

---

# MySQL Cluster

— 基础

简朝阳

2010.06.24

# 个人简介

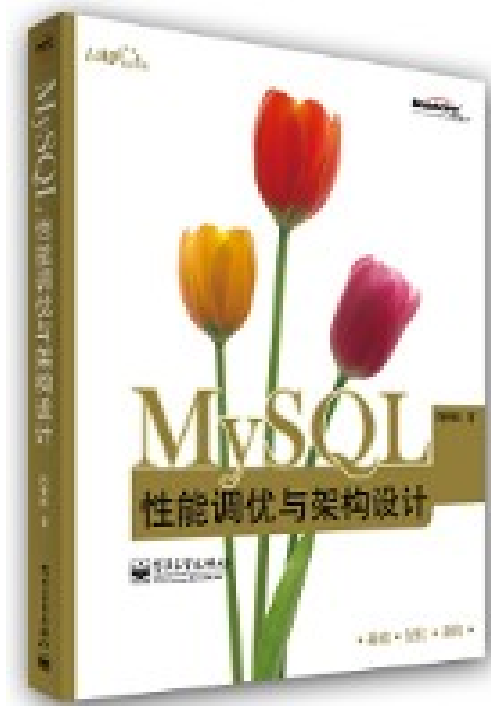
简朝阳 (sky000)

MySQL / Oracle DBA

<http://isky000.com>

<http://twitter.com/sky000>

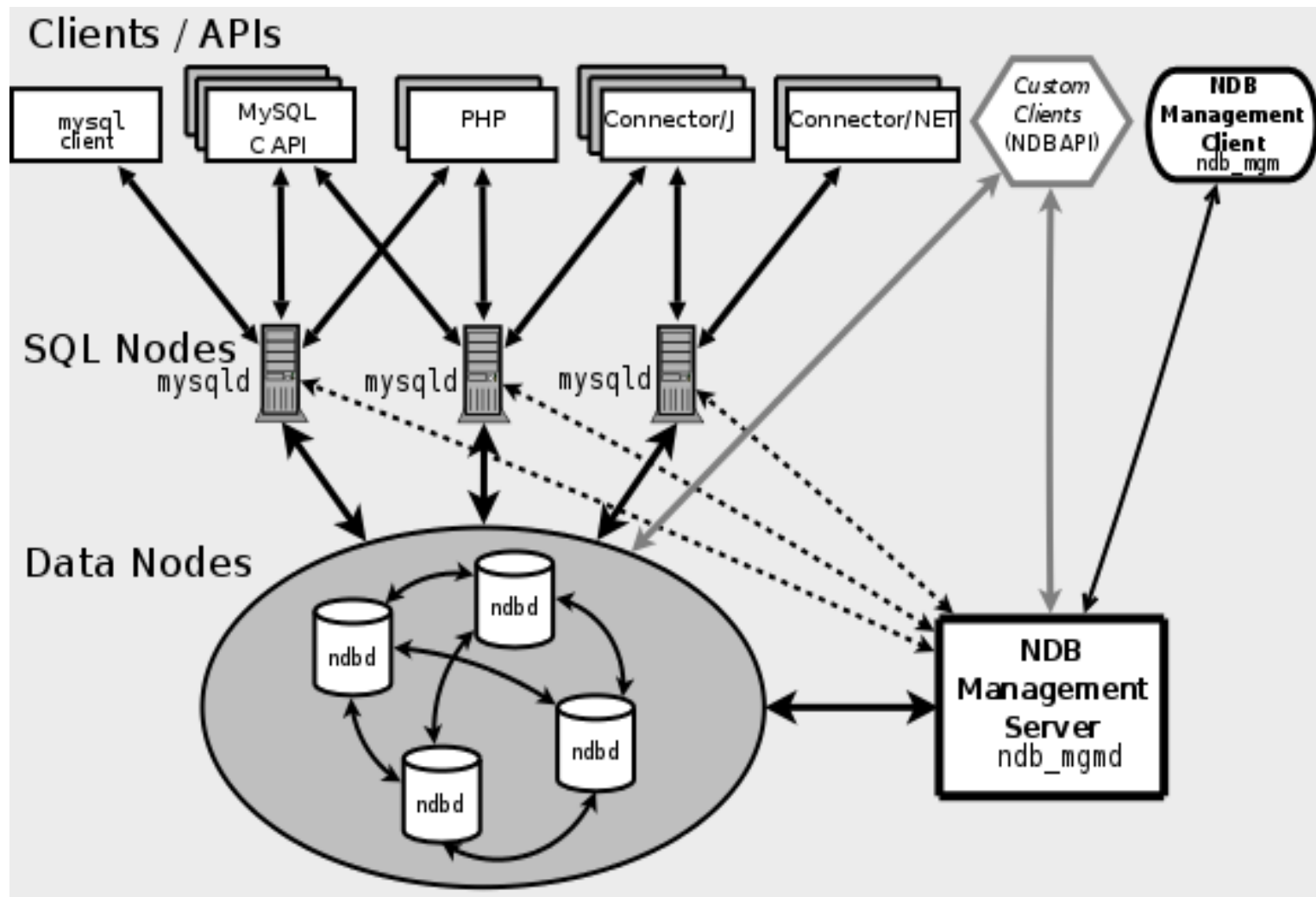
[mysqlr@googlegroups.com](mailto:mysqlr@googlegroups.com)



