
数据采集

技术要点剖析

技术是一个快速迭代的循环。有趣的是，其对于数据采集系统的进化而言是呈现双向作用的：一方面，技术的涌现使得数据采集系统的各类指标和性能得到了质的提升，开拓了很大的可能性；而另一方面，一些传统的数据采集系统或架构很容易就会被直接替换或淘汰，这给工程师们又带来了严峻的考验与挑战。

本章节通过分析和总结多家数据采集设备提供商的技术趋势展望报告等文献，从五个方面为您阐述构建下一代数据采集系统所需掌握的技术要点。

- I/O 微创新
- 摩尔定律在数据记录仪的作用
- 光纤数据采集与传输
- 极致稳定系统
- 海量数据与数据管理

1 I/O 微创新

“除了满足最基本应用的反馈信号外，至少，现有系统中的传感器和I/O都必须能传递其它信息。 ”

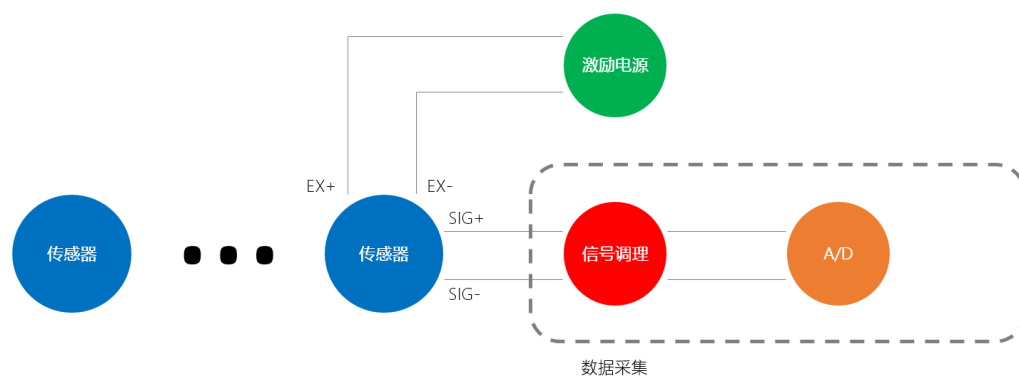
—— David Edeal, 技术产品总监, MTS传感器公司

“微创新”是一种脱胎于IT领域的方法论，它是一种从客户的角度出发，以提高用户体验的角度，不断地去做微小的改进，通过点滴的积累，最后达成巨大改善甚至颠覆。

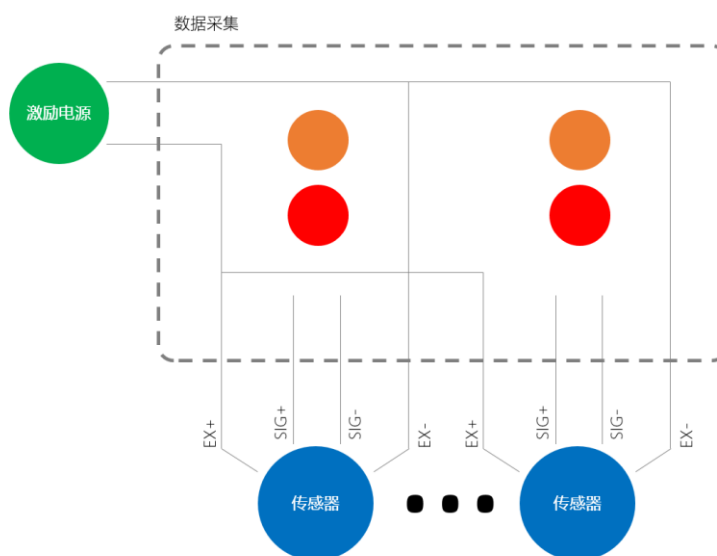
对于一个多通道数据采集系统来讲，其中传感器与数据采集设备接口的连接（即我们所说的I/O）这一环节往往都受到工程师及设备厂商的忽略。但是，随着几百甚至几千通道的大规模系统项目的不断出现，I/O环节中所发生的任何微小不便，都会得到指数级别的放大，而针对此环节的任何微小改善也可能让工程师的体验获得质的提升。由于这个原因，近期有一些具备先见之明的数据采集设备商已经开始将着眼点放在了I/O环节中，很多微小创意的产生都令人兴奋不已，但需要强调的是，这些努力其本质是为了提高工程师的使用体验，从而高效完成项目。

在一些通道数并不多的数据采集系统中，工程师们一般会采用如下图所示的接线拓扑：对于每一个传感器，在将其输出引线接入数据采集设备之外，还需并行地从激励电源设备（如精密电源）上并联上每个传感器。可以想象的是，这种在通道数少的情况下看不出任何瑕疵的方法，却完全满足不了大通

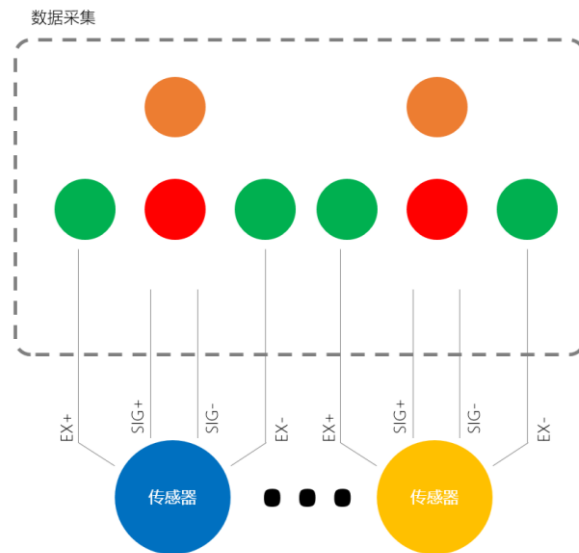
道数时所面临的接线压力。



正因为用户产生了改进这种方法的潜在需求，因此“微创新”开始了。下图是某些数据采集设备商做的一个巧妙的针对性改变：将外部激励电源的引线事先引入到数据采集设备内部，通过简单的路由，再统一引出到接口层。这样的一种小改进甚至都没有造成成本的上升，但却可以让工程师不用频繁在激励电源和数据采集设备之间切换接线，减少出错几率，大幅度方便了客户的接线体验。



