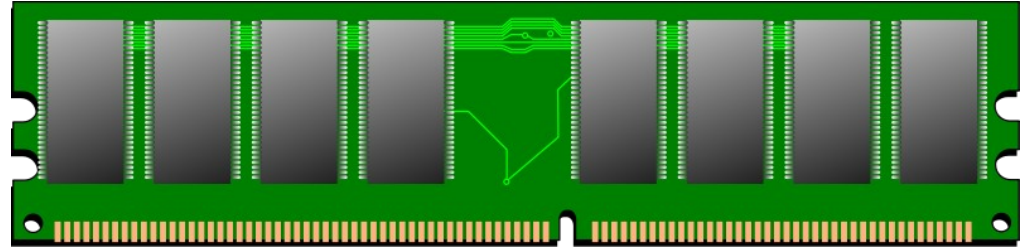
- known as NDB
- Clustered database engine
- Particularly suited for applications that require the highest possible degree of uptime and availability
 - High Availability
 - mission-critical 24/7 usage and heavy Web or logging usage.

MySQL Cluster

- Enables clustering of both in-memory and on disk bases.
- Shared-nothing architecture.
- No single point of failure.
- Storage engine called **NDB** or **NDBCLUSTER**
- Latest NDB 5.5.20-ndb-7.2.5
 - That's mean **NDB 7.2.5, MySQL 5.5.20**
- Both GPL License and Commercial

In-memory Vs. Disk-Bases

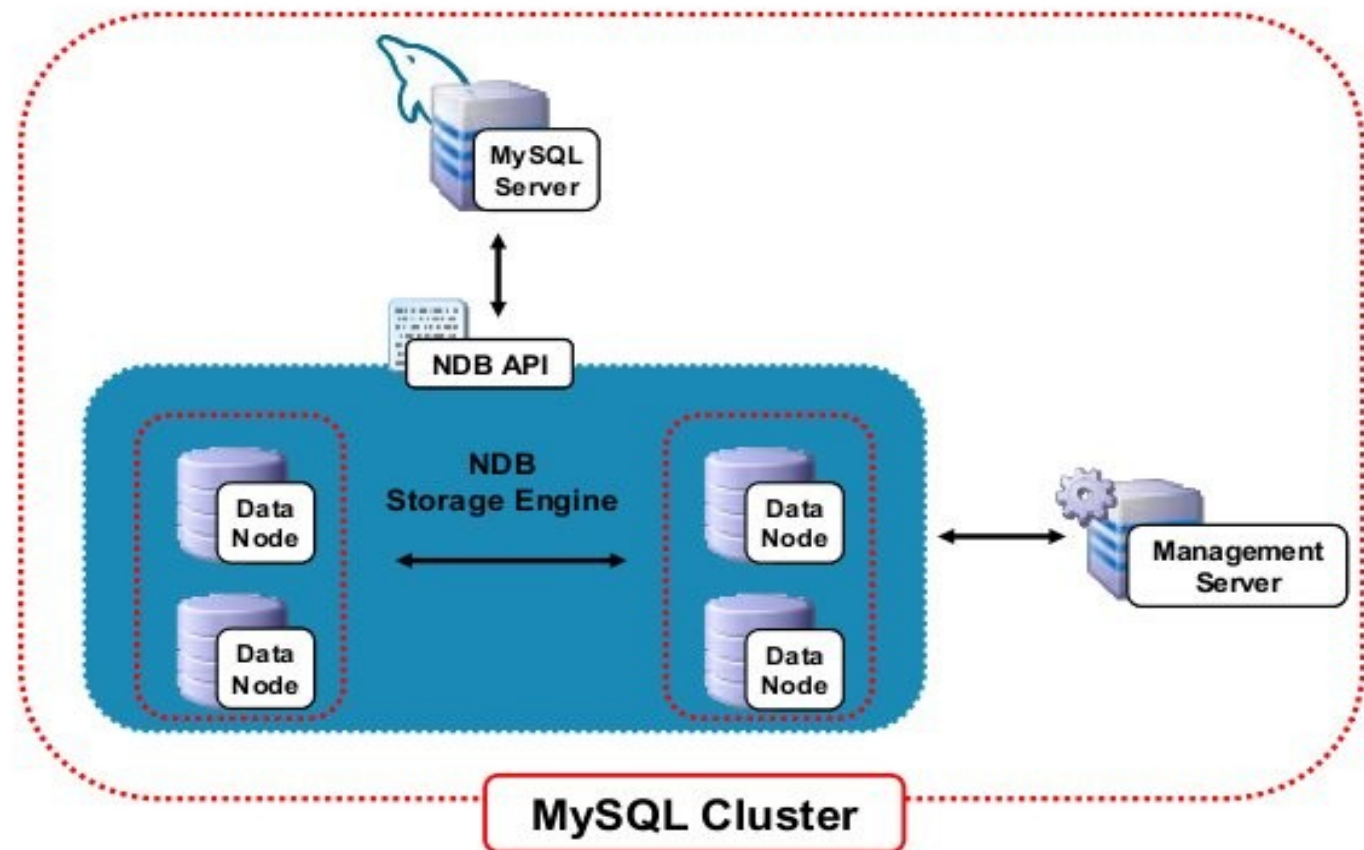
- In-Memory
 - Fastest than Disk
 - Required more memory
 - Focus on real-time application
- Disk-Base (Support on ndb-7.x up)
 - Slower then Memory
 - Working with larger data
 - Cheapest in cost



