



SequoiaDB分布式集群容器化部署实战



杜蓉

个人简介



杜蓉

前网易高级DBA，中国科学院大学计算机硕士。曾负责网易生产数据库及数据库平台运维开发工作，乐于在分布式数据库、Hadoop/Spark、大数据等领域学习和探索。



CONTENT



分布式集群架构



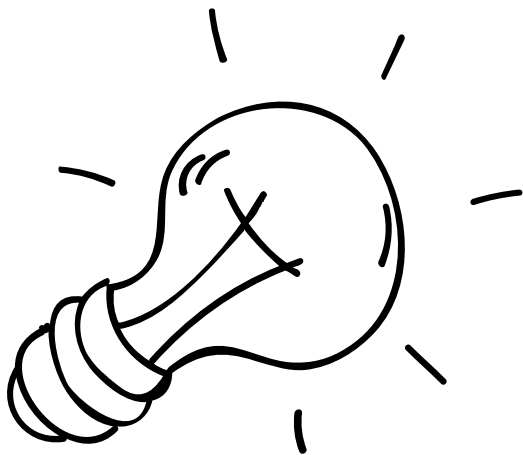
弹性扩展/高可用



容器化部署实战

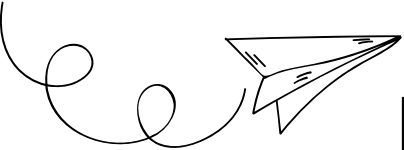


巨杉数据库生态工具

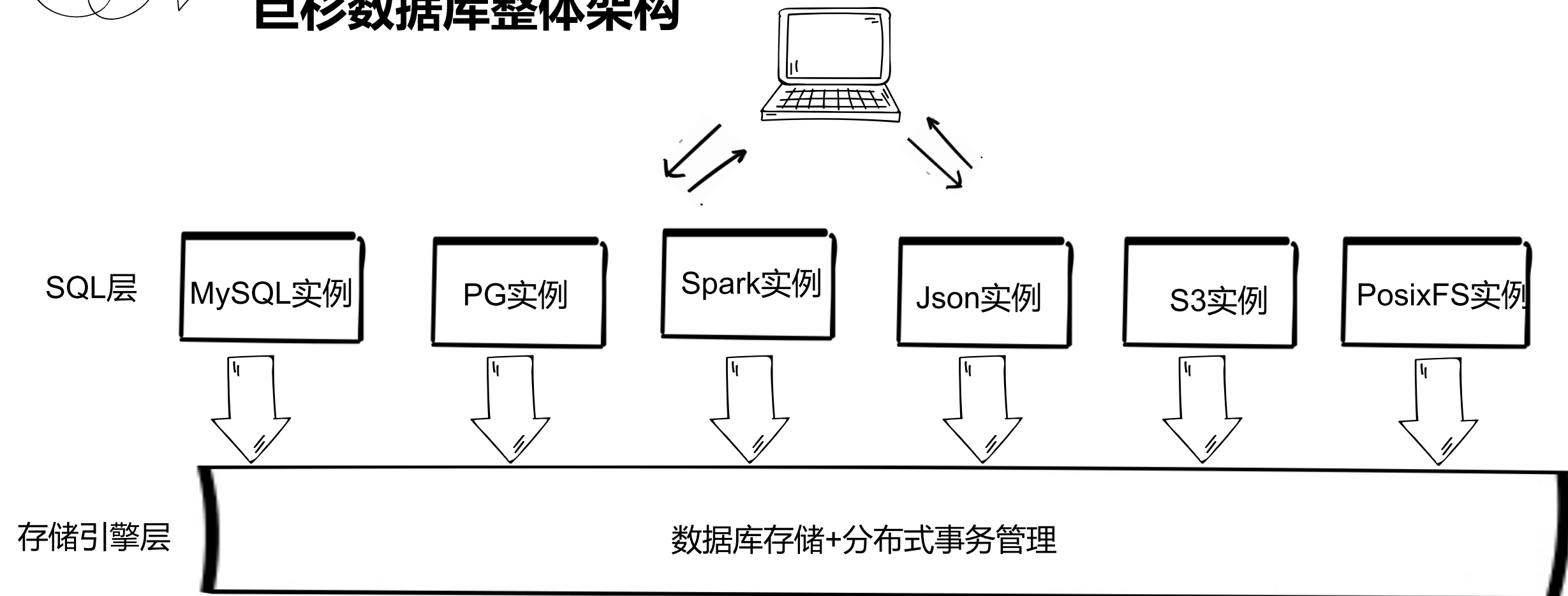


PART 01

分布式集群架构

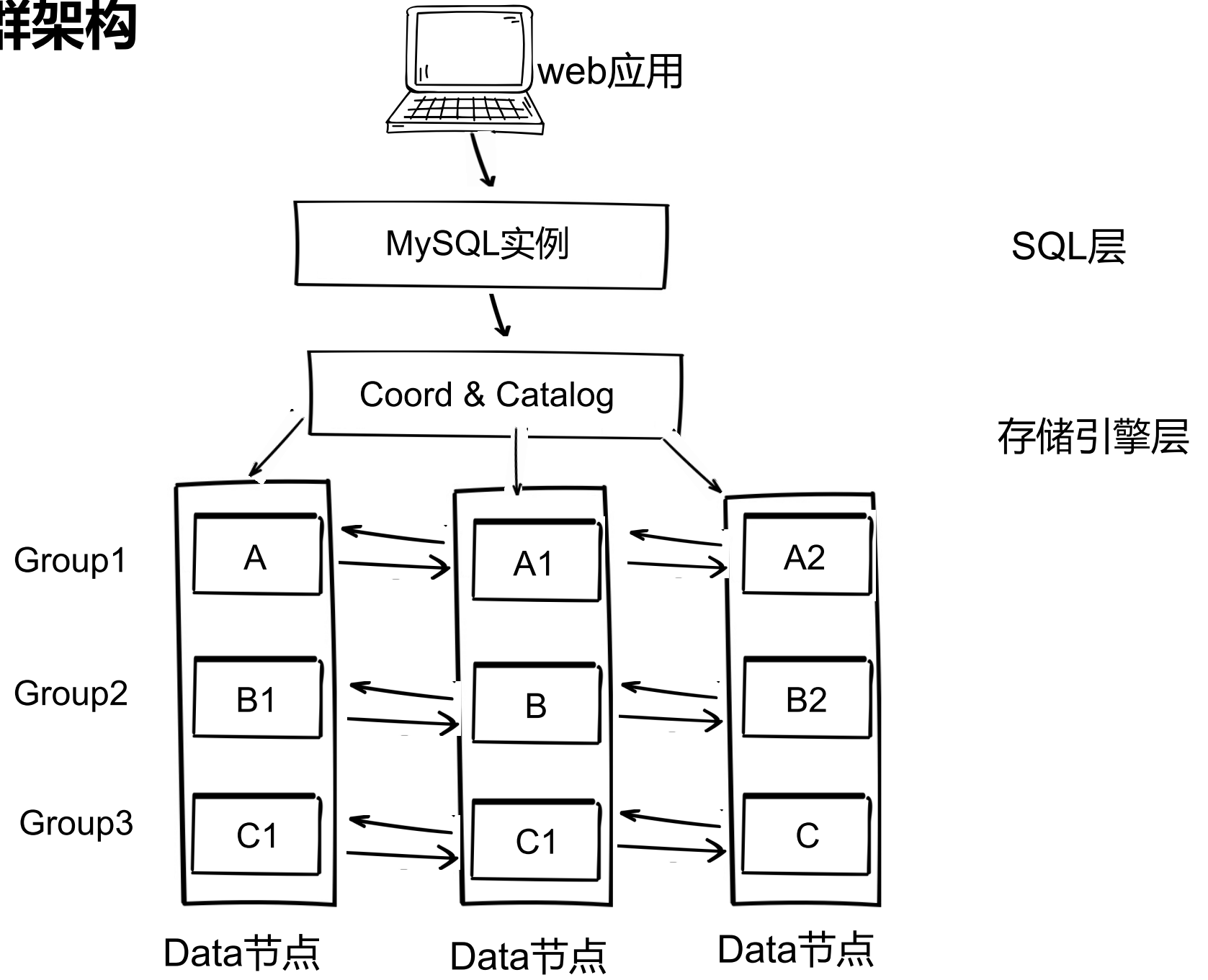


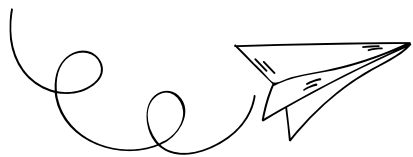
巨杉数据库整体架构





分布式集群架构





分片原理

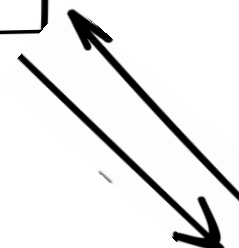
CategoriesID	MovieName	MovieDirector	MovieActor	MovieDesc	MovieData	MovieTime
6	何以笙箫默	无名	摩尼叔	23额度o	2012-02-22 0...	20
1	速度与激情	张艺谋	不可谓	科ouee	2012-02-18 0...	120
6	阿科未婚夫	无名	ndk	jkduw	2012-02-22 0...	20
1	井底蛙电脑	张艺谋	马拉	看到	2012-02-18 0...	120
6	洛带古镇	无名	农户	鸡窝	2012-02-22 0...	20
1	测试名字	张艺谋	潘长	摩尼叔	2012-02-18 0...	120
6	洛带古镇	无名	无名	。罚款	2012-02-22 0...	20

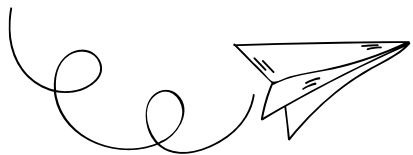


Targetpartition=DHT (ROW-PartitionKey)