# 主题

- 架构介绍
- 关键概念
- 基础特性
- Q & A

# 架构预览 - 架构图



# 架构预览 - SQL Node

- **mysqld 进程**

一个完整的MySQL Server 程序，同时可以使用其他存储引擎

- **互相无通信**

- **与 Manage Node 通信**

汇报自身状态信息，取得各节点状态信息以及 Data Node 的连接信息

- **与 data Node 通信**

与数据节点进行交互进行数据存取操作

# 架构预览 - Data Node

- **ndbd 进程**

ndbcluster 存储引擎进程，管理节点上所有数据和日志

- **互相有通信**

各存储引擎节点之间数据同步，根据管理节点指令进行 HA 维护

- **与 Manage Node 通信**

汇报自身状态信息，取得各节点状态信息以及 SQL Node 的连接信息

- **与 SQL Node 通信**

与 SQL 节点进行交互，接受数据存取操作指令

# 架构预览 - Manage Node

- **mgmd 进程**

负责整个集群（包括 SQL Node 和 Data Node）的基础协调工作，维护集群配置信息，维护管理集群日志，管理指令分发

- **一般单节点，可以集群化**

可以在多个物理主机上部署管理程序，但必须保证所有管理节点配置完全一致

- **与 Data Node 通信**

负责调度，管理各个 Data Node，进行配置分发，获取状态信息

- **与 SQL Node 通信**

对各个 SQL Node 维护状态信息，进行配置分发

# 关键概念

## ➤ Node

一个 ndbd 线程，一个 mysqld 线程，一个 mgmd 线程，就是一个 Node

## ➤ Node Group

数据互为冗余的一组 Node

## ➤ Partition

数据按照某个 Key 被进行水平拆分成多份后的一份，拆分规则可以由 MySQL Cluster 自行完成，也可以通过手工指定。每个 Partition 数据仅仅只会存放在一个 Node 上（冗余备份除外）