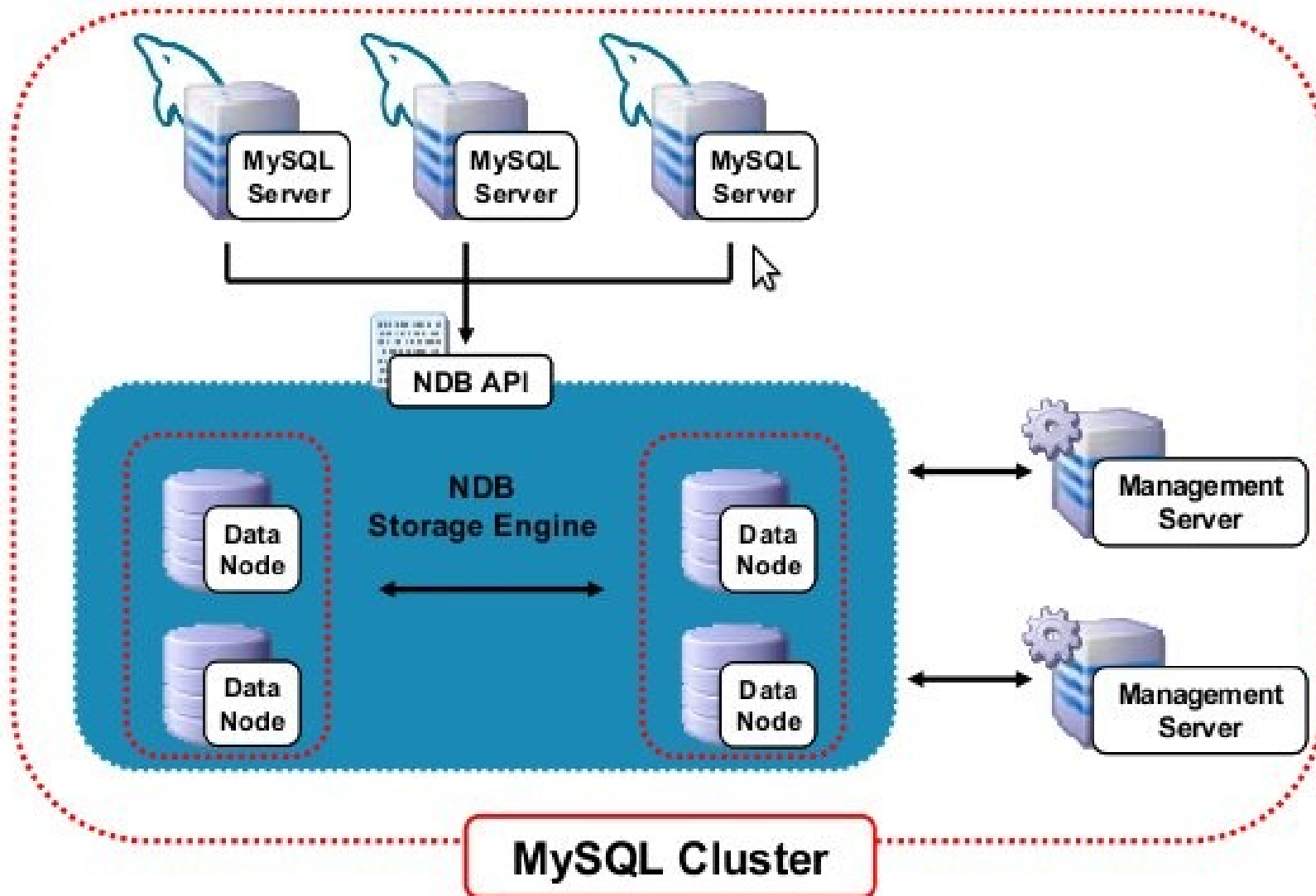
Images from <http://www.openclipart.org/>

