



开放数据中心委员会
Open Data Center Committee

[编号 ODCC-2022-05002]

分布式存储技术与产业分析报告

开放数据中心标准推进委员会

2022-04 发布

版权声明

ODCC（开放数据中心委员会）发布的各项成果，受《著作权法》保护，编制单位共同享有著作权。

转载、摘编或利用其它方式使用 ODCC 成果中的文字或者观点的，应注明来源：“开放数据中心委员会 ODCC”。

对于未经著作权人书面同意而实施的剽窃、复制、修改、销售、改编、汇编和翻译出版等侵权行为，ODCC 及有关单位将追究其法律责任，感谢各单位的配合与支持。

www.ODCC.org.cn

编制说明

本报告由中国信息通信研究院云计算与大数据研究所牵头撰写，在撰写过程中得到了多家单位的大力支持，在此特别感谢以下参编单位和参编人员：

参编单位（排名不分先后）：

中国信息通信研究院、烽火通信科技股份有限公司、中国移动通信集团有限公司、长城超云（北京）科技有限公司、深圳市江波龙电子股份有限公司

参编人员（排名不分先后）：

郭亮、李洁、王少鹏、孙聪、赵精华、李传宝、杜希、王娟、冯轶、杨志勇、伍瑞、李锐、李东辉、孙振、孔燕斌

项目经理：

孙聪 suncong@caict.ac.cn

www.ODCC.org.cn

前言

云计算、大数据、人工智能、5G 等新一代信息技术快速发展，对存储的需求不断提高，作为信息系统不可或缺的组成之一，存储在支撑数字经济发展中，扮演着至关重要的角色。

当前，随着数字化转型的不断深入，整个社会加速向数字化、网络化、智能化方向发展，由此带来的海量数据对存储提出了新的要求。传统存储虽有技术成熟、性能良好、可用性高等优点，但面对海量数据，却在可扩展性、成本等方面存在明显不足，而分布式存储技术则有较好的表现。

分布式存储技术从出现开始就吸引了众多企业的关注，成为存储产业发展的重要方向。分布式存储技术的发展历程具有“迎合应用需求”和“技术创新突破”的特点，随技术的演变而迅速发展，总体呈现高性能、易于扩展以及智能化等优势。

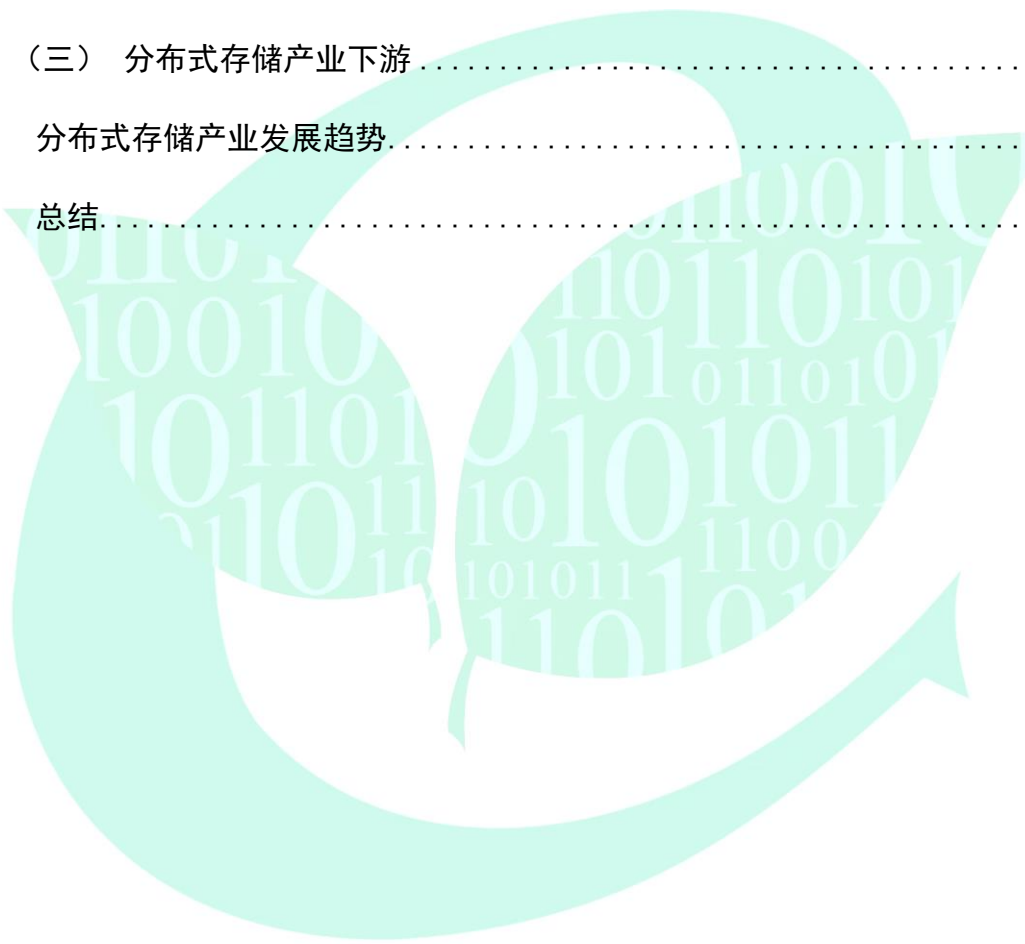
本白皮书基于全球视角和我国现状，对分布式存储市场趋势和产业政策进行了梳理，从技术创新、产业图谱等维度对分布式存储产业的发展进行了重点分析，并对我国分布式存储发展进行了展望。由于时间仓促，水平所限，错误和不足之处在所难免，欢迎各位读者批评指正。如有意见或建议请联系 dceco@caict.ac.cn。

www.ODCC.org.cn

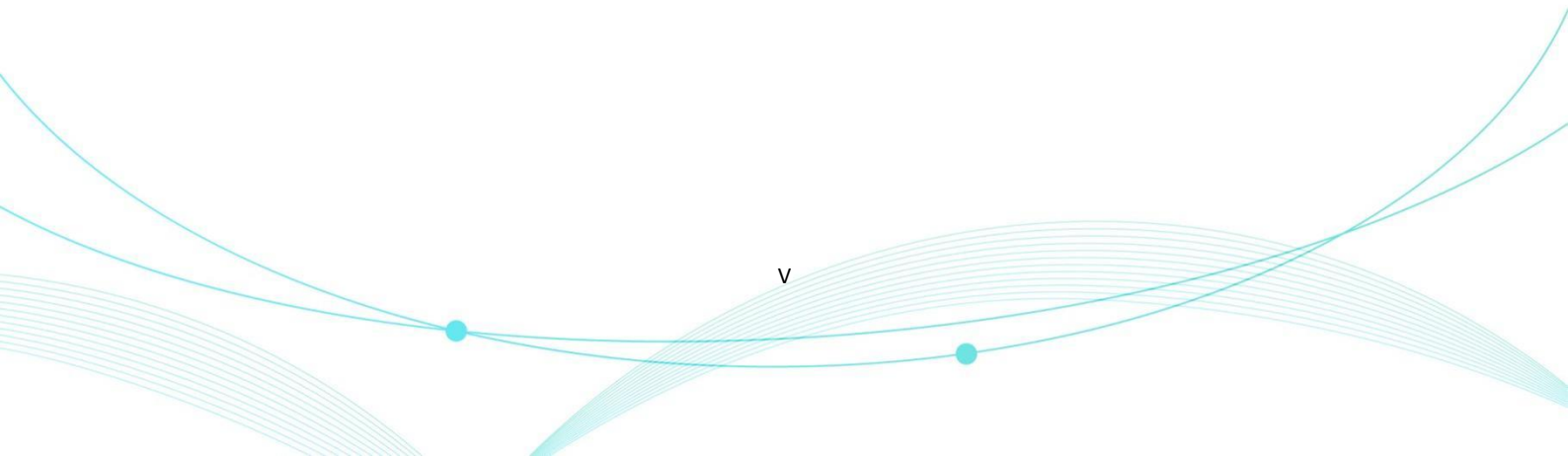
目录

版权声明.....	I
编制说明.....	II
前言.....	III
一、 分布式存储概况.....	1
(一) 分布式块存储.....	1
(二) 分布式文件存储.....	1
(三) 分布式对象存储.....	2
二、 分布式存储产业发展分析.....	2
(一) 全球分布式存储分析.....	2
1. 市场趋势.....	2
2. 产业政策.....	5
(二) 我国分布式存储分析.....	5
1. 市场趋势.....	5
2. 产业政策.....	7
三、 分布式存储技术分析.....	7
(一) 关键技术分析.....	8
1. 存储组网与协议.....	8
2. 存储部件.....	14
3. 系统软件.....	17
(二) 技术创新.....	20
1. 网络方面.....	20

