

The Research and Application of Cloud Computing in Chemical Safety Production

Xinjian Lu^{1,*}, Qing Chen², Zengqiang Sun¹

¹Nanjing Chemcyber Technology Company Ltd., Nanjing, China

²School of Mathematics and Physics, Xuzhou University of Technology, Xuzhou, China

Email address:

xinjian.lu@chemcyber.com (Xinjian Lu), chenqxzit@gmail.com (Qing Chen), zengqiang.sun@chemcyber.com (Zengqiang Sun)

To cite this article:

Xinjian Lu, Qin Chen, Zengqiang Sun. The Research and Application of Cloud Computing in Chemical Safety Production. *Science Discovery*. Vol. 5, No. 4, 2017, pp. 244-250. doi: 10.11648/j.sd.20170504.11

Received: April 6, 2017; **Accepted:** May 12, 2017; **Published:** May 20, 2017

Abstract: In this paper, a high frequency, real-time, mobile interconnection cloud technology solution is proposed for the high risk production process of petrochemical industry, combined with the internet of things and cloud computing. The solution involved early warning, continuous optimization and so on, what bring not only convenience but also data security barrier for the enterprises. Furthermore, the solution has been applied to some enterprises and got a good evaluation.

Keywords: Chemical, Safety, Cloud Computing

化工安全生产中的云计算探索与应用的研究

陆新建^{1*}, 陈卿², 孙增强¹

¹南京科思倍信息科技有限公司, 南京, 中国

²徐州工程学院数学与物理科学学院, 徐州, 中国

邮箱

xinjian.lu@chemcyber.com (陆新建), chenqxzit@gmail.com (陈卿), zengqiang.sun@chemcyber.com (孙增强)

摘要: 本文针对石化行业的高危生产过程, 结合物联网、云计算, 提出了一套高频、实时, 移动互联的云技术解决方案, 涉及石化生产装置的安全预警、持续优化等内容, 为企业带来管理便捷的同时, 也为安全生产增设了一道数字屏障。本文描述的解决方案, 已应用于石化企业, 并得到了用户的好评, 具有推广价值。

关键词: 化工, 安全, 云计算

1. 引言

实时数据库在上世纪末开始应用于石化企业, 在当时的计算、网络和应用需求下, 实时数据库+关系数据库的模式发挥了重要作用, 促进了大量应用, 如广为大家熟知的制造执行系统 (Manufacturing execution system, MES) [1]。但经过近20年的使用, 这种实时数据库+关系数据库

的结构, 至少在三个方面无法适应技术的发展和应用的升级。

第一个方面是点数受限。软件企业之前的商业模式是以出售软件和工程为主, 所以无论是集散控制系统 (Distributed Control System, DCS) 的数据采集能力, 还

是实时数据库的存储能力，点数据容量许可都是天价商品，并和软件企业的经济效益直接挂钩。这种模式直接导致生产企业虽然投资了实时数据库，但很多当时看来不重要的数据没有被采集，如控制器的设定值、控制阀的开度、控制器的自控状态等，而这些信息对于判断异常是非常有用的，也间接导致这方面的应用偏少或效率偏低。

第二个方面是数据形式单一。实时数据库+关系数据库的模式，要求大多数进入数据库的数据为整型和浮点型，而文本或图片涉及较少，音频和视频文件更是很少进行数据库存储。这种单一的数据格式，严重地限制了多数据格式应用的开发，如难以实现通过噪音和振动识别设备运行的早期故障。

第三个方面是计算能力薄弱。实时数据库+关系数据库的模式，决定了数据库的主要功能是存储数据，几乎所有围绕生产数据进行的计算，都是将数据从数据库中取出来，完成计算后，再返回数据库或存入应用系统数据库。这种模式带来了多义性，比如很多生产管理软件，都涉及平稳率计算，但各自的计算方法不同，导致结果不同，并带来管理困惑。

