

中国存力白皮书(2022年)





前言

纵观全球,新一轮的科技革命和产业变革正在全方位的改变着社会生产生活,积累数据资源、保障数据安全、做大做强数据产业,已经成为全球主要国家共同的战略选择。

党的十九届四中全会首次将数据作为生产要素参与分配,数据的作用受到国家高度重视,作为数据载体的存储也越来越重要。云计算、大数据、人工智能、高性能计算等新兴数字产业和前沿技术的加速应用,数据总量大幅增长,数字经济持续繁荣,起到重要支撑作用的存储需求也在不断扩大。不同应用场景需求对数据存储容量、数据传输速度、存储性能等各方面要求不断提升,产业需求持续推动存储技术演进,需求更新与技术迭代共同加速存储市场进一步释放增长空间。

本白皮书综合考虑存储的容量、性能表现、安全可靠、绿色低碳等特性,分析全球存储行业的产业现状,研究数据中心存储技术发展趋势,分析中国各省份存力发展规模和质量。总体来看,我国数据存力发展总体水平与发达国家相比还有一定差距,存储容量、先进存储设备应用、存算均衡等方面都需要进一步加强。我国区域存力情况与数据中心和算力的发展高度相关,东部各省份数据存力发展水平较高;中西部各省份虽冷数据存储业务占比较高,但其整体存储规模和存储质量都低于东部。最后,在我国数据存储产业发展趋势进行展望的同时,也从技术、标准和产业方面对我国数据存力高质量发展提出建议。

白皮书仍有诸多不足,恳请各界批评指正。如有意见建议请联系中国信通院云大所数据中心团队,邮箱:dceco@caict.ac.cn。



目 录

背景	·与意义	1
1.1	数据存力的定义与内涵	1
1.2	数据存力的重要作用	2
数据	存储产业发展分析	4
2.1	政策分析	4
2.2	产业分析	6
2.3		9
中国		12
3.1	我国数据存力发展情况分析	12
3.2	区域数据存力发展情况分析	12
3.3		
3.4	数据中心存力发展情况分析	15
中国		
4.1	Carlly related	
4.2	数据存力高质量发展建议	20
	1.1 1.2 数据 2.1 2.2 2.3 中国 3.1 3.2 3.3 4	1.1 数据存力的定义与内涵



背景与意义

1.1 数据存力的定义与内涵

在国家数字经济大战略下,数据作为生产要素,已成为重要的资产。海量数据呈指数级增长,数据流动加速,存储作为承载数据的关键设施,其重要性更加凸显。数据存储是一个复杂的系统,包括复杂的硬件、复杂的软件和介质。各行各业在数字化进程中涉及到的应用创新对数据的采集、存储、传输、交易和服务等环节提出了更高的要求,不仅强调存储容量足够大,也对存储的功能、可靠性、安全性、绿色性等提出了要求。

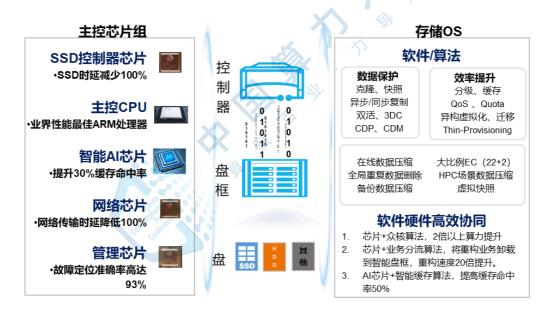


图 1 数据存储系统

数据存力成为数字经济背景下,全社会全产业发展的重要基础。数据存力的概念和内涵也更加丰富, 是**以存储容量为核心,包含性能表现、安全可靠、绿色低碳在内的综合能力**。

存储容量是应对数据爆发式增长的基础。进入大数据时代,移动智能设备、云端、物联网终端等产生的海量数据需要存储、分析、计算,数据存储的整体需求快速增长,中国信通院发布的《数字中国发展报告(2021年)》指出,2017年到2021年,我国数据产量从2.3ZB增长至6.6ZB。数据"存不下"的问题日益严重,不断增长的数据总量使存储容量成为数据存力的核心诉求。为此,业界积极探索存储容量提升方



案,一方面,通过存储硬件结构性改造和高科技散热技术增加硬盘密度,从而增加单位空间的数据存储容量,另一方面,软件纠删码等技术的应用也可降低冗余度、提高磁盘利用率。

数据存储性能是数据有效处理的关键。在涉及国计民生的重要领域,如政务、金融、交通、电信、医疗等行业,对存储读写速度快、非易失性、无机械装置、抗震性好、适温宽、功耗低、传噪少等提出了较高要求。数据存储性能是充分释放数据要素价值的基础,是数据有序流转、运营流通和场景应用的有力支撑。在新兴信息技术领域如人工智能,数据存储性能如读写速度,决定了人工智能算法能否快速获取足量训练数据,从而不断优化算法打造出敏捷模型。

数据安全可靠是发挥数据价值、保证数字社会健康发展的保障。数据安全可靠性是指保障数据在其生命周期内,所有数据完整、一致和准确的程度,尤其是在自然灾害、电力等设施失效、设备自身失效、人为破坏、勒索病毒、网络攻击等影响下,仍可以实现关键数据系统的"三不两永远"——数据不泄露,数据不被篡改,数据不丢失,业务永远在线,访问永远合规。目前,数据中心行业实现存储安全可靠的方式主要包括容灾和备份:"容灾",即保障业务系统在灾难条件下的连续性、"韧性";"备份",即保障数据在灾难条件下的完整性。

绿色低碳是数据存储未来需考虑的必要因素。为响应国家双碳战略,落实工信部、发改委等政策文件提出的"新建大型、超大型数据中心PUE达到1.3以下,绿色低碳等级达到4A级以上"要求,数据中心存储正在向绿色节能低碳方向发展。数据存力设备使用过程中的电力能耗和冷却能耗是数据中心能耗的重要组成部分,单位存储容量能耗也逐渐成为数据中心综合能耗水平和可持续发展能力的评估要素之一。数据存储的绿色低碳发展也逐渐被企业和政府重视起来,一方面,使用绿色节能的存储设备能为企业降低总体拥有成本;另一方面,通过低能耗存储设备的使用达到降低碳排放的目的,获得环境和社会价值。

