

T/CASME

中国中小商业企业协会团体标准

T/CASME XXXX—XXXX

数据库 3D 可视化系统规范

Specification for database 3D visualization system

（征求意见稿）

2023 – XX – XX 发布

2023 – XX – XX 实施

中国中小商业企业协会 发布

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由武汉乐吾乐科技有限责任公司提出。

本文件由中国中小商业企业协会归口。

本文件起草单位：武汉乐吾乐科技有限责任公司、武汉嘉诚嘉图自动化科技有限公司、XXX。

本文件主要起草人：吴克兵、吴克伟、高浩然、郝望栋、XXX。

数据库 3D 可视化系统规范

1 范围

本文件规定了数据库3D可视化系统规范的术语和定义、缩略语、功能要求、性能要求、运行测试。本文件适用于数据库3D可视化系统的设计开发。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 22239 信息安全技术 网络安全等级保护基本要求

GB/T 25000.51 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第51部分：就绪可用软件产品（RUSP）的质量要求和测试细则

GB/T 25068.2 信息技术 安全技术 网络安全 第2部分：网络安全设计和实现指南

GB/T 29831.3 系统与软件功能性第3部分：测试方法

GB/T 29832.3 系统与软件可靠性第3部分：测试方法

GB/T 29833.3 系统与软件可移植性第3部分：测试方法

GB/T 29834.3 系统与软件维护性第3部分：测试方法

GB/T 29835.3 系统与软件效率第3部分：测试方法

GB/T 29836.3 系统与软件易用性第3部分：测试方法

GB/T 36073 数据管理能力成熟度评估模型

3 术语和定义

下列术语和定义适用于本文件。

3.1

数据库 database

支持一个或多个应用领域，按概念结构组织的数据集合，其概念结构描述这些数据的特征及其对应实体间的联系。

3.2

可视化 visualization

通过创建图像、图表或动画来传达信息的技术。

4 缩略语

CPU 中央处理器（Central Processing Unit）

CSV 文本分隔值（Comma-Separated Values）

API 应用程序接口（Application Program Interface）

POP 开放平台（pctowap open platform）

ECS 云服务器（Elastic Compute Service）

JDBC Java数据库连接（Java Database Connectivity）

ODBC 开放数据库连接（Open Database Connectivity）

SQL 结构化查询语言（Structural Query Language）

NAS 网络附属存储（Network Attached Storage）

JSON 一种轻量级的数据交换格式 (JavaScript Object Notation)

5 功能要求

5.1 总体要求

5.1.1 系统的功能性应符合 GB/T25000.51 中 5.3.1 的规定,能够快速的对大量数据的收集整理,并经过筛选后精准的归类,节约人力成本,提高数据整理的效率和准确率。

5.1.2 数据库主要功能是以库或表为单位,基于集群形式的架构形态,用传统数据库常见的操作模式,主要面向分析型的业务类型,提供海量数据的存储和计算能力,能够对数据进行全生命周期的处理。

5.1.3 系统的易用性应符合 GB/T 25000.51 中 5.3.3 的规定。

5.1.4 系统可靠性应符合 GB/T 25000.51 中 5.3.5 的规定,系统可按照设定的数据分标准进行精准分类。

5.1.5 系统的维护性应符合 GB/T 25000.51 中 5.3.7 的规定。

5.1.6 系统的效率应符合 GB/T 25000.51 中 5.3.10 的规定,通过系统高速运算识别,高效率的进行数据的收集整理。

5.2 数据库功能要求

系统数据库还应满足如下功能要求:

- a) 数据应源自互联网,可采用全息存储、专用软件、专用芯片或编程数据处理器压缩存储海量数据等方法,通过局域网进行,依赖NAS技术实现网络存储。
- b) 应具备数据处理的基本功能,包括数据的导入、存储、计算和分析等;
- c) 应具备容错能力,当节点、进程、网络出现故障时,系统仍能对外提供正常服务,数据不丢失;
- d) 应具备安全能力,包括对用户进行身份认证的能力,按人员和使用场景进行操作权限分级管理的能力,按库、表、行、列为级别进行数据权限分配的能力,数据的加解密的能力,以及进行安全审计的能力;
- e) 应具备开放性和兼容性,支持ODBC、JDBC等开发接口,兼容关系型数据库数据导入和连接操作,具备和大数据批处理平台、大数据流处理平台友好的对接形式,兼容异构芯片和多种操作系统;
- f) 应具备集群规模弹性伸缩的能力,包括以在线或离线形式增加或减少集群服务节点,系统对外服务不出现异常,数据不丢失;
- g) 应具备运维管理能力,支持集群资源和任务的监控,集群部署升级、配置管理、节点启停、服务启停、用户管理等功能;
- h) 宜采用分布式存储和并行计算框架,结合Apache Spark (计算引擎)、MapReduce (基于集群的高性能并行计算平台)、Hadoop (分布式系统基础架构)等多种分布式计算引擎,对各类结构化、半结构化及非结构化的信息资源进行快速的分布式计算,提供基于关联、聚类、分类、预测等类算法库以及可视化组件、拖拽式的数据挖掘分析开发工具包。

5.3 可视化功能要求

系统的可视化功能还应满足如下功能要求:

- a) 应满足互联网爆发的大数据需求,快速的收集、筛选、分析、归纳、展现决策者所需要的信息,并根据新增的数据进行实时更新;
- b) 应以散点图、热力图、柱形图、饼图、折线图、地理图、雷达图、箱线图、关系图、矩形树图等主要的图表形式对数据进行可视化展现,让数据结果一目了然;
- c) 应提供海量数据的地理轨迹、地理飞线、热力分布、地域区块、3D地图、3D地球,实现地理数据的多层叠加,此外还有拓扑关系、树图等异形图表供自由搭配;
- d) 应适配非常规的拼接分辨率,创建的可视化应用能够发布分享,作为对外数据业务展示的窗;
- e) 应支持接入包括分析型数据库、关系型数据库、本地CSV上传和在线API等,支持动态请求。

5.4 拓展要求

数据库3D可视化系统还应满足以下要求:

- a) 硬件方面：系统运行应具有良好的纵向扩展性（通过在单机中增加CPU、内存、硬盘等资源提高系统处理性能）和横向扩展性（通过增加服务器数量提高系统的处理性能）；
- b) 网络方面：应使用能通过增加网络设备达到提高网络传输性能的目标；
- c) 软件方面：应能支持通过参数化配置应对业务功能的一般性变化，提供预留接口以支持与未来其他系统的连接，尽量减少业务扩充对系统的影响，采用统一的数据接口规范，通过数据交互中间件与其他业务系统以及外部系统进行数据交换，以实现各系统之间的数据互操作，应实现平台与其他外部系统互联操作，系统在设计时应根据具体情况定义接口规范。

5.5 多数据源要求

系统多数据源要求见表1。

表1 多数据源要求

类型	关系型数据源	文件类	API类	其他
数据源描述	主流关系型数据库支持RDS for SQLServer（云数据库）、MySQL数据库	支持CSV文件、静态JSON	支持POP API、API网关	数据代理服务、ECS

6 性能要求

6.1 数据处理要求

系统的数据处理应满足以下要求：

- a) 数据处理在针对普通数据的维度、度量进行直接分析和观察数据的整体效应以外，还应具备分析平均数、最大值、最小值、总体标准偏差、样本标准偏差、总体偏差、样本方差分位数等需求。
- b) 图形化页面渲染完成时间应不大于 3 s。
- c) 数据导入加载完毕后数据应能以结构化数据形式对外显示并进行处理操作。
- d) 数据批处理计算应以数据集合为单位，对已存储的数据进行批量计算，支持的数据转换应包括数据清洗、数据内容和数据结构的转化，支持的数据分析应包括基于 SQL 的数据查询、统计、分析等能力。