显然，在实时数据库+关系数据库的传统模式下，围绕生产数据开发的应用越多，再次形成信息孤岛效应的现象就越明显。当粗放管理的效益挖掘越来越困难时，企业的生产管理必然走到精细管理层面。而精细管理，必然对数据处理的范围和频率提出截然不同的要求，比如将早期报警的监测频率提高到1分钟以内。在大部分应用以8小时为最高频率的实时数据库+关系数据库的模式下，1分钟以内的计算频率是很难实现的，即便能够实现，效率和代价也是难以想象的。那么，有没有一种全新的结构，能够提供全局规划，满足数据处理频率高、数据服务趋于个性化的要求呢？答案就是工业大数据+云计算[2-5]。

2. 化工安全生产中的云计算探索

近年来，云技术得到了快速发展，在互联网、金融以及离散制造业等领域得到了广泛应用，并积累了大量的技术和经验，很多开源系统使软件企业在云计算方面的低成本探索和实践成为可能。相对于上述行业，化工为传统行业，虽然大型石化企业在早年的信息化投入很大[6-7]，数据采集和应用较多，但对涵盖中小型化工的整个行业来说，云技术的应用落后于互联网、金融和离散制造业。那么，有没有这样一种开源系统，经适度定制后可应用到化工等连续生产的流程工业呢？这正是本文探索的目的。

经过仔细对比和潜心实践，我们发现这种解决方案是存在的，典型过程如下：

数据采集：基于现有物联网技术和应用多年的实时数据采集技术，实现工业过程控制数据的实时采集，采集频率为15秒，主要针对生产过程数据，同时可以采集现有关系型数据库中的数据，如质量分析数据。

数据存储：采用灵活的方式对采集的数据进行云端存储，可以考虑本地存储和云端存储的混合方案，兼顾计算性能和数据存储的历史周期。

数据计算：采用可扩展的冗余架构，进行数据的高频、实时计算，以满足各种数据应用需求。既涉及传统的统计分析计算，也为将来进行机器学习奠定基础。

数据展示：结合移动互联，将计算结果展示给需要的人，而不是为一个部门或组织进行深度定制。在个性化技术的支撑下，将数据处理做成一种服务，使每一个用户成为数据的拥有者、加工者和受益者。

详细说明如下：

2.1. 物联网

相比于离散制造业，化工等流程工业由于生产过程高度危险，故传感器和控制技术相对先进。大型石化企业，早年都会直接采用Honeywell等国外厂商的昂贵的控制系统和传感器，以保证生产过程得到有效监视和控制。

对于云计算和存储而言，解决的第一步是数据采集。数据采集在化工行业因为传感器齐全相对简单，一般采用工业标准接口——OPC接口。只要OPC接口的容量足够，化工生产装置DCS中的数据都可以实现高频采集，通常满足早期生产异常监测的采集频率为15秒。由于数采技术成熟、应用可靠，本文不作详细讨论，重点是云技术分析。

2.2. 云存储与计算

云计算是分布式处理（Distributed Computing）、并行处理（Parallel Computing）和网格计算（Grid Computing）的发展[8]，是通过网络将庞大的计算处理程序自动分拆成无数个较小的子程序，再交由多台服务器所组成的庞大系统，经计算分析之后将处理结果回传给用户。通过云计算技术，网络服务提供者可以在数秒之内，处理数以千万计甚至亿计的信息，达到和“超级计算机”同样强大的网络服务。

云存储是在云计算（cloud computing）概念上延伸和发展出来的一个新的概念，是指通过集群应用、网格技术或分布式文件系统等功能，将网络中大量各种不同类型的存储设备通过应用软件集合起来协同工作，共同对外提供数据存储和业务访问功能的一个系统。当云计算系统运算和处理的核是大量数据的存储和管理时，云计算系统中就需要配置大量的存储设备，那么云计算系统就转变成为一个云存储系统，所以云存储是一个以数据存储和管理为核心的云计算系统。