1.2 数据存力的重要作用

(1)数据作为关键生产要素,推动数字经济高速发展

数据作为数字经济时代最核心、最具价值的生产要素正在加速渗透,成为全球经济增长的新动力、新引擎,深刻地改变了人类社会的生产和生活方式。5G联接、人工智能、云计算、区块链、产业互联网等ICT新技术、新模式、新应用,无一不是以海量数据为基础。随着数据量指数级增长,数据分析算法和技术迭代更新,数据创新应用和产业优化升级,数据对社会变革的影响更加深远。以数据为关键要素推进数字产业化和产业数字化,推动数字技术与实体经济深度融合,能够为数字经济提供持续动力;充分利用海量数据要素,大力发展数字产品制造业、数字产品服务业、数字技术应用业、数字要素驱动业等,能够为经济发展培育新的增长点;将数据要素与其他生产要素有机结合,能够提升其他生产要素的匹配效率、激发其他生产要素的创新活力,进而提高生产质量和效益,推动国民经济质量和水平实现整体跃升。



(2) 数据存储作为数字基础设施核心部分,支撑经济社会平稳运行

数据价值的发挥,需要与算力、存储和网络等相关设备结合,以基础设施的形式来服务经济社会。数据存储是算力基础设施的核心组成,是信息的起始和终点,优质的数据存储基础设施和产业生态不仅会赋能数据存力发展,而且还会驱动更大范围的社会经济衍生影响。在影像采集领域,诸如交通、医疗、制造等行业在不断扩展其对高清影像的应用,在产出大量4K、8K视频、上亿像素的高清图片作为数字生产材料的同时,需要更大的存储空间。在高性能计算领域,诸如气象预测、卫星遥感、灾害模拟等重要行业,需存储大量数据,保证性能容量可横向扩展。在实时信息处理领域,诸如金融、医疗等行业,需数据存储系统来满足快速增长的在线并发访问、业务全天候在线的需求。

(3)数据存储惠及企业,助力行业数字化转型

近年来,越来越多企业开始重视数据价值,采集、留存并分析生产和运营中的数据,增加了数据存储、管理、使用的需求。各行各业纷纷通过数字化转型提升企业运营效率,不仅可提高企业自身的运作效率,也能助力提升供应链的可控性,为企业经营决策提供更多支持。一方面,数据存储一定程度上解决了数据流通不畅问题,上下游各用户可以将需求和供应信息通过线上平台传递。制造业上游厂商如服装代工厂、布料厂商等在生产供货时可通过数据判断趋势,调整企业的生产制造策略,实现资源的有效配置;银行风控授信、保险公司、物流公司等也因为更多的精准生产和需求数据的输入能更好地开展业务,为其业务保驾护航,实现产业链的持续完善。另一方面,各企业因为自身生产效率的提升,降低了残次品率和订单延迟交付率,生产周期变短,供应链也更加稳定。



数据存储产业发展分析

2.1 政策分析

数据存储对国家数字经济发展有举足轻重的地位,各国为形成数据存储产业发展思路,强化存储产业的自主可控,打造数字经济供给侧的竞争与可持续发展能力出台了各项政策,以保证数据产业链的采集、存储、传输、交易和服务各环节高效可靠。

(1) 全球各国纷纷出台数据要素相关的国家战略

美国2020年发布《联邦数据战略》,明确将数据作为长期战略资源开发,聚焦点从"数据即技术" 到"数据即资产"的转变,明确政府需发挥其作为数据提供者、使用者的角色的重要作用。

英国2020年发布《国际数据战略》,明确"数据是世界现代经济的驱动力,打造世界领先数字经济,数字产业是最主要的支撑保障,希望公众成为蓬勃发展的数字经济中的积极参与者"。

沙特阿拉伯2020年发布《沙特数字经济政策》,明确"从传统的经济活动转向更大的以数据为依据的活动,突出数字经济的中心地位"。

新加坡2020年发布《研究、创新和企业2025计划》,明确"继续支持战略和新兴技术的发展,并加强数字能力向工业的转化。其目的是实现新加坡的智慧国家雄心,并利用数字经济中的增长机会继续加强数字经济、数字政府和数字社会的建设"。

德国2021年发布《联邦政府数据战略》,明确"着力打造数据文化以确保公平和防止数据垄断,增强德国数字能力,把德国打造为欧洲数据共享和创新应用领域的领导者"。

(2)全球各国布局半导体产业发展

美国高度重视半导体存储芯片的战略价值,持续完善政策体系,不断丰富行业标准,使存储产业规范化、标准化。美国政府以政策为抓手,鼓励本土半导体厂商进行技术创新,完善专利保护法,对半导体存储相关技术进行全面的保护。

韩国半导体产业在国内需求有限、国际竞争激烈的环境下,保持良好的发展态势和强大的产业优势。 韩国通过"BK21"及"BK21+"等计划对大学、研究所进行专项支援,在2016年时推出半导体希望基金, 投资半导体相关企业。2021年5月,韩国政府正式公布"K半导体战略",预计未来10年投资约4500亿美元。

日本为半导体产业保持蓬勃发展制定了应对策略。2021年6月,日本政府对半导体、数字基础设施及



数字产业做出综合部署,制定了《半导体数字产业战略》。2021年底,发布《半导体产业紧急强化方案》, 致力于提高日本半导体企业的营收。

(3) 我国将数据提升到生产要素地位

习近平总书记在中共中央政治局就实施国家大数据战略进行第二次集体学习时指出,我们应推动实施国家大数据战略,加快完善数字基础设施,推进数据资源整合和开放共享,保障数据安全,加快建设数字中国,更好服务我国经济社会发展和人民生活改善。

2019年11月,党的十九届四中全会在《中共中央关于坚持和完善中国特色社会主义制度推进国家治理体系和治理能力现代化若干重大问题的决定》中首次明确将"数据"列为生产要素,提出了"健全劳动、资本、土地、知识、技术、管理、数据等生产要素由市场评价贡献、按贡献决定报酬的机制",反映了党中央对信息技术发展的时代特征及未来趋势的准确把握,也凸显了数字经济时代数据对于经济活动和社会生活的价值。