MySQL Cluster Components

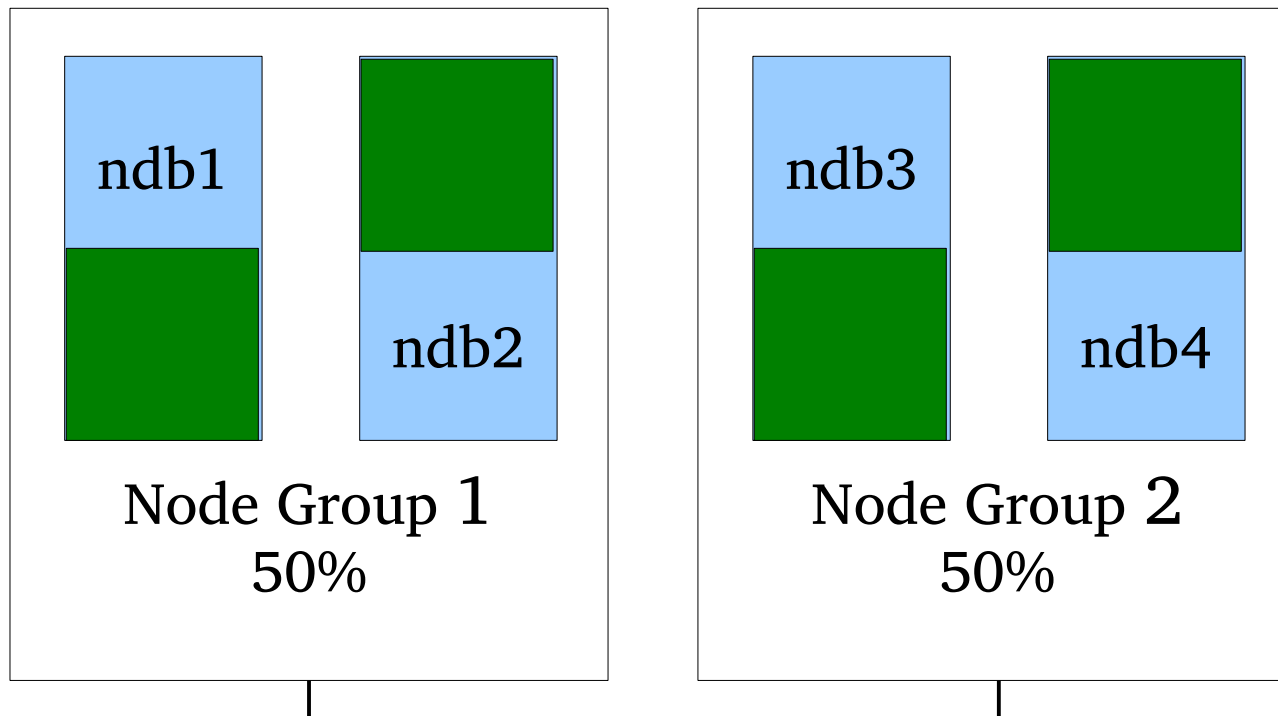
- Management node (MGM node)
- Data node
- SQL node



Scaling Out Configuration



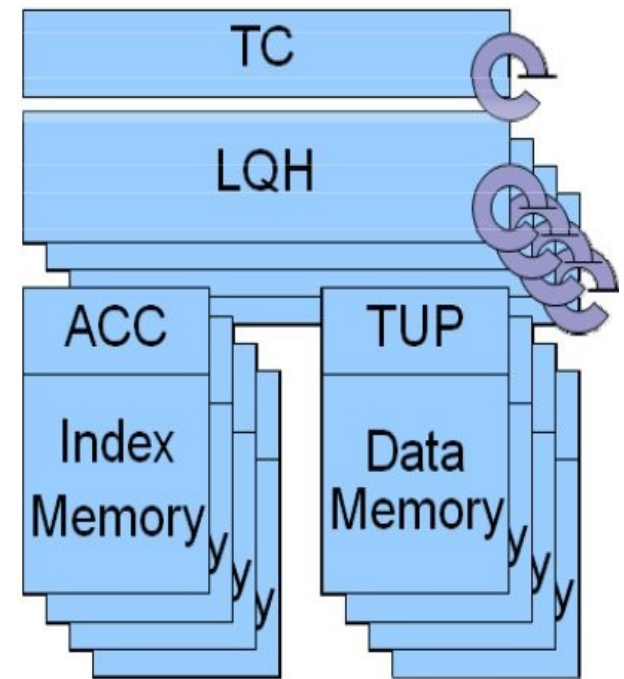
2 Node Groups Data Partitioning



Multi-thread Data node

- Support Maximum 8 threads per machine
 - Transaction Coordinator (TC) 1
 - Transporter 1
 - Replication 1
 - Local Query Handler (LQH)
 - Access manager (ACC)
 - Tuple Manager (TUP)

MaxNoOfExecutionThreads	LQH Threads
2	1
4	2
8	4



Advantages

- High availability and Load balanced
- High Throughput
- Scalability
- Can deploy with low cost hardware