## ➤ Replica

数据被水平拆分成多份后，在冗余机制的保证下，每份数据会有冗余，成为多份，互为冗余的每份被称为 Replica

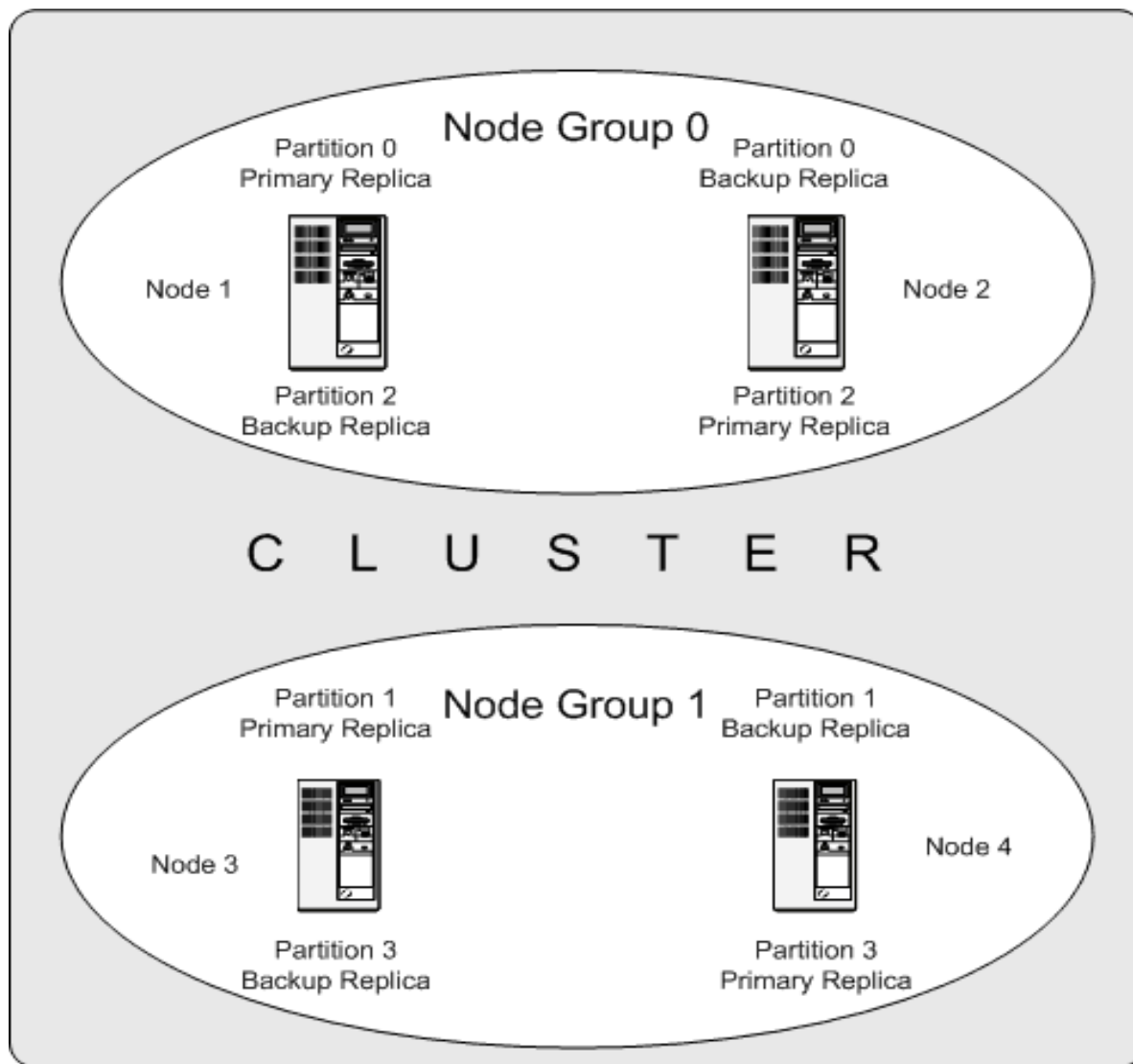
## ➤ CheckPoint

LCP (local checkpoint): 仅仅只保证单个 Node 上的数据被 flush 到磁盘，各个节点之间并发进行

GCP (global checkpoint): 保证整个集群所有节点的数据被 flush 到磁盘，一般是事务在所有节点都完成等，且每个节点的 redo 都已经 flush 到磁盘之后发生