但是，这种方法的一个局限性在于需要整个设备中所有通道的激励电压要求都相同，而这对于多通道数据采集应用而言是不切实际的，因为系统需要测量不同的物理量，因此就要选用不同的传感器。因此，设备又演变成下图所示的接口方式：数据采集设备内部自带了对单个通道的激励电压功能，通过总线，上位机可以控制设备对每一通道产生特定的激励电压，从而对每一个不同的传感器分别对待。

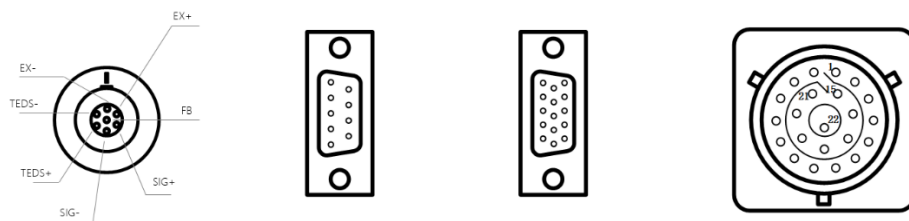


可以看到，即使只是一个看似微小的激励电压接线问题，通过层层改善，最终也能成功地将问题妥善解决。但是“微创新”的故事并没有结束，此时，工程师们又开始对一个已然成熟颇久的技术产生了老瓶装新酒的想法。

这个技术就是“TEDS”（Transducer Electronic Data Sheets），该协议是1994年由IEEE和NIST学会联合发起，目的是希望用标准接口解决传统传感器集成的问题。简单来说，TEDS的主要功能是将每个传感器都贴上电子标签，当时提出此标准的人们无法预料的是，这个技术现在对于大规模的数据采集系统

而言是极具意义的。下面的几个实际例子可以形象地说明这点：

- 就工程师的使用体验而言，当线缆插入设备的某个通道时，如果设备可以马上检测出并做相对应的指示（例如屏幕中某个通道的灯被点亮），那么在心理上会由于这一反馈而得到满足感。虽然这是一个看似微不足道的小体验，但是这种及时反馈的方式可以有效防止工程师没有插紧或者插错等不当操作。
- 在首次架设整套数据采集系统时，工程师们可以一次性地对每个通道写入传感器的各种信息，这样操作之后，在以后每次工程师启用系统接线时，只需要随意地将通道接入数据采集设备接头上即可，而不用做到必须和以前一一对应；这是因为数据采集设备已经自动解析了每一个插入通道的信息（包括传感器类型，激励要求，通道位置等），从而完全不影响最终的采集结果。
- 由于每个通道都已经预先写入了传感器的信息，因此在插入通道的同时，系统就可以自动地对该传感器进行个性化校准和标定，或者载入特殊热电偶的分度表等，整个过程完全自动化，而且出错率极小。



图：io7 I/O接口标准

对于数据采集设备供应商而言，即使是I/O接口这样一处小地方，也存在着如

此多的微创新点，这是一件令人鼓舞的事实。因此，作为多通道数据系统解决商，华穗科技（HUASUI）也针对I/O接口部分提出了io7这样的一个接口标准（上图），其包含了传感器信号、激励、TEDS以及激励电压回调等引线定义，从而将数据采集通道的接口方式得到了统一，最大程度上提高工程师接线的体验。

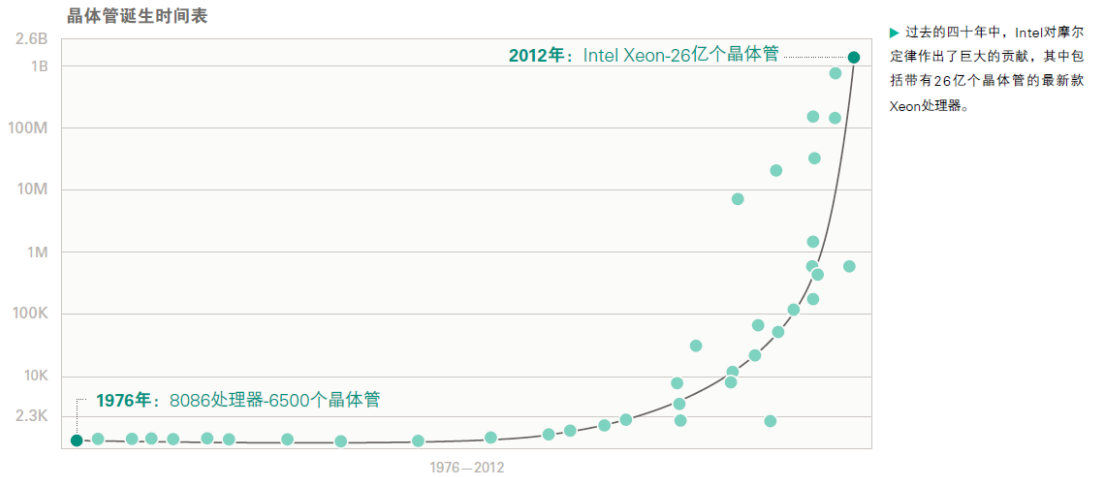
2 摩尔定律在数据记录仪的作用

“我们预计，在未来的数据记录系统需求中，不仅需要通过网络、服务器或PC采集数据，而且还需要提供智能，用于辅助决策过程。”

