ICS 35.240
CCS L 67

NATIONAL STANDARD OF THE PEOPLE'S REPUBLIC OF CHINA

GB/T 44216-2024

Information technology-Big data-Techinical requirements for integrated batch and streaming computing

信息技术 大数据 批流融合计算技术要求

Issued on: July 24, 2024 **Implemented on: February 1, 2025**

Issued by: State Administration for Market Regulation;
Standardization Administration of the PRC

Table of Contents

1 Scope

This document specifies the technical requirements for big data batch and stream fusion computing, including. technical requirements, scalability requirements, compatibility requirements, performance indicators, Standard, etc.

This document is applicable to the design, development and deployment of batch-stream fusion computing systems.

The evaluation of batch-stream integration products and services shall be carried out in accordance with this reference.

2 Normative references

GB/T 35295-2017

3 Terms and definitions

The terms and definitions defined in GB/T 35295-2017 and the following apply to this document.

3.1 batch processing

Decompose a large job into multiple tasks and assign them to multiple nodes for processing respectively, and then summarize the results of the decomposed multiple tasks.

To obtain the calculation framework for the final analysis results.

3.2 stream processing

For large-scale computing scenarios with high-speed concurrency and high timeliness requirements, it can be used for real-time, high-speed, instantaneous and other characteristics.

A computing framework for real-time processing of streaming data.

3.3 integrated batch and streaming computing

A computing framework that can support both batch processing and stream processing.

3.4 Scatter-gather map-reduce

The required computation is divided and distributed across multiple nodes, and the overall result is the processing of large data sets composed of the results of each node. form.

Example. MapReduce (a computing model that includes two computing processes, Map and Reduce) uses a disperse-gather processing approach.

3.5 Tenant

One or more cloud service users with shared access to a set of physical and virtual resources.

3.6 multi-tenancy

The allocation of physical or virtual resources ensures that multiple tenants and their computing and data are isolated and inaccessible from each other.