# 关键概念 - 图示



# 功能特性

## ➤ In Memory

所有的 Index 全部在内存中，MySQL 5.1 以前的 Cluster 所有的 Data 都在内存中，之后的未被索引的字段可以不在内存中，但是索引了的字段也需要在内存中

## ➤ SQL 与 存储引擎分离

mysqld 和存储引擎可以在不同的物理主机上，可以根据二者对硬件要求的差异定制硬件，如 mysqld 主要消耗 CPU，而 ndbcluster 主要消耗 IO

## ➤ 高可用

由于数据在同一个 Node Group 中的多个 Node 上会有冗余拷贝，所以单个节点失效后，对系统不会有任何影响，而且是完全透明，就如同磁盘 raid 一样，数据拷贝的数量，也可以灵活设定

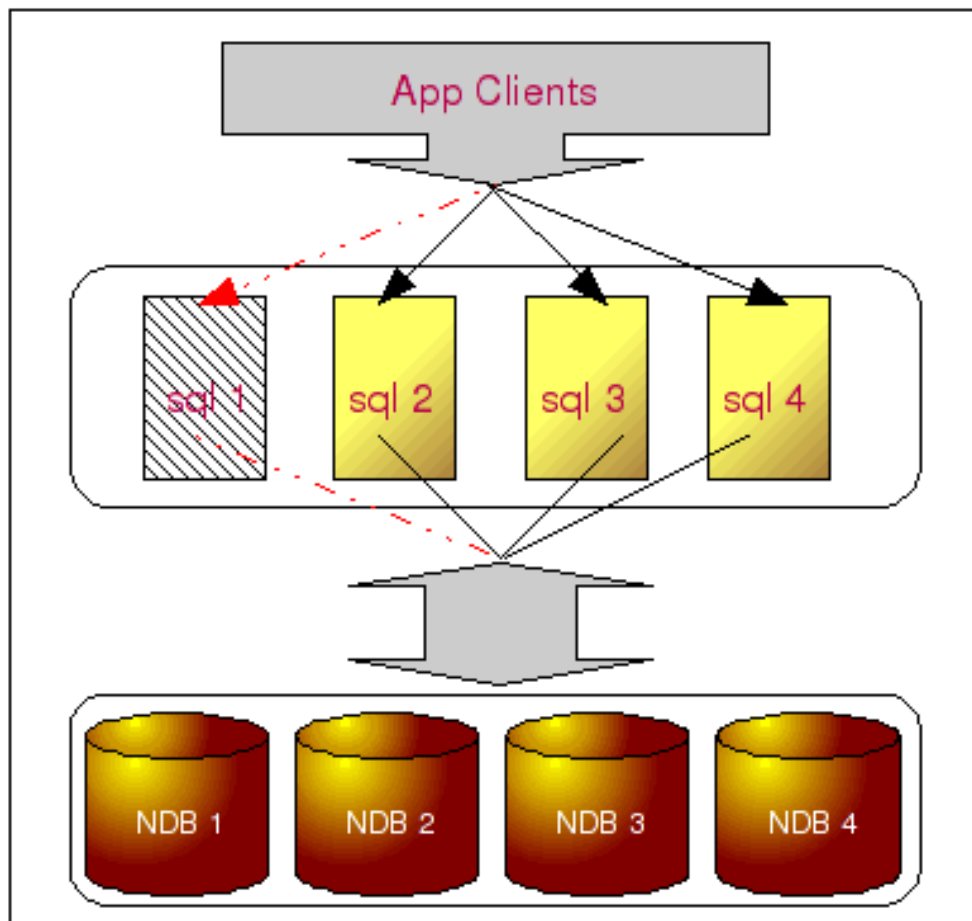
## ➤ 可扩展

SQL Node 可以通过增加 Node 让处理能力呈线性增长，Data Node 也可以达到相同的效果