—— Mariano Kimbara, 高级趋势分析专家, 弗若斯特沙利文咨询公司

数据记录仪是一种从传感器获取测量结果，并将这些结果存储起来备用的电子仪器，其常用的测量包括温度、压力、电流、速度、张力、位移和其他一些物理量。在过去的二十年中，数据记录系统的智能变得更加分散，处理单元更接近传感器和信号。由于这一变化，远程数据采集系统和记录器可以更多地集成决策过程，而不是像过去一样，仅仅收集数据。

随着摩尔定律（下图）的不断发展，将创造性能更强、成本更低、功耗更低、尺寸更小的处理器，而未来的数据采集和记录系统将充分利用这一技术，变得更智能、且功能更丰富。



当前，在业界已经开始出现一些集成Intel、ARM、Xilinx等公司的最新芯片和IP的高性能记录系统，其系统基于异构计算架构设计，并为处理器结合了可编程的逻辑，从而满足更高端复杂的测控应用需求。当前市场上一些高性能数据记录仪的例子包括：

- NI CompactRIO
- HBM QuantumX CX22W
- 华穗科技 ioBox L Series
- Yokogawa WE7000

在很多情况下，数据记录的智能化应用已经开始扮演举足轻重的角色。例如，在结构监测中，传统的数据记录系统的功能仅仅是将每个数据点记录到磁盘上，即便被测的物理现象没有明显的变化，也会进行记录。这样就会产生兆字节甚至千兆字节数据，通过离线才能进行分析和筛选。

相比之下，随着系统变得越来越智能化，我们就能够在迅速适应一定条件下连续监视机器或结构。例如，系统可以缓慢地监测加速度或结构负载，并将连续将存储的数据覆盖写入磁盘，等待触发条件的满足。一旦条件得到满足，系统就能改变自身的行为，例如在短时间内，以更高的速率进行采样，同时确保触发前后的数据记录都没有丢失。此外，随着处理能力进一步的提高，信号处理功能就能嵌入在数据记录系统的多核处理器或可编程的逻辑中。这些更高级的系统能够在线分析数据，从而以更快地速度得出更有意义的数据结果，消除传输或存储大量数据的无畏浪费。

在使用具有更多智能和处理能力的的数据记录系统时，系统中运行的软件将是区分供应商技术优劣的主要指标。其中一种思路是，工程师可以使用可快速实现系统配置及得到测量结果的交钥匙工具，如：HBM公司的 Catman或者华穗科技（HUASUI）的ioPanel数据记录软件。交钥匙工具的不足之处是在功能上往往不够灵活；另外一种思路是，工程师和科学家们可以利用一个基于文本的编程工具（如Microsoft Visual Studio）或图形化编程方法（如NI LabVIEW系统设计软件），在这些系统中对处理器进行编程控制。编程工具为这些数据记录系统提供了最大限度的自定义功能，包括采用更丰富的信号处理方法，以及能够嵌入任何类型的智能算法。但相比于交钥匙工具，它们有一个陡峭的学习曲线。

各种应用和行业都需要在他们的数据记录系统中具有更多的智能功能。例如汽车、交通运输和电力公司等行业已经开始使用高性能的数据记录系统。在电力方面，我们知道，电网在发生着巨大的变化，公共设施建设部门正在投

入大量的资源使其更加智能。让电网变得更智能的方法之一就是集成更多的测量系统和设备。电能质量分析仪就是这样的一个设备。一个典型的电能质量分析仪可以采集和分析电网中的三相电压，以此来计算国际标准中定义的电压质量。电压质量由频率、电压电平变化、波动、三相电压的不平衡、谐波谱、总谐波失真以及信号电压的电平来描述。在这种应用中，需要进行大量的分析和高速测量，一个传统的数据记录系统无法提供所需要的处理能力。华穗科技（HUASUI）的工程师们使用LabVIEW和具有嵌入式处理器和FPGA的CompactRIO嵌入式的采集系统，开发了一个灵活的、高性能电能质量分析工作站 ioBox Station（见示意图）。在这个系统中，处理器可以完成诸如高级浮点运算处理、高速流盘以及网络连接的任务。系统中的FPGA允许系统中增添额外的处理单元，来执行自定义的I/O定时和同步任务，以及应用中所需的任何高速数字处理任务。



图： ioBox Station 数据记录工作站

总结而言，随着我们生活的世界变得越来越复杂，用于监测和记录未来的机械、基础设施、电网、和车辆中电子和物理数据的系统也需要不断发展。芯片和IP供应商正在努力优化处理组件的性能、功耗和成本。现在，还需要数据采集公司使用足够直观、灵活且智能的高性能数据记录系统来捕获任何类型的数据。使用更智能的数据记录系统，我们就能够从任何来源获得更智能的数据，并且能够对所建设的系统改善性能、质量，并对其进行维护。

下一代数据记录仪的特性与功能

存储	<ul style="list-style-type: none">● 集成了高速总线（如PCI Express）、速度更快的处理器可以加快数据流盘的速率● 闪存存储将继续增加密度和降低成本，可以在高端数据记录器的非易失性存储介质中，实现数百GB甚至TB的存储能力● 供应商将提供数据记录系统和云之间的无缝集成，使用户能够在任何地方访问数据
处理	<ul style="list-style-type: none">● 由于处理器上拥有更多的内存和更高的时钟周期，数据记录系统将集成更高级的信号处理，后处理过程对PC的依赖将变得越来越少● 更多的处理技术，如对应于高端数字信号处理的异构处理和可编程逻辑，将在数据记录系统中发挥各自所长
I/O 速率和定时	<ul style="list-style-type: none">● 记录系统将采用更快的I/O速率（从1 MHz到超过