2020年4月,中共中央、国务院发布了《关于构建更加完善的要素市场化配置体制机制的意见》,数据作为新型生产要素正式出现在官方文件中,将数据纳入要素范围,意味着数据与资本、劳动等传统要素拥有了相同的地位和角色,这对于探索建立统一的数据标准规范、支持构建多领域的数字基础设施,推动壮大我国数字经济具有重大现实意义。《"十四五"国家信息化规划》、《"十四五"信息通信行业发展规划》中,也强调了加快推进数字基础设施建设,统筹布局绿色智能的数据与算力设施,强化数据采集、数据存储、加工处理、智能分析等能力。

2020年5月发布的《中共中央国务院关于新时代加快完善社会主义市场经济体制的意见》中提出,要加快培育发展数据要素市场,建立数据资源清单管理机制,完善数据权属界定、开放共享、交易流通等标准和措施,发挥社会数据资源价值。

2022年1月发布的《十四五数字经济发展规划》明确提出"数字经济是促进公平与效率更加统一的新经济形态,坚持应用牵引、数据赋能。坚持以数字化发展为导向,充分发挥我国海量数据、广阔市场空间和丰富应用场景优势,充分释放数据要素价值,激活数据要素潜能,以数据流促进生产、分配、流通、消费各个环节高效贯通,推动数据技术产品、应用范式、商业模式和体制机制协同创新。"。

(4) 东数西算、新型数据中心为数据存力发展带来了新的机遇

2021年,《全国一体化大数据中心协同创新体系算力枢纽实施方案》和《新型数据中心发展三年行动计划(2021-2023年)》等数据中心政策文件的发布,为数据中心产业营造了良好的发展环境,激发了数据存力的需求。随着我国东数西算工程的落地,新型数据中心行动计划的实施,数据中心成为产业界各方讨论的热点话题,数据存力协同算力、网络资源共同发展,新型基础设施才能充分发挥数据要素的价值。新型数据中心是具备安全可靠能力、提供高效算力服务、赋能千行百业应用的新型基础设施,其算力服务的提供需要数据存储能力的支撑。如工业互联网、人工智能等新兴技术的可靠稳定应用离不开数据



存储,存储效率和能力达不到要求的时候,算力也很难发挥作用。同时经济社会运作中产生的各类数据与物理世界完成交互后,最终处理的结果又会回到数据存储系统进行保存,指导模型、算法的精确性和可用性提升,更好支撑下一次的数据分析。

2.2 产业分析

由于数据存储的形式在不断变化,全球数据存储产业也从单纯注重数据存储的容量,向效率、经济性、可靠程度、节能等方面不断延展,以更好应对不断增长的数据总量以及数据多样性、实时性等对存储提出的新要求。

(1)磁性存储市场占比较高,半导体存储发展迅速

现代数据存储根据存储介质的不同,主要分为光学存储(CD、DVD、蓝光存储)、磁性存储(磁带、软盘、机械硬盘)和半导体存储三类,其中半导体存储在数据存储领域的应用越来越广,是存储技术发展的趋势和方向之一,包括易失性存储器 DRAM和非易失性存储器 NAND FLASH、NOR FLASH等。

光存储因技术原因应用场景受限。作为产生时间最早的数据存储技术,光存储极易受摩擦等外部作用而损坏,在存储容量、存储密度及存取速率等方面受到一定程度的限制。但光存储技术因其长寿命、安全可靠和低成本等优点,被业界认为适用于冷数据存储,归档类数据的长期备份使用。

磁性存储应用广泛。磁带存储在海外发展较早,诸如微软Azure、亚马逊AWS、谷歌等全球知名的云服务商已大量采购磁带,用来对关键数据做备份归档以及冷数据的长期保存。软盘存取速度慢,容量较小,但可装可卸、携带方便,适用于需物理移动的小文件。机械硬盘以其低廉的价格以及更高的储存容量占据了一定市场份额,但存在储存时间不长,能耗高,存储密度不大等问题。

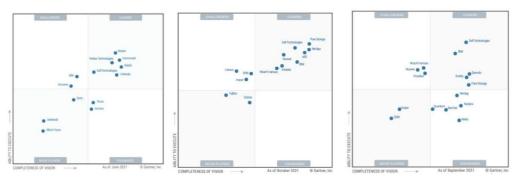
半导体存储已成为存储的重要媒介。近年来,以半导体为代表的新一代全闪存技术飞速发展,半导体闪存是新一代存储介质,由于具有断电不丢失数据、高性能快速读取、大容量、绿色节能等特点,被企业的信息系统大量应用,已逐步成为存储产业的主要媒介。在全球整个半导体市场中,存储器占据了约三分之一的规模,是半导体最大的细分市场之一,半导体正在成为业界主流的存储器。世界各主要厂商正在加速新一代全闪存技术的创新。Gartner 2021年全球存储市场报告显示,全闪存储市场快速增长,市场规模同比增长13.9%,全闪存储在全球存储市场占比首次超过50%。

(2) 我国存储技术水平突围成效显著,与美企仍有一定差距

在全球数据存储的市场份额中,美国企业处于垄断地位。Gartner 2021年发布的存储相关的魔力象限报告,在企业级备份和恢复解决方案的魔力象限全是国外企业,包括Commvault、Rubrik、Veeam、Cohesity、Veritas Technology、Dell Technology等美国公司,占据数据保护、备份与恢复的主要市场;在主存储魔力象限上,Pure Storage、Dell Technology、Net App、HPE、IBM等美国公司在全球占比较



高,日本Hitachi Vantara、以色列Infinidat也处于领导者地位,仅有我国企业华为技术有限公司进入,凭借其优异的产品能力,在亚太地区有广泛应用;**在分布式文件系统和对象存储等魔力象限上**,也都以Dell Technology、IBM、Pure Storage、Qumulo等美国公司为主。