协调节点

按照分区键切分并判断分区





数据读写

MySQL实例

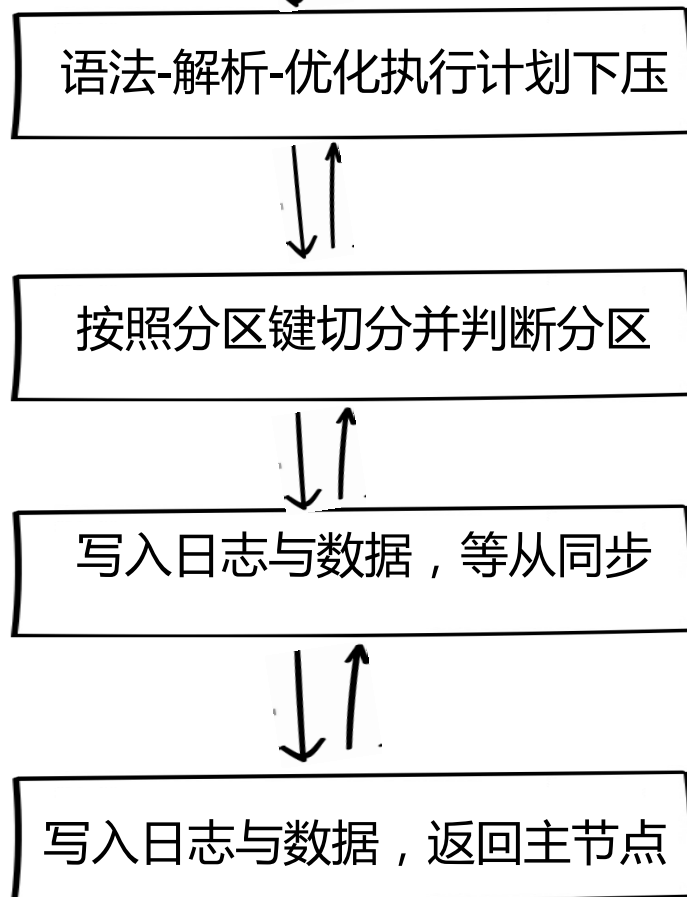
协调节点

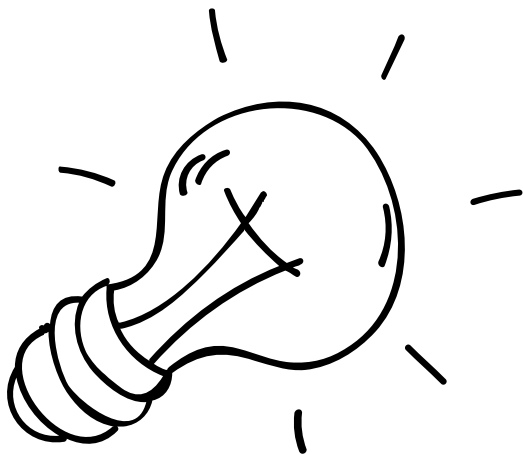
数据主节点

数据从节点



Insert/select 读写SQL

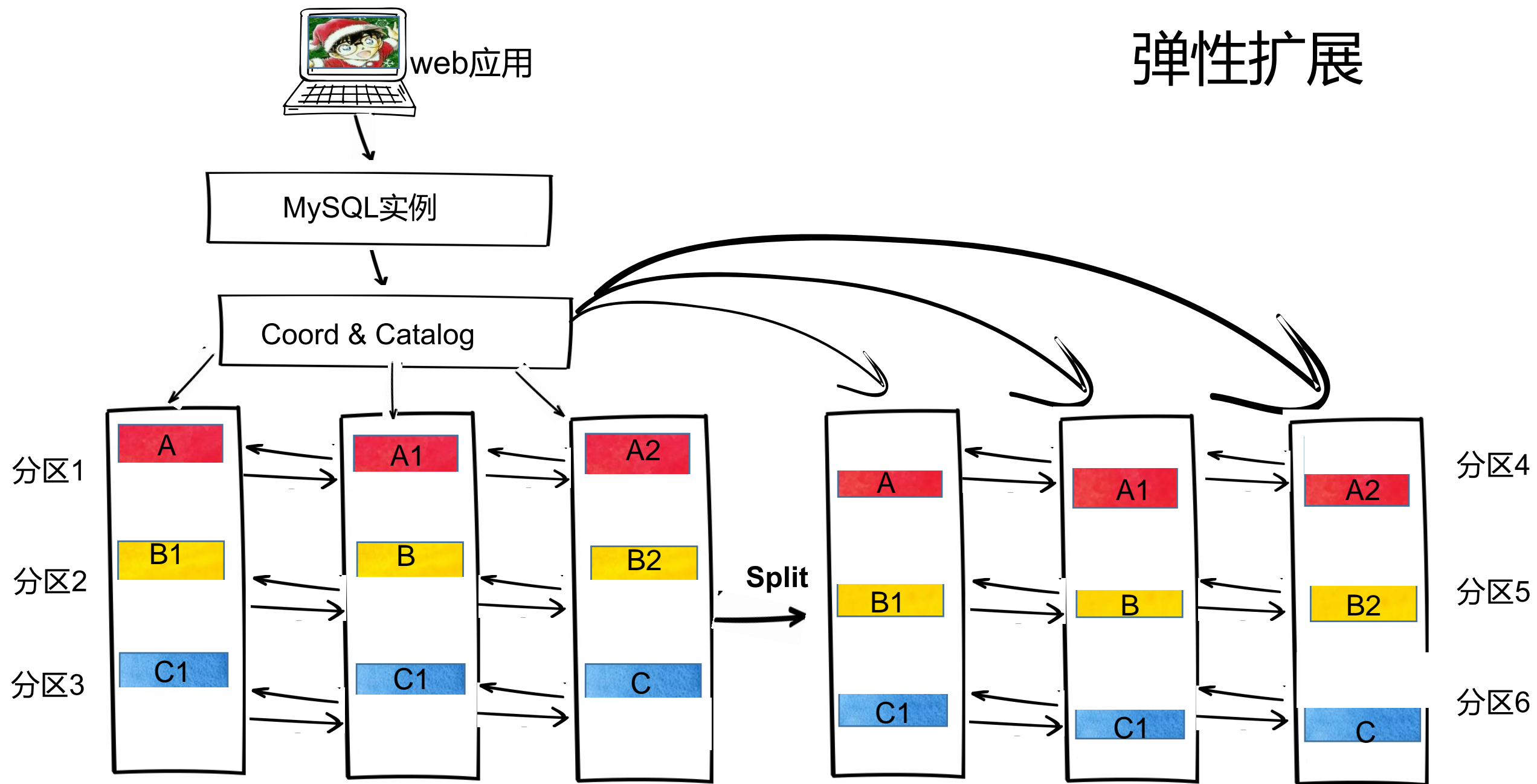




PART 02

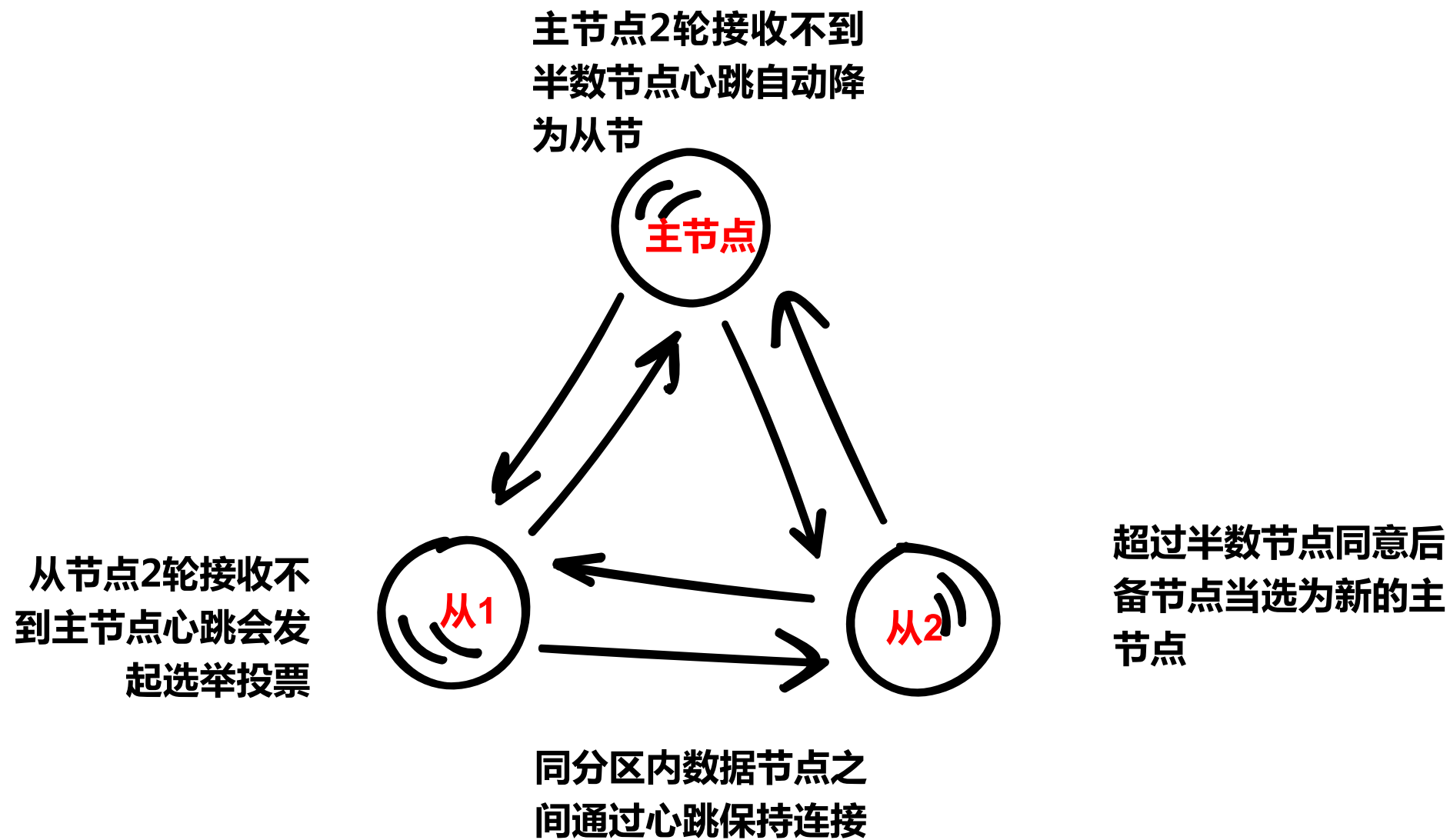
弹性扩展·高可用

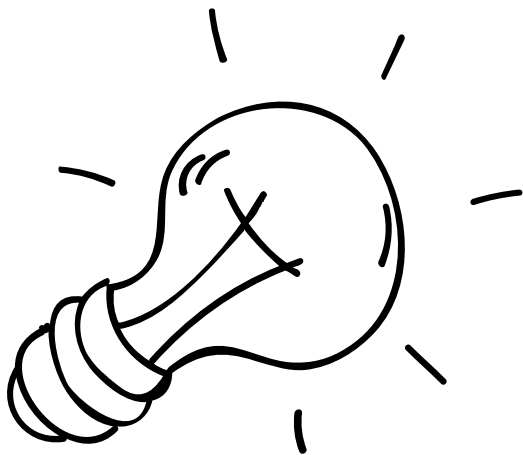
弹性扩展





高可用

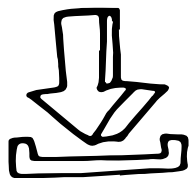




PART 03

容器化部署实战

The deploy for SequoiaDB cluster with Docker images



巨杉数据库下载部署

安装包

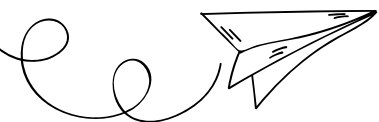
