



# 中华人民共和国国家标准

GB/T 37722—2019

---

## 信息技术 大数据存储与处理系统功能要求

Information technology—  
Technical requirements for big data storage and processing systems

2019-08-30 发布

2020-03-01 实施

---

国家市场监督管理总局  
中国国家标准化管理委员会 发布

## 目 次

前言 .....	Ⅲ
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 缩略语 .....	2
5 概述 .....	2
6 大数据存储子系统功能要求 .....	3
6.1 基本要求 .....	3
6.2 分布式文件存储 .....	3
6.3 分布式结构化数据存储 .....	3
6.4 分布式列式数据存储 .....	3
6.5 分布式图数据存储 .....	4
7 大数据处理子系统功能要求 .....	4
7.1 基本要求 .....	4
7.2 批处理框架 .....	4
7.3 流处理框架 .....	4
7.4 图计算框架 .....	5
7.5 内存计算框架 .....	5
7.6 批流融合计算框架 .....	5

## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

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这些专利的责任。

本标准由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本标准起草单位:华为技术有限公司、中国电子技术标准化研究院、浪潮电子信息产业股份有限公司、上海计算机软件技术开发中心、勤智数码科技股份有限公司、深圳市金蝶天燕中间件股份有限公司、新华三技术有限公司、中兴通讯股份有限公司、杭州中奥科技有限公司、天津南大通用数据技术股份有限公司。

本标准主要起草人:赵华、符海芳、卫凤林、张群、苏志远、赵江、陈敏刚、刘振宇、蔡立志、刘宇峰、李正、林琳、潘子健、吴文峰、张东涛、朱松、沈贝伦、陆韵、武新、张绍勇、李冰、尹卓、孙嘉阳。

# 信息技术

## 大数据存储与处理系统功能要求

### 1 范围

本标准规定了大数据存储与处理系统的分布式文件存储、分布式结构化数据存储、分布式列式数据存储、分布式图数据存储、批处理框架、流处理框架、图计算框架、内存计算框架和批流融合计算框架等的功能要求。

本标准适用于大数据存储与处理系统的设计、开发和应用部署。

### 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 35295—2017 信息技术 大数据 术语

### 3 术语和定义

GB/T 35295—2017 界定的以及下列术语和定义适用于本文件。

#### 3.1

##### **图数据库 graph database**

一种应用图理论存储实体及其之间关系信息的非关系型数据库。

注 1: 图数据库的数据模型由节点及边组成(即节点间关系)。

注 2: 图数据库支持图查询、图遍历及图分析等功能,适用于复杂关系的探索与发现。

#### 3.2

##### **批处理 batch processing**

将一个大型作业分解成为多个任务交由多个节点分别处理,再将分解后多个任务处理的结果汇总起来,得出最终的分析结果的计算框架,具备高可用、高扩展、高并发等能力。

#### 3.3

##### **流处理 stream processing**

能够对具有实时、高速、无边界、瞬时性等特性的流式数据进行实时处理的计算能力。

#### 3.4

##### **图计算 graph processing**

以“图论”为基础的对数据的一种“图”结构的抽象表达,以及在这种数据结构上的计算模式。

注: 在图计算中,基本的数据结构表达包括:节点、边、权重等。

#### 3.5

##### **内存计算 in-memory processing**

优先使用内存对数据进行计算、分析的一种数据处理技术。