- 2. 硬件设备..... 21
- 3. 软件系统..... 21
- 四、 分布式存储产业图谱..... 22
 - (一) 分布式存储产业上游..... 24
 - (二) 分布式存储产业中游..... 25
 - (三) 分布式存储产业下游..... 26
- 五、 分布式存储产业发展趋势..... 28
- 六、 总结..... 29



www.ODCC.org.cn



一、分布式存储概况

数据中心作为新基建的重要“底座”，是助推数字经济发展的关键力量。在国家战略的指引下，推进数据中心产业高质量发展，成为全行业“十四五”时期的重要任务。当前，随着数字化转型的不断深入，整个社会加速迈向数字化、网络化、智能化，加快了数字化转型的步伐。与此同时，大数据、人工智能、5G等新型技术的应用普及，让“数据洪流”更加汹涌，数据中心存储技术的创新和发展迫在眉睫。

存储设备基于存储架构的差异，主要分为集中式存储和分布式存储。集中式存储采用“控制框+硬盘框”架构，已经演进多年，系统的稳定性和增值功能逐渐强大。分布式存储基于通用存储服务器，与集中式存储相比，可扩展性较强，系统更易管理。据IDC发布的《数据时代2025》报告显示，预计2025年全球数据量达到175ZB，超过2020年数据量的3倍¹，数据已成为第五大生产要素。面对爆炸式增长的数据存储需求，尤其是非结构化数据，业界一致认为分布式存储是未来存储的发展趋势。

（一）分布式块存储

分布式块存储是通过分布式块存储软件将通用存储服务器的本地存储资源组织起来，构建全分布式存储池，支持大规模横向扩展的存储产品，为上层应用和客户端提供SCSI、iSCSI等标准的访问接口协议，满足云资源池及数据库等场景的存储需求。

（二）分布式文件存储

分布式文件存储采用分布式文件存储软件，将通用存储服务器的本地存储资源整合成统一的资源池，对外提供统一命名空间。通过NFS、CIFS、POSIX等

¹ 《ODCC 2020 开放数据中心峰会亮点剧透之全闪存阵列技术规范》

标准文件协议接口访问，具备大规模横向扩展能力，为用户提供非结构化数据共享存储资源，应用于视频、音频、高性能计算、视频监控等多业务场景。

（三）分布式对象存储

分布式对象存储通过分布式对象存储软件，组织通用存储服务器的本地存储资源，构建分布式存储池，为上层应用提供 S3 等标准访问接口协议。分布式对象存储具有高速直接访问和数据共享等优势，不但提供了高可靠性、跨平台性以及安全数据共享的存储体系结构，而且适用于存储海量的、非结构化的、读多于写的数据。

表1 分布式块、文件、对象主要技术指标对比

存储类型	常用接口	支持操作	冗余策略	特性	应用耦合程度
分布式块存储	SCSI、iSCSI	创建、删除、扩容、挂载、卸载卷等操作	多采用副本	-	业务与块存储部署在同一局域网
分布式文件存储	NFS、CIFS、FTP、POSIX	目录可以嵌套或重命名等；文件支持多用户并发读写、随机修改、重命名等。	多采用EC	-	业务与文件存储部署在同一局域网
分布式对象存储	S3 V2 REST API、S3 V4 REST API、openstack swift	支持对象常规操作、分段上传、多版本、生命周期管理等；一般不支持修改写。	多采用EC	支持丰富的语义、用户可自定义对象属性	应用与对象存储可不在同一局域网

二、分布式存储产业发展分析

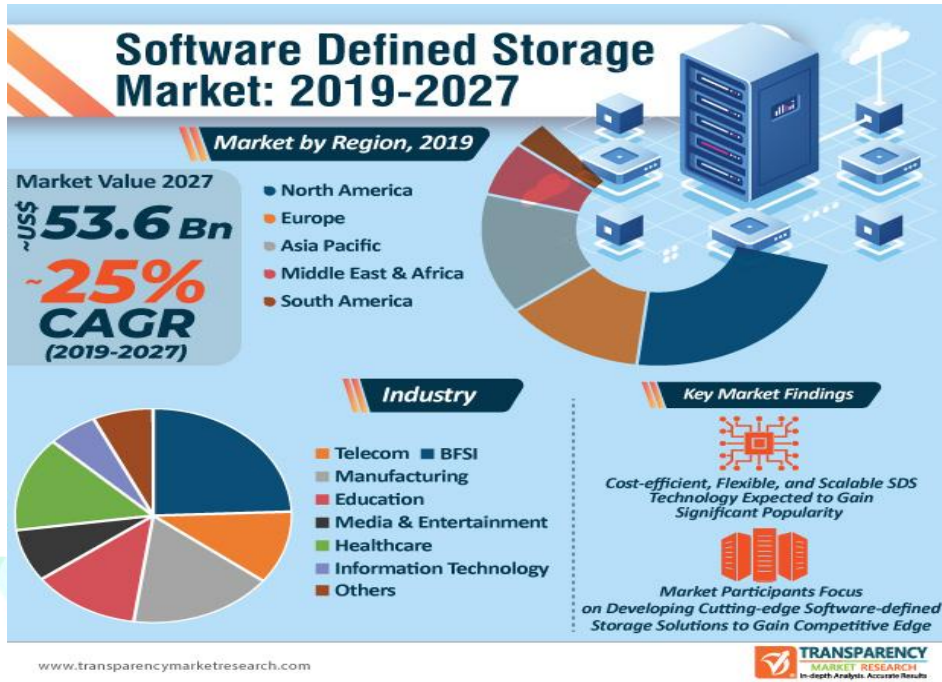
（一）全球分布式存储分析

1. 市场趋势

全球数据总量急剧增长，数据价值不断释放。随着云计算、大数据和人工智能等技术的发展，信息技术已经广泛融入到生产、生活、交易、社会治理等方面，数据的产生和应用更为广泛。在数据产生方面，据 IDC 预测，2018 年到 2025 年之间，全球产生的数据量将会从 33ZB 增长到 175ZB，复合增长率达到 27%，到 2025 年全球数据总量将达到 163ZB，相当于 2016 年的十倍。在数据总量快速增长的同时，数据资产的价值也在不断凸显，如通过大数据分析刻画用户画像，为用户提供更加优质的服务，实现精准营销；利用数据分析对生产参数进行精细控制，制造出质量更高的产品，实现精益生产等。