tar,gz包
版本SequoiaDB3.2.1

Vm镜像

Vmware station
导入虚拟机镜像

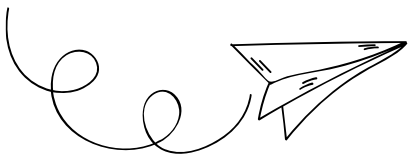
Docker镜像

拉取镜像
运行容器
搭建分布式集群

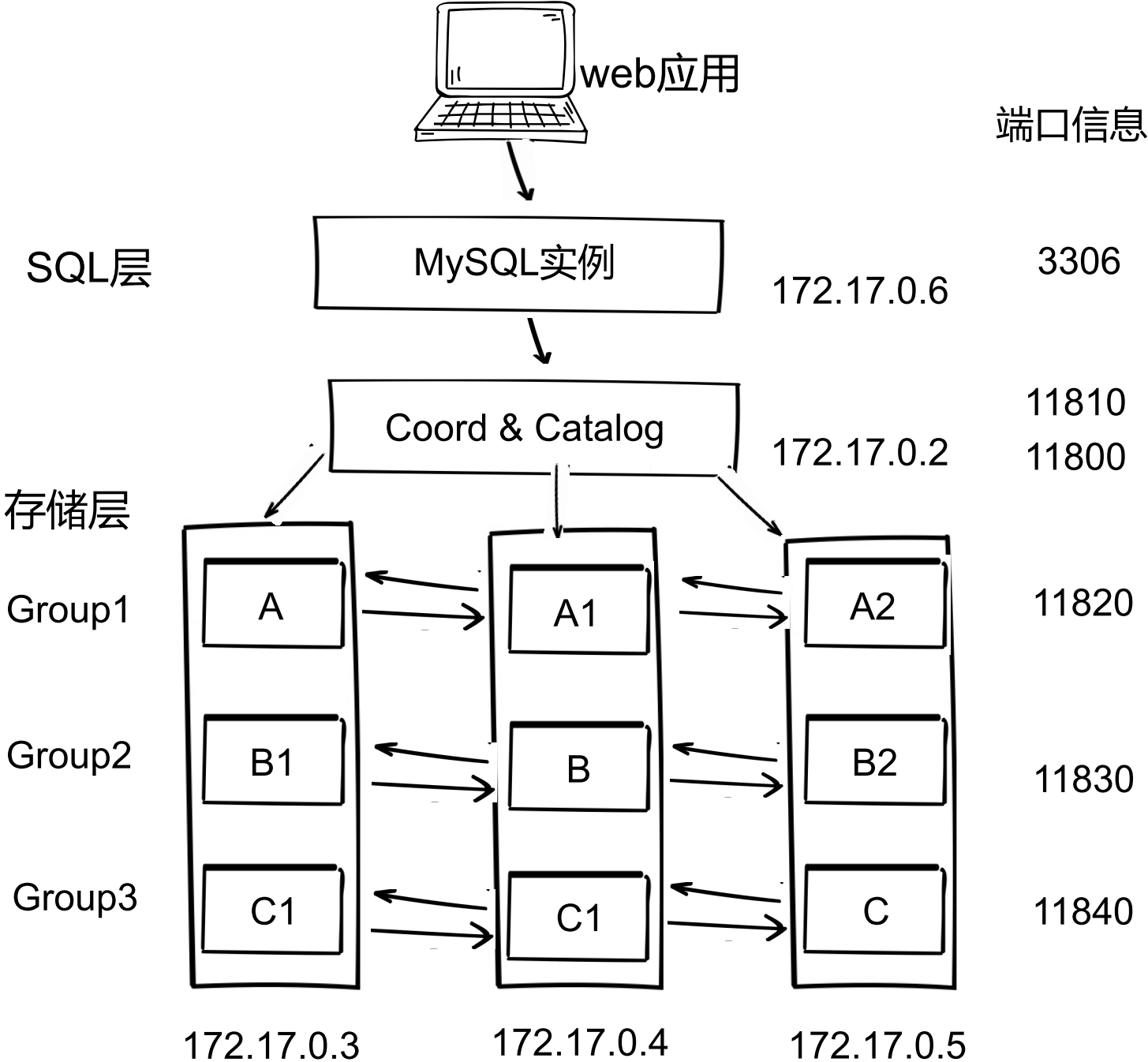


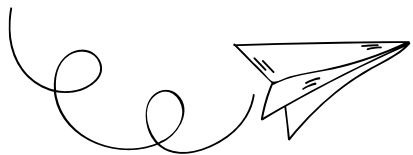
Docker集群信息表

主机名	IP	分区组	数据及功能	部署软件
协调节点 +编目节点	172.17.0.2	Coord+catalog	协调管理 存系统信息	SDB
Data节点1	172.17.0.3	data	A主节点 B1副本 C1副本	SDB
Data节点2	172.17.0.4	data	B主节点 A1副本 C副本	SDB
Data节点3	172.17.0.5	data	C主节点 A2副本 B2副本	SDB
MySQL节点	172.17.0.6	MySQL	SQL层功能	SequoiaSQL- MySQL