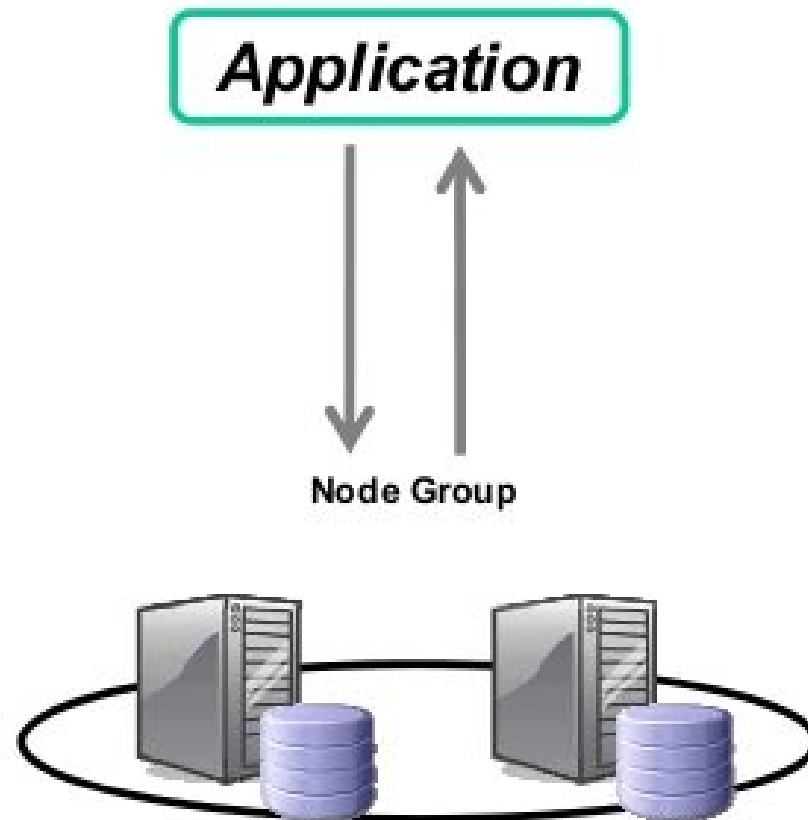
Disadvantages

- Need many number of machines
- Difficult in configuration
- Not work with complex query
- Required high technical skill

Disk Bases Table

- Beginning with MySQL 5.1
- Disk based data does not support variable sized storage.
- A disk-based table stores its data in a tablespace and stores undo data in a log file group

