

一. 简介

1. 应用场景:

- 1) 推荐系统: 实时缓存.
- 2) 网站统计: 实时销量、流量统计.
- 3) 监控预警系统、金融风控系统.

2. 特性

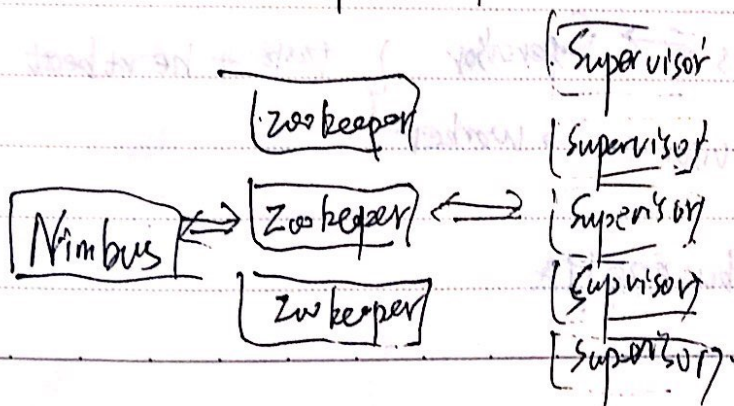
- 1) 适用场景广泛: 实时处理和更新, 持续并行化查询, 满足大量需求.
- 2) 可扩展性强: 扩展计算任务.
- 3) 数据无丢失: 保证每条消息都得到处理.
- 4) 异常健壮: 集群.
- 5) 容错:
- 6) 语言无关:

二. 原理.

1) 主从架构: Hdfs Yarn Storm Nbase spark, Flink.

2) 对称架构: zookeeper, kafka.

Storm 架构:



Storms组件:

① Nimbus

1) 接收 client topo 代码, 拆分 task, 将 task 存入 zookeeper.

2) 将 task \rightarrow Supervisor 映射 \rightarrow zk

3) 故障/异常检测

② Supervisor

1) 从 Nimbus 读取拓扑代码. $zk \xleftarrow{\text{task}} \xrightarrow{\text{task}} \text{Nimbus}$

2) worker 执行.

3) 监控 worker

③ Worker

1) $zk \xrightarrow{\text{task}} \text{worker}$,

2) Executor 多个线程. 执行 task.

④ Zookeeper

1) Nimbus \leftrightarrow Supervisor } task + heart beat

2) Supervisor \leftrightarrow worker }

3) Nimbus 高可用 HA



Zookeeper 简介.

1. 分布式开源、分布式协调服务框架.

基于 chubby.

2. 核心功能.

1) 文件系统: { 1. 每个目录 = znode
2. znode \leftrightarrow data
3. 持久化、持久化顺序、临时、临时顺序.

2) 通知机制

{ 1. 客户端监听
2. znode 节点变化 (修改/删除/子节点添加删除) 通知客户端

3. 安装部署

archive.apache.org/dist/zookeeper

注: 之前已经搭建过 zookeeper 集群, 需要请视 /tmp/ 下的软件版本.

bin/zkServer.sh start/status.

4. 命令访问:

\$ bin/zkCli.sh.

[zk] creat /test ""

get /test

set /test 222.

rmdir /test



5. zk集群配置部署.

~~zk.conf~~ zk.cfg.

server, <id> = hostname: 2888:3888

x3 {
 server.2 = {
 server.3 = {

↘ leader选举算法.

\$ cd zk/data.

\$ vi myid

1.

\$ scp -r zookeeper-3.4.5 / root@2/3 : /opt/modules.

\$ vi myid

2.

\$ vi myid

3.

主机名配置.

\$ cat /etc/sysconfig/network

NETWORK=yes

HOSTNAME=hostname.

主机名与IP的映射关系.

\$ cat /etc/hosts

192.168.0.101 hostname1

.102. hostname2

.103 hostname3



2. 集群环境.

1) 清理 data, `rm -rf ./data`, `vi myid`.

`$ bin/zkServer.sh start X3`.

2) 配置.

`$ vi storm.yaml`

`storm.zookeeper.servers:`

- "hostname1"

- "hostname2"

- "hostname3"

`nimbus.seeds: ["hostname1", "hostname2"]`

`$ scp -r storm root@2/3: /opt/module`

3) 启动

host1 | `$ nohup bin/storm nimbus &`

`$ nohup bin/storm supervisor &`

host2 | `$ nohup bin/storm supervisor &`

`$ nohup bin/storm UI &`



④

\$ bin/zkCli.sh -server ~~localhost~~ hostnam1:2181 --in host1上zoo host2

\$ bin/zkCli.sh -server ~~localhost~~ hostnam1:2181 --in host1

[zk] connect hostnam3:2181 --in host3

Storm 安装配置

1. 不修改配置

nohub
设置交互。
要shell关闭后
不合并进程。

\$ bin/storm.sh nimbus \$

\$ bin/storm.sh supervisor \$

\$ bin/storm us \$

hostname: 8080

集群概述: slots 桥, worker

主节点概述

从节点概述

\$ bin/storm logviewer