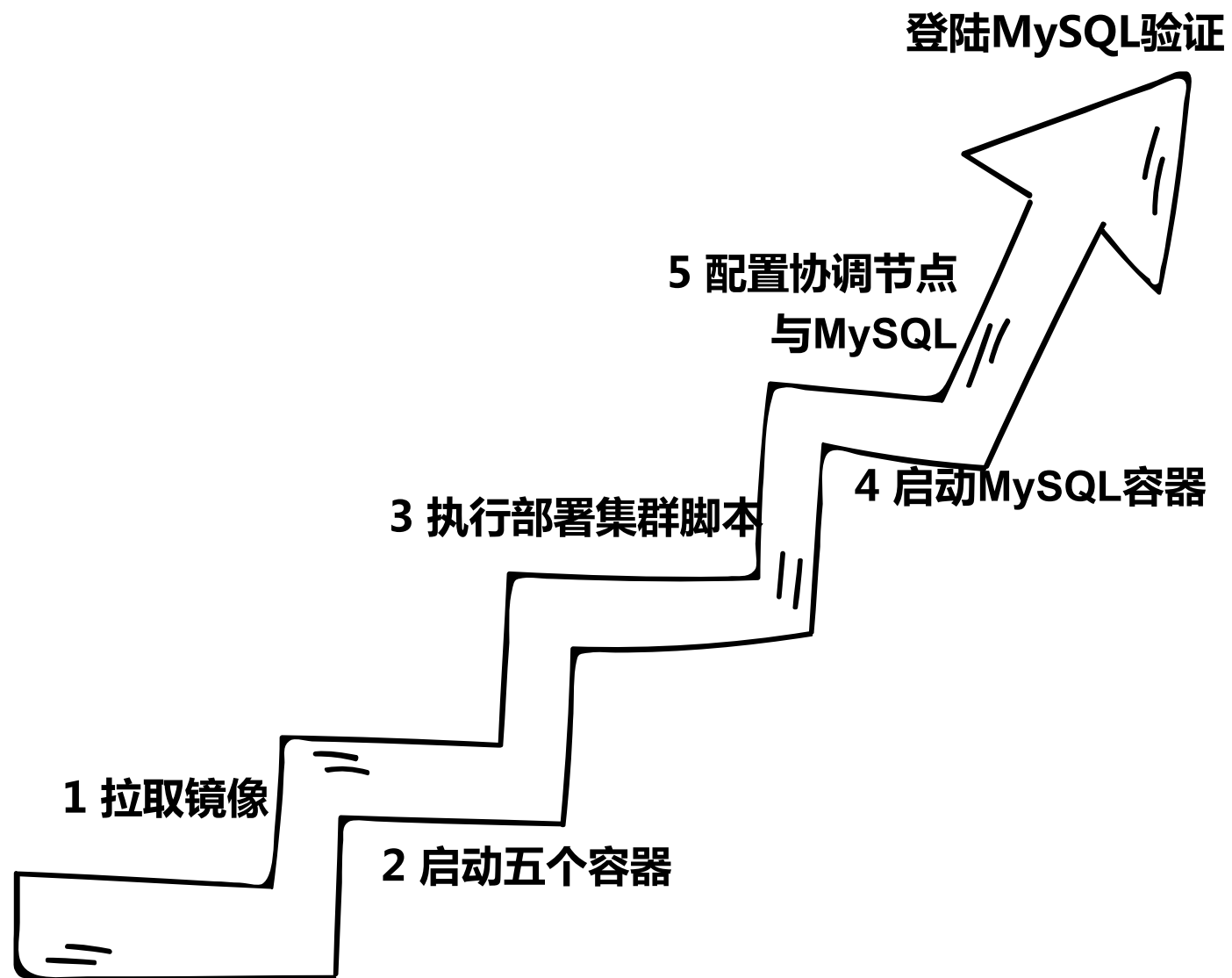


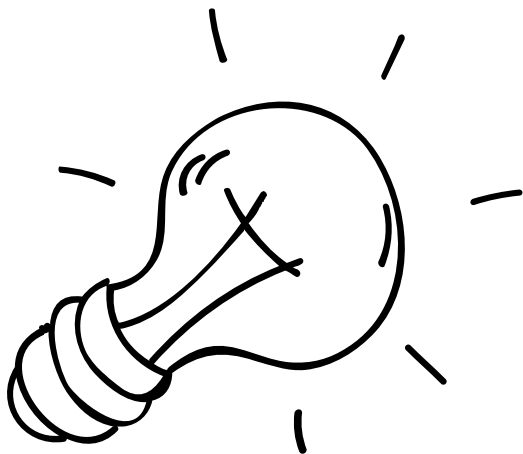
演示架构





Docker容器化部署步骤

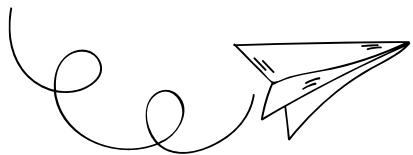




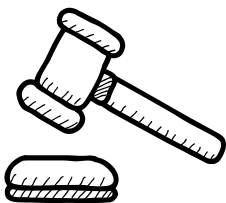
PART 04

巨杉数据库生态工具

The database tools for SequoiaDB system



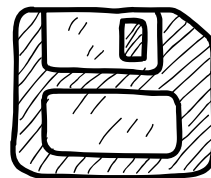
生态工具



兼容MYSQL客户端工具

图形化管理工具

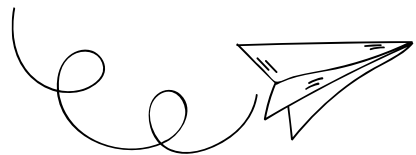
命令行监控管理工具



数据复制工具

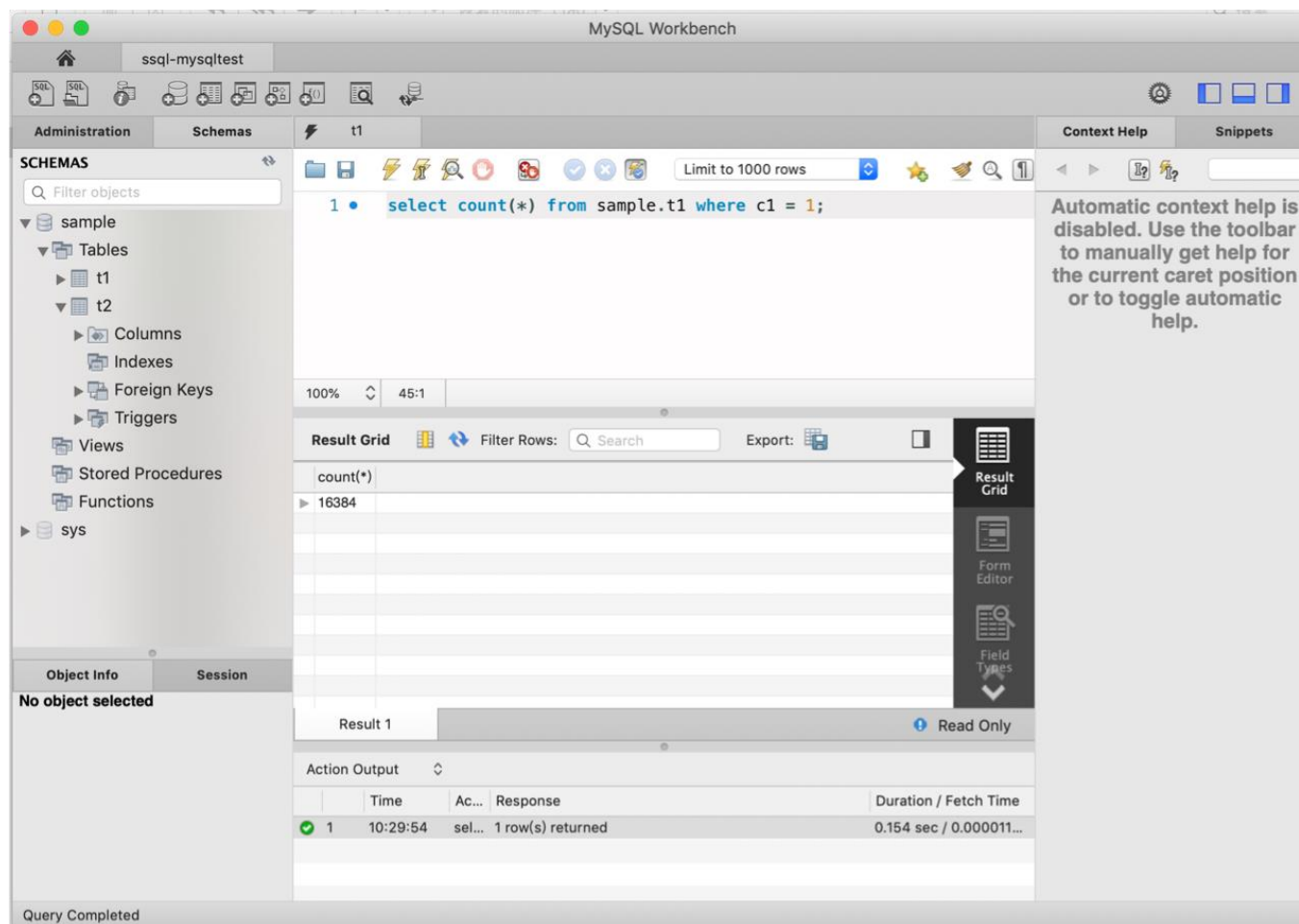
导入导出工具

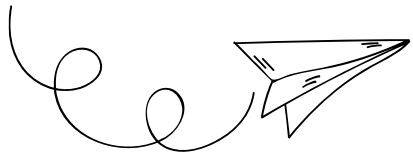
备份工具



100%兼容MYSQL

兼容MYSQL的访问工具 navicate/MySQLworkbench





SAC图形化监控工具

SequoiaDB SAC

admin

数据库实例 3

分布式存储 1

主机 3

创建实例

删除实例

重启实例

配置实例存储

集群

sdbDemo

Demo cluster

操作

一键部署

创建集群

删除集群

MySQL | 创建时间: 2019-06-07-08.49.05

正常运行...

MySQL | 创建时间: 2019-06-07-08.49.05

一共 1 台主机

CPU 0%

内存 21.91%

磁盘 18.82%

存储集群

MySQLInstance → SequoiaDB1

MySQLInstance1

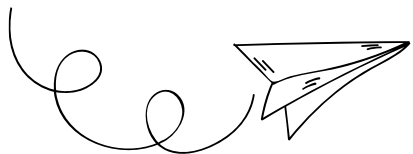
MySQL | 创建时间: 2019-06-07-08.50.07

正常运行...