	<p>1 GHz) , 因为它们的处理器更快 , 可以处理更多的数据</p> <ul style="list-style-type: none"> ● 充分利用异构计算技术和可编程逻辑 , 可以实现更多的自定义定时功能
应用软件	<ul style="list-style-type: none"> ● 软件工具 , 包括交钥匙工具和可编程工具 , 都可以直接在移动设备上运行 , 并可以进行自定义 ● 供应商将给软件用户带来如下的体验 : 易于使用的交钥匙系统 , 而且能够快速自定义系统 , 以满足每个用户独特的应用需求
可视化	<ul style="list-style-type: none"> ● 与移动设备 (如智能手机和平板电脑) 的集成将变得很常见 ● 技术领先的供应商将使用户能够在触摸屏上使用基于Web的软件快速创建自定义的用户界面 ● 按钮和旋钮将消失 , 数据记录系统中的每一个界面 , 无论是使用智能设备还是一个集成的显示器 , 都将实现多点触控功能
连接能力	<ul style="list-style-type: none"> ● 随着联网功能的增加和处理器能力的增强 , 数据记录系统将有更多的连接和同步选项 , 如Wi-Fi、蜂窝 (GSM、3G、4G等) 、千兆以太网、IEEE 1588等
尺寸和坚固性	<ul style="list-style-type: none"> ● 低功耗处理技术 , 如英特尔的Atom和ARM处理器 , 将使高性能记录器尺寸更小 , 并具工作温度

范围更大

- ◆ 说明：基于摩尔定律和先进的处理技术的高性能数据记录系统未来可能拥有的能力。

3 光纤数据采集与传输

“将光纤技术和电气测量系统相结合，这与将两者分开的传统方法截然不同，如是的融合势必解决更多的实际问题。”

—— Alfredo Cigada，教授，米兰理工大学

诚然，各种商业可用技术的高速发展对测试测量行业产生了巨大的影响，光纤技术也是其中之一。

前香港中文大学校长高锟在20世纪60年代提出了采用光路来实现数据通讯的方式，自此，从1977年光纤系统首次商用安装以来，各通信公司就已经开始使用光纤链路替代旧的铜线系统。从传统的双绞线、电缆+双绞线阶段，目前对于高速、大数据量以及超远距离的数据布线而言，基本已经进入了光纤这个阶段。其主要的原因，也是显而易见的：

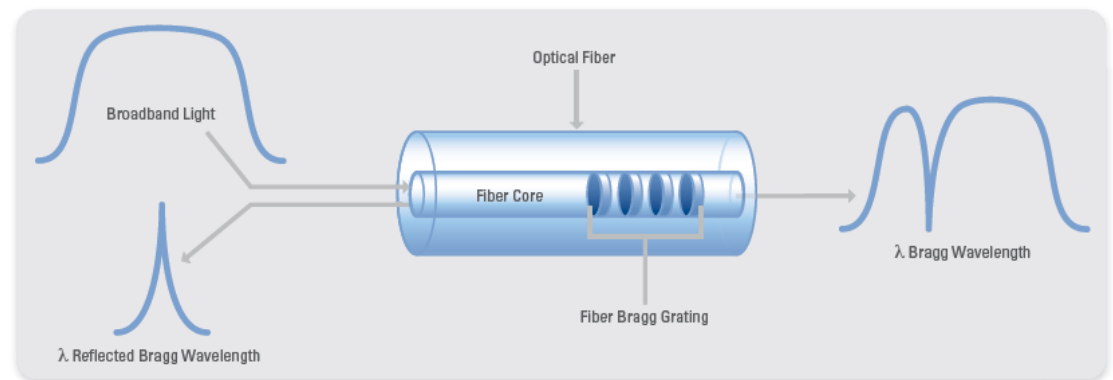
- 就传输速率而言，相对于铜线每秒1.54MHZ的速率，光纤网络的运行速率达到了每秒2.5GB；
- 从带宽来看，光纤具有较大的信息容量；
- 极低的维护成本；
- 抗干扰性强，安全性能高

与此同时，在测试测量领域，特别是大通道数的数据采集系统中，光纤技术

的爆发态势直接开拓了传统测量方法所无法企及的应用。

- 光纤传感

数十年来，电学传感器是将物理和机械现象转化为数据采集系统能够测量的模拟信号的标准机制。尽管它们无处不在，电学传感器仍然有很多局限性，电磁干扰（EMI）、易爆性、对噪声的高敏感以及布线麻烦等缺陷使它们在很多应用中不够安全，不适应实际情况而导致无法使用。



图：FBGs 传感器原理

相比之下，光学传感器通过使用光学的标准的光纤取代电学的铜线，消除了很多困扰电学传感器的挑战。其中最为广泛采用的光学传感器技术就是基于光纤布拉格光栅(FBG)技术。FBGs是最本质的点传感器，它在光纤芯内具有周期性的折射率变化。FBG，一般5毫米长，有选择地反射以布拉格波长为中心的很窄的频带，同时透射所有其它的波长。反射的布拉格波长作为温度与应变的函数而变化，这使利用FBGs作为传感器成为可能。工程师们可以将多个FBG传感器沿一根光纤呈环状排布，因为每个FBG都被制造为独一无二的标称波长，所以它能够运行在唯一的光学光谱频带。与电学传感系统不同，每个光学通路能够测量很多FBG传感器，极大地减少了测量系统的尺寸，重量以及复杂性。此外，因为FBG传感器调节波长而不是光强，您可以将其部署很

远的距离——长达10千米——并且不会对测量精度造成任何影响。

目前已经有越来越多的设备供应商开始专注于光纤传感的研究方向，例如微光光学公司（MicronOptics）一直致力于将光纤传感技术应用于各类结构健康监测项目中，在美国Chulitna大桥健康监测项目中，在光纤上成功布置了几百个应变传感器；而NI公司也推出光学传感解调器，实现对FBG应变和温度传感器的同步采样，将光纤传感引入PXI这个测试测量标准平台之中。

