系统性能指标和测试结果说明

(一) 性能测试报告

测试目标

运营商手机上网记录查询系统案例，以某运营商为例，日均上网记录数近 10 亿条，每月数据量近 9TB，移动互联网用户快速增加，智能终端迅速普及、户均流量显著增长，上网记录数据将进一步猛增，每 6 个月，流量翻一番，如此大的数据量已经超越了传统关系型数据库可管理的容量上限，关系型数据库上对大规模数据进行操作会造成系统性能严重下降。通过本测试，验证星环科技成熟稳定的商用 Hadoop 平台，是否可以有效解决数据采集、加载、存储、查询、分析等问题。

测试内容

1）存储节点数和存储量验证；

2）并发加载数据的效率验证；

3）分别选取简单查询（短信话单查询），单表统计（某天某客户通话次数），大表关联统计（统计指定用户的上网记录）三个应用场景验证产品性能。

测试环境

软硬件环境配置如下：

表 9-1　服务器配置

服务器	推荐配置及说明
节点数量	300+台服务器

CPU	两路 6 核处理器 2*E5-2620
内存	64GBECCDDR3
硬盘	2 个 600G 的 SAS 硬盘，15000RPM，RAID1，作为系统盘 12 个 2TB 的 SATA 硬盘，7200RPM，不做 RAID1
网络	双电口万兆（10Gbps）以太网卡

部署环境如下：

表 9-2　集群配置

PC 服务器	300 多台
NameNode 节点	3 台
DataNode 节点	278 台
Zookeeper 节点	7 台
集群监控节点	1 台
入库服务节点	24 台
Web 查询应用服务节点	20 台

网络拓扑情况如下：

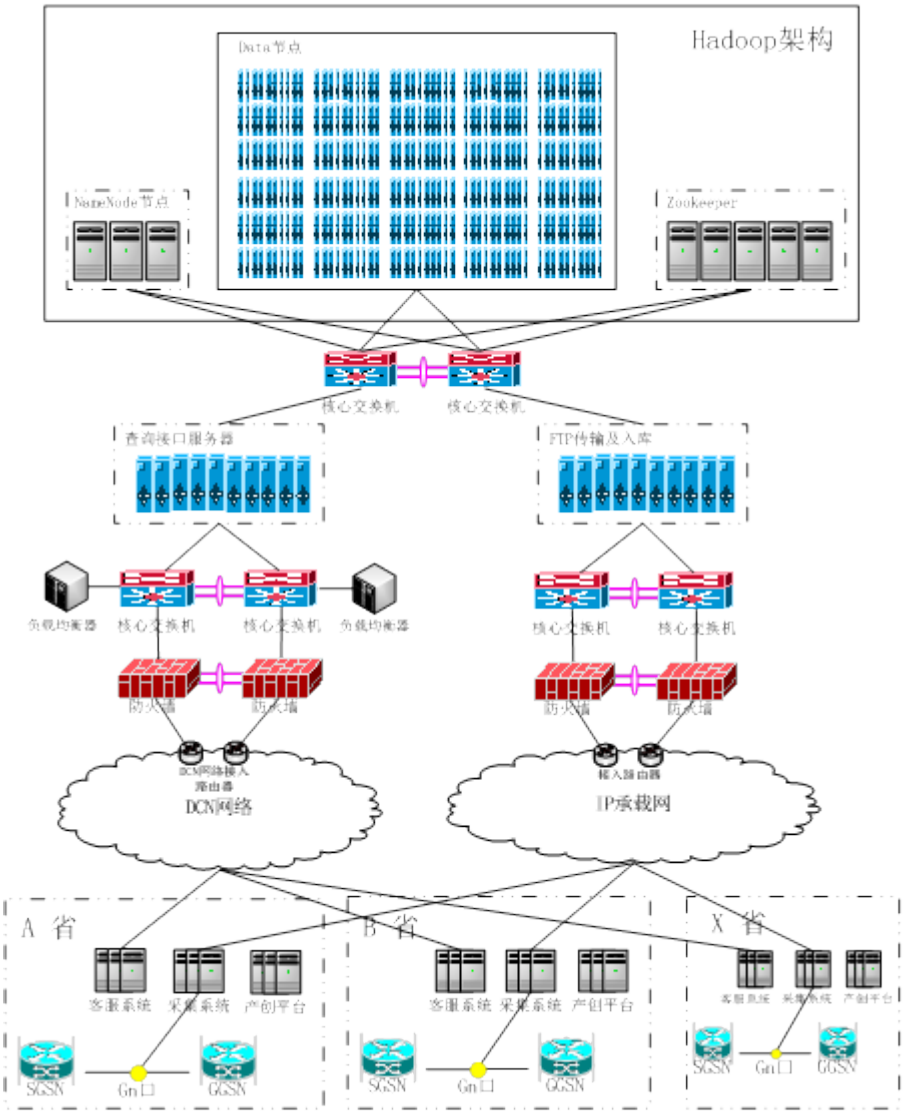


图 9-1 拓扑结构图

测试过程和结果

1）现有 HDFS 集群已被占用 10.5PB，3 个副本，压缩率在 1/3 左右，因此实际 HBase 表数据也已经有 3.5PB 左右。目前数据存放 6 个月，每天导入日志数据在 21TB 左右，每月导入新增日志数据量为 630TB，近一个月为常用热数据，数据量增长较快。

2）并发加载数据的效率

Transwarp Hyperbase 集群每秒平均达到 1500 万记录/秒，峰值时达到 5000 万/秒，集群导入性能没有问题。

3）支持并发查询数目：远高于 100000 请求/秒

上网记录查询速度：不高于 1 秒（含用户访问查询页面的时间）

场景一：短信话单查询

图 9-3 话单查询表

测试相关表数据量	表名	条数
	CDR_GSM_13	3114402798
测试语句	SELECT*FROMCDR_GSM_13WHEREUSER_ID=?;	
场景说明	使用程序查询指定 USER_ID 的短信话单	
TranswarpSQL 性能	并发:5000 单 SQL 平均执行时间:12ms	
Transwarp Hyperbase API 性能	并发：>30W 单 SQL 平均执行时间：3ms	

场景二：某天某客户通话次数：

测试相关表数据量	表名	条数
	CDR_GSM_13，cdr_gsm_stat	3114402798， 430293346
测试语句	SELECTcount(*)FROMCDR_GSM_13C， cdr_gsm_statGWHEREc.user_id=g.user_idandg.type =’1’andg.date=’20151212’g.user_id=?;	
场景说明	统计某天某客户通话次数	

TranswarpSQL 性能	并发:2000 单 SQL 平均执行时间:130ms
Transwarp Hyperbase API 性能	并发：>10W 单 SQL 平均执行时间：90ms

场景三：关联统计相关测试，统计制定用户的上网记录

图 9-4 上网记录表

测试相关表数据量	表名	条数
	CDR_GSM_13，cdr_gsm_stat	3114402798， 430293346
测试语句	SELECTcount(*)FROMCDR_GSM_13C， cdr_gsm_statGWHEREc.user_id=g.user_idandg.type= ’1’ andg.user_id=?;	
场景说明	使用程序统计指定 USER_ID 的上网记录	
TranswarpSQL 性能	并发:3000 单 SQL 平均执行时间:150ms	
Transwarp Hyperbase API 性能	并发：>10W 单 SQL 平均执行时间：80ms	

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。

如要下载或阅读全文，请访问：

https://d.book118.com/608026034135006116