一共 1 台主机

存储集群

©2019 SequoiaDB. 版本: 系统时间: 17:42:56 系统状态: 良好



SAC图形化监控工具

SequoiaDB SAC

admin

数据

监控

部署

策略

配置

MySQLInstance

#	表名		存储引擎	行数	表大小	创建时间	描述
1	customer		SequoiaDB	60000	70.25 MB		
2	district		SequoiaDB	20	960 KB		
3	history		SequoiaDB	60038	14 MB		
4	item		SequoiaDB	100000	25.88 MB		
5	new_order		SequoiaDB	16940	3.75 MB		
6	oorder		SequoiaDB	60000	23.63 MB		
7	order_line		SequoiaDB	600152	218.5 MB		
8	stock		SequoiaDB	200000	143.5 MB		
9	warehouse		SequoiaDB	2	896 KB		

1 / 1

显示 9 条记录，一共 9 条

数据库列表:

hsdb

test

wordpress

information_schema

mysql

performance_schema

sys

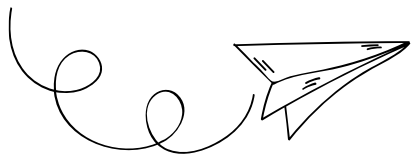
操作

创建数据库

删除数据库

创建数据表

删除数据表



SAC图形化监控工具

SequoiaDB SAC

admin

数据

监控

部署

策略

配置

SequoiaDB1

节点列表

刷新

修改配置

重启

NodeName	catalogaddr
sdbserver001:11810	sdbserver001:11810,sdbserver001:11813,sdbserver001:11823
sdbserver001:11820	sdbserver001:11820,sdbserver001:11813,sdbserver001:11823
sdbserver001:11830	sdbserver001:11830,sdbserver001:11813,sdbserver001:11823
sdbserver001:11840	sdbserver001:11840,sdbserver001:11813,sdbserver001:11823
sdbserver001:11850	sdbserver001:11850,sdbserver001:11813,sdbserver001:11823
sdbserver001:11823	

修改配置项

普通

高级

自定义

numpreload

maxprefpool

maxreplsync

logbuffsize

sortbuf

hjbbuf

syncstrategy

页面预加载代理数据

数据预取代理池最大数量

日志同步最大并发数量

复制日志内存页面数

排序缓存大小

哈希连接缓存大小

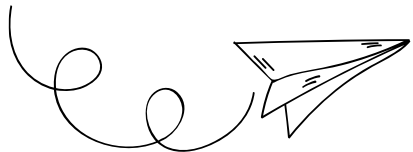
副本组之间数据同步控制

确定

取消

transisolation	weight
1	
1	
1	30
1	10
1	20

©2019 SequoiaDB. 版本: 系统时间: 17:45:35 系统状态: 良好



命令行管理工具 sdbtop

```
sequoiadb — sdbadmin@c7401:~ — ssh 218.17.39.139 — 111x32
refresh= 3 secs          version 3.2.2          snapshotMode: GLOBAL
displayMode: ABSOLUTE    Main Window          snapshotModeInput: NULL
hostname: localhost      filtering Number: 0
servicename: 11810       sortingWay: NULL sortingField: NULL
usrName: NULL
```

```
##### For help type
# # # # # sdbtop -h: us
# # # # #
# # # # #
#####
```

SDB Interactive Snapshot Monitor V2.0
Use these keys to ENTER:

window options(choose to enter window):

```
m : return to main window      s : shd
c : show collection spaces     t : shd
d : show database state        h : hel
```

options(use under window above):

```
G : show options only          g : fi
n : filter by specified node    r : set
A : sort column by ascending order D : som
C : specify filter condition    Q : can
N : skip specified lines ahead  W : shd
Tab : switch display Model      <- : mov
-> : move right                 Enter: to
ESC : cancel current operation  F5 : ref
q : quit
```

Licensed Materials - Property of SequoiaDB
Copyright SequoiaDB Corp. 2013-2016 All Rights Reserved

```
sequoiadb — sdbadmin@c7401:~ — ssh 218.17.39.139 — 111x32
refresh= 3 secs          version 3.2.2          snapshotMode: GLOBAL
displayMode: ABSOLUTE    Sessions           snapshotModeInput: NULL
hostname: localhost      filtering Number: 0
servicename: 11810       sortingWay: NULL sortingField: NULL
usrName: NULL            Refresh: F5, Quit: q, Help: h
```

SessionID	TID	Type	Name	QueueSize
1	1	88922 Task	DATASYNC-JOB-D	0
2	1	21445 Task	DATASYNC-JOB-D	0
3	1	193724 Task	DATASYNC-JOB-D	0
4	1	89012 Task	DATASYNC-JOB-D	0
5	1	21501 Task	DATASYNC-JOB-D	0
6	1	193755 Task	DATASYNC-JOB-D	0
7	1	21500 Task	DATASYNC-JOB-D	0
8	1	193754 Task	DATASYNC-JOB-D	0
9	1	89013 Task	DATASYNC-JOB-D	0
10	1	193756 Task	DATASYNC-JOB-D	0
11	1	89014 Task	DATASYNC-JOB-D	0
12	1	21502 Task	DATASYNC-JOB-D	0
13	2	88958 LogWriter	""	0
14	2	21455 LogWriter	""	0
15	2	193735 LogWriter	""	0
16	2	89015 LogWriter	""	0
17	2	21510 LogWriter	""	0
18	2	193784 LogWriter	""	0
19	2	21503 LogWriter	""	0
20	2	193757 LogWriter	""	0
21	2	89045 LogWriter	""	0
22	2	193811 LogWriter	""	0
23	2	89088 LogWriter	""	0
24	2	21555 LogWriter	""	0

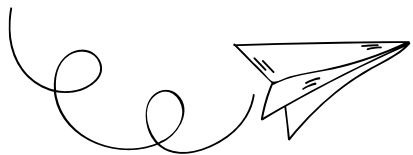
1:~ — ssh 218.17.39.139 — 111x32

```
3.2.2          snapshotMode: GLOBAL
ase           snapshotModeInput: NULL
              filtering Number: 0
              sortingWay: NULL sortingField: NULL
              Refresh: F5, Quit: q, Help: h
```

```
IndexRead      DataWrite      IndexWrite
0.667          90894.000      90893.667

Insert          ReplUpdate      ReplDelete
90894.000        0.000          0.000

Read            ReadTime        WriteTime
2.000           0.000          0.000
```



