GB

NATIONAL STANDARD OF THE

PEOPLE'S REPUBLIC OF CHINA

ICS 35.240
L 67

GB/T 38675-2020

Information technology - General requirements for big data computing systems

信息技术 大数据计算系统通用要求

Issued on: April 28, 2020 **Implemented on: November 01, 2020**

Issued by: State Administration for Market Regulation;

Standardization Administration of the PRC.

Table of Contents

Foreword 3
1 Scope 4
2 Normative references 4
3 Terms and definitions 4
4 Abbreviations 4
5 Hardware requirements 5
5.1 Hardware composition 5
5.2 Computing unit 6
5.3 Computing acceleration unit 6
5.4 Storage unit 7
5.5 High-speed interconnection unit 8
5.6 Management monitoring unit 9
5.7 Power supply unit 9
5.8 Structure and heat dissipation unit 10
6 Software requirements 10
6.1 Software composition 10
6.2 Operating system 10
6.3 Virtualization software 11
6.4 Resource management software 12
6.5 Communication management software 12
7 Network requirements 13
8 Security requirements 13

Information technology - General requirements for big data computing systems

1 Scope

This Standard specifies the hardware, software, network and security requirements of big data computing systems.

This Standard applies to the development, design, operation and maintenance of big data computing systems.

2 Normative references

The following documents are indispensable for the application of this document. For the dated references, only the editions with the dates indicated are applicable to this document. For the undated references, the latest edition (including all the amendments) are applicable to this document.

GB/T 21028 Information security technology - Security techniques requirement for server

GB/T 35295-2017 Information technology - Big data - Terminology

3 Terms and definitions

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

3.1

Big data computing systems

Software and hardware systems which provide basic computing support for the storage, processing, and analysis applications of big data systems.

4 Abbreviations

The following abbreviations apply to this document.

API: Application Programming Interface

computing acceleration unit, storage unit, high-speed interconnection unit, management monitoring unit, power supply unit, and structure and heat dissipation unit.

5.2 Computing unit

The computing unit is required as follows:

a) There shall be corresponding measures to ensure availability;

b) It shall support multiple processors;

c) It shall support hardware status monitoring management and remote management;

d) It shall support the expansion of internal components of the system;

e) It shall support multiple virtualization methods;

f) It shall support the configuration of the amount of memory of the virtual machine when the operation is stopped;

g) It shall support the configuration of the amount of extended memory when the operation is stopped;

h) The memory of each node shall not be less than 8GB;

i) It shall support extended DDR3, DDR4 or above memory version standards;

j) It should support memory error checking/correction;

k) It should support memory mirroring/hot standby;

l) It should support reading and writing skills above PB;

m) It should support the expansion of external nodes of the system;

n) It should support horizontal and vertical online expansion under normal operation;

o) It should support a hybrid memory architecture.

5.3 Computing acceleration unit

The computing acceleration unit is required as follows:

a) It shall provide dedicated computing resources, to assist the computing unit in computing processing;

f) It should support advanced features such as compression, deduplication, cache acceleration, quotas, and QoS;

g) It shall support storage node-level dynamic expansion and scaling and storage data rebalancing;

h) It shall support online expansion of SATA, SAS, PCIe interface SSD, HDD, etc., and support hot swap;

i) It should support the expansion of storage systems connected to protocols such as FC, iSCSI, NAS;

j) It shall support the expansion of the amount of memory when the operation is stopped;

k) It should support online upgrades;

l) It should support reading and writing capabilities above PB;

m) Availability shall be guaranteed by corresponding measures;

n) Enterprise-level hard disk shall be configured to ensure reliability;

o) It shall support data reconstruction of failed hard disk.

5.5 High-speed interconnection unit

The high-speed interconnection unit is required as follows:

a) It shall provide the interconnection between the computing unit, storage unit and computing acceleration unit;

b) It shall provide the interconnection between different accelerator cards in the local computing acceleration unit;

c) It shall support network protocols such as TCP/IP to realize the interconnection between nodes;

d) It should support protocols such as IPoIB and RDMA to achieve high-speed interconnection between nodes;

e) It shall support at least 100M network connection between nodes;

f) It shall support the interconnection between the expanded computing unit, storage unit, and computing acceleration unit when the operation is stopped;

g) It shall support bandwidth above 1 Gbps;

monitoring system;

d) It shall have power-on sequence control function;

e) It shall have dynamic adjustment technology of power load;

f) It shall support cabinet-level centralized power supply;

g) It shall support N+N or N+M power supply redundancy mode.

5.8 Structure and heat dissipation unit

The structure and heat dissipation unit is required as follows:

a) The heat dissipation method shall support air cooling; should support liquid cooling and conduction heat dissipation;

b) The heat dissipation efficiency shall be monitored by sensors and can be seen in the management monitoring unit;

c) Modular design shall be adopted;

d) Centralized heat dissipation technology shall be used;

e) The fan shall support tool-free hot swap;

f) It shall support the front air inlet and the rear air outlet.

6 Software requirements

6.1 Software composition

The software part of big data computing systems includes operating system, virtualization software, resource management software and communication management software.

6.2 Operating system

The operating system is required as follows:

a) It shall support mainstream operating system types and major versions;

b) It shall provide automated batch deployment capabilities of operating system;

c) It shall provide a man-machine interface;

d) It shall support computing acceleration units and drivers; provide

6.4 Resource management software

The resource management software is required as follows:

a) It shall support the monitoring and management of virtual machine resources;

b) It shall support the security management of virtual machine resources, including but not limited to rights management, access control, etc.;

c) It shall support dynamic resource allocation according to workload; provide resource allocation capabilities for task types and scheduling queues of big data systems;

d) It shall provide deployment and monitoring agent software; in accordance with standard interfaces, provide the big data system with relevant information on the operating status of resources and perform various monitoring tasks;

e) It should provide automatic deployment tools that support multiple resource modes, to automatically add new physical or virtual resources to the infrastructure;

Example: Bare machine, virtual machine, cloud service.

f) It should support the unified management and pooling of multiple resources.

6.5 Communication management software

Communication management software is required as follows:

a) It shall provide standardized message and communication interfaces; support reliable data queues, transmission and reception of data;

b) It shall support point-to-point and store-and-forward models;

c) It shall support the collection and transmission of operational data of various infrastructure resources through a publish/subscribe mode;

d) It shall support different communication protocols for the transmission of different types of data;

Example: TCP is used for structured data; non-TCP is used for large files such as videos and images.

e) It shall support the encryption and decryption of transmitted data.

This is an excerpt of the PDF (Some pages are marked off intentionally)

------ The End ------