数据来源: Gartner

图2 2021年度 Gartner企业级备份和恢复解决方案、主存储、分布式文件系统和对象存储魔力象限排名

近年来我国在数据保护领域也涌现出很多优秀的企业,包括华为、爱数、鼎甲、新华三、浪潮等。IDC 数据显示,在数据复制与保护系统市场,国内企业在本土市场占有率逐步提升,2021年华为、爱数、鼎甲和新华三占据42%的市场份额,占有40%的其他类别中也包含了大量的国产厂商。在软件定义存储 (SDS) 及超融合存储 (HCI) 系统市场,软件定义存储 (SDS) 及超融合存储 (HCI)系统市场前五大供应商均为国内企业,市场份额占比均超过70%。



数据来源: IDC

图3 2021我国数据复制与保护、软件定义存储、超融合存储系统市场份额

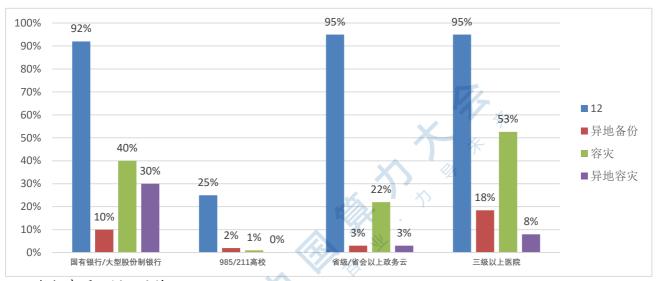
(3)不同行业灾备水平参差不齐,亟需建设典型引领场景

我国在经济社会高质量发展中逐渐重视数据安全,对数据存力的可靠提出相应要求。可靠的数据存力是算力基础设施建设的基础,数据中心产业已开始重视数据安全和容灾备份建设,以应对各类数据安全事故。日前,《数据安全法》已经发布并开始实施,保护数据安全成为数据存储发展的关键,存力的可



靠性建设将会加速,以应对数据安全风险和数字经济时代的挑战。

我国多数行业的信息化灾备建设体系亟待完善,普遍存在"有备份、无容灾","缺乏异地灾备"的现象。数据安全风险要求较高的行业,如金融、运营商等行业,具有信息化程度高和信息安全要求高等特征,行业灾备体系建设相对完善,典型灾备应用行业示范意义大。根据灾备设备出货量统计,金融行业,如国有银行/大型股份制银行,数据备份建设比例达到约92%,但异地备份的比例仅约10%,业务系统容灾比例不到40%,异地容灾低于30%。



数据来源:调研统计

图4 国内各行业数据灾备覆盖率

其他行业信息化发展起步晚,灾备体系建设也相对滞后。全国31个省(直辖市、自治区)的政务云只有5%部署同城灾备,异地灾备部署比例更低,大部分省级政务云只是部署了简单本地备份,缺乏数据中心容灾能力,难以满足核心系统业务连续性和数据安全的要求。教育行业,以高校为例,数据灾备比例非常低,容灾建设比例不到1%,只有个别学校如哈工大部署了异地容灾系统,同地校区备份的比例平均在25%左右,异地备份仅2%左右。医疗领域的灾备系统体现出水平参差不齐的特点,三级以上医院几乎都采用数据备份系统,普及率达到95%以上,具体应用技术手段有所不同,部分医院在系统升级换代时会同步考虑容灾技术。二级以上医院在地域条件允许的情况下会采用远程备份,实施方案主要为远程硬拷贝等传统方式,总体占比不高,部分有条件的二级医院会建设容灾系统,只有少数多院区医院考虑建设数据异地容灾。



2.3 技术分析

事物的演进普遍遵循螺旋式发展的规律,数据存储产品和方案的演进同样如此,业务需求和技术进步是螺旋式发展的两个主要驱动力。从八、九十年代的初步电子化、信息化,到互联网+,以及当前全面推进数字化,未来的数字孪生、元宇宙,数据的规模成千倍增长,移动互联网发展导致并发访问量百倍增大、业务7*24连续服务、毫秒级响应体验。

存储技术和产品从初始的磁带和磁盘开始,到DAS、SAN、NAS等专业存储设备,以及目前的全闪存存储、分布式存储等,在性能、可靠性、容量这三个核心方向上迭代发展,再加上访问协议、存储网络等技术的演进,数据存储能力也在快速提升。

(1) 存储介质技术升级在提高存储容量同时降低单位存储成本

存储介质由传统磁介质转向闪存介质。目前,利用GMR巨磁阻效应磁头、SMR瓦楞式堆叠磁盘技术的企业级机械硬盘(HDD)单盘容量可达到18TB。磁通控制-微波辅助磁记录(FC-MAMR)、微波辅助交换-微波辅助磁记录(MAS-MAMR)、热辅助磁记录技术(HAMR)、石墨烯涂层等新技术的研究和应用,在未来几年可实现单盘容量的翻番增长,满足温、冷数据近线存储的需要。同时,伴随半导体介质生产工艺提升迅速,3D NAND Flash已经实现了176层生产工艺,同时闪存颗粒的存储密度从SLC(Single-Level Cell)、MLC(Multi-Level Cell)已经过度到TLC(Trinary-Level Cell)为主,正在向QLC(Quad-Level Cell)演进。NAND Flash技术和工艺的提升,SSD单盘容量已经达到30TB以上,在同等应用场景下,SSD数据访问比磁盘快100倍左右,吞吐量大100倍,单盘的IOPS大1000倍以上,在提供高性能的同时降低了成本,可以更好为热数据提供在线存储。

热辅助磁记录、晶格介质存储、全息存储、DyCo5材料存储等高密度关键存储技术持续研发创新,更好地响应了持续增长的数据存储容量和存储密度的需求。热辅助磁记录技术在写入存储介质时,通过聚焦光束对最小磁化区域进行加热,解决少量热量对纳米级比特区域的干扰,有效利用热量,确保在纳米级比特单元在加热时仍保持磁化状态;晶格介质存储通过光蚀刻微影的方式在晶格介质上划分出统一的网格磁性单元,实现单比特占用空间少,存储密度高;全息存储通过消费光学和人工智能技术,将数据存储为可重写的全息图,借助商用高分辨率相机技术和深度学习技术,将HSD设备负责性迁移到云端,缩短读写时间,提高实时访问效率;DyCo5材料存储针对超高密度热辅助数据存储设备,提出高效节能的解决方案,使得数据写入时可实现快速磁化,且材料耗能更少、性能更优。