2.2.1. 时间序列数据库选择

降低成本，扩大应用的主要途径是选择开源数据库。目前还没有能够完全满足化工生产数据存储的开源数据库，现有的开源数据库，都需要进行一定程度的定制。目前主流的两个开源数据库是：时间序列数据库OpenTSDB和大数据系统Spark。

OpenTSDB 是时间系列数据（库），它基于HBase存储数据，充分发挥了HBase的分布式列存储特性，支持数百万每秒的读写，它的特点就是容易扩展，灵活的tag机制[9]。它的架构如下图[9]所示：

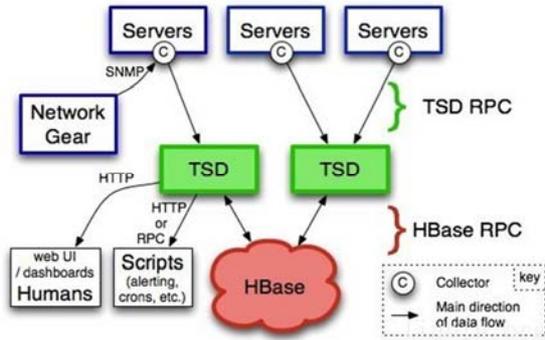


图1 OpenTSDB架构图。

OpenTSDB最主要的部件是TSD,它是接收数据并存储到HBase处理的核心所在。而带有C (collector) 标志的Server,是数据采集源,它的主要作用是将数据发给 TSD 服务。

Apache Spark, Spark Streaming是核心Spark API的一个扩展,它不会像Storm那样一次一个地处理数据流,而是在处理前按时间间隔预先将其切分为一段一段的批处理作业。Spark针对持续性数据流的抽象称为DStream (DiscretizedStream),一个DStream是一个微批处理 (micro-batching) 的RDD; RDD是一种分布式数据集,能够以两种方式并行运作,分别是任意函数和滑动窗口数据的转换[10]。Spark架构图如下所示[10]。

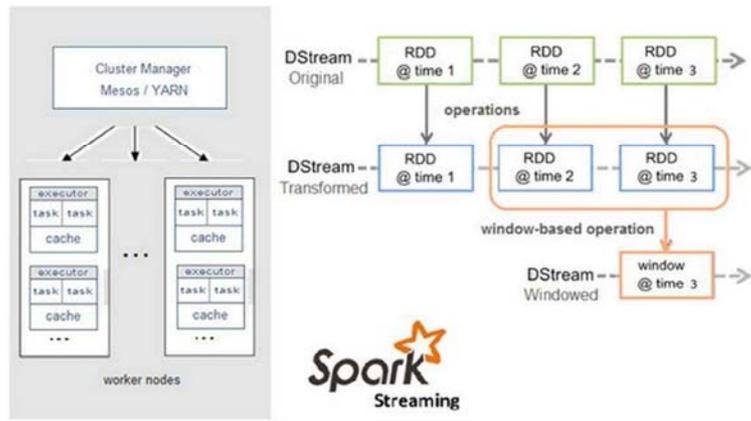


图2 Spark架构图。

适用性分析。在化工安全生产的云计算平台中,因OpenTSDB部署和开发的代价较小,早期建议采用OpenTSDB。而Spark有不同类型数据的存储能力,如前文提到的音频文件。所以,后期需要升级到Spark。另外Spark自带Spark-R功能,使其在数据库内部实现大规模计算成为可能。

前文提到,目前有一种开源架构的数据库能够完全适用于化工生产数据,这是由化工生产数据的多噪声、耦合、易缺失等特性决定的,因此任何一种数据库应用到这

个领域,都需要进行完善的事务处理功能开发,确保在各种条件下,所有的存储和计算都可以有效完成。

2.2.2. 系统结构分析

本文给出的、适用于化工安全生产管理的云计算架构如下图,该架构的核心是可扩展性。在可扩展性的支持下,系统才能够适应未来的采集容量、存储、计算和服务人群不断增加的需求,并能够为每一个用户提供个性化的计算服务。

