

面向应用的系统用例规约

一、项目背景

《中国制造2025》以智能制造为突破口和主攻方向。所谓智能，就是能独立管理制造过程。智能工厂是利用各种现代化的技术，实现工厂的办公、管理及生产自动化，达到加强及规范企业管理、减少工作失误、堵塞各种漏洞、提高工作效率、进行安全生产、提供决策参考、加强外界联系、拓宽国际市场的目的。

传感器（监控数据）是智能工厂的灵魂。智能工厂包括生产制造、物流、服务三个内部环节。传感器监控数据在各环节均发挥着重要的作用。

在生产制造环节，要求（1）透明的生产过程，综合的追溯功能；（2）保障多工种实时协作的安全，同时实现工业自动化控制；（3）产品质量的优化提高。在物流运输环节，要求（1）支撑场内外物流运输系统，实现端到端透明；（2）分析产业链供求关系，物流过程行为统计及优化。在服务环节，通过传感器的监控数据实现远程维护。

智能工厂的以上需求全部依赖监控数据的存取来实现，因此，在智能工厂设计中，监控数据存储系统位于设备的自动化层和生产制造的管理系统之间，作为两者联系的桥梁。前者着眼于数据采集实时性应用物联网自动化实现监控、报警和反馈控制，后者要求历史数据的快速回放来完成统计分析、工艺优化、设备预维护等管理行为。

所以，对于监控数据存储的需求始终存在，在电力、石油、化工等领域，企业搭建生产实时监控系统的方式是配备专门的实时数据库（如PI）。而现在，随着硬件成本降低，传感器需求量增多，同时云服务越来越便宜，监控数据存储系统上云成为大势所趋。

作为通用的第三方服务，允许不同用户根据自身场景需要进行个性化定制、形成柔性化产线是重要的。为此，我们应该充分调研不同行业对监控数据的具体操作，着重解决通用问题，同时对特殊行业的需求加以改进。

二、智能工厂需求细化

上文介绍了智能工厂三个内部环节相较于传统工厂的更高要求和特性，总的来说就是更高级的自动化和更全面的生产管理，分别体现了监控数据的实时性和完整性。下面我们将分析智能工厂的功能及实现方式，明确云数据库在系统中的角色，从而具体化对于监控数据的操作及要求。

实时性：DCS分布式控制系统集成云时序数据库

(1) 真实生产模型及实时性要求

智能工厂监控数据上云在实时性方面的应用包括实现安全保障、允许管理人员实时操作、以及利用工业物联网自动化实现自动控制。为了探究以上三个应用对实时性的要求（如数据采集和控制的执行间隔），我们调研了火电厂的生产规模、工艺要求及解决方案。

生产规模	一个火力发电厂通常配备4到8台汽轮发电机组，初始建造时2台同时建造，一台35万千瓦汽轮发电机组有1700个测量点和620个操作项目。
工艺要求	针对不同信号的控制任务要求不同，模拟量信号，执行周期最大不超过250ms，数字信号，执行周期最大不超过100ms。针对模拟和顺序控制回路这些对处理有快速要求的回路，处理执行能力要达到125ms/1次和50ms/1次。
解决方案	DCS分布式控制系统。

在工业自动化控制领域，自动化控制系统经历了从集中型控制系统CCS到分布式控制系统DCS，再到现场总线控制系统FCS的演变，每次技术进步，都在向高可靠、低时延的特性迈进。目前普遍使用的是分布式控制系统，它具有多年的实践经验积累，并且能够集成更加现代化的现场总线控制系统，因此学习DCS对于了解实时数据在工业控制中的角色至关重要。

(2) DCS系统功能要求

(1) 数据采集和存储功能：模拟输入信号的扫描时间，数字信号的扫描时间，历史数据存储（包括调整趋势、历史趋势）的种类、点数和周期，输入信号的处理功能。

(2) 控制功能：具备完整的监测、调节和顺序控制功能及设备运行监视、联锁保护等功能，反馈控制功能的最大能力（输入输出的种类、数量、逻辑运算能力、顺序步骤，执行周期），过程控制语言能力（如ST、FORTRAN, BASIC, C或其他的控制语言），先进控制软件的要求。

(3) 显示功能：CRT尺寸、颜色，画面刷新时间，画面种类与数量。显示画面包括总貌画面、流程图画面、控制分组画面、调整画面、趋势画面、过程报警画面、系统报警画面、系统状态画面、操作日志画面、历史查询画面等，并具有窗口（多个、拉伸、重叠）显示功能，汉字功能以及触屏功能。