(2) NVMe打造主机和存储设备间高速通道

存储通信协议的发展,充分发挥了存储介质变化和网络带宽提升所带来的效益。NVMe(Non-Volatile Memory Express, Non-Volatile Memory,非易失性存储器)是由NVM Express非盈利组织发布的规范,定义了主机软件通过PCIe总线与非易失性内存通信的方法,是固态硬盘的协议标准,其通过与CPU的PCIe



通路直接通信、更深的队列深度、操作系统协议堆栈的精简等一系列手段来全面提升固态硬盘的性能表现。NVMe系列标准规范了SSD访问接口,确保NVMe在网络架构、接口管理、服务器和系统管理规范化和明确化。它是所有形式(U.2、M.2、AIC、EDSFF)的PCle固态硬盘(SSD)的行业标准。除了NVMe基本规范之外,还包括用于规范网络架构上使用NVMe命令的NVMe over fabric(NVMe-oF),以及用于服务器和存储系统管理NVMe/PCle SSD的NVMe管理接口(NVMe-MI)规范。随着使用非易失性存储器高速(NVMe)接口规范的技术从HDD发展到SSD,再发展到SCM,存储性能得到了巨大提升。通过NVMe访问存储介质所花时间相比以前的硬盘技术减少了1000倍。

NVMe协议正逐渐成为主流,固态硬盘接口也从传统的SATA/SAS过渡到PCIe,采用NVMe协议后的PCIe接口固态硬盘可以进一步释放固态硬盘的性能,提高响应速度。随着NVMe的发展,其也逐步被应用于存储网络。当前NVMe over Fabrics有两类主流的方式,分别是NVMe over FC和NVMe over RoCE。其中,NVMe over RoCE以开放的以太网为基础,具备大带宽、低时延的优势,随着技术的逐渐成熟,已经规模化落地实践。NVMe over RoCE在性能、成本等方面的优势将促使架构逐步转型,存储网络逐步转向全IP化,在获得更高的传输性能的同时,具备以太网运维管理相对简单的特性。

目前,PCle 6.0规范已经正式发布,预计最高写入速度为11GB/s,最高读取速度为12GB/s,而PCle 5.0 SSD设备将在2022年正式上市,PCle 系列规范带宽速度如下所示。

插槽 宽度	PCle 1.0 (2003)	PCIe 2.0 (2007)	PCle 3.0 (2010)	PCIe 4.0 (2017)	PCle 5.0 (2019)	PCIe 6.0 (2021)
x1	0.25GB/s	0.5GB/s	~1GB/s	~2GB/s	~4GB/s	~8GB/s
x2	0.5GB/s	1GB/s	~2GB/s	~4GB/s	~8GB/s	~16GB/s
x4	1GB/s	2GB/s	~4GB/s	~8GB/s	~16GB/s	~32GB/s
x8	2GB/s	4GB/s	~8GB/s	~16GB/s	~32GB/s	~64GB/s
x16	4GB/s	8GB/s	~16GB/s	~32GB/s	~64GB/s	~128GB/s

表1 PCle 系列规范带宽速度

数据来源: PCI-SIG

PCIe市场规模持续扩大,服务器设备是规范迭代升级的主要受益者。企业级SSD的发展使得PCIe在2021年获得更大发展机遇,并在未来5年将持续增长。服务器厂商先后推出众多支持PCIe 4.0的服务器主板,可使用NVMe硬盘。当PCIe 5.0实现商用后,市场规模将得到进一步扩大,大幅降低存储在整个计算机系统中,造成的性能瓶颈,而NVMe存储设备或将成为受益最大的设备;到2023年,服务器设备或将率先实现PCIe 6.0应用,为人工智能、数据中心、物联网、航空航天和军事等领域提供有效支撑,实现对高性能计算中心提供基础存储设备支撑。



(3) 存储系统工程和算法演进保障数据存储安全可靠

存储是数据安全的最后一道防线,随着企业数字化的深入,数据的安全风险变得复杂多样,从数据的增长到架构的复杂度,如何保证数据的可靠性和安全性是数据存储的关键。在存储系统层面,利用多控/集群架构、关键部件冗余、热插拔、多路径软件、负载均衡、在线扩容、在线升级等技术,系统可用性大大提升。在算法层面,算法的改进提升了固态硬盘的使用寿命,由于固态硬盘的使用寿命由NAND闪存的擦写次数决定,固态硬盘使用寿命是用户最主要的担忧。但随着写入算法的不断演进,可以在软件层面最大限度地延长固态硬盘的使用寿命,在实际使用和替换周期来看,固态硬盘的实际故障率低于机械硬盘、使用寿命并不会成为其瓶颈。在数据保护方面,存储双活已经成为保障关键业务系统连续性的主要方案之一,并可以扩展至"两地三中心"容灾方案,避免在自然灾害、电力等设施失效、设备自身失效、人为破坏、勒索病毒、网络攻击等影响下的数据恢复能力和业务系统的连续性。





中国数据存力发展情况分析

3.1 我国数据存力发展情况分析

我国数据存力发展总体水平与发达国家相比还有一定差距。据罗兰公司测算,美国单位GDP存储容量高于我国,我国存储在经济发展中的支撑作用有待加强。规模上,据中国信通院统计,2021年我国存储总量容量已达800EB,出货量年均增长速度已达50%。在先进存储设备上,半导体闪存具有断电不丢失数据、高性能快速读取、大容量、绿色节能等特点,其使用占比正不断攀升。



数据来源:中国信通院

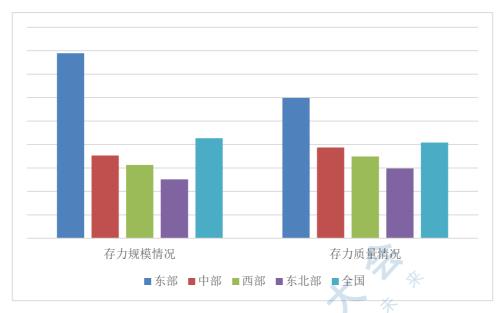
图5 2021年我国存储容量增量

3.2 区域数据存力发展情况分析

数据存力发展水平与数字经济规模相关,数字经济规模越大,数据资源越丰富。东部沿海地区经济相对发达,数字经济规模较大,因此东部数据存力在规模上具有较大优势。此外,数据存力发展水平与数据中心和算力的建设情况相关,东部仍是数据中心建设运营主要区域,数据存力发展规模和质量均位于全国前列。中西部数据存力发展水平相差不大,结合《中国综合算力指数(2022年)》测算分析,中西部数据存力总分、规模及质量得分均略低于全国平均水平。东北部数据存力规模和质量相对其他区域发展