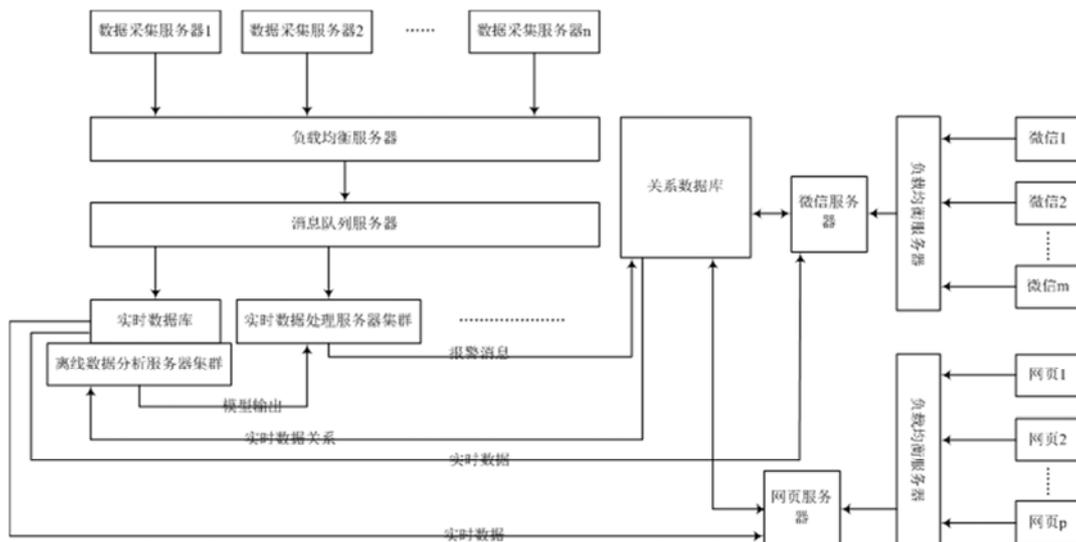


图3 化工安全生产管理的云计算架构图。

如上图，采集的生产数据，经过负载均衡服务器、消息队列服务器，进入实时数据库。这里的实时数据库，如前文所言，早期可以选择OpenTSDB，如需深入发展各种应用，则需要替换为Spark，或者选择其他更为合适的平台。而数据采集、负载均衡，以及消息队列，则需要自行开发。

系统中的事件存储，可采用开源的关系数据库，如PostGre SQL。PostGre SQL关系数据库性能良好，可以满足系统事件、消息、各种配置，以及用户个性化设置的存储和调用。

展现形式方面，除了必不可少的网页，还需要重点考虑移动应用，实现数据与用户个人的无缝连接。如果不能实现这种直接的联系，则无法达到云计算的效果，甚至遭到弃用，造成投资浪费。



图4 移动互联网趋势截图。

2.4. 个性化

个性化在很多企业应用中并不会被提及，原因可能是此前的应用计算量小，个性化需求不明显。另外一个重要原因是个性化会导致系统结构复杂度显著增加，开发难度随着增加。

但随着工业社会逐步进入智能时代，大规模计算和个性化是应用发展的必然趋势。其实在化工生产的安全管理方面，个性化应用存在巨大的需求空间。比如同一个报警，不同层面的管理者，需要不同宽度的报警范围，这样既可以将异常处理控制在一个有效的范围之内，又不至于因为信息过多而导致管理效率降低。

借助微信进行化工生产安全管理，使个性化实现成为可能。而个性化能够最大限度地降低系统的维护量和企业内部管理员的工作量，使每个管理者按自己的方式提取数据，为自己的工作服务。作为高危行业，化工装置的传感器越来越丰富，特别是新近设计和筹建的装置，连续量和

2.3. 移动互联

移动互联单列一节的主要原因有两点。一是互联互通的刚性需求，移动应用已成为人们生活、工作和学习不可或缺的组成部分。二是微信的易用和使用人数众多。微信虽然是一种社交工具，但其企业功能和开发能力亦非常强大。基于微信的应用可以通过扫码、无安装的方式直接使用，方便快捷，导致国内很多移动应用，都提供了微信形式。