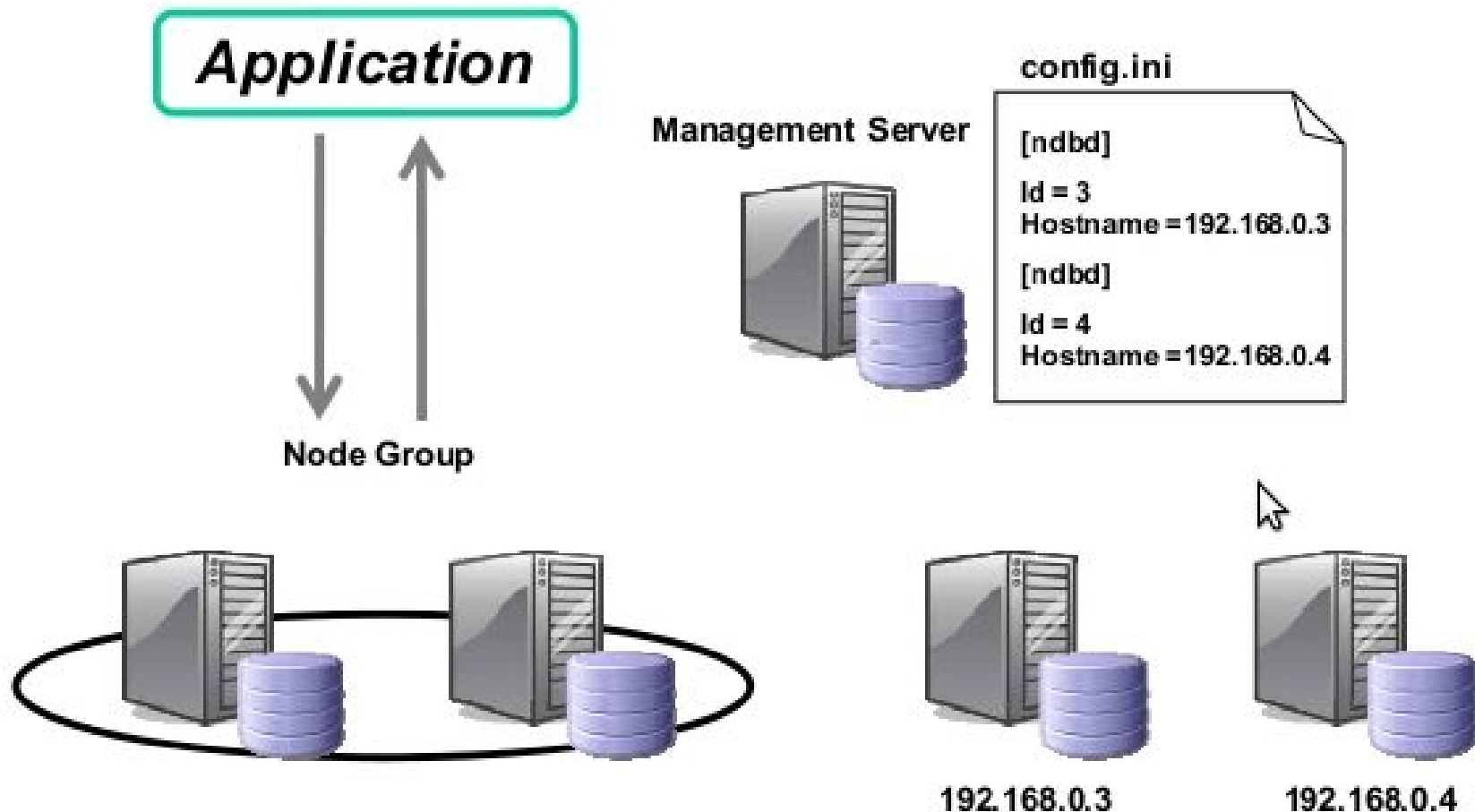
On-line add node 1



"author" Table

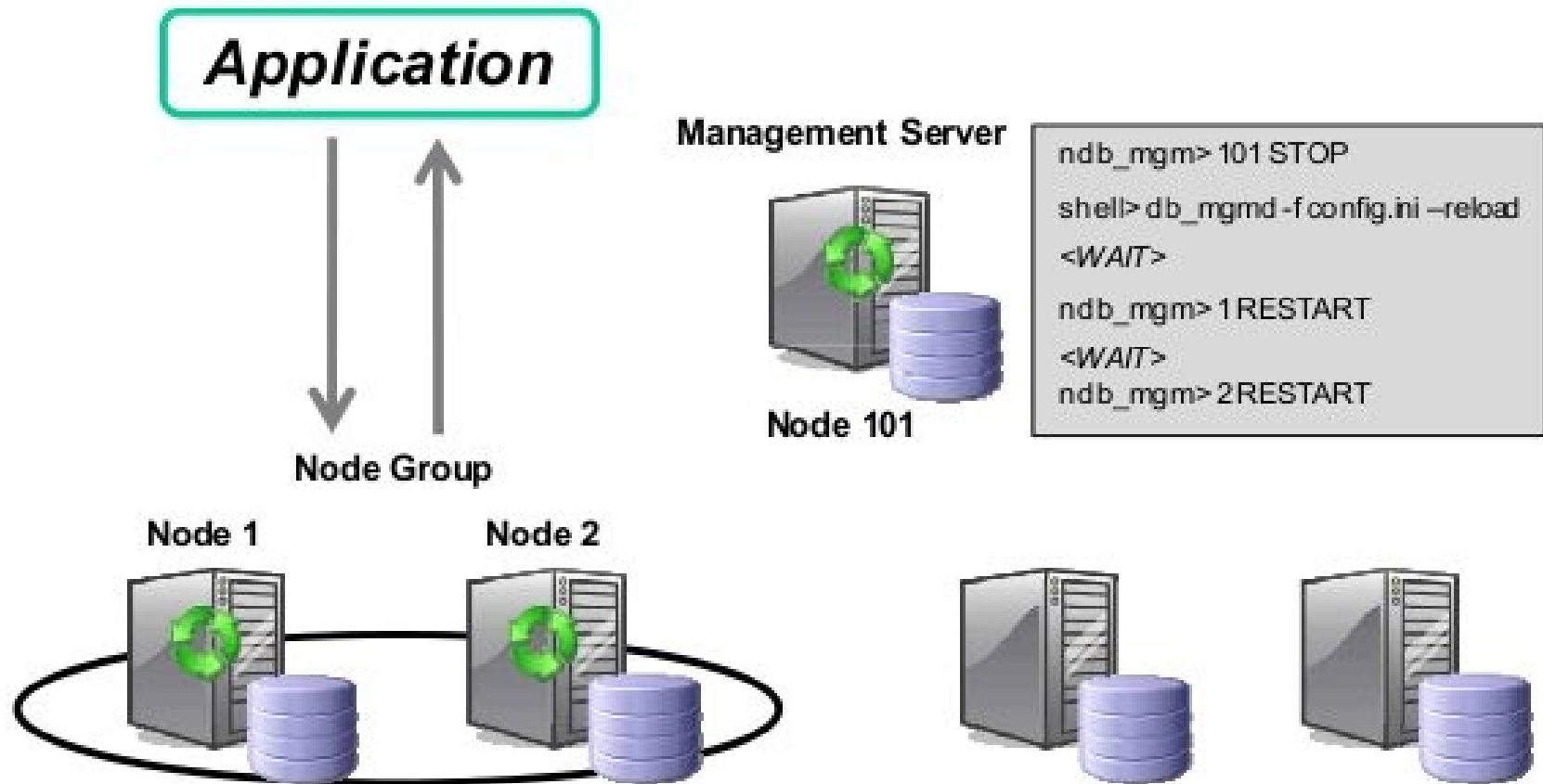
authid (PK)	fname	lname	country
1	Albert	Camus	France
2	Ernest	Hemingway	USA
3	Johann	Goethe	Germany
4	Junichiro	Tanizaki	Japan