较差,指数值也小于全国均值,发展空间较大。



数据来源:中国信通院

图 6 2021年我国区域存力发展情况

东部省市,数据存力发展水平全国领先。北京、上海、广东及周边省份数据计算需求较多,数据中心市场规模不断上升,为数据存力营造了良好的发展环境,其数据存力发展也受到重视。这些省份对数据处理的时效性和安全性要求较高,在热数据进行计算和存储时,追求数据存取和读写速度,需求带动本地及周边存力发展。规模方面,东部省份中广东省存储规模最大,2021年的存储总体容量超过90EB。随着5G、云计算、人工智能等新一代信息技术快速发展,信息技术与传统产业加速融合,数字经济蓬勃发展,北上广等发达地区对数据计算和存储的需求快速攀升,也将带动本地及周边存力规模的增长。质量方面,全国整体先进存储占比相差不大,平均先进存储占比约13.5%,东部部分经济发达的省份如北京和河北省,先进存储占比超过15%。一线城市对数据处理的时效性和安全性较高,热数据计算需高性能的数据存储,新的业务场景也对存储质量有更高的要求,一线城市在存储规模日益扩大的背景下,也将更加关注先进存储、存储性能等存储质量。

中西部省市,数据存力发展水平总体低于东部,但部分核心省份数据存力发展良好。湖北辐射带动中部和长江中游地区,支撑长江经济带发展,数据存储需求多,其数据存力发展水平较好。作为西部大开发重要平台的重庆和四川经济较发达,对数据的存储、转移和处理需求多,这些需求驱动了作为数据载体的存力快速发展,数据中心产业发展较好,数据存力水平相对较好。受"东数西算"政策影响,北京、上海、广州、深圳等一线城市的部分数据存储应用需求逐步向西部地区转移,推动了当地数字经济以及信息化产业的发展,如贵州省、新疆、内蒙古、甘肃省和陕西省,这些大数据产业蓬勃发展地区,充分发挥本



地大数据资源价值,数据中心产业发展迅速,受政策牵引和自身资源禀赋的影响,数据存储规模不断增长,但存储质量有待提升。部分中西部省份资源充足,但其网络带宽小、跨省数据传输费用高、算力需求有限等现状也不能忽视,受制于算力发展,数据存力发展水平较差。随着"东数西算"划定的枢纽节点落地,贵安数据中心集群、和林格尔数据中心集群、庆阳数据中心集群和重庆数据中心集群,未来数据存力发展有了新的机遇。

3.3 省际数据存力发展情况分析

目前,数据存储总体容量集中度高,容量排名前7的广东、江苏、上海、北京、河北、浙江、贵州省存储总体容量占全国的52%,北上广的存储总体容量超230EB,约占总容量的30%。为提高数据中心标准机架的效率,为算力提供更好的数据支撑,需关注数据中心单机架的存储容量即单机架存力指标。全国各省单机架存力水平普遍在120TB至180TB之间。湖北、广东、北京、浙江和江苏的单机架存力超过150TB,其他省份单机架存力在130-150TB之间。未来,受东数西算政策影响,响应"加快推动数据中心向西大规模布局,特别是对于后台加工、离线分析、存储备份等对网络要求不高的业务,率先向西转移,由西部数据中心承接"的政策,预计八大枢纽节点省份数据存力将继续扩大规模。



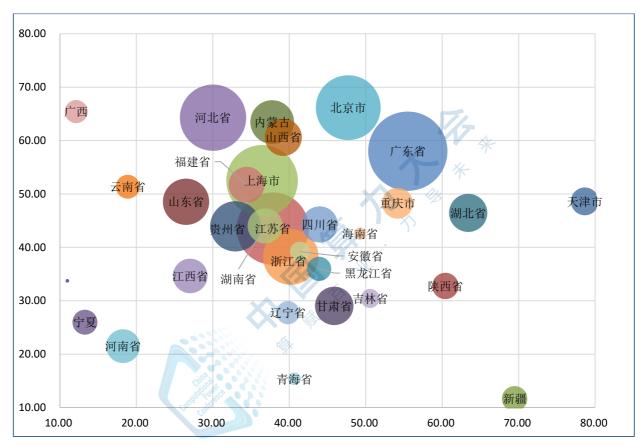
数据来源:中国信通院

图7 2021年我国部分省份存力规模情况

全国总体数据存储质量有待提高。存算均衡方面,根据《中国综合算力指数(2022年)》测算,大部分省份存算均衡得分在25-55之间,天津、新疆、湖北等省份存算资源相对均衡,得分大于60分,我国存算均衡整体与发达国家相比还有一定差距。**存储先进性方面,**各省先进存储占比相差不大,全国平均先



进存储占比约13.5%,对存储质量要求较高的一线城市先进存储占比最高也仅17%,除新疆和青海等省份,其他省份先进存储占比得分均在25-65之间,广西虽然存储规模整体较小,但先进存储设备的选择和应用占比较高。**存储性能方面,**北上广及周边地区的存储性能总能力相对较高,广东、江苏、河北、北京和浙江的存储性能总量均超过16 GIOPS,海量数据驱动了存储性能的提升;其他省份存储性能相差不大,部分数据经济规模较小的省份,存储性能得分小于10分。



注:横轴表示存算均衡情况、纵轴表示存储性能情况、气泡大小为先进存储占比情况数据来源:中国信通院

图8 2021年我国部分省份存算均衡、存储性能及先进存储占比情况

3.4 数据中心存力发展情况分析

全国数据存力发展的落地应用依托于各省数据中心存储设备的部署,为加快存力的高质量建设,在评价数据中心的存力时,需要综合考虑存储容量、性能表现、安全可靠、绿色低碳等维度,如存储容量、容量增长率、先进存储占比等具体指标。



