**时间序列模型**

时间序列数据库主要用于处理带时间标签（按照时间的顺序变化，即时间序列化）的数据，带时间标签的数据也称为时间序列数据。

每个时序点结构如下：

**timestamp:** 数据点的时间，表示数据发生的时间。

**metric:** 指标名，当前数据的标识，有些系统中也称为name。

**value:** 值，数据的数值，一般为double类型，如cpu使用率，访问量等数值，有些系统一个数据点只能有一个value，多个value就是多条时间序列。有些系统可以有多个value值，用不同的key表示。

**tag:** 附属属性。

**时间序列数据库OpenTSDB**

OpenTSDB运行在Hadoop和HBase上，其充分利用HBase的特性。通过独立的Time Series Demon(TSD)提供服务，所以它可以通过增减服务节点来轻松扩缩容。

特点：

1．Opentsdb是一个基于Hbase的时间序列数据库（新版也支持Cassandra）。

2．其基于Hbase的分布式列存储特性实现了数据高可用，高性能写的特性。受限于Hbase，存储空间较大，压缩不足。依赖整套HBase, ZooKeeper。

3．采用无模式的tagset数据结构(sys.cpu.user 1436333416 23 host=web01 user=10001)结构简单，多value查询不友好

4．HTTP-DSL查询

**时间序列数据库Influxdb**

Influxdb是业界比较流行的一个时间序列数据库，特别是在IOT和监控领域十分常见。其使用go语言开发，突出特点是性能。

特点：

1．高效的时间序列数据写入性能。自定义TSM引擎，快速数据写入和高效数据压缩。

2．无额外存储依赖。

3．简单，高性能的HTTP查询和写入API。

4．以插件方式支持许多不同协议的数据摄入，如：graphite，collectd，和openTSDB

5．SQL-like查询语言，简化查询和聚合操作。

6．索引Tags，支持快速有效的查询时间序列。

7．保留策略有效去除过期数据。

8．连续查询自动计算聚合数据，使频繁查询更有效。

Influxdb已经将分布式版本转为闭源。所以在分布式集群这块是一个弱点，需要自己实现。

**InfluxDB和OpenTSDB对比**：

与InfluxDB相比,OpenTSDB的主要差异有:

1. OpenTSDB的Metric需要预定义数据结构,而InfluxDB的Measurement是无Schema的。

2. OpenTSDB只支持秒级时间戳,InfluxDB支持更高精度。

3. OpenTSDB仅支持数值型Value,而InfluxDB的Field支持更丰富的类型。

4. InfluxDB有更加简单的部署方式,OpenTSDB较重且复杂。



综上：

InfluxDB更适合：性能高，精度高，数据量较小(亿级)，多查询，存储类型不止是数值型Value，需要复杂查询和高可扩展性的场景

OpenTSDB更适合：数据量较大，大数据服务基础，分布式集群需求，多数据写入和存储，频繁扩容缩容，已有Hadoop生态集成需求的大规模环境的场景。

4tv5doIUpAzc74a716kGcteBiYpImKgG8wE3t\_bp3lkUCq2Ucdnq0SS\_RkIKMUcnULN77QRPFYksHq88GJokbg==