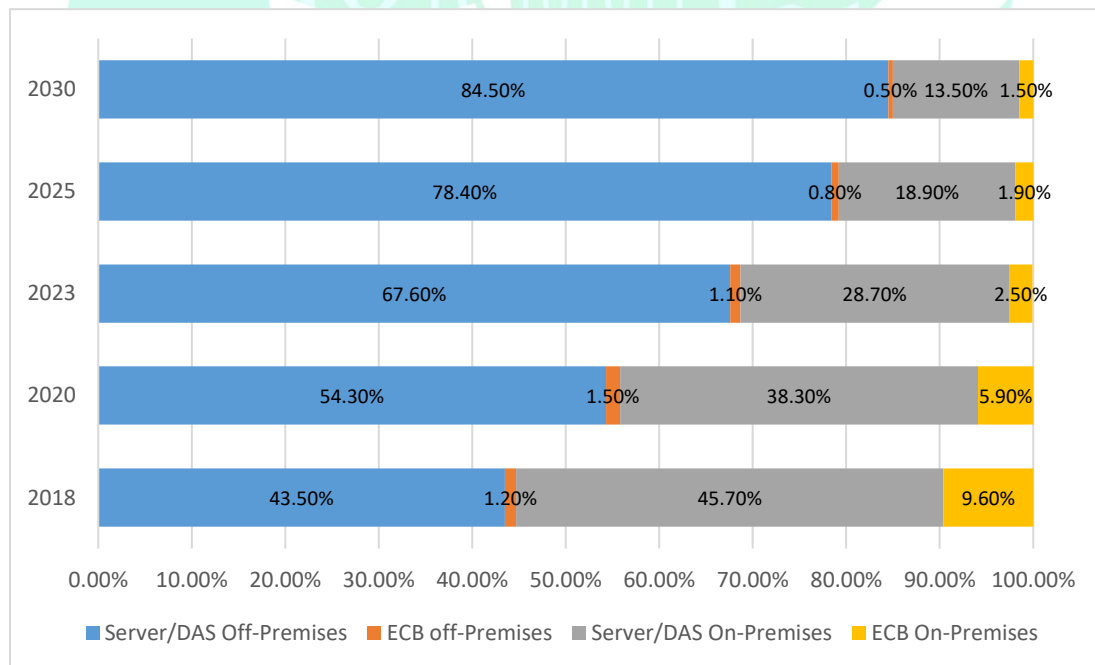
云端及边缘市场需求不断扩大，市场规模占比显著提升。近年来，移动终端及智能感应终端设备快速增长，且终端产品存储容量也在不断提升，但不可否认的是云端和边缘端依然是数据存储的重心。IDC 数据统计，2022 年云端存储占比将达到 47.6%，超过边缘存储成为最大的数据存储位置，到 2023 年云端数据存储容量将超过边缘存储和端侧数据存储的总和，占比达到总量的 50.4%。当前，云边端协同发展成为数字经济发展的关键趋势，世界主要国家都在加速推进数据中心、边缘数据中心的建设与发展。在云边协同战略发展的背景下，分布式存储将迎来新的发展机遇。全球分布式存储厂商开始进一步关注到互联网数据中心和边缘数据中心对存储产品的需求，并结合特定业务场景、建筑布局及运维管理需求开展存储产品研发设计。

软件定义的存储市场规模持续攀升。随着大数据、云计算以及人工智能的发展，推动了存储技术的重大进步。软件定义存储技术与网络连接存储等传统存储技术相比，具备自动化、灵活、可扩展的特性，成为存储市场的新兴力量。据 Transparency Market Research 发布的《Software Defined Storage Market》表明，从 2019 年到 2027 年，全球软件定义存储市场年复合增长率达到 25%，市场规模将达到 536 亿美金。北美和欧洲的存储市场受到政策引导和资金扶持，将持续保持高速增长，其中北美仍是软件定义存储的主要市场，如图 1 所示。同时，根据 Gartner 2020 年存储战略路线图预测，到 2024 年，全球存储容量的 50% 将作为 SDS 部署在本地或公共云上。与 2020 年相比，2024 年的大型企业将会存储在本地，边缘或公共云的文件或对象存储的非结构化数据将增加两倍，如图 2 所示。



数据来源: Transparency Market Research

图1 软件定义存储市场预估



数据来源: Gartner

图2 全球各种存储类型市场占比

2. 产业政策

美国政府发布多项政策培育和推动高端半导体技术发展。半导体存储是现代分布式存储产业发展的重要方向，美国政府高度重视半导体存储芯片在下一代信息技术中的应用前景，鼓励国内半导体厂商进行技术创新。2022年2月，美国发布《2022年美国竞争法案》（America COMPETES Act of 2022），强调对半导体芯片产业领域的支持和补贴。除了政策支持外，美国的行业协会也积极通过商业形式加强了存储产品的测试和认证，如美国存储性能委员会（Storage Performance Council）推出了SPC存储设备测试规范，对存储设备性能进行了全方面的评估，有助于推动美国及全球存储产业发展。

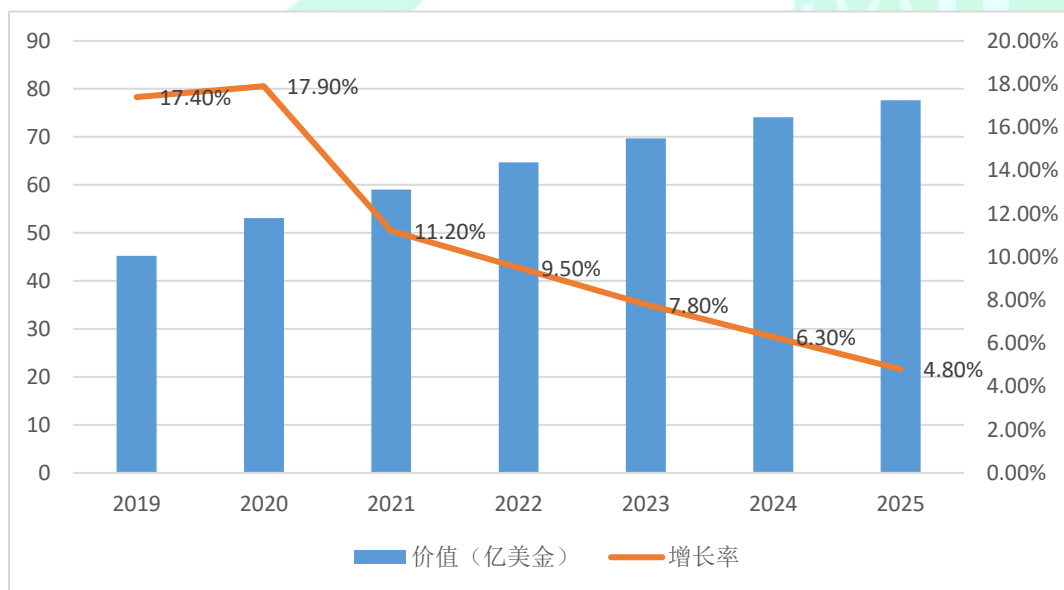
日韩政府积极出台产业支持政策，加速分布式存储产业发展。依托于美国存储产业发展，依靠人力成本及精细管理的优势，日本企业快速进入存储产业赛道。2021年5月，日本政府成立了旨在强化日本半导体产业的“推进半导体战略议员联盟”，构建半导体存储产业链，推进和美国的联盟合作，提升日本企业的竞争实力。2021年底，在批准的预算修正案中“半导体产业基盘紧急强化一揽子方案”获得7740亿日元的预算，涵盖半导体生产、半导体设备、5G通信等。频繁的半导体产业政策预示着日本半导体存储产业链将迎来高速发展新时代。与此同时，韩国半导体产业在20世纪90年代末进入到快速发展时期。2020年10月，韩国政府颁布了《人工智能半导体产业发展战略》，在“政府+大企业”的战略推动下，加强了“资金+技术+人才”的深度融合，为韩国半导体存储产业发展提供了强大的推动力。2021年5月，韩国政府发布了旨在主导全球半导体供应链的“K-半导体战略”文件，指出了政府和企业将在京畿道和忠清道建设半导体产业集群的规划战略，构建集半导体设计、原材料、零部件、尖端设备等生产体系，推动全球半导体供应链加速发展。

（二）我国分布式存储分析

1. 市场趋势

我国分布式数据市场规模持续攀升，大容量、高性能存储设备增长最快。随着我国新基建的全面开展，以及数字化转型的不断深入，数据中心、云计算、

工业互联网等建设规模将保持快速增长的态势，对于分布式存储的需求也将快速增长，为存储市场发展提供不竭动力。据 IDC 数据显示，2021 年我国存储市场规模达到 59.046 亿美元，比 2020 年市场规模增长 17.9%，未来 5 年存储市场仍将保持增长趋势，如图 3 所示。政府发布“十四五”规划和“新基建投资计划”，鼓励用户投资数据中心和基础设施，推进企业数据化转型，实现业务和应用数字化，增加数据消费和存储支出，拓宽我国存储市场规模。从细分领域看，为应对数据激增和时间敏感业务的挑战，适用于大规模数据存储、可扩展性较强的分布式存储成为支撑我国存储市场持续增长的存储技术。



数据来源：IDC

图 3 中国存储市场预测

新应用场景成为推动分布式存储发展的有力抓手。据 IDC 发布的《中国软件定义存储 (SDS) 及超融合存储 (HCI) 系统市场季度跟踪报告，2020 年第四季度》报告显示，2020 年软件定义存储与 2019 年相比增长了 51.7%，超融合存储系统增长了 36.5%。随着各行业新兴技术的不断涌现，激发了新应用场景和多元化产品的兴起，促进了中国存储市场的发展。从新应用场景角度来看，金融行业出现了如“远程柜台”、“人脸/指纹识别”等新型金融应用，实现了结构化数据向非结构化、多类型数据的转变；教育行业积极部署信息化和网络辅助教育，教学模式多样化如“线上教学”、“混合式教学”等形式对存储容量要求也不