相比于短信，微信不仅不会增加软件企业的额外费用负担，而且能够让使用者获得短信无法提供的数据趋势分析。尤其是后者，对得到消息的生产管理者来说，获得过去4小时的趋势图，远比获得一条短信提醒包含的信息量大，确认的代价低，更便于快速做出决策。如图4展示了查询或推送到管理者微信端的实时趋势和综合异常监测趋势。

状态量都很多。按照传统的方式，通过人来关注是非常困难的，且需要耗费大量精力，而信息泛滥还会带来管理效率的降低。个性化则让管理者实现数据为自己服务。

2.5. 计算方法

云计算应用中算法是核心部分，不同行业对计算方法和效果的要求不同，对于连续生产的流程工业，尤其是面向化工生产过程管理的数据计算，依据变量的数量和计算的复杂程度，主要有以下三类：

2.5.1. 单变量简单检测

单变量简单检测，主要是针对关键控制参数的运行上下限进行检测和推送，包括报警分析，经济、质量目标的跟踪分析。

以报警为例，目前业内报警管理的主要模式是在操作控制层面（现场控制），以及工厂内部局域网层面进行重

要参数的报警管理。一般情况下，管理人员离开工厂后，就无法跟踪报警发生情况，无法浏览历史趋势，以及评估报警的严重程度。

云计算和应用的结合，使得管理人员即使离开办公室，也可以掌握报警动态，并可根据当前的工作重点，调整报警项目、报警范围、报警接收的时间段，以及实现报警分享，进而提高管理效率。

2.5.2. 单变量复杂检测

单变量简单检测，为管理人员的安全生产管理解决了时空限制问题。如果需要深度分析，获取仪表直接测量数据背后的价值，则需要更多的复杂检测，本文探索了一些典型应用，如下：

速率跟踪：在化工生产过程中，有两类测量值需要进行速率跟踪。一类是连续量，如反应器的温度、机泵的出

口压力、运转设备的定子温度等，这些参数的特征是在一个典型值附近连续变化。如显著偏离这个典型值，则可能有异常发生。针对这类数据，在触发上下限报警之前，可以对其变化速率进行跟踪，一般可以获得5-15分钟的提前报警时间，帮助操作人员判断并提醒及时进行异常处理。

第二类是周期量，主要是液体或压缩气体储罐的液位。这类数据会在一个很大的范围内周期性地变化，所以很难在运行过程中触发常规的上下限报警，而辅助的速率报警有助于发现运行中可能的异常。

速率报警的常用计算方法是统计学上的最小二乘法，相比于移动平均法，最小二乘法处理连续变化的工业数据的准确性高，误报率低。

以原油罐液位为例，DCS数据曲线（蓝色）和最小二乘回归的速率跟踪曲线（红色）如下图所示。

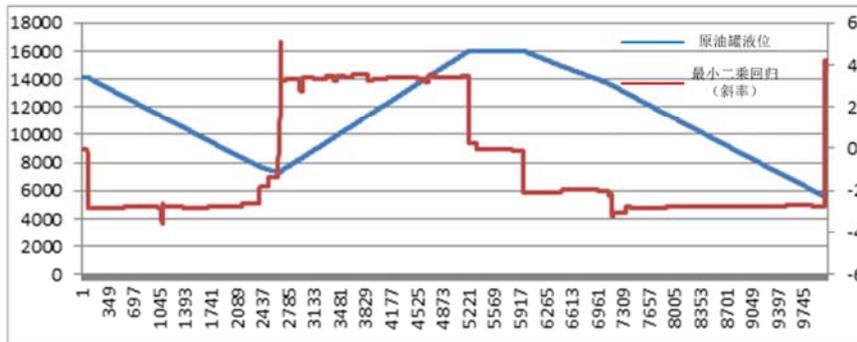


图5 速率跟踪曲线图。

由图5可知，在原油罐状态由付料到收料，再到平稳的过程中，相比液位DCS数值的变化，相应的最小二乘回归计算的斜率值从负值变为正值，再到零。甚至付料速率的变化也能够从相应的斜率值变化中反应出来，且变化更明显，反应更快。