- 数据传输

Google公司在2012年推出了“Google Fiber”服务，利用光纤传输技术，实现高达1G的网速，从本质上即将改变人类的生活；同样，在目前的各类大型分布式数据采集应用中，光纤也已经逐渐成为主流的传输方式。凭借其独特的技术优势，使用光纤传输可以有效解决以下实际工程问题：

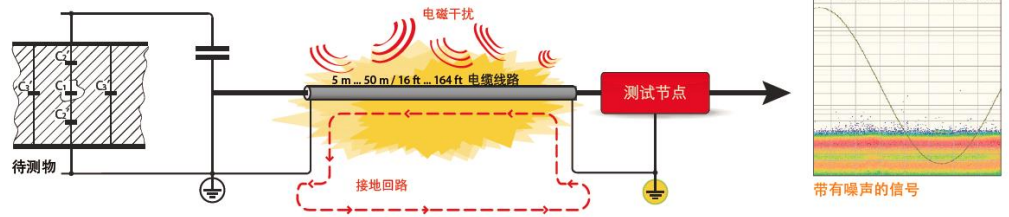
- 传输距离

在分布式的数据采集项目中，如果采用的是USB或者以太网总线，那么其数据的传输距离都会受到一定的限制条件，例如以太网的理论最大传输距离是100米（实际是达不到的）。因此对于需要传输距离较长的应用来说，采用光纤就可以很好地解决这个问题，其理论传输距离可高达20公里。

- 环境对信号的干扰

实际现场环境对信号的影响是不容忽视的因素之一，例如在电力监测等项目现场，都会具有电磁干扰等因素，在用传统的电缆传输电信号的时候，会产生一定的干扰；相比电路，采用光路的传输则完全不受电的干扰和影响（如下图）。

传统数据传输方式



光纤数据传输方式



图： 光纤数据传输方式的优势

■ 高压隔离与安全性

在电力、能源或其他有高电压、电流的测试环境下，线缆连接的安全性是需要考虑的重要因素。不同于网线和其他金属电缆，光纤本身绝缘，不具有导电性，即使穿越布满高电压设备的环境，也不存在被击穿的风险，不会对光纤两端的人员和设备带来危险。这是其他线缆通信方式所不具备的优点。

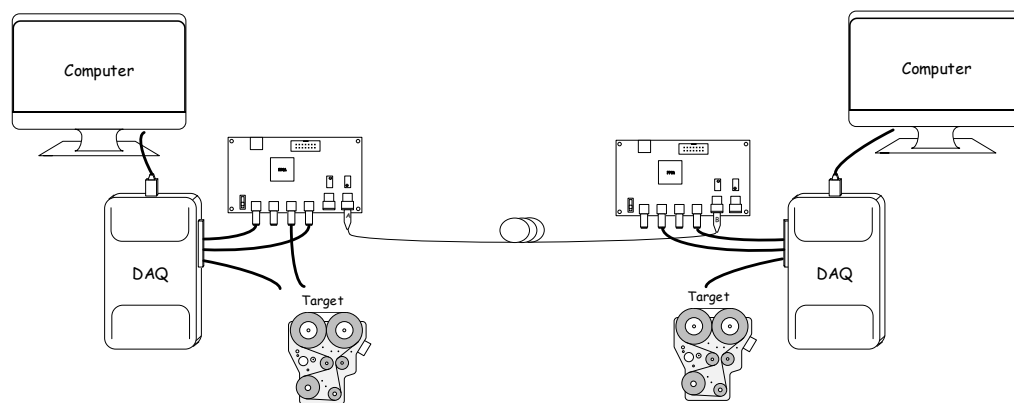
● 同步

时钟同步是多节点或者模块化的数据采集系统的基本问题，业界最为常用的解决方法是在多个采集设备中用线缆共享时钟信号。但由于线缆的传输距离限制以及传输延时使得这一传统技术在长距离、高精度测试系统中无法使用（根据CAT-5e标准，线缆最大传输距离应小于100米，每米最大传输延时为4.98ns）。

基于GPS的同步授时是一个不错的选择，其时钟源来自GPS卫星信号，技术较为成熟，目前这项技术也已经广泛用于以太网、电视传播等领域，现在也开始在测试测量领域崭露头角。但是这种方式对GPS依赖性过高，无论从系统

的长期稳定性以及信息安全性等角度，都不是一劳永逸的解决之道。

选用光纤来解决时间同步的问题，在世界范围内已经有了一些技术的成果，例如在粒子冲撞机等大物理领域，最新诞生出叫作White Rabbit的同步技术，虽然目前还没有完全成熟，但是理论上可以实现在10km空间尺度中纳秒级的精确度；华穗科技（HUASUI）也将光纤技术应用于时间同步授时，并成功开发出光纤同步模块。该模块无需过多配置即可使用，其工作形式类似于一根普通的线缆，线缆输入端和输入端的采样时钟能保证亚微秒级别的同步精度，线缆输入端连接华穗光纤同步模块，经过长距离光纤和连接到另一个采集节点的光纤同步模块，同步模块采用基于FPGA的光纤同步技术，可以保证公里级别的高精度时钟同步。



图：光纤同步模块工作原理

技术的革新势必带来应用的革命，不可否认的是，光纤技术在测试测量领域将逐渐扮演日益重要的角色，工程师们如何把握这样的技术趋势，从而提高系统的构建效率和稳定度，这是一个值得讨论的话题。

4 极致稳定系统

“测控系统的挑战不仅在于更大的数据吞吐，更要求系统的低功耗、稳定性以及可部署性……”

—— 《全球数据采集技术发展趋势报告》

近来，一种观念正在悄然变化。

那就是工程师们对数据采集系统的稳定性需求逐渐达到了最高峰值。纵观整个测试测量行业，虽然设备成本依旧是需考虑的因素之一，但是我们可以看到，系统的总体可靠性、长时间工作稳定性以及对严酷环境的适应性逐渐成为系统选择和决策的关键考察指标；基本上，对于较多通道数的数据采集系统而言，几乎已经没有低端数据采集设备商的身影。