(4) 报警功能：模拟输入/输出信号报警（绝对报警、变化率报警），数字输入/输出信号报警，仪表停用（校验、停扫描）报警，系统部件（卡件、网络）故障报警，相邻报警时间分辨率，报警管理功能（如历史查询过滤、报警组设置、报警点组的优先级、报警抑制功能等）。

(5) 报表和屏幕拷贝功能：即时报表功能，定时报表功能（包括时报、班报、日报、月报、年报），报警汇总记录，操作记录报表，操作或参数修改打印，打印机屏幕拷贝功能。

(3) DCS系统架构及原理

分布式控制系统采用分层的方法，在结构上分为过程级、操作级和管理级。过程级主要由过程控制站、I/O单元和现场仪表组成，是系统控制功能的主要实施部分。操作级包括：操作员站和工程师站，完成系统的操作和组态。管理级主要是指工厂管理信息系统（MIS系统）。

过程级的基本单元为过程控制单元，又称现场控制单元，本质与PLC（可编程序控制器）类似，都接收传感器的I/O输入，进行运算后将I/O输出传给执行器，同时能够将输入的实时信号通过工业网络传给操作站。现场控制单元一般远离控制中心，安装在靠近现场的地方，其高度模块化结构可以根据过程监测和控制的需要配置成由几个监控点到数百个监控点的规模不等的过程控制单元。

操作站用来显示并记录来自各控制单元的过程数据，是人与生产过程信息交互的操作接口。典型的操作站包括主机系统、显示设备、键盘输入设备、信息存储设备和打印输出设备等，主要实现强大的显示功能（如模拟参数显示、系统状态显示、多种画面显示等等）、报警功能、操作功能、报表打印功能、组态和编程功能等。

综上，系统功能的自动保护、自动调节、自动控制由过程控制单元完成，遥控（开环控制或借助模拟式调节器的闭环控制）和程序控制由操作站完成。通常要求操作站操作的执行延时不超过 1 s。

(4) DCS系统的实时数据存储与用途

DCS 数据存储系统主要功能是通过保存 DCS 运行过程中产生的大量的实时数据，为整个系统的运行状况和故障预报提供数据来源和分析依据，在整个 DCS 中处于操作级层次，因此数据存储系统的主要使用者和应用场景也为置于这些操作站、工程师站或者服务器内的应用。

根据DL/T1083-2008《火力发电厂分散控制系统技术条件》标准，历史数据要求存储 30 天。

(5) 使用/集成时序数据库

根据以上对DCS的介绍，可以看出DCS重点在于控制，大部分控制执行过程不关注历史数据，只关注当前输入的数字/模拟量。历史数据在实时强的控制中参与较少，主要作用在于参与操作站的显示、报警、趋势分析、报表打印以及对实时性要求较低的生产控制（更综合的数据输入，较长的执行周期）。

因此，在DCS中集成时序数据库，与时序数据库本身写多读少、无事务要求、海量高并发持续写入、基于时间区间聚合分析以及基于时间区间快速查询的假设一致。联系其它时序数据库的典型应用，如采集设备状态参数来进行分析、调度、生产管理、故障监测及安全报警，操作站所需实现的趋势分析、报警管理和实施操作与之类似。

在DCS中集成时序数据库，将过程控制层数据上传云端，令操作层和管理层成为云数据库的上层应用，在保障实时控制的同时延长历史数据的存储时间，从而实现全系统综合信息管理。

(6) 关于实时性操作的结论

目前，过程控制层与云服务的结合仍比较困难，将云数据库作为闭环控制的节点，会引入可靠性低、延时不确定的商用网络，因此现场的过程控制层仍是必要的。在此基础上，考虑到操作站只是现场工业网络中的极少数节点（一个火电厂通常有4个操作员站，1个工程师站），实时数据上云是通过增加成本投入来获得更强的信息管理能力。

其它涉及实时数据的系统如PLC+组态软件建立的控制系统、数据采集系统（SCADA）等也可借助云时序数据库实现，且对于实时数据的操作要求较低，在此不讨论。

有关实时性操作的具体功能要求参照DCS系统功能要求中由操作站完成的相关功能，执行具体参数如下表所示。

用例	频率	并发量	延时	数据操作
实时数据采集（batch）	1次/秒	1000	< 1s	批写入，batchsize取1000
实时数据采集（不batch）	执行间隔 0.1s ~ 0.2s	100	< 50ms	写入
显示功能	刷新间隔 1s ~ 10s	< 400	< 1s	基于时间区间查询，过滤，聚合，join，持续查询
绝对值报警	1次/秒	100	< 1s	查询某传感器最新数据
变化率报警	执行间隔 10s ~ 1min	100	< 1s	基于时间区间查询，过滤，聚合
仪表停用或网络故障报警	执行间隔 10s ~ 1min	100	< 1s	基于时间区间查询，过滤