不变跟踪：不变跟踪是化工生产中应对仪表失效的常用方法。当仪表失效时，在输出上表现为一个恒定的数据。但在化工生产中，对于前文中提到的连续量和周期量，数据恒定几乎是不可能的。因此可以通过不变跟踪，有效地检测一些关键变量是否出现了仪表异常，如运行中的锅炉

液位、收付作业中油罐的液位，当这些变量数值恒定时，通常意味着有异常发生，需要及时处理。

不变跟踪的常用计算方法是统计学上的方差计算。当方差为零时，表明监测对象的数据恒定，没有发生变化，需要确认。不变跟踪也可以反过来应用，即监测一些关键控制阀的开度是否出现了变化，如外界燃料气进装置流量控制阀的开度变化，以便即时掌握装置的能源消耗变化。

以原油罐液位为例，DCS数据曲线（蓝色）和方差计算曲线（红色）如下图所示。



图6 方差计算曲线图。

由图6可知,在原油罐状态由付料到收料,再到平稳的过程中,随液位DCS数值的变化,相应的方差计算值出现变化。如付料过程中,液位逐渐下降,此时方差不为0且基本维持不变,表示控制阀阀位无变化;在收料过程中,液位逐渐上升,此时方差不为0且基本维持不变,表示控制阀阀位无变化;在由付料变为出料的过程中,液位变化不明显,但方差出现明显变化,表示控制阀阀位出现明显变化;原油罐不收也不付的状态下,方差为0,表示控制阀阀位为零。甚至付料速率变化时,相应的方差计算值仍不为0,但数值有变化,且变化更明显,反应更快。

稳定跟踪:生产工艺不同、使用目的不同,稳定性的计算方法会出现显著不同。通常该计算方法主要分为两类,一类是评估过去一段时间装置或设备的运行是否稳定,这个过程如同实验室的数据分析,在工业生产管理上意义不大;另一类是即时稳定性评价,即通过统计方法,评价当前的生产工况、设备的运行是否平稳,进而指导操作调整。

对于跟踪过程中的测量值 X_i 及其平均值 X_{mean} ,即时性稳态计算的方法如下:

$$\sigma^2 = 1/(k-1) \sum (X_i - X_{mean})^2 \quad (1)$$

$$\delta^2 = 1/(k-1) \sum (X_{i+1} - X_i)^2 \quad (2)$$

$$\delta_r^2 = 1/(k-1) \sum T^2 \quad (3)$$

其中, T代表容忍度,默认值为0,当

$$\max(\delta^2 / \sigma^2, \delta_r^2 / \sigma^2) > 95\% \text{置信度} \quad (4)$$

表示不确定度为0,即状态稳定;否则,不确定度为1,表示状态不稳定。

以常压塔塔顶温度测量值为例,DCS数据曲线(蓝色)和稳定状态曲线(红色)如下图所示。

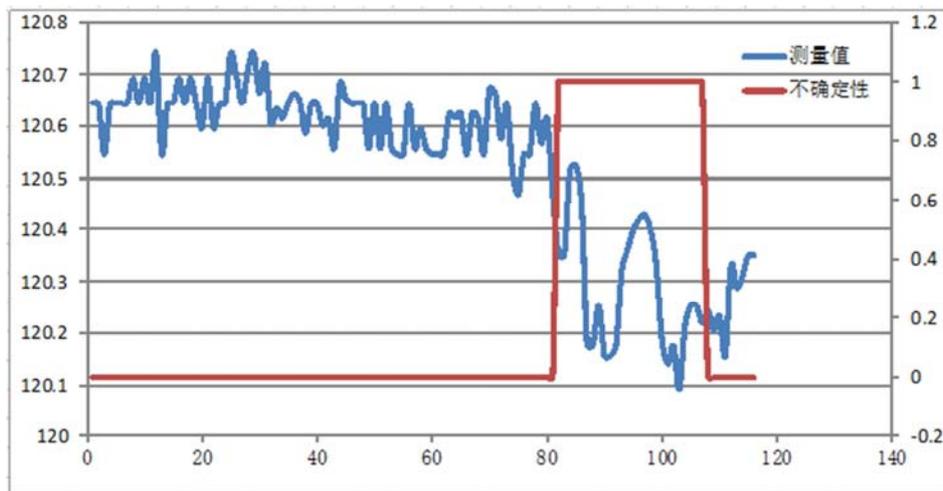


图7 稳定状态曲线图。

由图7可知,该测量值最初稳定,随后出现波动,最后达到新的稳定状态;相应的不稳定性值由0变为1,再变为0。

2.5.3. 多变量复杂检测

多变量复杂检测的目的是通过多个变量(可以是计算的二次变量)联合检测,避免单变量检测中容易出现的误判事件,提高诊断的准确性。多变量复杂检测在化工生产领域,常用的方法如下:

SPE检测: SPE(Squared Prediction Error)即使用过程实际值与模型预测值之间差值的平方进行检测。生产数据发生变化时,由训练集建立的模型已不再适用,样本远离了模型,采用SPE检测可直接反映模型的误差。

T²检测: 即使用样本与模型中心的距离进行检验。找出某生产运行模式下的异常点。

SPE检测和T²检测主要用于判断生产过程中的异常。

PCA监测: 生产过程中的变量常常是相互关联的,通过PCA(Principal Component Analysis)检测可对数据进行

降维,提取变量间的相互关联信息,减少需要分析的变量数目和复杂程度。由于应用普遍,本文亦不作详细讨论。

3. 应用与结论

南京科思倍信息科技有限公司针对上述难题和需求,开发了云服务平台——优惕云。优惕云涉及数据采集、云存储、云计算、移动应用等环节,目标定位于化工企业生产管理的安全、经济和便利等领域。其中安全功能,主要是通过大规模、高频、实时计算,实现生产异常的早期报警,缩短企业管理者的判断时间,提高异常事件的处理效果。平台中包含了科思倍的专利技术,能够实现处理上下限报警以外的单参数复杂统计报警,以及多参数组合报警,让企业的管理者在关注重要参数的同时,不遗漏偶发异常,提高安全生产的管理效率。

同时为了适应不同企业用户对数据存储位置的不同管理需求,优惕云提供多种方式的灵活部署。如下图,除了推荐的、实施快捷的优惕云方案外,还提供了完全企业私云[11-12]的部署方案。