断提高。针对多类型数据混合存储带来的存储性能问题，分布式存储系统能够有效解决。

2. 产业政策

新型基础设施建设推进不断加速，分布式存储助力新基建发展。2020年3月4日，中共中央政治局常务委员会召开会议，明确指出“加快5G网络、数据中心等新型基础设施建设进度”。2020年4月20日，国家发改委进一步明确新基建的范畴，包括信息基础设施、融合基础设施、创新基础设施。分布式存储作为重要的信息技术之一，是信息基础设施和融合基础设施的主要组成，是创新基础设施的重要赋能基础。以新基建中的数据中心为例，面向云计算、大数据、5G、工业互联网、移动互联网等应用需求，在《关于加快构建全国一体化大数据中心协同创新体系的指导意见》、《全国一体化大数据中心协同创新体系算力枢纽实施方案》、《新型数据中心发展三年行动计划(2021-2023年)》等政策助推下，全国数据中心建设规模快速稳定增长，拉动了我国分布式存储市场发展。同时存储设备及存储解决方案应用也进一步提升了数据中心基础设施的存储服务能力。

近年来，我国颁布多项政策法规支持分布式存储产业技术发展。2019年10月，工信部发布《关于政协十三届全国委员会第二次会议提案答复的函》中指出“工信部将继续支持中国工业半导体领域成熟技术发展，推动中国芯片制造领域良率、产量的提升。”2021年3月，《“十四五”规划和2035年远景目标纲要》不仅提出了加快数字新基建的建设目标，增强数据感知、传输、存储和运算能力，同时还提及了加快布局存储等前沿技术。2021年5月，国家发改委、网信办、工信部、能源局发布《全国一体化大数据中心协同创新体系算力枢纽实施方案》指出“加大分布式计算与存储等软硬件产品的规模化应用，支持和推广分布式数据操作系统等方面的平台级原创技术”。近年来，产业界对于分布式技术的研究也在逐步深化。众多存储产业政策支持文件出台，充分表明了我国对分布式存储产业及高端前沿存储技术发展的高度关注。

三、分布式存储技术分析

（一）关键技术分析

1. 存储组网与协议

（1）存储组网

存储系统根据组网方式的不同主要分为 IP（Internet Protocol）组网存储、FC（Fiber Channel）组网存储、IB（InfiniBand）组网存储等。

① IP 组网存储

IP 组网存储是采用以太网技术进行组网的存储设备，具备高兼容、低成本的优势，常见的速率包括 1Gb、10Gb、25Gb、100Gb 等。用户不但能够在 IP 环境下享受 SAN 环境优势，降低搭建 SAN 所需的光纤网络成本，构建 IP SAN 存储平台，提高用户体验。

② FC 组网存储

在信息化时代背景下，数据产生量呈爆发式增长，数据存储需求日趋增多，企业和服务商逐渐选择 FC 组网存储作为网络基础设施。FC 组网存储是采用 FC 光纤技术进行组网的存储设备，具备高效率、高灵活的特性，通过支持存储和服务器之间传输海量数据，为数据备份提供有效存储方式。

③ IB 组网存储

IB 组网存储是采用 InfiniBand 技术进行组网的存储设备，具备低延迟、高速率的特性，常见速率包括 40Gb、56Gb、100Gb、200Gb 等。IB 组网结构设计紧密，有效提高系统的性能、可靠性和有效性，缓解各硬件设备之间的数据流量拥塞。IB 作为一种“通用载体”技术，与其他以太网、LAN 以及 WAN 等标准兼容，具备解决大型集成问题的能力，促进应用领域的创新，满足用户的多样需求。

(2) 存储协议

①SATA 协议

SATA 是一种基于行业标准的串行硬件驱动器接口，以连续串行的方式传输数据，主要用于 SATA 主机与大容量存储设备之间的数据传输。SATA 协议通信模型如下图所示，分为应用层、命令层、传输层、链路层和物理层，每层有各自功能定义并彼此协作完成存储命令，具体每层的功能如表 2 所示。

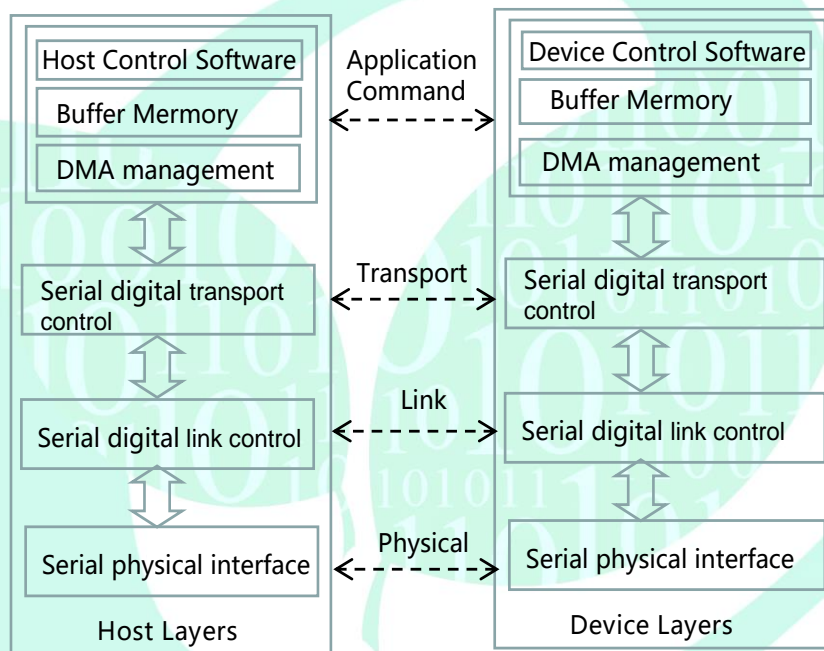


图 4 SATA 模型结构

表 2 SATA 协议通信各层功能

类别	定义和功能
物理层	<ol style="list-style-type: none"> 1. 具备以 1.5Gbps、3.0Gbps 或者 6Gbps 的速率传输 DNRZ 串行数据流； 2. 发送端和接收端提供 100 欧姆的终端匹配电阻； 3. 串化从链路层接受的并行数据进行发送； 4. 接受以及发送带外信号。
链路层	<ol style="list-style-type: none"> 1. 接收传输层传送的数据（Dword），计算数据的CRC校验值，将帧头、帧尾和CRC校验信息添加在数据上； 2. 接收对方链路层的数据信息，进行CRC校验，移除数据中的帧头、帧尾和CRC

	校验信息； 3. 帧流量控制。
传输层	1. 负责组帧和解帧。 2. 基于应用层的请求组建帧； 3. 将需传递的数据帧以及数据帧的内容发送到链路层； 4. 将传输状态报告到应用层。
命令层	1. NON_Data命令 2. PIO命令 3. DMA命令 4. NCQ命令
应用层	1. 接受来自主机端的命令，根据命令的要求将自身的信息发送给主机端。 2. 接收来自传输层的数据帧送入写FIFO中或将来自总线的数据保存在读FIFO； 3. 通知传输层构造数据帧。

②SAS 协议

SAS 接口由国际信息技术标准委员会 (INCITS) 的 T10 技术委员会开发和维护。SAS 标准将 SAS 架构分为 Physical 层、phy 层、链路层、端口层、传输层和应用层，每层有各自功能、定义，如表 3 所示。