On-line add node 2 update config.ini

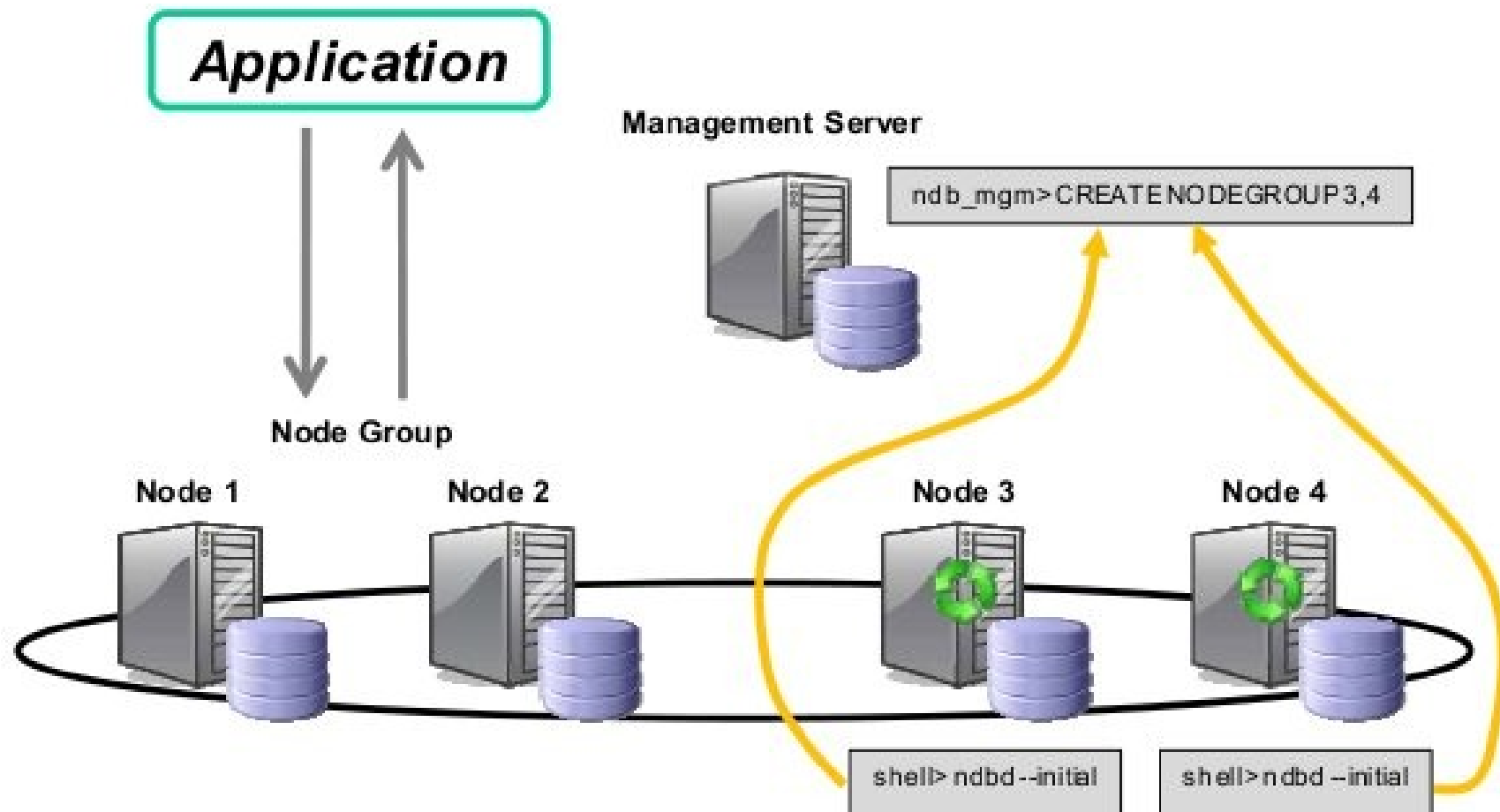


On-line add node 3

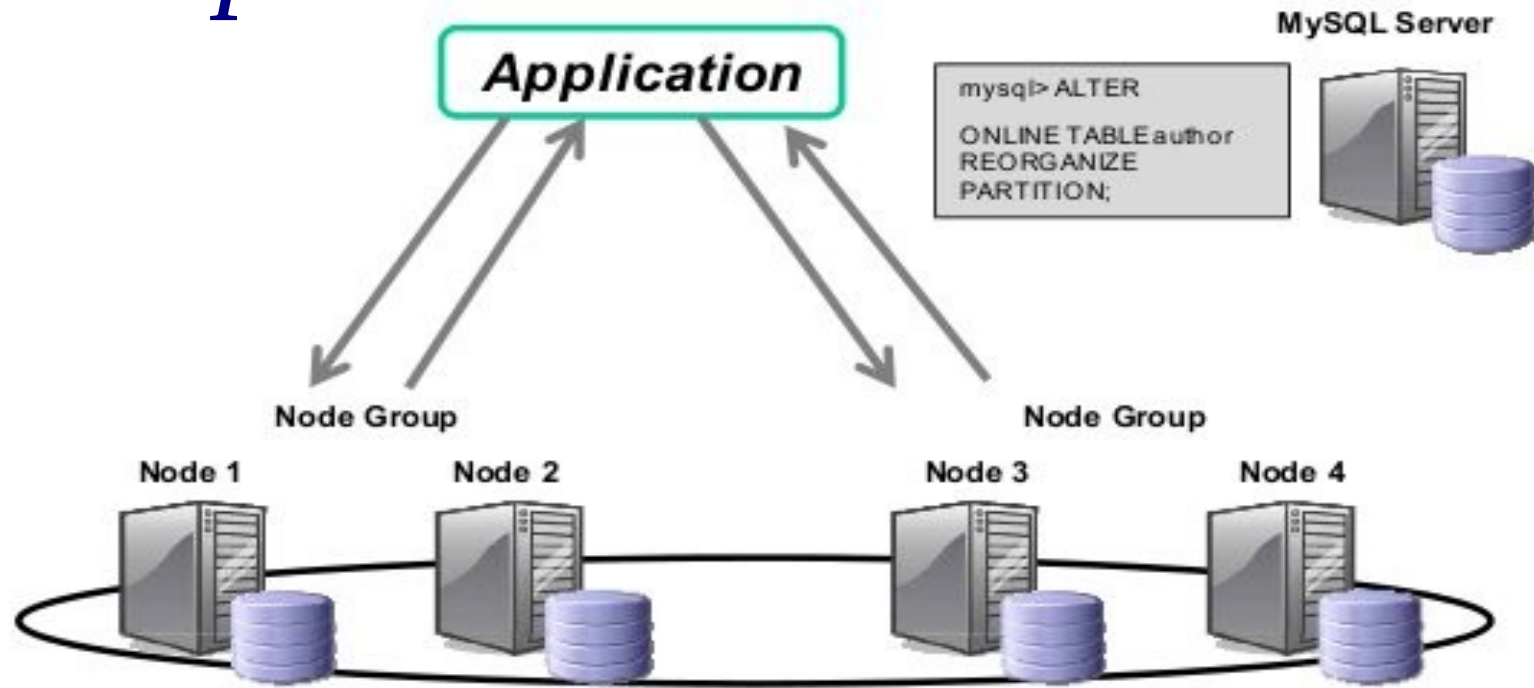
Rolling Restart



On-line add node 4 Repartition Data



On-line add node 5 Data Repartitioned



authid (PK)	fname	lname	country
1	Albert	Camus	France
2	Ernest	Hemingway	USA
3	Johann	Goethe	Germany
4	Junichiro	Tanizaki	Japan

authid (PK)	fname	lname	country
2	Ernest	Hemingway	USA
4	Junichiro	Tanizaki	Japan

Cluster Backup & Restore

- Two mechanisms
 - NDB native backup
 - The mysqldump program
- NDB native backup not restorable to clusters running earlier or later versions.

Limitation

- Cannot create indexes on NDB table columns that use any of the TEXT or BLOB data types.
- Does not support FULLTEXT indexes
- The NDBCLUSTER storage engine supports only the READ COMMITTED transaction isolation level
- There are no partial transactions, and no partial rollbacks of transactions.
- MySQL Cluster does not handle large transactions well
- TRUNCATE & LOAD DATA INFILE are non-transactional and cannot be rolled back.

CMON Web Monitoring

