

内容大纲

- ▶ HBase简介
- ▶ HBase与传统数据库的对比
- ▶ HBase适用场景

HBase简介

- ▶ Apache HBase™ is the Hadoop database, a distributed, scalable, big data store.
- ▶ NoSQL(NoSQL = Not Only SQL), 意即“不仅仅是SQL”。
- ▶ Hadoop与MapReduce解决了海量数据的存储与批处理的问题。
- ▶ HBase解决了海量数据的随机、实时读写问题

HBase与传统数据库区别

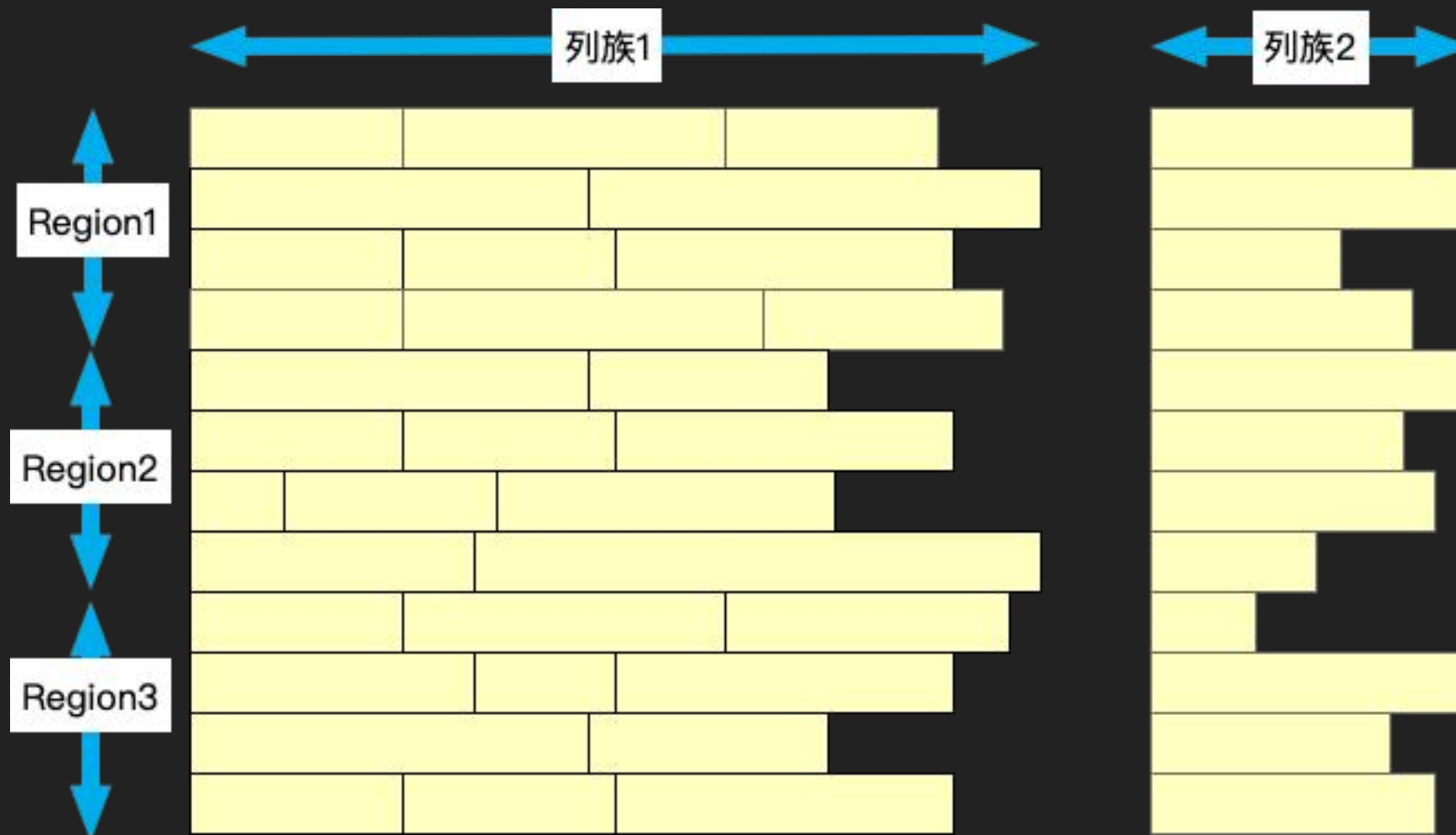
关系型数据库

模式预先定义、固定列、适合结构化数据存储

HBase

无模式、稀疏矩阵、适合非结构化数据存储

HBase与传统数据库区别



HBase 适用场景

- ▶ 云服务：短信、通话记录、照片等。
- ▶ 用户数据：订单、画像等。
- ▶ 时序数据：监控数据等（openTSDB）。
- ▶ 时空数据：气象数据等（geoHash）。
- ▶ 对象存储：新闻、网页等。
- ▶ 日志数据：淘宝鹰眼。

GeoHash

地点	经纬度	Geohash
鸟巢	116.402843,39.999375	wx4g8c9v
水立方	116.3967,39.99932	wx4g89tk
故宫	116.40382,39.918118	wx4g0ffe