www.ODCC.org.cn

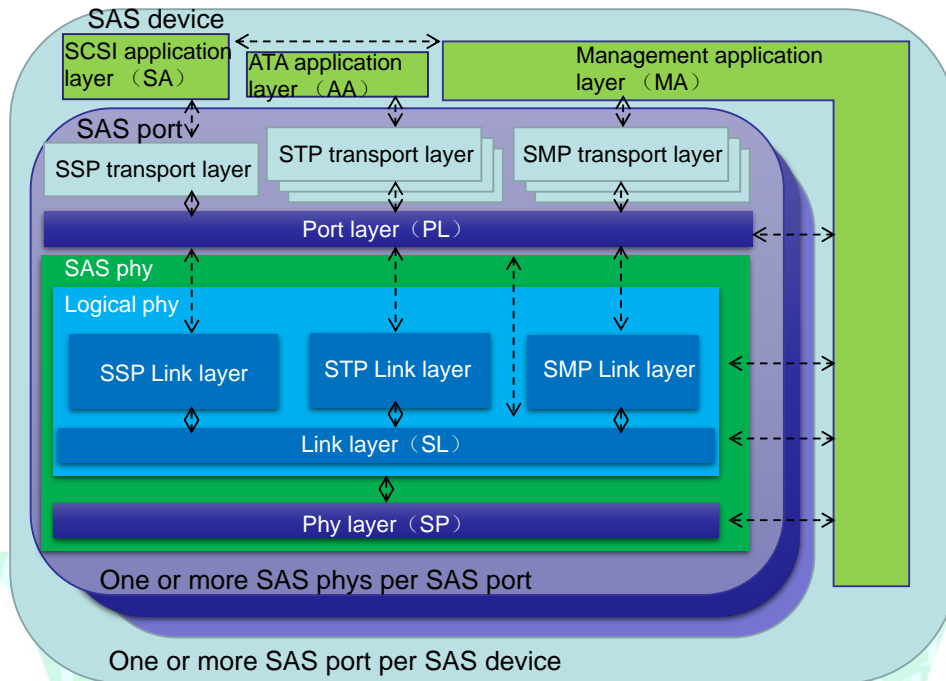


图 5 SAS 协议分层

表 3 SAS 协议各层功能

类型	定义和功能
物理层	<ol style="list-style-type: none"> 1. 对于线缆、接头、收发器等硬件进行了定义； 2. 完成传送数据的线缆和连接器的底层物理信号处理，多个 PHY 捆绑成一个逻辑接口，数据并行地在多个 PHY 中传输。
phy 层	<ol style="list-style-type: none"> 1. 包括 00B 和速率协商、8b/10b 编码、供电/复位序列； 2. 完成 8bit/10bit 编码处理
链路层	<ol style="list-style-type: none"> 1. 实现帧传输、原语和 phy 层连接管理； 2. 处理各种协议 primitives (原语)、地址帧格式、链路层对接 PHY 层和传输层执行设备识别、硬复位流程、链路建立与管理。
端口层	<ol style="list-style-type: none"> 1. 宽端口处理，包括如何请求、中断、如何选择建立连接； 2. 完成宽端口组合和处理； 3. 翻译和传递收发数据、命令和命令确认。

传输层	<ol style="list-style-type: none"> 1. 如何将所传输的命令、状态、数据封装在 SAS 帧中，以及如何分解 SAS 帧； 2. SSP、STP、SMP 协议在传输数据之前都需要首先由 Initiator 向 Target 端发起连接请求，同时请求也被映射到了链路层，链路层会建立相应的连接。
应用层	<ol style="list-style-type: none"> 1. 在不同类型的应用下使用 SAS 的细节； 2. 具备 SCSI 应用层 (SSP)、ATA 应用层 (STP) 以及管理应用层 (SMP) 的功能。

③ iSCSI 协议

iSCSI (Internet Small Computer System Interface, Internet 小型计算机系统接口) 技术将现有 SCSI 接口与以太网络 (Ethernet) 技术结合，具备高安全的优势。基于 TCP、IP 的协议连接 iSCSI 服务端和客户端，通过封装后的 SCSI 数据包在通用互联网传输，实现 iSCSI 服务端映射为存储空间并提供给客户端。iSCSI 主要分为 SCSI 层、iSCSI 层以及 TCP/IP 三层，如表 4 所示。

表 4 iSCSI 协议各层功能

类型	功能
SCSI 层	<ol style="list-style-type: none"> 1. 建立 SCSI CDB (命令描述块)； 2. 将命令快传递到 iSCSI 层； 3. 接收来自 iSCSI 层的 CDB； 4. 将接收到的数据返回到应用。
iSCSI 层	<ol style="list-style-type: none"> 1. 为便于在 TCP/IP 协议的网络上进行传递，对 SCSI CDB 进行封装； 2. 完成 SCSI 到 TCP/IP 的协议映射。
TCP/IP 层	<ol style="list-style-type: none"> 1. 对 IP 报文进行路由和转发； 2. 提供端到端传输。

④ InfiniBand 协议

InfiniBand 一种开放标准的高带宽、低延迟网络互联技术。InfiniBand 网络以极低的 CPU 负荷消耗，构建高性能数据中心内部高速网，通过远程 InfiniBand 互联设备，实现数据中心与数据中心之间的高速数据传输。InfiniBand 也是一种分层协议，模型如下图 7 所示，主要为应用层、命令层、网络层、传输层和物理层，每层负责不同的功能，下层为上层服务，不同层次相互独立，如表 5 所示。

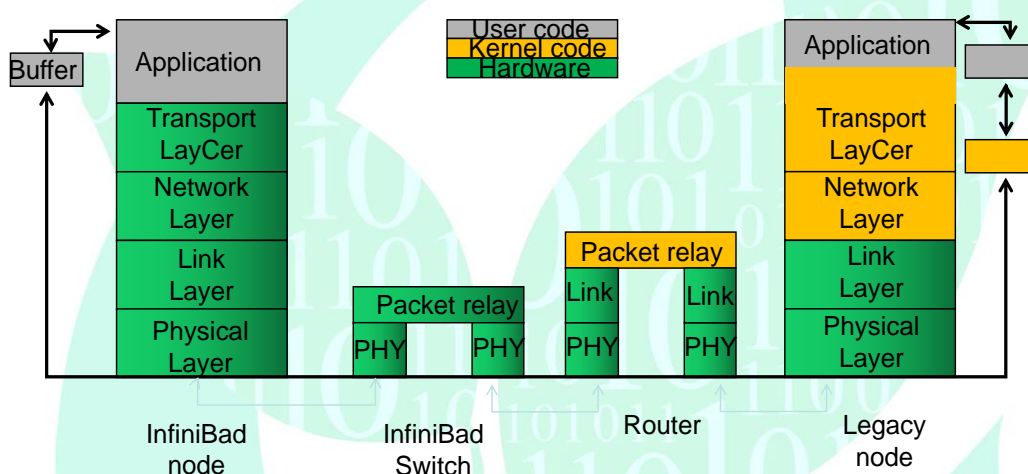


图 6 InfiniBand 协议分层

表 5 InfiniBand 协议通信各层功能

类别	定义和功能
物理层	1. 定义电气特性和机械特性； 2. 构建有效包的信令协议。
链路层	描述数据包的格式和数据包操作的协议
网络层	1. 实现子网间的数据路由； 2. 数据包中包含全局路由头GRH，用于子网间数据包路由转发。
传输层	1. 负责报文的分发、通道多路复用、基本传输服务 2. 处理报文分段的发送、接收和重组 3. 将数据包传送到各个指定的队列(QP)中；

	4. 指示队列如何处理该数据包。
上层应用层	1. 根据用户类型提供上层协议 2. InfiniBand主要支持SDP、SRP、iSER、RDS、IPoIB以及uDAPL等上层协议

2. 存储部件

(1) 机械硬盘

机械硬盘主要由盘片，磁头，盘片转轴及控制电机，磁头控制器，数据转换器，接口，缓存等几个部分组成。机械硬盘根据接口、转速、规格、气体、磁碟磁录方式以及扇区的不同分为多种类型，如表6所示。

表 6 机械硬盘类型

接口	SATA、SAS
转速 (rpm)	5400、7200、10.5k、15k
规格	2.5寸、3.5寸
气体	空气盘、氦气盘
磁碟磁录方式	PMR (CMR)、SMR、MAMR、HAMR、HDMR
扇区	512n、512e/4Kn、4Kn

随着新型技术的不断发展，硬盘界也酝酿了一场全新的革命浪潮，其中磁盘为机械硬盘更新换代中的主要技术。磁盘的关键技术主要涉及提升容量和效能和管理功能，其包含的主流技术为新型磁头技术、S. M. A. R. T (Self-Monitoring、Analysis And Reporting Technology)、SPS (Shock Protection System, 震动保护系统) 等。同时，随着科学技术的不断进步以及信息技术的加速发展，未来将会催生出更多的新兴技术。

①新型磁头技术

MR (Magneto—Resistive Head, 磁阻磁头) 为新型的磁头技术, 与传统技术相比, 具有更高的实际记录密度、记录数据, 增加硬盘容量, 提高了数据吞吐率。GMR (Giant Magneto Resistive, 巨型磁阻磁头) 是利用电阻值不断随磁场变化的原理, 读取盘片上的具体数据, 通过磁阻效应和多层薄膜结构, 满足更高的存储密度需求。