产生如此现象并非空穴来风。首先，随着技术的发展，数据采集设备采样速率和通道数都获得了巨大的提升，而这种提升所给整个系统带来的压力并非线性而是指数般的关系，因此整体系统的出错率也会成倍增大（试想一下对8个通道和对80个通道进行布线的区别），以前可以忽略不计的小概率事件现在几乎已成必然；其次，数据采集系统所在的工作环境相比以前也复杂了很多，以前的系统基本上都是静止而独立的放置在实验室中的机柜中，通过较长的电缆连接到被测物上（真正的试验地点）的各种传感器；而现在随着分布式的架构以及将数据采集设备尽可能接近传感器的布线原则逐渐主流，数据采集系统会放置在无人的室外区域（例如热带雨林环境监测）或者直接在被测物所在的试验点工作（例如会产生高温、腐蚀气体的航空发动机试验台架）；最后，工程师们对数据所具有的价值也得到了意识上的更新，在一次

成本异常昂贵的试验（例如汽车破坏性碰撞安全试验）中如果数据采集系统出现些许的问题（不必要的延迟或者漏点），那么极有可能会造成本次试验的完全失效。

虽然上述三点现状成困于数据采集系统应用领域的不断开拓，这是好事，但是讽刺意味的是，由于事物的两面性，这反过来又向工程师们提出了更巨大的挑战。因此，针对这一发展趋势，一系列的对新一代数据采集系统稳定性方面的要求正在无形之中融合；借鉴于电视行业中采用“极清”来表示“高清”的下一个发展阶段，这里我们将这种理想系统暂命名为“极致稳定系统”。

虽然目前尚未形成明晰的标准，但我们可以确信的是，以下几个因素会是实现“极致稳定系统”的主要考虑方向：

- 防护性

承载实际数据采集工作的各种硬件设备是整个多通道数据采集系统中的核心，如果它们在某些工作环境的突然变化（例如温度过高）下发生失效，那么势必会造成整个系统的瘫痪。因此，努力提高这些硬件设备的防护性是很容易想到的一点。

所谓防护，对于电子设备来说，主要就是通常意义上的“三防”，即防尘、防水与防振。对于前两者而言，国际较为通用的是IP（Ingress Protection）等级，其由两个数字所组成，第一个数字表示防尘；第二个数字由表示防水，数字越大表示其防护功能越好。例如通过IP67标准的产品表示其可以完全防止粉尘进入且可防止浸水时水的侵入。对于防振，主要会看的是设备在工作

状态时能够承受多少量级的往复运动和冲击，此外还有从多少米下做的特定跌落测试等等。



图：ioBox L 数据记录仪

观察手机的发展轨迹，不难发现，以前的人们可能都不会想到，现在触摸屏已经是手机不可或缺的组件，而当时流行的翻盖样式的手机现在已经杳无踪迹。同样，在不久之前，具备三防特性会被认为是设备高端的标志，一般这类产品只会用在军用级别的领域，且价格极其昂贵；而现在，具备IP等级的数据采集设备已全面进入工业领域，价格适中，甚至某些数据采集设备商所提供的产品系列都是默认三防，例如华穗科技（HUASUI）推出的ioBox L数据记录仪（上图）无论是结构还是接插件就能完全做到IP67的防护等级，从而非常适用于特种运输等应用领域。

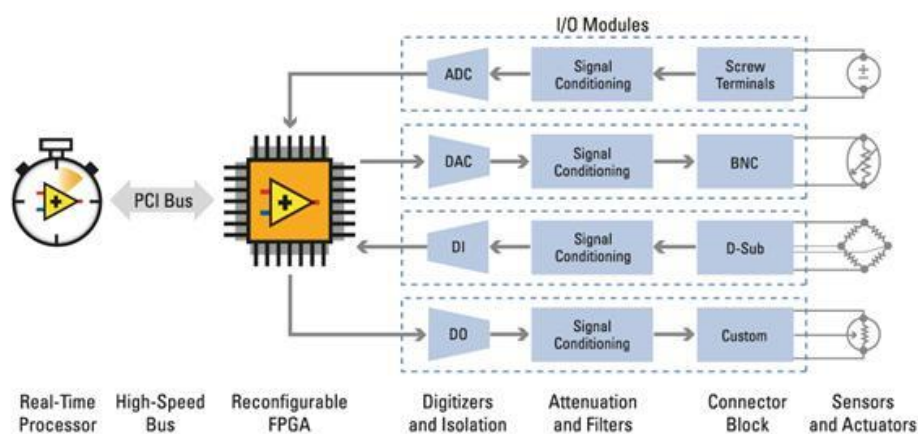
- 确定性

很多多通道数据采集应用都需要系统具备高度的确定性，这里我们所要讨论

的范畴其实已经和硬件无关，而是软件层面的问题。

所谓确定性，就是指在软件中是否能够始终如一地在已知时间长度内完成一项任务。例如，如果某设备多次或循环执行相同任务时执行时间的变化或抖动范围越大，或者受外界交互影响（例如鼠标操作或者新打开一个文件夹等），那么其确定性就越差，从这个角度来看，类似Windows 7或者Mac OSX等一般的操作系统其实是完全不具备确定性可言，而实时操作系统（例如VxWorks、uC-OS/2等）可以通过非常精确地对您应用程序任务进行时序控制来实现确定性，从而确保关键应用程序执行始终如一，耗时固定。因此，在数据采集系统架构中，我们一般会将实际采集操作这些对确定性要求很高的部分放在下位机的实时系统中来完成，而将系统配置、急停、显示等功能放在上位机来做。