- CMON is a daemon that aggregates information from MySQL Cluster that earlier was only accessible from the cluster log or the management client, such as: cluster state, node state, backup statistics, Statistics and Cluster events (cluster log basically)
- CMON can also start ndbd nodes and make decisions on how they should be started (with or without --initial).
- CMON also does analysis and suggests changes to mysql variables in case they are wrongly set.

CMON Screen

DASHBOARD TOOLS LIBRARY CONTACT REGISTRATION GLOBAL SETTINGS

Replication Cluster
'default_name'(1/3)

MySQL Servers

- app01
- app02
- app03
- storage01

(page refreshes every 10th seconds)

Replication Cluster - default_name

Current Cluster Load

queries/sec	Inserts/sec	selects/sec	updates/sec
0	0	0	0

Master

Hostname	MySQL Status	Host Stats	Server Stats	Version	Serverid
app02		ping: ok cpu util: 0% uptime: 6days 19h 28m	Queries/s: 0 DB Connections: 4 Uptime: 15m 48s	5.5.16-log	2

Relay Master

Hostname	MySQL Status	Replication Status	Lag	Master Hostname	Host Stats	Server Stats	Version
app01		Running	0	app02	ping: ok cpu util: 0.5% uptime: 3days 1h 1m	Queries/s: 0 DB Connections: 4 Uptime: 16m 13s	5.5.16-log