6.2 运维要求

系统应该具备对系统各模块完备的运维管理能力，包括但不限于以下内容：

- a) 系统部署：能够以图形化或脚本等友好的部署方式，向导式安装部署分布式分析数据库，使其能够正常提供服务；
- b) 集群资源监控告警：能够将系统实时和历史硬件资源的使用情况，以图表等可视化形式进行展示，同时设定异常阈值，当出现异常时，能够以多种方式进行告警通知；
- c) 系统服务管理：能够对整体系统服务、子模块服务、执行的应用进行显示以及批量启动、停止操作；
- d) 用户管理：能够在统一的管理页面实现用户的创建、修改和删除，并配置密码，同时为不同级别的用户分配针对不同库、表的操作权限。

6.3 兼容性要求

系统应该对上层应用开放友好的开发接口，包括但不限于以下内容：

- a) JDBC和ODBC接口的支持；
- b) 支持不同型号或架构的芯片等进行异构计算；
- c) 支持多种主流的操作系统；
- d) 对分布式批处理平台、分布式流处理平台的兼容性。

6.4 安全性要求

系统除了信息安全应符合GB/T 22239的规定、网络安全应符合GB/T 25068.2的规定、数据安全应符合GB/T 36073的规定以外，还应具备防止恶意的访问，防止关键数据的泄漏的安全保护技术，包括但不限于以下内容：

- a) 系统各个模块对用户以密码或秘钥方式进行身份认证的功能；
- b) 至少按照管理员、审计员和操作员三种角色进行分级权限管理的功能，针对操作员按照数据库、表、行列等不同粒度进行权限控制；
- c) 支持多种类型的数据加解密方式；
- d) 支持对于操作进行审计的功能，包括用户操作、任务执行操作、运维操作、权限操作等。

6.5 容错要求

系统应该支持各个组件在出现故障后系统进行恢复能力，包括但不限于以下内容：

- a) 存储数据能够以全量、增量、加密、压缩等形式进行在线或离线备份以及恢复；
- b) 系统各组件能够在插拔硬盘、网络故障、节点故障、应用故障等常见故障发生时，进行主备节点切换，不影响正在执行的业务进度，同时保障数据正确性。

6.6 性能测试要求

系统应该能够满足特定应用场景中对大数据批处理的性能要求，性能测试主要考虑以下内容：

- a) 测试需基于一定资源冗余的硬件环境，包括服务器节点规模、网络链路、内存总量、CPU总量、磁盘总量等；
- b) 测试需定义良好的测试业务场景，包括合理规模的数据量、与实际业务相符的数据模式和操作模式；
- c) 测试需定义合理的衡量指标，例如单节点吞吐量、系统总吞吐量、平均时延、最大时延、执行时间等；
- d) 常见的测试场景包括多类型查询、混合负载、压力测试、稳定性测试等。

7 运行测试

7.1 测试环境

系统测试时集群节点应大于3，并保持测试环境的网络通畅。

7.2 测试范围

测试范围应包括系统功能测试和和非功能性测试。对集成后的系统从业务以及功能角度进行测试并需要对同一信息平台进行非功能性测试，非功能性测试包括性能测试、负载测试、压力测试等。其中，性能测试是通过性能测试工具模拟正常场景、正常峰值、异常负载对产品的各项性能指标进行测试。从客户端、网络、服务器三方面，采用LoadRunner（负载测试工具）等专业辅助工具进行性能测试。

7.3 测试方法

7.3.1 功能性

系统功能性应按照GB/T 29831.3的方法进行测试。

7.3.2 可靠性

系统可靠性应按照GB/T 29832.3的方法进行测试。

7.3.3 可移植性

系统可移植性应按照GB/T 29833.3的方法进行测试。

7.3.4 维护性

系统维护性应按照GB/T 29834.3的方法进行测试。

7.3.5 效率

系统效率应按照GB/T 29835.3的方法进行测试。

7.3.6 易用性

系统易用性应按照GB/T 29836.3的方法进行测试。