用例	频率	并发量	延时	数据操作
即时报表	20次/天	无	< 1min	基于时间区间查询，过滤，聚合，join
定时报表	1次/天	无	< 1min	基于时间区间查询，过滤，聚合，join

(7) Batch可行性思考

现场控制单元可根据过程监测和控制的配置成由几个监控点到数百个监控点的规模不等的过程控制单元。现场控制单元与操作站通过工业网络进行数据传输。

操作站对实时数据的操作实时性要求较低，显示、报警、趋势分析、实时控制的延时要求不超过 1 s，执行周期大于或远大于。因此从应用的角度看，实时数据并非实时可见。

综合上述两点，过程控制单元可以适当应用batch策略，降低overhead，减少写入的并发数。

完整性：历史数据记录的价值

对于数据完整性的需求及相关操作根据部分企业智能工厂设计方案给出。

以智慧环保物联网监控系统建设项目为例，除实时监控外，完整的监控数据（1）提供高可见性，提高政府行政效率、满足公众知情权；（2）管理决策上，有利于对环境风险作出科学评价，通过模型和评价体系解决环境管理问题；（3）提供环境预报预测的能力。

在流水工厂的生产环境中，提供高可见性，实现透明的生产过程，实现产品跟踪、端到端透明；进行生产过程分析优化，识别导致生产效率低的因素，进行资源配置；进行设备的预测性维护，利用统计知识，结合人工智能方法，确定可能发生的故障并及时维修或更换。

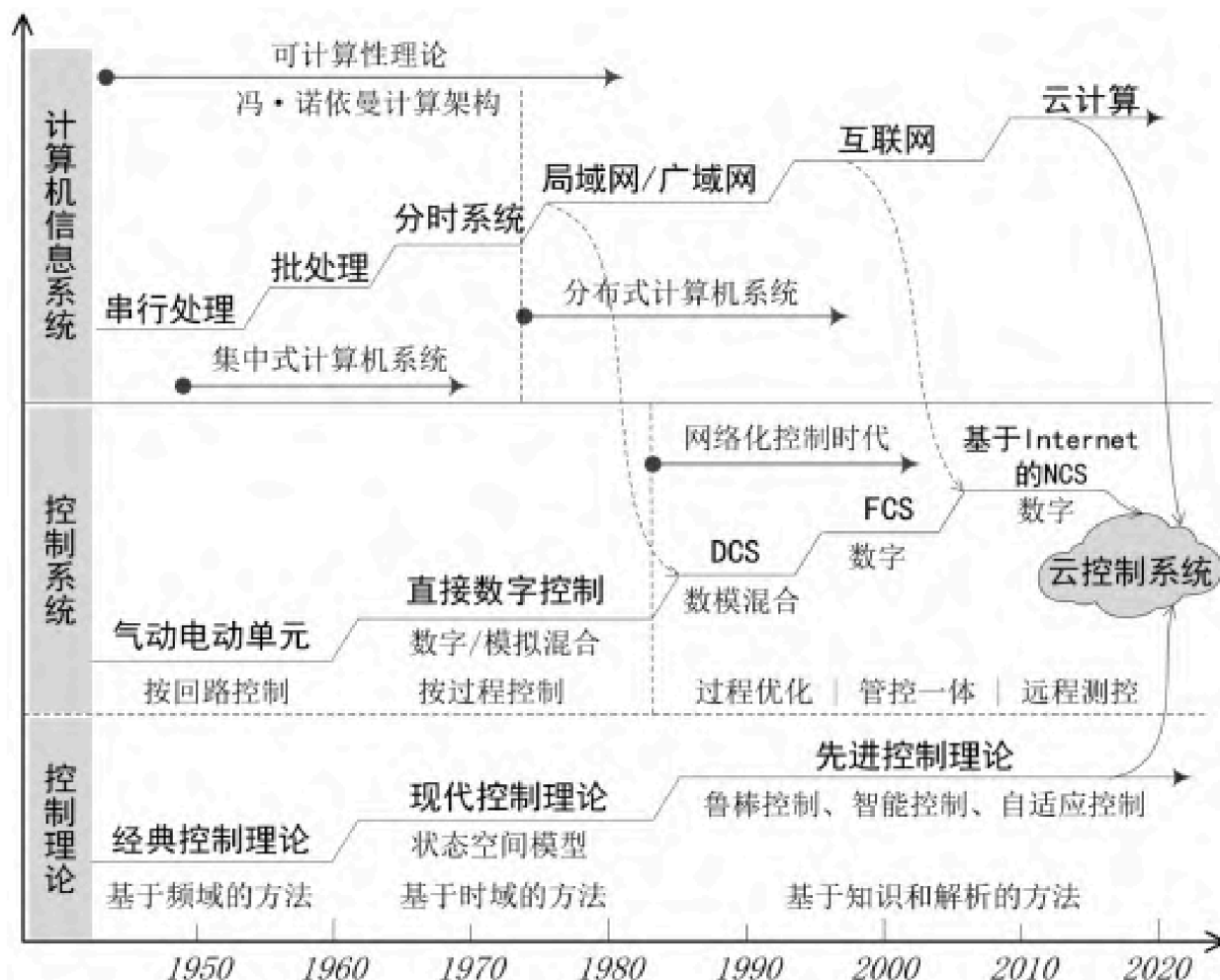
随着数据科学领域的进步，完整地保存生产监控数据具有巨大的潜在价值，在特殊情况下会出现当下难以预料的使用方式。