当然，除了使用实时系统之外，有些厂商为工程师们想得更远。NI公司创造性地将实时系统与FPGA两者有效的融合，提出了可重置I/O，即RIO的系统架构，如下图所示：



图：可重置I/O架构

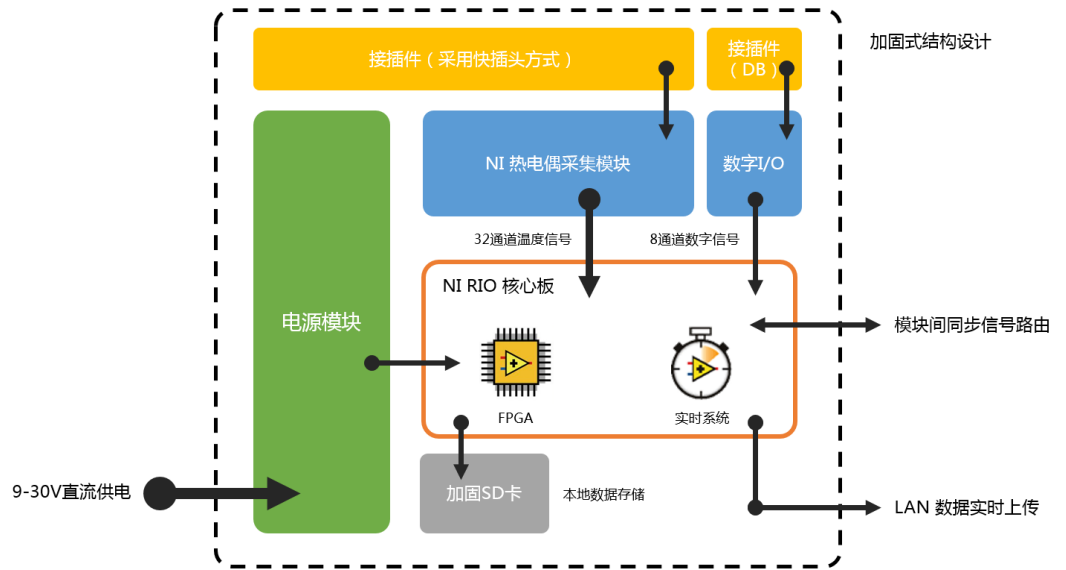
由于篇幅原因，这里不做过于细节的描述，这种架构的关键点在于引入了FPGA，它可以帮助处理器分担密集型的任务，具备极高的吞吐量，此外它的确定性决策和操作可以直接与硬件I/O连接，从而实现更多的高性能访问，例如互锁、条件触发、时序、同步等，控制速率最快可以达到ns级别。相比单一的实时操作系统，这样的架构将有更多的可塑性和想象空间。

- 数据冗余

作为一个稳定工作的多通道数据采集系统，最关键的就是其采集数据不出差错，不漏点，不多点等，特别是长期数据采集的情况下（例如连续3-4个月的长期监测），要满足这样的严苛条件，其实是很困难的。这个时候，采用冗余的方法是一个简单有效的思路。

冗余这个理念并不新鲜，通常指通过多重备份来增加系统的可靠性。例如在工程项目中，有时工程师会为某些关键设备提供UPS备用电源，这样当突然断电的情况出现时，可以让设备继续正常工作一段时间，从而不影响整个系统，这就是冗余最简单的应用。

下图是华穗科技（HUASUI）ioBox Rack 机柜式数据采集设备的系统构架图，其中就专门为数据的安全做了冗余设计。整体系统通过NI模拟输入采集模块对传感器输出信号进行调理，通过LAN口实时传输到远程上位机中，同时数据也在系统本地储存到加固SD卡中，起到双重数据安全缓冲作用。这样，即使上位机出现任何突发性问题，数据还是会存储在内部SD卡中，因此也不会造成整个试验的失败。



图：数据冗余设计

综上所述，考虑整体系统的稳定性，是工程师在构建多通道数据采集系统之前就必须想到的。我们相信，“极致稳定系统”在不久的将来必定会发展成为一种规范和标准，而以上所讨论的防护、确定性以及冗余设计等都极有可能成为各类测试设备的“标准配置”。

5 海量数据与数据管理

“当前的主要差距不再是关于谁可以采集更多的数据；而是关于谁可以快速理解所采集的数据。”

—— Matt Wood, 首席数据科学家, 亚马逊公司

过去，硬件的采样速率受到模数(A/D)转换速度的限制，在物理上限制了能够采集到的数据的多少。但时至今日，硬件供应商已经加快了数据采集速

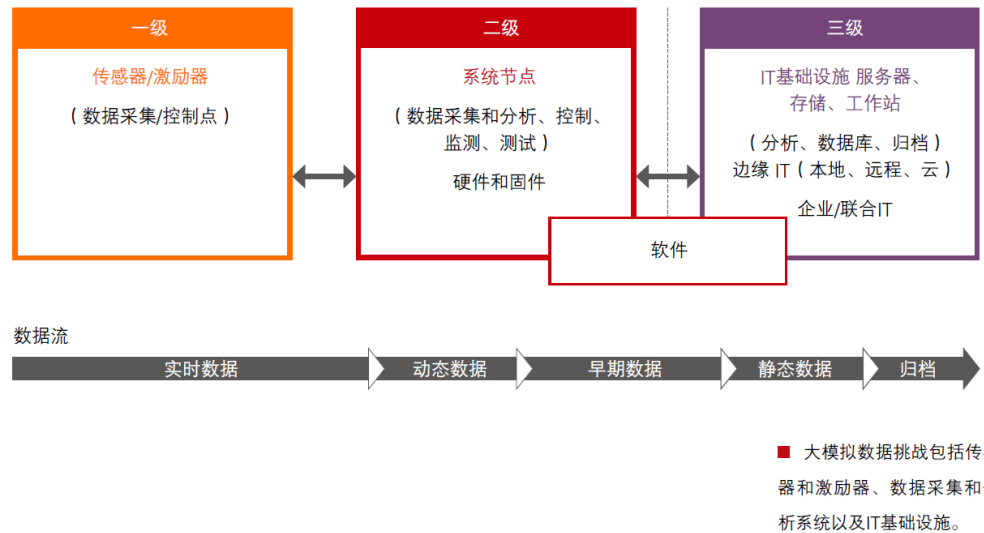
率，使工程师和科学家迅速地突破了速率和分辨率的障碍，他们最终引发了新一轮的数据膨胀浪潮。简单地说，在采集应用中，硬件不再是限制因素；对所采集到的数据如何进行有效的管理才是未来的挑战。