备份工具

数据库实例级备份

MySQL/PostgreSQL备份

集群级备份

全量离线备份

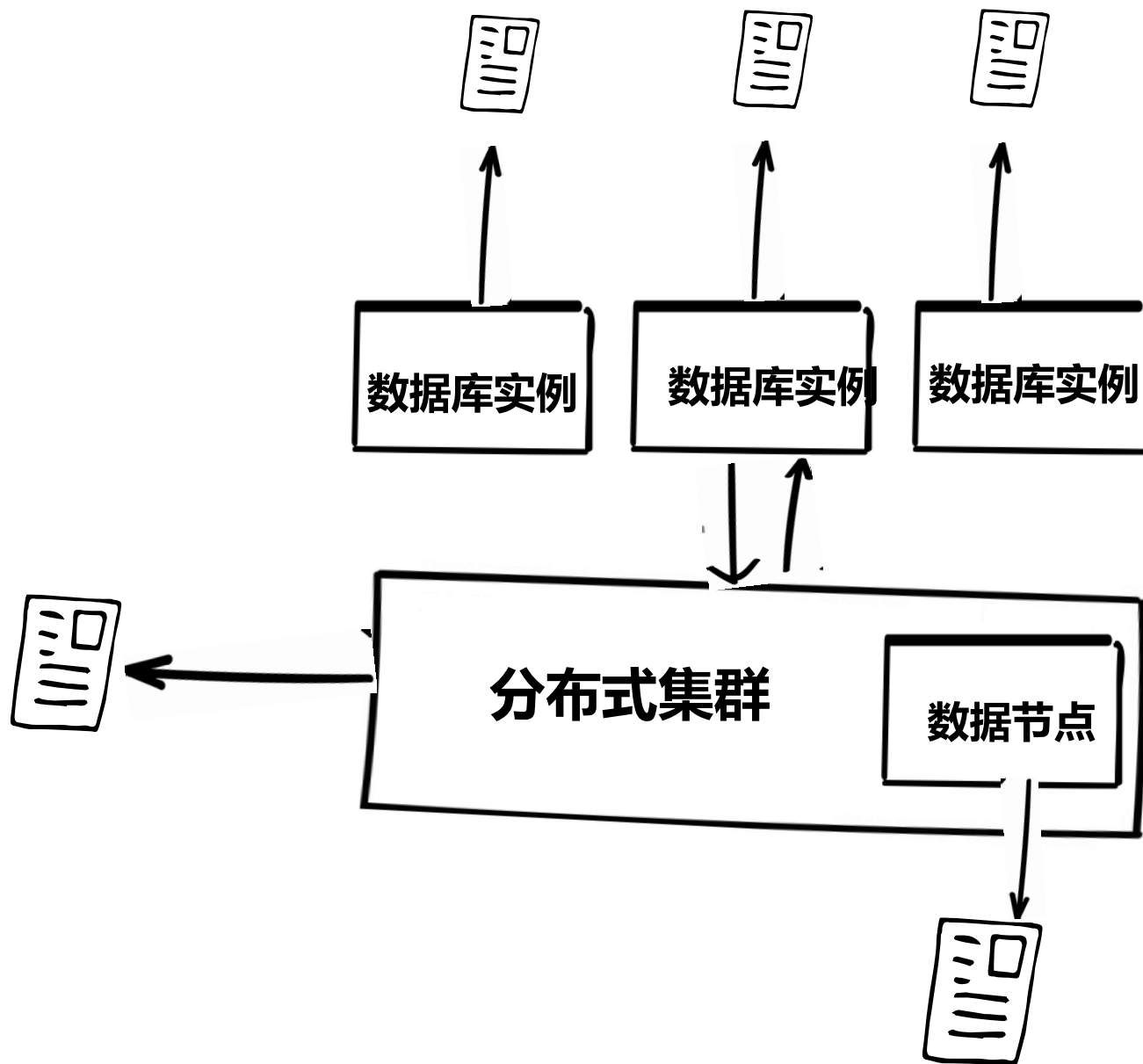
全量在线备份

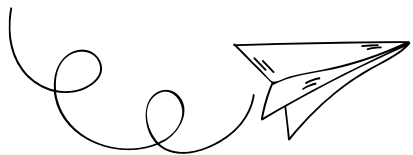
增量在线备份

文件系统级备份

读节点文件系统全量备份

静态文件增量备份





数据导入导出

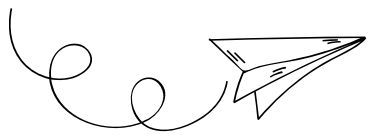
标准开源工具mydumper/myloader

标准图形化工具 Kettle

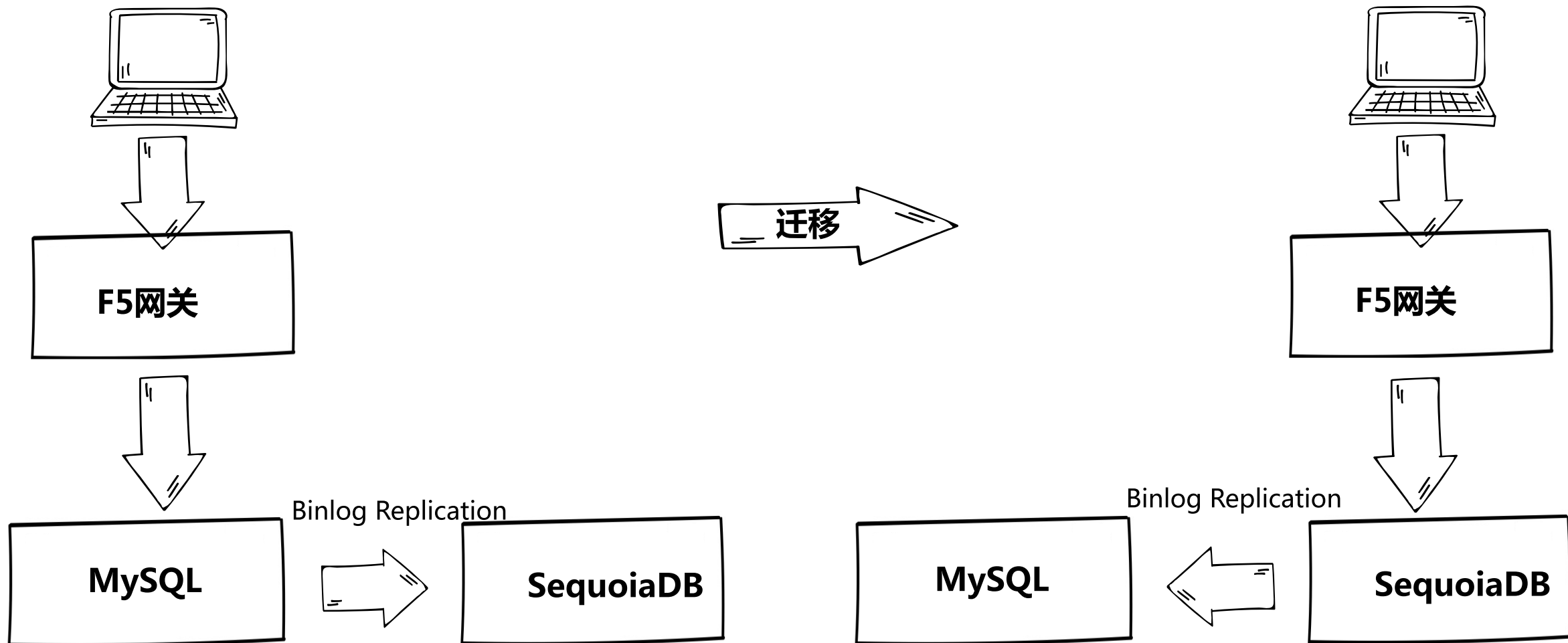
MySQL实时复制Binlog replication

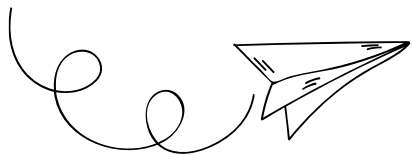
Oracle实时复制Golden Gate

巨杉导入导出sdbimpert/sdbexpert



数据双向复制

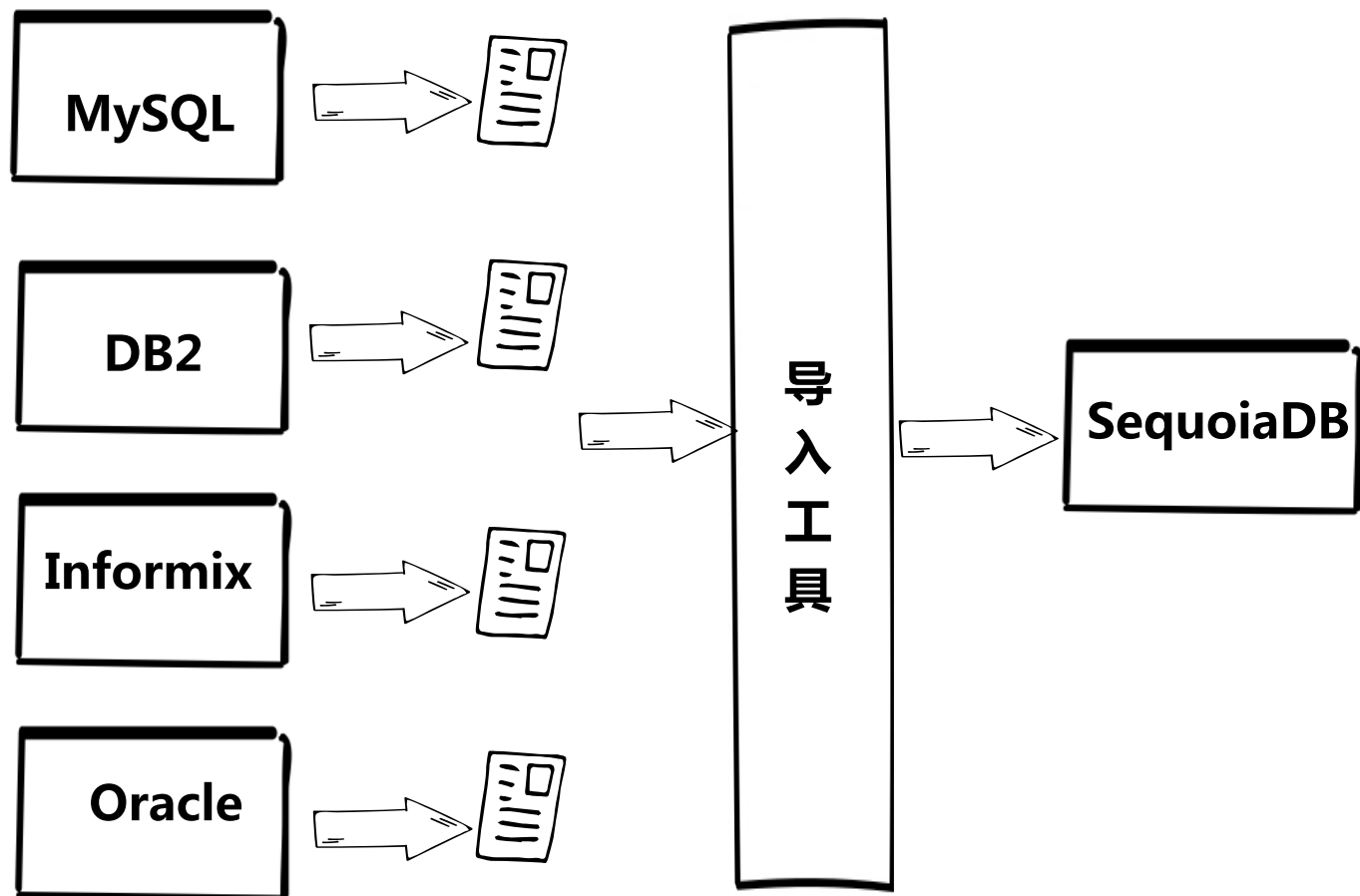


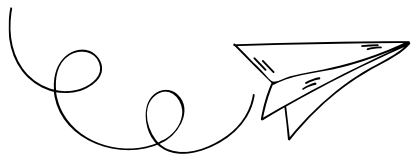