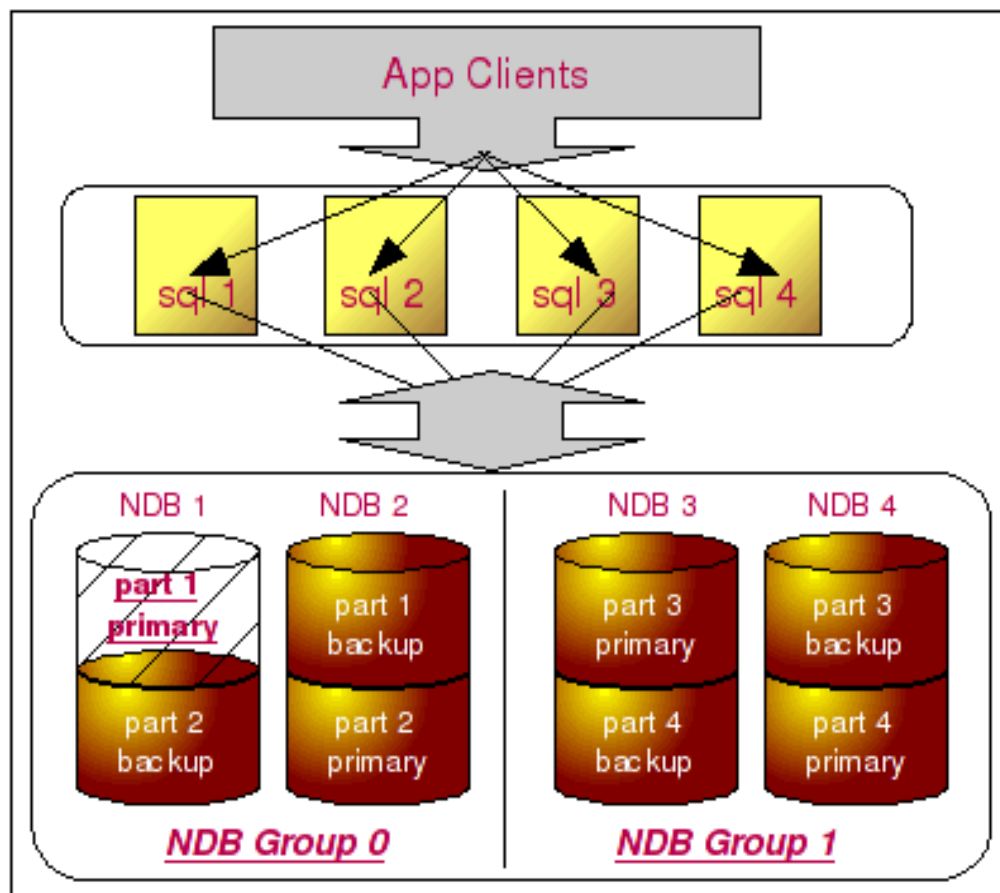
## ➤ 在线热备

NDB Cluster 存储引擎模前支持在线热备份，可以方便的进行在线备份操作

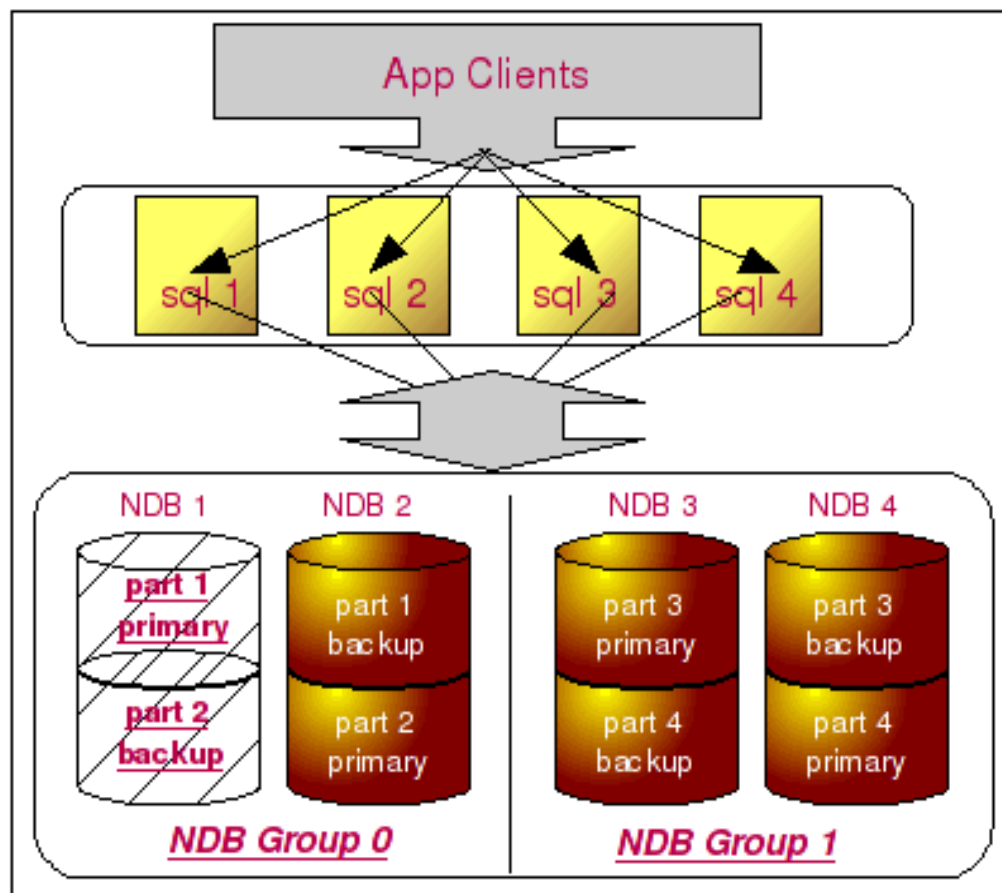
# 功能特性 - 高可用



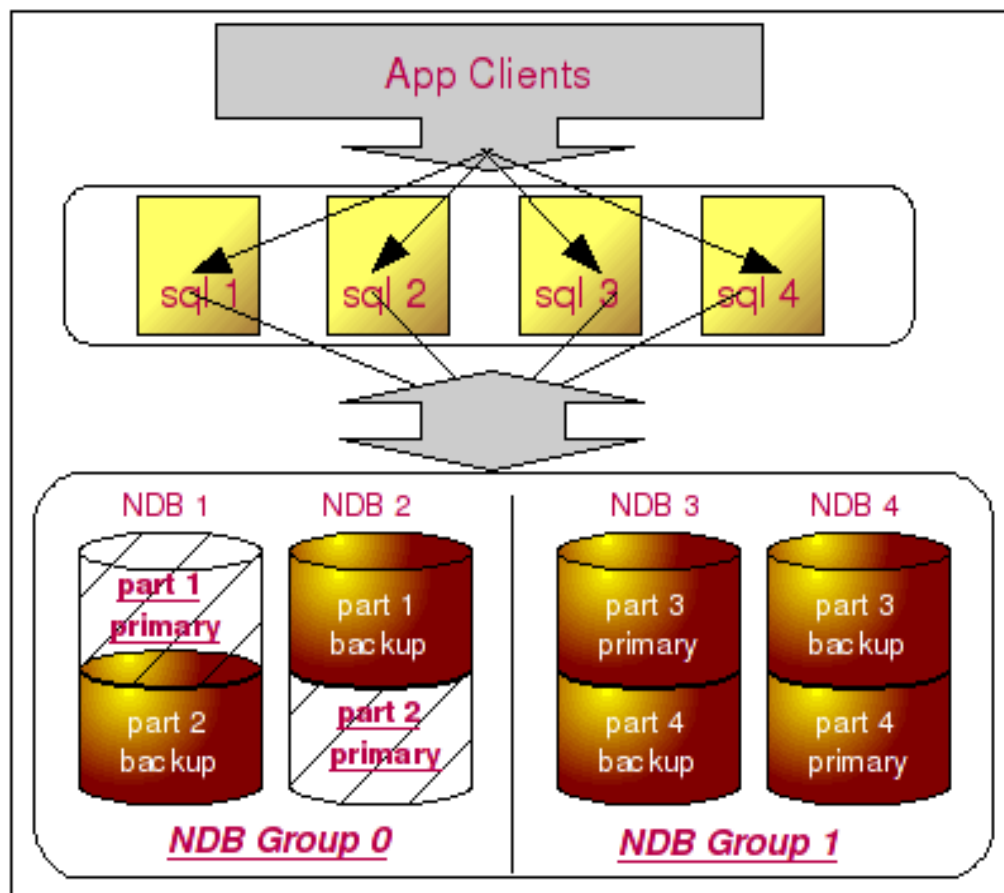