计算技术的进步——包括微处理速度的提高和硬盘储存容量的增加，以及硬件和软件成本的降低——极快地引发了数据的大爆炸。特别是在测量应用中，工程师和科学家们可以在一天中的每秒采集大量的数据。在欧洲核子研究中心（CERN）的大型强子对撞机运行的实验中，仪器每秒可以生成40 TB的数据。一架波音喷气发动机在工作过程中，系统每30分钟将会生成10TB的运行信息。对于跨越大西洋的单程飞行过程中，一个四引擎的大型喷气式客机可以生成640 TB的数据。算一算每天有超过25,000次的航班，您就会理解了为什么会存在如此大量的数据。这就是“海量数据”的由来。

技术研究公司IDC最近进行了一项针对数字数据的研究，其中包括世界范围内的测量文件、视频文件和音乐文件等。这项研究表明，可获取的数据量每两年就会翻一番。这个事实类似于电子领域最著名的定律：摩尔定律。如果数字数据的产生速度继续模仿摩尔定律，那么一个公司或组织能否取得成功的关键将在于它能否快速地将所采集的数据转变为有用的结论。

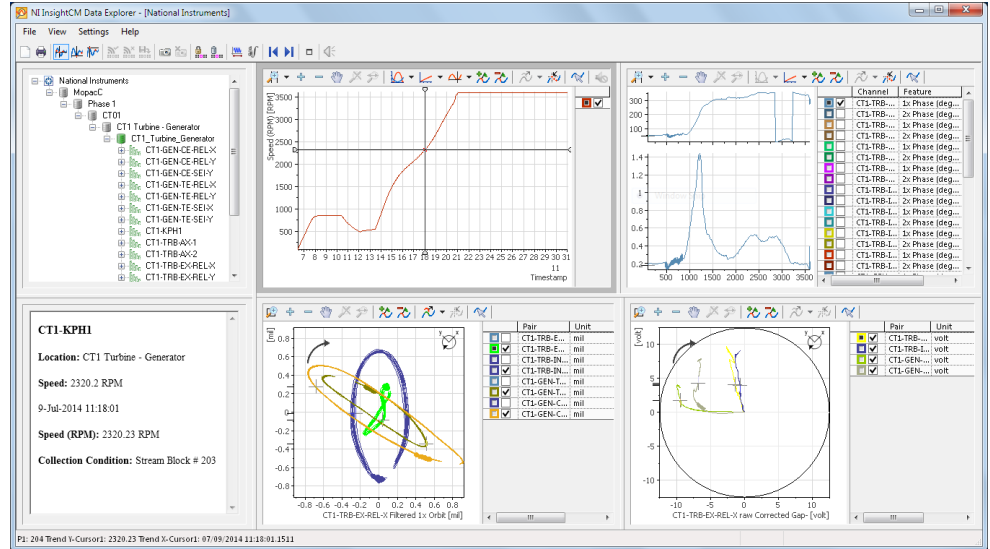
海量数据的出现为数据分析、搜索、整合、报告以及系统维护带来了新的挑战，它们必须能够满足与数据的指数型增长速度保持同步。数据的来源很多。但是，其中工程师和科学家最感兴趣的是真实世界生成的模拟数据。它是从振动、射频信号、温度、压力、声音、图像、光、磁和电压等测量中收集而来。通常数据的特点由四个V组成——数量(volume)、种类(variety)、速

度(velocity)和价值(value)。但另一个“V”——可视化(visibility)，正逐渐成为一个关键的决定性特性。也就是说，全球企业一个日益增长的需求就是访问不同地区的商业、工程和科学数据。这就需要云等互联信息技术(IT)系统与数据采集系统(DAQ)紧密连接，这是目前数据采集领域所引起的革命性的理念突破。



具体来说，工程师正在寻找如上图所示的三级解决方案体系架构，以构建一个统一的集成解决方案，从而改善前端传感器的实时数据捕获和后端IT基础设施的数据分析。数据流始于第一级的传感器，被第二级的系统节点捕获。这些节点执行初始的实时、动态和早期数据分析。被认为重要的信息通过软硬件“边缘”传输至传统IT设备。第3级的IT基础设施（服务器、存储和网络设备）负责管理、整理和深入分析早期数据或静态数据。最后，对数据进行存档以备后用。通过数据流的不同阶段，不断发展的海量数据分析领域正在为人们提供前所未有的有用信息。例如，NI公司与Duke Energy能源公司共同开发的InsightCM（下图）状态监控系统，将硬件监控转变为软件和网络

级，实现了总共3万多个传感器，200多个节点，60个不同地域的智能电网监控。



图：NI InsightCM Enterprise 界面

可以预见，历史最久、速度最快、数据量最大的海量模拟数据正拥有着日益重大的科学、工程和商业意义。为了挖掘这一巨大的资源，开发人员正在转向基于工具和平台的解决方案，这些工具和平台应能够相互集成，而且能够与其他合作伙伴提供的工具和平台相集成。由于这个三级大模拟数据解决方案可解决科研、产品测试以及机器状态和资产监测等关键应用领域的难题，其需求正在不断增长。