②S. M. A. R. T 技术

S. M. A. R. T 技术是一种自动的硬盘状态检测与预警系统, 通过硬盘内的检测指令, 对硬盘硬件的运行情况进行监控, 并分析电路、磁盘以及磁头的不同状态。当设备出现异常时, S. M. A. R. T 技术通过发出警告、自动降速以及备份数据的措施, 保障硬盘数据的安全。

③SPS 技术

SPS (Shock Protection System, 震动保护系统) 防震技术主要基于用户在安装或其他非操作状态下, 硬盘发生震动、撞击时所设计的保护技术。当硬盘遇到撞击时², 磁头和磁头臂在盘片上停泊, 硬盘其他的部分能够将所受到的冲击能量吸收, 提升了硬盘的抗震性能, 最大限度的降低硬盘在运输、使用过程中受到的震动损坏, 减少硬盘数据丢失和出现错误的概率。

(2) 固态硬盘

固态硬盘主要由固态电子存储芯片阵列制成, 与HDD相比具备高可靠性、高性能、低功耗以及大空间等方面优势, 在车载、电力以及网络监控等领域广泛应用。固态硬盘涉及的关键技术主要为先擦后写、FIT、垃圾回收、预留空间、磨损平衡以及Trim功能等。

² 《浅析计算机硬盘技术发展与应用》

表 7 SSD、HDD 对比

优势	HDD	SSD
可靠性	机械部件，振动时面临更大的数据丢失风险。	无机械部件，更可靠。
性能	平均 2ms 延迟 175-250IOPS	延迟 20us 低高达 850000IOPS
功耗	发热量散热成本高	发热量散热成本低
空间	接口单一	接口丰富

①FTL

FTL (Flash Translation Layer 闪存转换层) 技术针对基于块映射、页映射以及混合映射等三种不同的映射方式，完成 Host 的逻辑地址到 Flash 物理地址的转换。FTL 在垃圾回收 (GC)、数据保持 (Data Retention) 处理、磨损平衡 (Wear Leveling)、读干扰 (Read Disturb) 处理、坏块管理等方面具有主要应用，成为决定 SSD 在性能、耐用性、可靠性等方面的核心技术。

②GC

GC (Garbage Collection, 垃圾回收) 是将现存数据重新转移到其他闪存位置，同时彻底擦除无效的数据的过程。由于 NAND 闪存设备不能覆盖现有数据，所以操作系统删除文件时，无效的数据仍然保留在硬盘上，导致产生大量的失效文件即数据垃圾。GC 能够有效删除大量的数据垃圾，提升 SSD 的利用效率，成为影响 SSD 寿命和性能关键技术之一。

③OP

OP (Over provisioning 预留空间) 指用户不可操作的容量，为实际物理闪存容量减去用户可用容量，通常被用于优化操作，包括 WL (Wear leveling 磨损均衡) 和坏块映射等。通过增加 OP 可以提高 SSD 的 WL (Wear leveling 磨损均衡)，降低写放大，提高随机写的性能，增加闪存使用寿命。通过 OP 比例

分析 WA（写入放大）越小，则闪存损耗越小，从而闪存使用寿命延长。因此，OP 越大写放大越小，SSD 写性能越优。

④磨损平衡

磨损平衡技术是确保 NAND 中的所有块以近乎相同程度进行磨损，依托于 P/E 次数均衡的原则，当新数据被写入到 NAND 时，可有效避免 NAND 块被持续使用磨损，达到磨损均衡的目的。磨损平衡分为动态均衡和静态均衡两种形式。当冷数据被存储于 NAND 块中时，其他 NAND 块被持续使用，P/E 次数越来越高，静态均衡机制将利用存储冷数据 P/E 次数低的 NAND 块。当满足静态均衡条件时，静态均衡机制将块上的冷数据搬移至新的块，而原来 P/E 次数低的块将用于频繁更新的热数据，达到均衡磨损的目的。

3. 系统软件

分布式存储技术随着新兴技术的加速演进，呈现了多种技术形式，如 HDFS、Ceph、GFS、GlusterFS 及 Swift 等技术成为当前存储技术的发展主流。同时，随着云计算、大数据等技术的不断发展，更多新的分布式存储技术及平台仍在涌现。

(1) Ceph

Ceph 经过多年的发展，受到众多存储和云计算厂商的重视，成为了应用最广泛的开源分布式存储平台。Ceph 利用存储节点计算能力，计算得出该数据存储的位置，将数据分布均衡。同时，Ceph 采用了 CRUSH、HASH 等算法，杜绝了传统的单点故障问题，实现水平扩展的性能。

Ceph 的特点如下：

- (1) Ceph 支持对象存储、块存储和文件存储服务；
- (2) 采用 CRUSH 等算法，数据分布均衡，并行度高；
- (3) 支持多个副本，并保持多副本间的强一致性；
- (4) 去中心化，MDS 之间地位相同，无固定的中心节点。

(2) GFS

GFS (Google File System) 适用于大量的顺序读取和顺序追加，注重大文件的持续稳定带宽。2013 年，Google 公布了 Colossus 项目，作为下一代的 Google 文件系统，与前一代的 GFS 相比，Colossus 增加了主控服务器集群，重新平衡了旧的冷数据，并在磁盘上进行新数据的分发。

GFS 的特点如下：

- (1) 适合大文件场景的应用以及适数据访问延时不敏感搜索类业务；
- (2) 中心化架构，只有 1 个 master 处于工作状态，Colossus 有改进后，将单一主控服务器改造为多主控服务器构成的集群，将所有管理数据进行数据分片后分配到不同的主控服务器上，增加可用性；
- (3) 缓存和预取，通过在 client 端缓存元数据，尽量减少与 master 的交互，通过文件的预读提升并发性能；

(3) HDFS

HDFS (Hadoop Distributed File System) 是适用于硬件上的分布式文件系统。作为是 Hadoop 的核心子项目，HDFS 具备海量数据存储、高容错性以及大文件存储的优势。

HDFS 的特点如下：

- (1) 适用于大文件、大数据处理，处理数据达到 GB、TB、甚至 PB 级别的数据；
- (2) HardLoop1.X 版本支持 NameNode 及 Secondary NameNode，SecondaryNameNode 会周期性的将 fsimage 和 edits 合并，然后将最新的 fsimage 推送给 NameNode；
- (3) HardLoop2.0 版本支持 Standby NameNode，采用主从备份架构，对 Active NameNode 进行热备份；
- (4) 多副本保护提高容错性；

(5) 不支持并发，同一时刻只允许一个写入者或追加者。

(4) GlusterFS

GlusterFS (Gluster File System) 是一种 Scale-out 的文件存储系统，具备强大的横向扩展能力，支撑数 PB 存储容量以及处理数千应用客户端的优势。

GlusterFS 的特点如下：

- (1) 采用无中心对称式架构，不存在元数据服务器瓶颈，元数据存在于文件的属性和扩展属性中；
- (2) 采用弹性哈希算法，具有很好的可扩展性；
- (3) 采用模块化堆栈架构，可实现灵活配置。

(5) SWIFT

SWIFT 最初是由 Rackspace 公司开发的分布式对象存储服务，采用完全对称、面向资源的分布式系统架构设计，所有组件都具备扩展性，降低因单点失效而影响整个系统的可用性的效率，主要用于解决非结构化数据存储问题。

SWIFT 的特点如下：

- (1) 原生的对象存储，不支持实时的文件读写、编辑功能；
- (2) 完全对称架构，无主节点，无单点故障，易于大规模扩展，性能容量线性增长；
- (3) 数据实现最终一致性，不需要所有副本写入即可返回，读取数据时需要进行数据副本的校验。

本文根据存储方式、主控服务器、元数据节点数量、分块大小、是否开源等 6 个维度分析 HDFS、Ceph、GFS、GlusterFS 及 Swift 等技术的差异性，如表 8 所示。

表 8 主流技术对比

维度	Ceph	GFS	HDFS	GlusterFS	SWIFT
存储方式	块/文件/对象	文件	文件	文件	对象
系统架构	去中心化	中心化	中心化	去中心化	去中心化
主控服务器/ 元数据节点 数量	多个	1	2（主备）	无	多
分块大小	默认 4M（可调）	64M	128M	默认条带 128K	存储对象
是否支持数 据冗余	多副本	多副本	多副本	多副本	多副本
是否开源	开源	非开源	开源	开源	开源