耒り	数据中心存力评价框架
ᄍ	数1店中小1千八斤0mm未

一级指标	二级指标	三级指标
	存量	数据中心存储容量
存储容量	增量	存储容量增长率
	· · · · · · · · · · · · · · · · · · ·	存力使用率
性能表现	平衡	存算均衡
注肥水火	敏捷	先进存储占比
		灾备覆盖率
安全可靠	安全可靠	RTO时间恢复目标
		RPO恢复点目标
绿色低碳	经济	单位存储拥有成本
一	绿色	存储设备能耗水平

上表是数据中心存力评价基本框架,包含4个维度、7个方面和10个具体指标。从存力定义出发,数据中心存力评价框架从"存储容量、性能表现、安全可靠和绿色低碳"四个维度进行评价。其中,存储容量包含了存量和增量,存量体现了数据中心提供的总体存力体量对于业务产生各类数据的有效保存和支撑的能力;增量反应的是用于衡量投入能否保障存力匹配数据生产要素增速,并支持前沿技术部署持续提升竞争力。性能表现包含了平衡和敏捷,平衡反应的是算力、存力两大核心关键资源之间的协同利用效率;敏捷衡量数据存储领域先进生产力闪存的部署情况,能否满足数据的保存和调用效率的要求。安全可靠反应企业抵御数据破坏性事件并在事后快速恢复的能力。绿色低碳包含了经济和绿色,经济指企业存储投入的成本情况,涵盖存储设备外的人工、能耗等总体拥有成本;绿色衡量数据中心存储设备总能耗水平。

如某股份制商业银行总行的数据中心,存储存量超30PB,存力容量使用率约为70%,规划年容量增长率为5%,可满足该行各信息系统如国内国际结算系统、网上银行系统、基金托管估值系统、债券交易管理系统等产生的即时数据存储和转移需求。此外,因银行日常业务运作对数据存储调用的时效性、准确性和安全性的要求较高,该数据中心大面积部署半导体全闪存,全闪存容量占比近60%;针对重要业务数据基本实现灾备,整体灾备覆盖率约20%,RTO时间恢复目标即灾难发生后,从IT系统宕机导致业务停顿之刻开始,到IT系统恢复至可以支持各部门运作,业务恢复运营之时的时间间隔仅5分钟,RPO恢复点目标为0,即灾难发生并不会丢失数据。该数据中心在保证存储规模和性能的基础上,也兼顾了绿色低碳,其单位存储运营成本为3.5万/台,存储设备能耗水平约35W/TB。



目前,为充分发挥新型数据中心对数字经济的赋能和驱动作用,需协调发展存储、算力和网络,新型数据中心已经开始重视存储规模和存储质量。规模方面,百度云计算(阳泉)数据中心存储容量超过6EB,可存储的信息量相当于30多万个中国国家图书馆的藏书总量,为百度搜索+信息流、智能云、智能驾驶等业务提供有力支持;中国联通贵安云数据中心拥有5000TB高效存储和3000TB对象存储,可支撑国家部委、政府机关、国际知名云计算厂商、大型互联网厂商、银行保险机构等多家客户。质量方面,中国电信南京(吉山)云计算数据中心采用HDD与SSD相结合的存储方式,其中SSD占比较高,目前该园区内存储能力可达50PB,可有效支撑各类云平台PB级别的业务场景,实现海量数据的存储及高性能的计算分析;中国雅安大数据产业园项目采用分布式存储技术,存储能力近300PB,满足400余家企业云上服务,该园区2020年数字经济及相关产业规模超过200亿元。



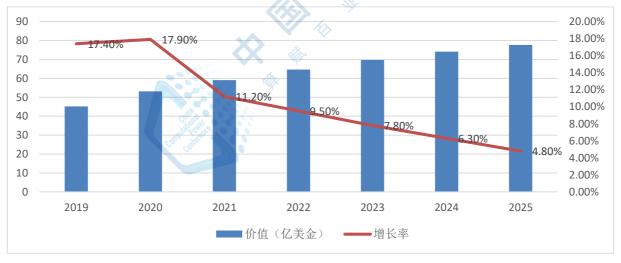


中国数据存力发展展望与建议

4.1 数据存力发展展望

(1)我国存储市场规模将持续攀升

据中国信通院《中国数字经济发展白皮书》显示,2021年我国产业数字化规模达37.18万亿元,占数字经济比重81.7%,成为数据经济发展的强大引擎。海量数据的计算和转移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。据IDC数据显示,2021年我国存储市场规模达到59亿美元,比2020年市场规模增长18%,预计未来5年存储市场仍将保持增长趋势,如下图所示。"十四五"规划和新基建投资计划等文件,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,也将进一步扩大我国存储市场规模。



数据来源: IDC

图9 我国存储市场预测

(2)非结构化数据存储需求带动分布式存储快速发展

5G、AI和大数据等技术,使得数据挖掘得到更为广泛的应用,面向未来,以视频、图片、文档为代表的非结构化数据占比将达到80%以上,成为分布式存储的主要消费者。**市场方面,**分布式存储与更多关键业务的融合成为存储市场主力,随着新兴业务和云存储的发展,除了云原生的应用,越来越多的应用



倾向使用更易于流动和归档的对象存储,分布式存储市场中对象数据份额将会超过块数据。**技术方面,**随着大数据、AI的发展,HPC应用已经广泛融合并且催生出新的应用场景,新的业务负载要求存储支持文件、对象及HDFS协议互通,数据仅存一份。HPC正在向HPDA(高性能数据分析)、HPC-Based AI方向演进。

(3)新型存储全闪存将成为温热数据存储的主流方式

根据数据全生命周期的热温冷差异,对存储设备的要求也不同,未来,温热数据的存储将逐步向高性能、高容量和高性价比三个方向演进。为了追求最极致的读写能力,适应业务系统日益提升的性能需求,SSD在存储阵列中的使用比例逐渐扩大,全闪存存储阵列开始出现在企业的数据中心,并呈现快速增长态势。全闪存存储阵列不仅是硬盘全部采用闪存盘,而是需要从系统层给予优化,最大程度激活闪存能力。目前业界已经涌现了许多新型内存型介质技术,如SCM、PCM、MRAM、ReRAM、FeRAM等。这些介质在性能、容量、成本、寿命、能耗、可扩展性等多个方面都有更强的性能表现,支持字节级访问和持久化,不需要再进行数据搬移,但其成本较高,或将成为热数据存储的主流存储方式。

