

如何定义问题？

问题的根本来源应该是课题要求，系统设计、技术方案选择都是围绕课题系统需求展开的。对于我们的项目来说系统需求包括过程边缘到企业云的传输可靠性，以及云数据库存取性能。

系统需求：

- 传感器连接的边缘网关设备可能在任意时刻断网，无法连接到云服务器传输数据，需要最终使数据成功上传
- 监控数据存储要尽可能快
 - 每个传感器数值记录请求延迟越低越好，不能超过 50ms
 - 系统需要满足每秒约 10k 个请求的处理
- 数据查询操作要能查询到已经记录的数据，且查询操作应尽可能快
- （选做）由于监控数据的量很大，需要对数据进行压缩存储

可靠性需要明确的定义作为实现目标，而性能需求本身既是问题也是目标。

定义正确

10000个传感器 1 秒采集一次，传感器所连接的边缘网关设备可能在任意时刻断网，所采集的数据完整上传。

定义可靠性

我们认为要求数据的完整性是本监控系统的创新点，目前没有较为通用的解决方案，而结合软硬件技术赋予系统新的特征，是我们的创新。

可靠性能够达到的水平取决于软硬件能力，因此我们将在系统设计过程中根据需要在合理条件下定义实现可靠性所依赖的假设以降低软硬件负担。

硬件层面：

根据半导体制造企业Analog的产品说明，工业以太网产品包括嵌入式交换机和物理层设备，产品手册中对稳健型的描述如图所示，其它参数如耐高温，抗压能力也有说明，这里我们可以假设除发生重大灾难外，数据可以传输到本地的大型存储设备中。

稳健性

在恶劣的工作环境下，能够承受外部辐射和传导噪声源。选择已经过测试并符合CISPR 32和IEC 61000-4-2至IEC 6100-4-6等EMI/EMC标准的产品。这种稳健性有助于避免冗长的重新设计周期，以便后期在开发流程中获得认证。

软件层面：

多备份的大型存储设备容灾能力提升，但会降低性能，将基于备份的存储集群作为存储模块。我们假设生产环境断开互联网的MTTR小于一天，令大型存储设备容量大于一天内产生的数据量，则可靠性等于存储模块的Availability。

因此得出可靠性的假设和目标：

假设：使用工业以太网；不发生重大灾难；生产环境断网后修复时间小于一天；存储模块、云数据库可用；

目标：实现系统正确。

对于性能及测试的看法

首先，每增加一个中间件（比如本地存储）势必会引入测试的复杂性，增大从传感器采集到上传云数据库延迟。但在此过程中，就算通过工业以太网实现局域网传输的延时保证，互联网网络延迟也是不可控的。同时为了给延迟指标意义，系统需求中所要求的数值记录请求延迟应为从接受请求到真正放入数据库（能够被查到）的时间。

所以此处性能指标应为云端时序数据库的写入延时、吞吐量和查询延时。

在阿里云服务器上用Docker部署influxDB进行测试，测试环境如图：

CPU	内存	带宽	版本号
4核	16G	1Gbit/s	Ubuntu 4.8.4- 2ubuntu1~14.04.3

用influxDB官方给的测试工具（go语言），使用个人电脑向部署在腾讯云的单机版influxDB发送HTTP请求，不断提高每秒请求数量。

测得写入最大吞吐量为：600000 p/s，远超系统需求。

测得千次查询平均查询时间：1.64ms。

此外，本测试场景与工厂大型存储设备向云数据库发送请求场景类似。

因此得出性能的假设和目标：

假设：使用单机版influxDB，服务器运行正常

目标：数值记录请求延时不超过50ms；

系统满足每秒10k写入请求的处理；

查询请求平均延时不超过2ms；

对大型存储设备角色的看法

系统中大型存储设备实现消息队列的性能决定了它能否接受所有边缘设备传来的数据，关系到系统正确性。

性能测试将在接下来展开。