（二）技术创新

分布式存储技术在网络、设备以及软件系统方面迭代创新，呈现出高性能、易于扩展、服务化和智能化等优势。

1. 网络方面

NVMe（NVM Express）协议凭借高吞吐、低延迟的优势成为主流的存储协议之一。HDD磁盘和早期SSD磁盘的传输协议一般采用AHCI（高级主机控制器接口，Advanced Host Controller Interface）。随着存储介质的不断演进，AHCI协议已无法满足高性能和低延时存储的需求，存储系统迫切需要更快、更高效的协议和接口，NVMe（NVM Express）协议应运而生。NVMe协议具备高吞吐、高IOPS以及低延迟的优势。在吞吐量方面，基于NVMe的驱动器的吞吐量可实现高达16Gbps，未来32Gbps或更高吞吐量产品的应用将成为未来发展的主流趋势。在I/O方面，许多基于NVMe的驱动器，其IOPS可以超过50万，部分可提供150万、200万甚至1000万IOPS。在低延时方面，许多驱动器的延迟低于20微秒，部分低于10微秒。因此，NVMe SSD凭借卓越的性能，在2018到2023年之间以38%的复合增长率加速增长，成为云客户、OEM厂商和企业存储数据的首选设备。

在网络协议层，随着闪存介质的普及，SCSI协议框架对性能的限制也越来越突显。NVMe和NVMe Over Fabric（NVMe-oF）技术的出现打破限制，推动IP化、

低时延化。基于IP网络的NVMe-oF技术不但前端存储网络基于IP直接与本地局域网连接，而且也能够直接与广域网相连。同时，利用NVMe-oF技术小于10us的超低附加时延优势，增加了替换后端SAS网络的可能性。在数据中心建设中使用统一的以太网进行构建，不但降低了整个数据中心的建设成本和运营成本，而且提高了云及大数据应用环境下的数据共享。如华为NoF+存储网络在联机事务以及分析处理过程的场景中，IOPS最高提升85%，拥塞时延最大降低46%，端到端故障切换时间<1s，有效提升了网络性能。

2. 硬件设备

全闪存储的高速发展，既是新应用对性能需求驱动的结果，也是闪存技术不断创新的成果。在需求侧，随着数据爆炸式增长，云计算、物联网、区块链等新技术发展不断积累，EB级容量、亿级IOPS以及数据智能管理加速了存储介质的变革。在技术侧，全闪存储与传统的存储相比，具备更快的读取速度、更低的功耗以及更低的故障概率优势，完成了对传统存储性能的全面超越。在产品侧，各大存储厂商均推出了全闪存储产品。

EDSFF (Enterprise & Data Center SSD Form Factor, 企业和数据中心固态硬盘规格) 以高密度、易管理、高效率的优势成为固态硬盘外观发展的新趋势。EDSFF 是面向云服务、企业数据中心的 NVMe SSD 新形态标准，与传统 SSD 相比具备高吞吐、低延迟、强扩展性的特性。同时在性能、散热、功耗等方面都做统一的改进，更易适用于大容量、高密度的全闪存储阵列。

3. 软件系统

随着存储介质、存储网络以及软件技术的发展，出现了一些新型存储软件系统，其中IPFS和LightbitsOS技术极具代表性和先进性。

(1) IPFS

以IPFS为代表的分布式存储系统为下一代基础设施提供存储新思路。

IPFS(The InterPlanetary File System, 星际文件系统)是一种旨在创建持久

的分布式存储和共享文件的网络传输协议的分发网络协议，与传统存储协议相比具备更安全、更快速、低成本的优势。IPFS网络中的节点构成分布式文件系统，通过分布式的多点传输，大幅度提高IPFS的传输速度。同时分布式存储系统与去中心化的区块链技术高效融合，使得中心化互联网架构下的数据存储的安全性得到有效保障，解决了用户协作的时效性问题。

(2) Lightbits OS

Lightbits Labs的LightOS™是一种软件定义的块存储解决方案，与传统技术相比，具备高性能、高灵活、高扩展的特性，通过采用NVMe/TCP存储解决方案，简化基础设施的运营和管理，大幅降低企业的生产成本。从高性能角度分析，LightOS具备冗余功能的解耦合存储解决方案，通过启动驱动器冗余以及服务器冗余的单台Light OS存储服务器，提供高达300万的4K随机读IOPS和80万的4K随机写IOPS。LightOS与传统技术相比，具备高IOPS和高带宽的优势，成为数据库、数据分析以及大量可扩展云原生应用的首选存储介质。从低成本角度分析，NVMe存储应用服务器通常利用率为15%-25%，而使用LightOS存储服务时，在容量和性能利用率方面，与NVMe存储服务器相比，具有极大的提升。

四、分布式存储产业图谱

本文根据各产业中企业之间的供给与需求的关系将分布式存储产业链划分为上游、中游以及下游。产业链上游包含存储芯片、元器件和核心软件等制造商，中游为分布式存储系统厂商，下游为各行业应用客户，主要包含为政府、金融、电信、医疗等应用领域，如图7所示。



图 7 存储行业产业链

在存储行业产业链条中，每个环节都有术业专攻的企业，各企业以自身核心优势为切入点逐步向存储行业产业链渗透并拓宽业务范围，积极构建生态圈。本文选取各环节代表性较强的企业，形成如下图谱。

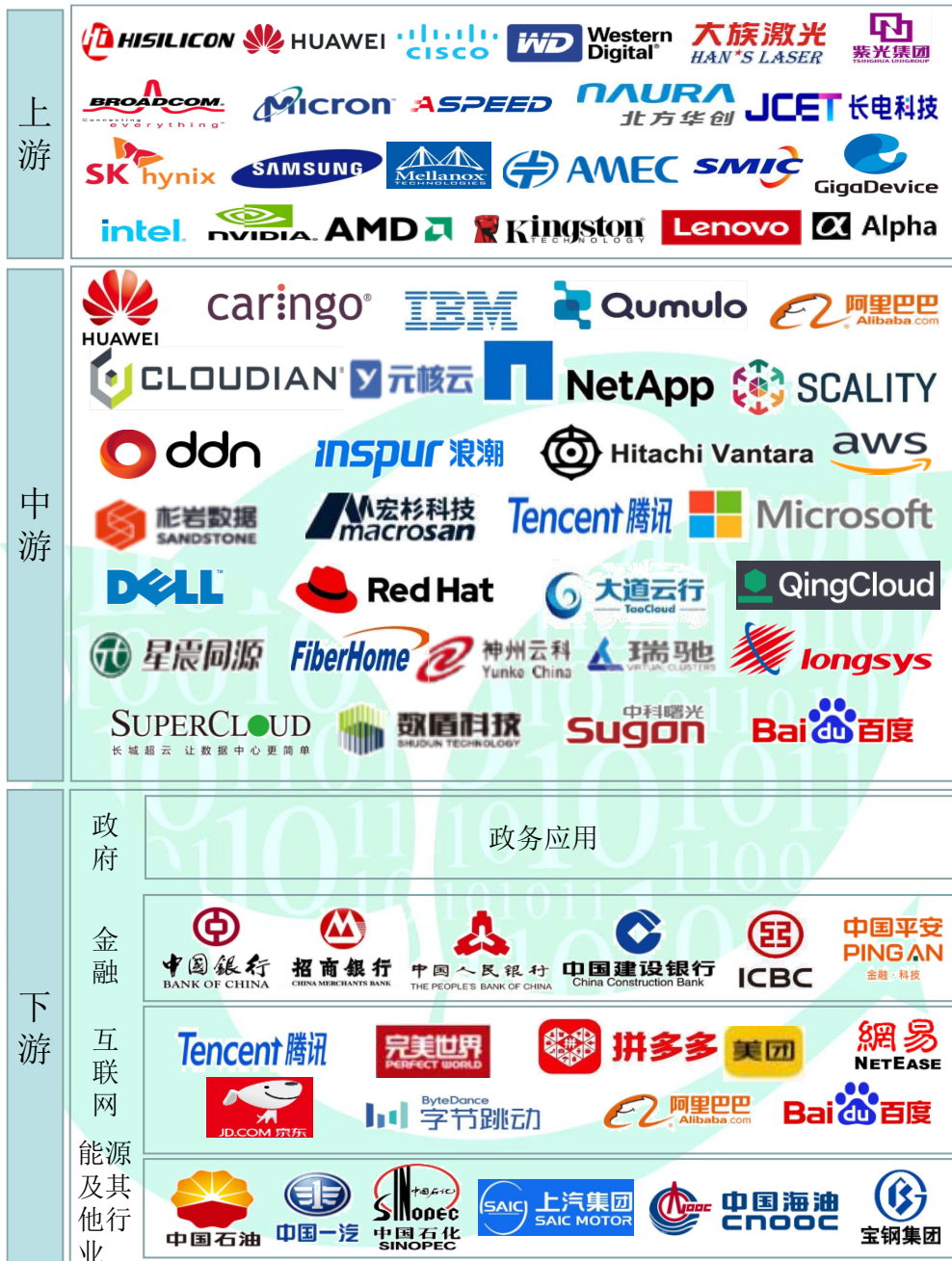


图 8 分布式存储产业图谱

(一) 分布式存储产业上游

分布式存储产业上游处于整个产业链的开端，是原材料、零部件制造以及生产的提供者。上游主要包括存储芯片、元器件以及核心软件等，具有基础性、原料性强的特点，决定着分布式存储产业的创新速度，成为产业链发展受限的主要环节。