# 功能特性 - 高可用



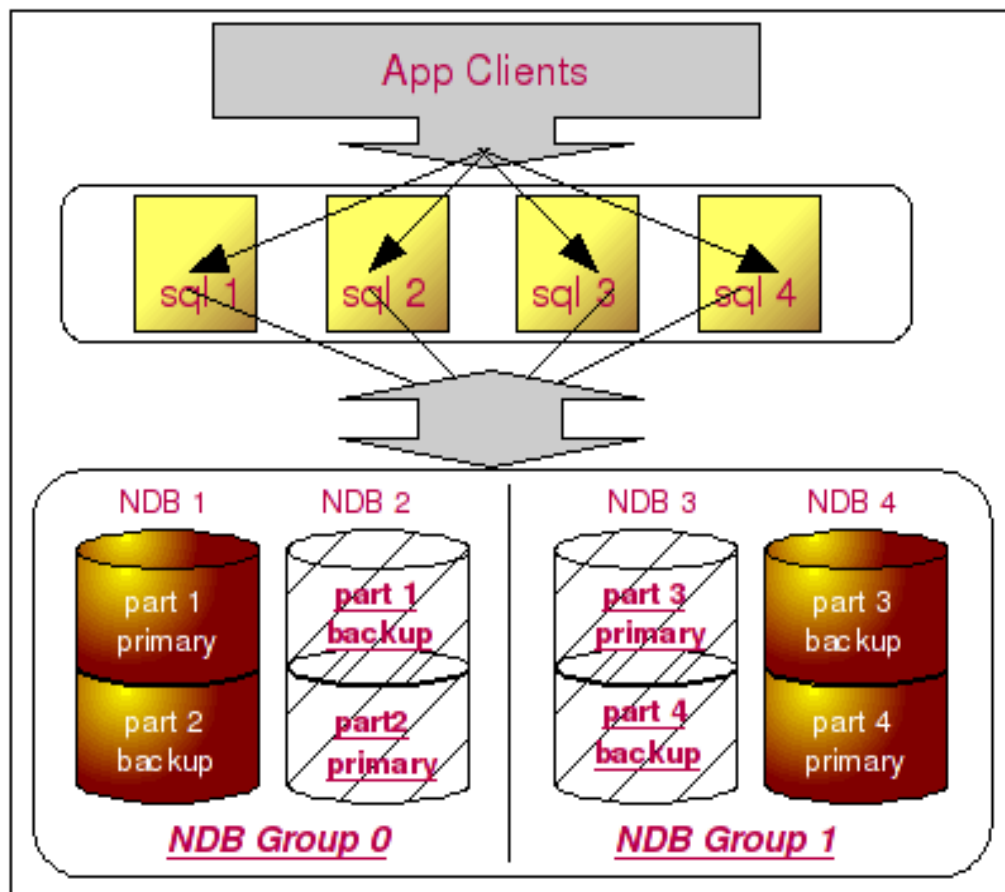
# 功能特性 - 高可用



# 功能特性 - 高可用



# 功能特性 - 高可用



---

# The End

Q & A

<http://iSky000.com>  
[mysql@googlegroups.com](mailto:mysql@googlegroups.com)