异步复制

异步复制策略

- 1 使用Json或CSV格式定期导出异构数据为文本文件
- 2 通过FTP等方式将文件传输至巨杉数据库客户端
- 3 通过sdbimprt工具将文本文件导入巨杉数据库
- 4 满足异构数据源T+1 数据复制策略，简单可靠

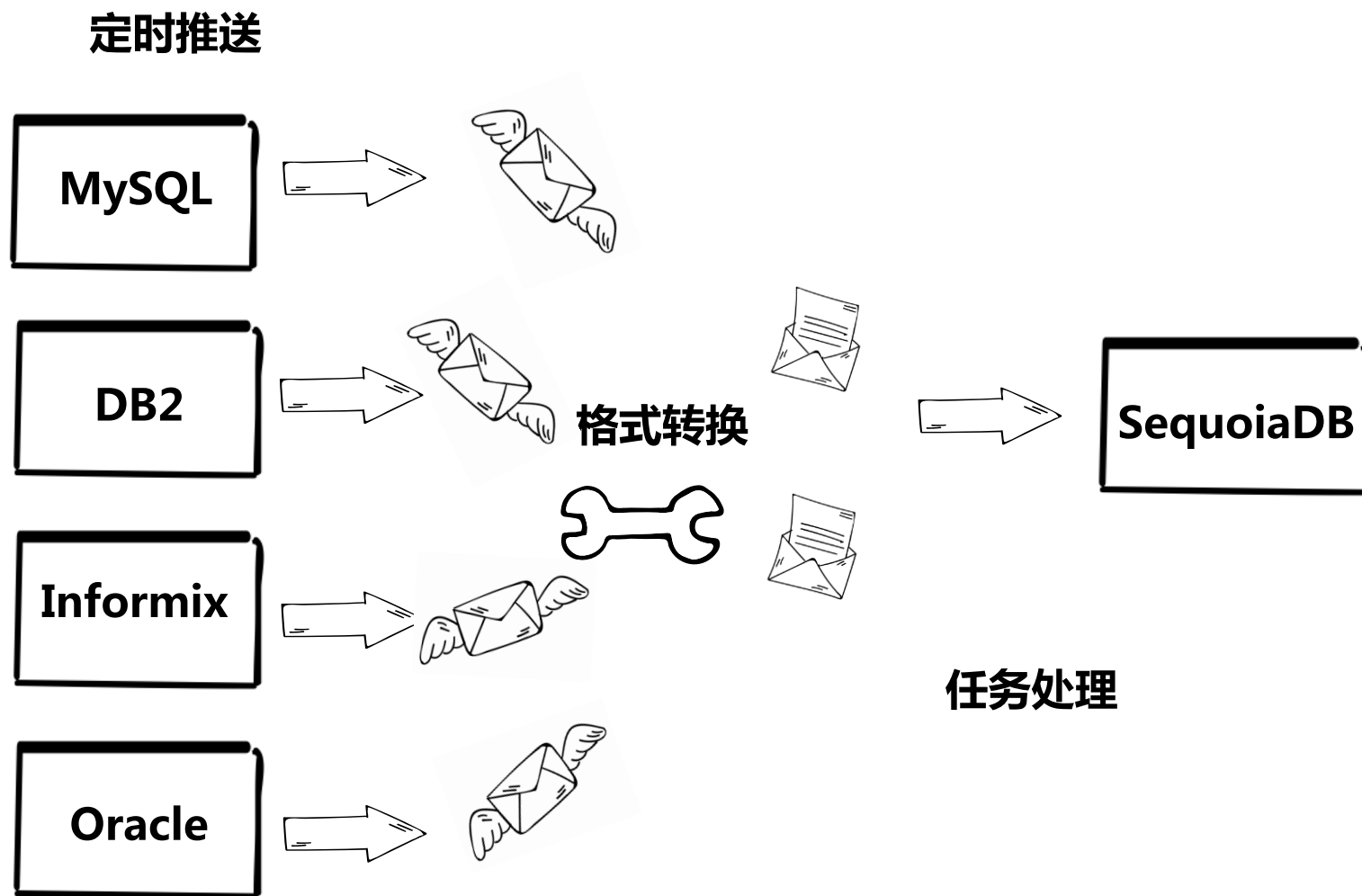




准同步复制

准实时复制策略

- 1 异构数据源使用相关工具将日志文件实时解析写入管道
- 2 通过ApacheStorm对管道信息监听并转换为标准的DML/DDDL命令
- 3 指令分发至多线程处理服务进行巨杉数据库的增删改查
- 4 满足异构数据源T+0的数据复制策略，秒级演示
- 5 当前中支持 Oracle Golden Gate (Oracle) DB2 IIE (IBM的Informix) 以及Cannel (对应MySQL)
- 6 对应当前不支持的数据库需要寻找开源的日志解析工具进行独立开发





总结



分布式集群架构



弹性扩展/高可用



容器化部署实战



巨杉数据库生态工具



欢迎关注巨杉数据库社区