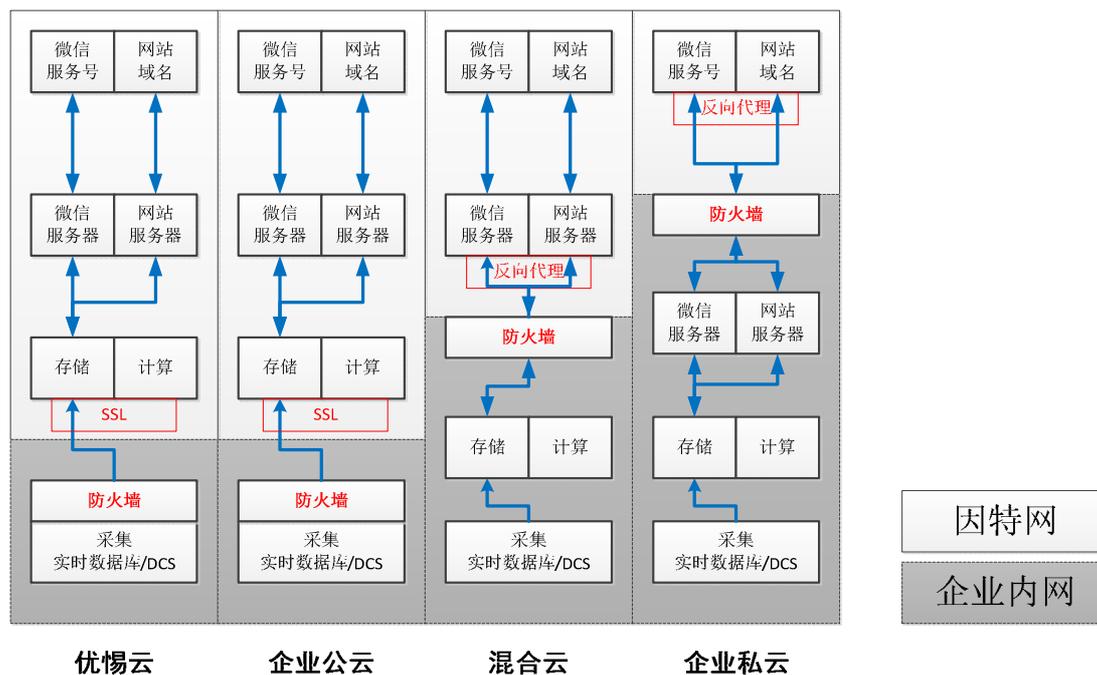


图8 优扬云不同部署方案示意图。

在化工生产企业面临越来越大的安全、环保压力的新形式下，优扬云的应用，得到了企业用户的认可，也在一定程度上，证明了云技术一定能够在化工安全生产中发挥巨大作用。

本文阐述了实时数据库+关系数据库的结构在化工生产中适应新技术发展和应用升级方面存在的限制；探索了适合高频率数据处理和数据服务个性化的云计算；并将计算结果通过手机微信进行展示，实现了预警信息、报表、优化的及时查询，使安全管理更加方便。该模式原理简单、成本低廉，容易扩展到其它流程行业。基于此类应用可通过信息化逐渐改变企业传统的生产方式，并通过持续提供服务，促进互联网企业与传统企业发展合作经济。

参考文献

- [1] 张连棠. 华北石化公司制造执行系统(MES)设计与应用研究[D]. 天津: 天津大学, 2010: 5-11.
- [2] 宋扬. 探析云计算在化工企业中的应用[J]. 中国新技术新产品, 2016,(15):27.
- [3] 张涛凤. 浅谈云计算在化工企业中的应用[J]. 科技风, 2014,(04):98.
- [4] 刘柳, 王云飞, 李勃, 曹妍明. 云计算技术在石油化工行业中的应用[J]. 计算机光盘软件与应用, 2013,(15):27-28.
- [5] 罗娜, 万锋, 叶贞成, 杜文莉, 祁荣宾, 钟伟民, 钱锋. 基于云计算和虚拟现实技术的化工过程虚拟工厂[J]. 化工进展, 2012,(12):2815-2819.
- [6] 尚巍, 张锐峰. 云计算在化工企业的应用[J]. 中国管理信息化, 2014,(03):58-60.
- [7] 曹卫民. 石化企业云计算数据中心设计[J]. 计算机与应用化学, 2013,(30):899-902.
- [8] 阿伦. 云计算技术在化工仿真教学中应用的展望[J]. 内蒙古石油化工, 2009,(23):76-77.
- [9] 网络资料: <https://www.oschina.net/p/opentsdb>
- [10] 网络资料: https://image.baidu.com/search/index?tn=baiduimage&ct=201326592&lm=-1&cl=2&ie=gbk&word=spark%20streaming%BC%DC%B9%B9%CD%BC&fr=ala&ori_query=spark%20streaming%E6%9E%B6%E6%9E%84%E5%9B%BE&ala=0&alatpl=sp&pos=0&hs=2&xthttps=111111
- [11] 马婷, 刘宁. 陕西煤业化工集团基于ASSA云的IT基础设施架构建设[J]. 陕西煤炭, 2015,(05):19-21.
- [12] 李传坤, 陶少辉, 王春利, 高新江. 基于云计算的石化过程安全管理平台研发[J]. 计算机与应用化学, 2014,(31):213-216.