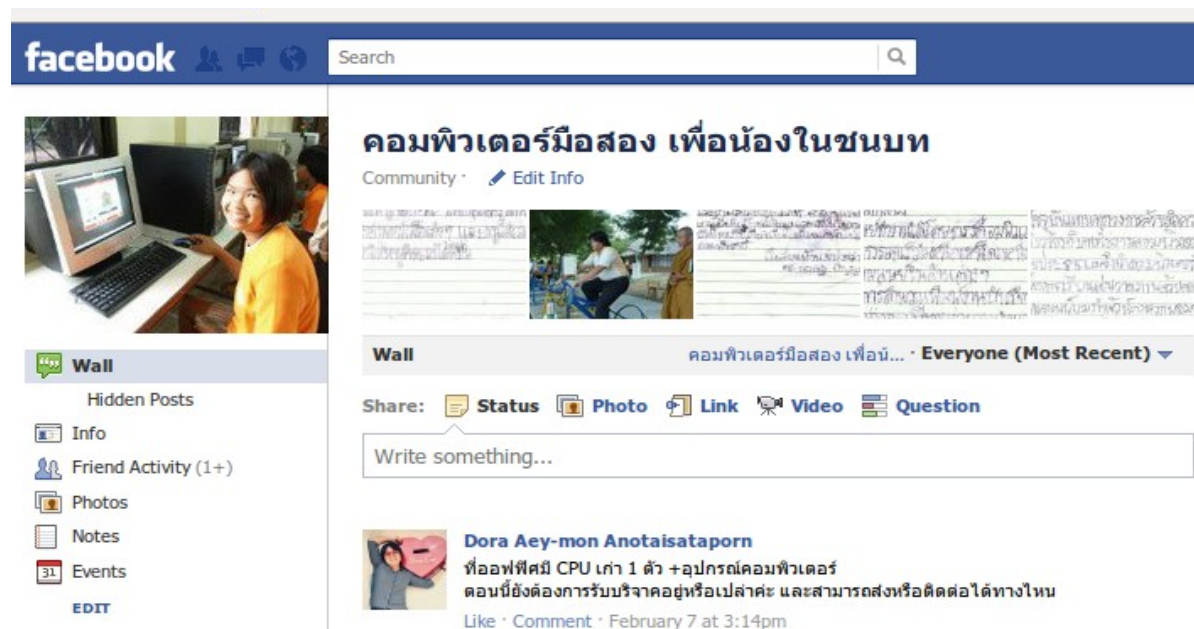
Slaves

Hostname	MySQL Status	Replication Status	Lag	Master Hostname	Host Stats	Server Stats
app03		Running	0	app02	ping: ok cpu util: 0% uptime: 7days 12h 37m	Queries/s: 0 DB Connections: 4 Uptime: 15m 48s
storage01		Running	0	app02	ping: ok cpu util: 1% uptime: 58m 11s	Queries/s: 0 DB Connections: 4 Uptime: 15m 48s

References & Recommended Site

- <http://dev.mysql.com/downloads/cluster>
- “MySQL Cluster 7.0: Architecture and New Features”.
A MySQL® Technical White Paper by Sun Microsystems, April 2009.
- Configurator for MySQL Cluster
 - <http://www.severalnines.com/config/>
- Principal Consultant, MySQL AB
 - <http://johanandersson.blogspot.com/>

“คอมพิวเตอร์มือสอง เพื่อน้องในชนบท”



เกี่ยวกับโครงการ

- กิจกรรมทำอะไร ?
 - รับบริจาคเครื่องคอมพิวเตอร์และอะไหล่ต่าง ๆ มาประกอบเป็นเครื่องที่สมบูรณ์ ติดตั้งระบบปฏิบัติการลินุกซ์เพื่อการศึกษา **Edubuntu** แล้วนำไปส่งมอบให้โรงเรียน พร้อมทำให้ทุกเครื่องเชื่อมต่ออินเทอร์เน็ตได้ ด้วยการวางระบบเครือข่ายให้กับทางโรงเรียน
- ทำในรูปแบบไหน ?
 - เป็นกิจกรรมที่รวมตัวกันทำในกลุ่มเพื่อน ๆ คนคอมพิวเตอร์ที่อยากไปเที่ยว อยากทำดี

เกี่ยวกับโครงการ (ต่อ)

- ทำมาแล้วกี่ครั้ง
 - จัดส่งมอบมาแล้วทั้งหมด 6 โรงเรียน รวมส่งมอบไปห้าสิบกว่าเครื่อง ติดตั้งระบบเครือข่ายให้ 5 โรงเรียน
- เครื่องที่บริจาคสเปกแรงแค่ไหน
 - เครื่องที่ส่งไป RAM 512 MB. ฮาร์ดดิสก์รวมกันให้ได้ 6 GB. ก็เอาละ ก็ตามคุณสมบัติขั้นต่ำที่ลง Edubuntu ได้ครับ (เราจัดต่ำกว่าที่เขากำหนด แต่ใช้ได้ครับ)







โรงเรียนบ้านป่าว่านสามัคคี
พื้นที่การศึกษาศรี...

โครงการ
คอมพิวเตอร์มือสอง เพื่อน้องในชนบท
ครั้งที่ ๔
โรงเรียนบ้านป่าว่านสามัคคี ต.ห้วยยาง อ.อนสาร จ.ชัยภูมิ



Thank you.

Tweet to me at @kittirak

Download this presentation at

http://www.clusterkit.co.th/pdf/MySQL_Cluster_Ossfes11.pdf