

目录

目录	1
产品概述	2
产品优势	2
稀疏性	2
海量数据实时存储	2
高可靠性	2
扩展性强	2
应用场景	2
海量数据实时存储	2
车联网场景	2
用户画像场景	2

产品概述

HBase服务是一款基于Apache HBase构建的，支持NoSQL数据库的分布式集群。具有海量数据的实时处理能力、高并发吞吐、以及水平扩展等能力，可支持消息/订单、Feeds流、推荐画像、小文件对象存储、车联网、物联网等场景。

HBase 是典型的 NoSQL 数据库，通常被描述成稀疏的、分布式的、持久化的，由行键、列键和时间戳进行索引的多维有序映射数据库，主要用来存储非结构化和半结构化的数据。因为 HBase 基于 Hadoop 的 HDFS 完成分布式存储，以及 MapReduce 完成分布式并行计算，所以它的一些特点与 Hadoop 相同，依靠横向扩展，通过不断增加性价比高的商业服务器来增加计算和存储能力。

产品优势

稀疏性

HBase服务解决了数据稀疏性问题。HBase是列存储数据库，数据库中的数据都是以字符串形式存储的，为空的列并不占用存储空间，很大程度上节省存储开销。

海量数据实时存储

HBase服务支持海量数据实时存储。HBase单表可以有百亿行、百万列，可以提供海量数据存储。同时，region切分、主键索引和缓存机制，可以保证HBase读取性能。

高可靠性

HBase服务具有高可靠性。HBase底层使用HDFS，多副本存储可以保证出现故障时自动恢复，同时HBase内部也提供WAL和Replication 机制。云盘机型数据盘也会提供三副本支持，若集群中单点故障时，服务无影响。

扩展性强

HBase服务可扩展性强。HBase的扩展是横向的，集群需要扩展时，只需要扩容添加节点到现有集群即可。HBase服务支持动态的实时扩展，对现有服务无任何影响。

应用场景

海量数据实时存储

HBase服务支持海量全量数据的低成本存储、快速批量导入和实时访问，具备高效的增量及全量数据通道。HBase服务数据支持高压缩比，支持云数据盘或者HDD盘，成本低廉；支持高并发吞吐，集群可横向扩展至千万级QPS保证性能。

车联网场景

HBase服务可以存储车联网中的行驶轨迹、车辆状态、车联监控等重要数据，支持网租车、网约车、车辆检测、物流运输等场景功能。HBase服务使用灵活，支持动态列并且列可以有多个版本，可以自定义标签等数据信息。

用户画像场景

画像通常用一些标签来刻画自然人/物的特征，而每一个自然人/物所拥有的标签集合是不确定的，数据更新非常频繁，这类数据被广泛应用于市场决策、推荐以及广告系统中。HBase服务的稀疏矩阵模型，非常适合非结构化数据的存储；支持行的任意更新，并且利用HBase自身的多版本机制，支持保存数据的多个历史版本。