用例	频率	并发量	延时	数据操作
生产过程查询、产品跟踪	1000次/天	低	秒级	基于时间区间查询，过滤，join
分析、优化调度	100次/天	低	秒级	基于时间区间查询，过滤，join，其他数学计算
故障检测、预测性维护	执行间隔 > 30min	低	秒级	基于时间区间查询

云控制系统：未来可能的发展趋势

在对要求实时性的数据操作进行调研的过程中，我们的出发点是探究云上数据怎样与自动化结合，给出具体操作配置的例子。

传感器、控制器和执行器是反馈控制系统的基本组件。显然，如果传感器数据上云，而控制器是云数据库的客户端，那么将会引入相当频繁的数据读取，不符合目前时序数据库的假设。这时需要重新量化数据库workload，进行重新设计。



云控制系统就是将传感器、云控制器、执行期组成反馈控制系统，将现行的由I/O模块进行信号通信的现场控制单元改为互联网通信，实现过程控制层软硬件解耦。优势在于节省大量成本，但面临的问题是实时性、可靠性、安全性难以保证。

《云控制系统研究现状综述》概述了近年来国内外关于云控制系统的70篇论文，包括研究方向、阻碍等。《Industrial Automation as a Cloud Service》首次提出将过程控制层转移到云端，分析了成本节约和实时性、可靠性的困难。上文中DCS集成时序数据库的方式是根据数据库上云的需要提出的，没有参考

其他工作，论述中存在一些不严谨的地方，而实际上，早在2012年，Invensys System公司便提出了将控制系统的L2层（操作层）转移到云端的方法。

近年，云控制系统在业界成为研究热点，但其概念还比较抽象、理论体系尚未形成，也许在不久的将来，云控制系统在软实时控制领域将被实际应用。

三、测试设计

结合火电厂规模、工艺、业务需求以及实现方式，具体化应用场景的系统测试如下表所示，用来指导测试所使用的负载模拟器的配置，并统计延时能否达到要求。

用例	频率	并发量	延时	数据操作
实时数据采集（batch）	1次/秒	1000	< 1s	批写入，batchsize取1000
实时数据采集（不batch）	执行间隔 0.1s ~ 0.2s	100	< 50ms	写入
显示功能	刷新间隔 1s ~ 10s	< 400	< 1s	基于时间区间查询，过滤，聚合，关联，持续查询
绝对值报警	1次/秒	100	< 1s	查询某传感器最新数据，条件逻辑
变化率报警	执行间隔 10s ~ 1min	100	< 1s	基于时间区间查询，过滤，聚合，条件逻辑
仪表停用或网络故障报警	执行间隔 10s ~ 1min	100	< 1s	基于时间区间查询，过滤
即时报表	20次/天	无	< 1min	基于时间区间查询，过滤，聚合，关联
定时报表	1次/天	无	< 1min	基于时间区间查询，过滤，聚合，join
生产过程查询、产品跟踪	1000次/天	低	秒级	基于时间区间查询，过滤，关联
分析、优化调度	100次/天	低	秒级	基于时间区间查询，过滤，关联，其他数学计算
故障检测、预测性维护	执行间隔 > 30min	低	秒级	基于时间区间查询

四、数据操作整理抽象

根据上表对实时数据操作的统计，得到时序数据的特性如下：

数据顺序追加；数据可多维关联；通常高频访问热数据；冷数据需要降维归档；数据主要覆盖数值，状态，事件。

时序数据库特性：

写入速率稳定并且远远大于读取；按照时间窗口访问数据、极少更新；存在一定窗口期的覆盖写；批量删除；具备通用数据库要求的高可用，高可靠，可伸缩特性；通常不需要具备事务的能力。