图 9 存储芯片、元器件、核心软件相关企业图谱

存储芯片市场规模不断扩张，推动存储产业规模加速发展。分布式存储上游经过多年发展，已经具备较高的技术成熟度，分布式存储元器件、核心软件等市场格局较为稳定，存储芯片作为存储运算的大脑，成为限制存储产业规模加速发展的主要因素。据研究机构数据分析，2023 年全球存储芯片市场的规模，预计将达到 2196 亿美元，较之前预估的 2022 年增加 392 亿美元，同比增长达 21.7%，其中三星、东芝、西部数据、美光以及 SK 海力士等企业占据市场 9 成以上的份额，我国企业占据市场份额较低。因此，近年来国家通过政策引导和产业资金扶持，鼓励存储芯片企业加强技术研究，实现中国存储芯片自主研发，如长江存储、长鑫存储、兆易创新、紫光国芯以及中芯国际等企业在存储芯片企业研发动力不断增强，尝试打破技术瓶颈，提高存储芯片技术水平，拓宽存储芯片应用场景，提升产品自给率，推动存储芯片行业的发展。据相关研究数据显示，到 2026 年，中国存储芯片市场规模有望达到 5598 亿元，占全球存储芯片消费市场比重达到 40%左右。

（二）分布式存储产业中游

分布式存储产业中游主要是分布式存储系统，包含存储整机、系统集成和存储服务，利用上游的芯片、元器件和核心软件组装成具有数据存储功能的存储系统。



图 10 分布式存储产业中游企业图谱

分布式存储产业中游企业受到上游政策变动的影 响，以及下游客户的高技术、高要求，导致竞争加剧。随着新兴技术的加速发展，市场需求发生颠覆性变化，中游企业面对当前市场格局采用多种竞争方式拉动企业的发展。华为和宏杉等公司通过打通上游，自主研发设计核心软件、控制器等部件，并与上游厂商其他通用部件进行组装，形成一体化模式，提高自身竞争实力；Cloudian、DDN (DataDirect Networks) 以及浪潮等各大厂商提供数据存储服务、云数据中心、云服务，为块、文件和对象存储提供解决方案；阿里、百度等大型企业，纷纷通过自建数据中心为客户提供集成的存储服务。中小型企业加入分布式存储数据中心的建设，对客户 提供数据存储和研发产品等服务。未来客户应加强参与标准制定和研发工作中，对不同软件运维和使用场景提供便捷的存储服务，提高企业对的数据隐私、安全、可靠以及性能的要求。

(三) 分布式存储产业下游

分布式存储产业下游处于产业链的终端，主要涉及政府部门、教育、电信、金融、互联网等应用行业，和整个存储产业的影响和互动愈加密切。



图 11 分布式存储产业链下游图谱

伴随信息化建设与应用的深入，市场持续高涨，分布式存储产业在政府、金融、电力行业、制造业及其他行业用户不断增长。针对政府层面，在国家政策大力引导和产业各界的共同推动下，数字政府市场加速扩大，政府用户对数据存储需求明显提升。国家“十四五”规划纲要提出，对提高数字政府建设水平作出战略部署，将数字技术广泛应用于政府管理服务，推动政府治理流程再造和模式优化，不断提高决策科学性和服务效率³。近年来，我国各地数字政府建设进程不断加快，并且已经进入全面提升阶段，预计 2025 年数字政府行业市场规模仍将稳定提升。针对金融领域层面，金融领域业务的数据量激增，导致数据存储需求不断提升。随着大数据、区块链和人工智能等创新前沿技术的快速发展，我国正经历着从金融互联网化向金融智能化的演进，业务越来越多样化，数据规模也呈爆发式增长，因此对海量数据的有效管理、业务敏捷上线等一系列的要求促使存储技术加速提升。针对其他领域层面，数字化转型在电力行业、制造业及其他行业不断深入，对数据存储的需求不断增加。为确保关键应用的稳定可靠，传统的存储技术已无法满足海量数据的存储需求，分布式存储技术有效的应对 PB 级别数据量的存储，为敏态 IT 提供数据底层。随着互联网信息

³ 《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》

技术的不断发展，电力行业、制造业等行业的改革进程加快，产生的新应用新业务也继续更多的数据存储，对分布式存储未来的需求也将进一步扩大。

五、分布式存储产业发展趋势

多层技术融合，拓展应用创新空间。针对企业数据存储服务需求更加复杂、精细和特征化的业务需求，分布式存储基于自动化以及易于拓展的特征，满足企业高可靠、高性能、高扩展的存储数据要求。在业务层面，分布式存储技术不限制上下层业务形式，提供块存储、文件存储、对象存储等服务，实现数据的协议互通。因此推动多层存储技术融合，降低行业应用开发部署门槛和成本，为存储行业应用提供基础性产品，提升数据存储服务产品力和创新力，提高行业应用创新效率。

生态逐步发展，产业规模持续增长。据中国信通院《中国数字经济发展白皮书》显示，2020 年我国产业数字化规模达 31.7 万亿元，占数字经济比重 80.9%，成为数据经济发展强大的引擎。海量数据的计算和存储均需存储技术提供数据服务，因此存储行业逐渐形成涵盖技术设备、系统平台、业务应用服务等全方位产业链体系。多方企业利用技术平台结合行业特色，联合打造服务完善高低搭配的企业梯队。IDC 数据统计，到 2022 年核心存储占比将达到 47.6%，2023 年云端数据存储容量将超过边缘存储和端侧数据存储的总和占比达 50.4%。

发展环境向好，政府加强扶持力度。分布式存储作为“十三五”时期战略性新兴产业重点培育的五大产业中“新一代信息技术产业”的重要组成部分，属于整个数据产业的基石，为推动整个产业的发展，国务院及有关政府部门先后颁布了一系列支持政策，加大研发投入力度，加快数据基础设施建设，加强数据安全保护，为存储行业发展提供了政策保障和扶持，营造良好的发展环境。2020 年 4 月 3 日，国家科技部表示加大研发分布式存储、先进计算，核心软件，宽带通信，人工智能等前沿技术，支撑国家新型基础设施建设。2020 年 10 月，《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》进一步明确“推进数据要素市场化改革、加快数字化发展”，将

“推动超大规模分布式存储技术创新”写入到“数字经济重点产业”的第一项。2021年5月24日，国家发改委、网信办、工信部、能源局发布关于印发《全国一体化大数据中心协同创新体系算力枢纽实施方案》的通知，在国家枢纽节点重点任务部分指出，要加大分布式计算与存储等软硬件产品的规模化应用，支持和推广分布式数据操作系统等方面的平台级原创技术。

六、总结

在国家政策重点关注、数据中心市场加速扩张、数据价值被进一步挖掘的背景下，各行业对海量数据存储系统的需求不断扩大，分布式存储产业未来发展空间广阔。近年来，国家明确将“新基建”作为未来经济增长的驱动力，新基建的核心任务是“加快5G网络、数据中心等新型基础设施建设进度”。分布式存储作为重要的信息技术之一，对加快数字化转型、促进数字经济发展具有重要的作用。相信，随着分布式存储技术创新不断深入，未来分布式存储产业也将迎来更大的发展机遇。

www.ODCC.org.cn



ODCC服务号



ODCC订阅号