(4)存储网络的以太化改造成为热点话题

存储网络是指连接应用主机和存储的高速网络,其基本要求是大带宽、低时延。传统以太网存在丢包问题,网络质量较差。FC网络则以其大带宽、低时延的特点,成为存储网络的主流解决方案。然而,FC 技术长期被欧美少数公司所垄断,技术壁垒高、产品价格高且生态封闭。近年来,无损以太网络快速发展,其与RDMA、NVMe等技术结合所形成的NVMe over RoCE,成为存储网络的重要解决方案,网络性能好且生态开放。

表3 无损以太网络和FC网络的对比

对比项	FC	无损以太	对比结果
性能	协议处理硬件卸载,低时延,大带宽(32Gb)	协议处理硬件卸载, 低时延, 大带宽(100Gb)	时延上相当;带宽能力上以太网有明显优势。
组网规 模能力	需要额外license和兼容性要求。 使用:需要专业工程师进行规划和 维护。 供应:博科、思科等国外垄断,国	使用: IT的网络工程师都具备配置管	以太网规模更大,规划 维护更灵活易用
主机适配器	由Qlogic、PMC等厂家提供硬件和 驱动,操作系统的兼容性非常全面	由Broadcom、Qlogic、Chelsio等厂家提供硬件和驱动、操作系统主要	·



(5) "存算一体"架构开始显现出独特优势

存算一体是计算单元和存储单元紧耦合的一种方式,即存储单元与计算单元合一,打破算力和存储的边界,有效改善功耗墙和内存墙的问题,相比传统冯诺依曼架构,预计能效将大幅提升。基于SRAM(Static Random-Access Memory,静态随机存取存储器)、NOR Flash(非易失闪存)等成熟存储器实现的存算一体,有望率先规模化商用,并将在端侧、边缘侧的人工智能推理运算中展现出10倍能效优势。基于ReRAM(Resistive randomaccess memory,可变电阻式内存)、PCM(Phase Change Memory,相变存储器)、MRAM(Magnetoresistive Random Access Memory,磁性随机存储器)等新型非易失存储器的存算一体探索不断加快,因其具有高性能、低功耗的特点,有望在数据中心侧实现突破。针对实时处理要求极高的最热数据的,存算一体化的性能好,可以打破存储墙,消除不必要的数据搬移延迟和功耗,并使用存储单元提升算力,成百上千倍的提高计算效率。但其价格相对较高,针对实时热数据、温冷数据合理规划、分类处理,才能达到整体架构总成本最低。

4.2 数据存力高质量发展建议

数据存储在国家信息技术和高科技领域的关键基础地位日益凸显,已然成为数字经济发展的战略要地。《"十四五"国家信息化规划》、《"十四五"信息通信行业发展规划》中,强调"加快推进数字基础设施建设,统筹布局绿色智能的数据与算力设施,强化数据采集、数据存储、加工处理、智能分析等能力"。存储作为承载数据的载体,已引起社会广泛关注。目前,数据存储还面临诸多挑战,一方面,海量数据呈指数级增长,数据在加速流动;另一方面,承载数据的关键数据基础设施如存储器、数据网络等还存在可靠性、安全性、可持续供应性等多方面的挑战和问题,亟待加以重视,出台政策推动相关产业发展。

为推动国家数据存力高质量发展,首先要在战略上提高对存储关键性的重视程度。从大数据时代全球竞争和国家数据安全的战略高度出发,以先进存储、存算均衡、灾备覆盖率等为牵引,尽快研究形成引导我国数据存储产业发展的具体政策和规划,打造一批存储产业和数据存储应用示范项目,加快数据存储设备和技术自主可控,降低存储技术和制造的对外依存度,提升我国在数据存储领域的国家竞争力。

技术方面,一是加快部署下一代存储技术,推动存储核心技术底层研发和技术攻关,提升存储介质、安全设计、数据防护等关键技术水平,打造全球存储的创新高地;二是鼓励联合攻关存储技术瓶颈,发挥行业创新领军企业的创新带领优势,广泛发动我国相关领域的科研院所、高校和领军企业,开展存储全产业链的产学研用协同,增强存储产业国际竞争力。三是进一步完善数据产业人才培养机制,建立多层次、多元化的人才培养目标,鼓励培养专业理论与行业知识兼备的复合型人才,鼓励地方政府引导企业完善人才激励机制,引进和留住高端存储专业人才,优化人才的地域和行业布局。



标准方面,应尽快研究建立存储的行业规范和标准,完善数据存储效率、数据保护、数据灾备、绿色低碳等数据存力关键标准体系。一方面,鼓励存储相关企业及研究机构进行标准和评估体系研究,鼓励行业龙头企业通过硬件开放、软件开源、使能生态伙伴等策略,给予合作伙伴尤其是创新型小微企业充分的发展空间,共同摸索形成存储行业标准;另一方面,推进数据存储产业国际交流与合作,加强与国际标准组织的合作,加速我国技术标准的国际化进程,积极参与数据存储国际规则和技术标准制定。

产业方面,一是构建先进存储应用生态,促进存储产业相关供应链融合;二是提升产业链供应链安全保障能力,构建产业集群梯次化发展体系;三是针对金融、医疗、教育等重点行业,打造示范应用项目,树立高效存储、先进存储、存算均衡、存储安全等方面的标杆。如在安全可靠上,推动政务云、国家一体化大数据中心、行业大数据中心、智慧城市等场景的灾备体系建设,全国范围重点打造一批数据灾备的示范基地、示范项目,实现标准和政策落地,带动千行百业,更好应对数据安全风险。

2022中国 算力 大会 ^{算 赋 百 业 · 力 导 未 来}



CAICT算力 公众号

地址:北京市海淀区花园北路52

电话: 010-62300095/18601152291

邮箱: dceco@caict.ac.cn