



# 先进数据存力

加速智能经济涌现的  
高性能引擎



## 执行摘要

展望未来，人类文明正加速从数字化迈向智能化。以大模型为代表的原生 AI 应用已走向千行百业，以其涌现的智能为带来组织生产力跃迁、使能产业升级、提升社会福祉，乃至强化各国竞争力。智能时代具备三大核心特征。一是数据、能源与技术的重要性提升，升级为智能时代三大核心生产要素，促进通用智能“涌现”，二是智能的涌现将端到端重塑企业数据治理范式、加速数据要素市场规模化发展，三是“通用 AI 智能体”将出现，并成为千行百业实现生产力跃迁的核心驱动力、使能人类文明精神世界的极大丰富。

数据存储产业自数字时代起就作为 ICT 核心使能技术之一，持续赋能产业与社会经济发展。在如今的智能时代，为应对全球数据量快速增长、数据要素市场蓬勃发展、原生 AI 应用智能涌现、网络安全环境日趋复杂、数据产业的低碳绿色发展需求深化等关键产业趋势，数据存储产业应从“存力”迈向“先进数据存力”。

先进数据存力是智能时代的重要组成部分与核心使能技术，我们认为先进数据存力在智能经济中将持续扮演“高性能引擎”的角色，可有力驱动数据要素规模化发展、使能社会经济智能化。为落实自身在智能经济中的定位、良好地适配智能经济的产业环境，我们认为先进数据存力的核心发展目标有三。一是集约化汇聚全域异构数据，二是高效安全可靠的数据存储，三是高效使能多元化应用生态。

从价值和影响力看，1 GB 先进数据存力投入预计能带来 60 元 GDP 的增长，其在赋能产业升级、提升社会民生福祉与强化区域竞争力三大领域可产生显著价值。

- 赋能产业升级方面，到 2030 年先进数据存力自身有潜力成长为成为千亿级美金产业；此外还将直接带动全球数据要素市场规模化发展，有望在 2030 年达至少 3000 亿级美金；还可作为数智化解决方案的组成部分，使能 10 万亿美金级智能经济的发展

- **提升社会福祉方面**，先进数据存力的应用有潜力带动全球及中国数据中心降碳 **16%与24%**，并作为以数智化解决方案的重要组成部分持续赋能政府治理与民生福祉水平的提升，到 2030 年预计为全球及中国社会分别带来 **3 万亿美金及 1 万亿美金**的经济效益
- **强化区域竞争力方面**，先进数据存力赋能新质生产力的发展。AIGC、元宇宙、脑机接口、量子信息、人形机器人、生物制造、未来显示、未来网络与新型储能等 **9 大未来产业**均需先进数据存力的使能、方可充分发挥其价值

虽然先进数据存力可有效赋能千行百业加速数智化转型，但不同行业企业所面临的情景与业务需求不尽相同。因此，**全球各区域与企业应体系化规划并推进先进数据存力的建设**、以正确且全面的认知对先进数据存力进行适度超前的持续投资、以应对智能经济时代的业务挑战。具体而言，我们建议企业从 **6 大维度出发综合考量先进数据存力的建设**：容量规划、资源利用率、性能要求、安全可靠防勒索、方案级 TCO 与原生 AI 应用赋能。

为使能全球各区域与各行业企业科学客观地评估自身在先进数据存力建设现状，评估未来建设先进数据存力的具体策略，我们于本白皮书发布了“**先进数据存力指标体系**”，并对**全球主要经济体进行了梳理分析**，并将全球各区域细分为“普惠闪存型”、“充足传统型”、“精益聚焦型”与“后发渐进型” **4 大类**

- **“普惠闪存型”**国家存力充足性以及闪存化率均较高，通过定位全球存储产业地位、出台激励与规范政策、推动联合创新以及鼓励头部企业采纳先进方案，全面发展国家先进数据存力产业，这些国家未来将继续强化在存储技术创新、数据安全与可持续性方面的投入，巩固其全球领导地位
- **“充足传统型”**国家存力充足性处于世界领先地位，但闪存化率低于全球平均。这些国家通常已有较好的存力容量，但闪存化进程缓慢，未来可考虑通过政策支持加速全闪化方案在数据中心场景的替换

- **“精益聚焦型”**国家闪存渗透率高，但存力充足性较低。这些国家未来需加强对新兴产业的支持，并通过普及先进数据存力技术，使更多中小企业和个人受益
- **“后发渐进型”**国家存力充足性和闪存化率均较低，发展较缓慢，但拥有“弯道超车”的潜力，这些国家可以通过直接投资全闪化方案，跳过传统存力技术阶段，实现快速发展

对有志于发展先进数据存力产业的国家，我们有三点核心建议。首先，一个国家需明确认知自身在全球产业格局中的位置，为自身在全球存储全栈产业生态中进行“战略定位”（如需全栈化发展、还是仅核心零部件的突破）。其次，基于产业的“战略定位”制定**发展目标、提供政策支持、鼓励技术创新**。尤其对中国这种旨在成为全球先进数据存力全栈技术创新引领者的国家而言，以政策支持与产学研协同等机制鼓励全栈技术创新十分重要。最后，各国在市場应用上应在公共部门（如政府及银行等国有部门）积极开展先进数据存力方案试点，推动先进数据存力应用的**落地与规模化发展**。

# 目录

先进数据存力：加速智能经济涌现的“高性能引擎” .....	1
执行摘要 .....	1
先进数据存力定位：迈向智能经济的“高性能引擎” .....	6
1.1 先进数据存力蓝图 .....	7
1.1.1 智能时代远景展望 .....	7
1.1.2 先进数据存力蓝图 .....	13
1.2 先进数据存力定位 .....	14
1.2.1 “先进数据存力 x 数据要素市场”：数据有序流通的“加速器” .....	15
1.2.2 “社会经济转型 x 先进数据存力”：通用智能涌现的“催化剂” .....	18
1.3 先进数据存力核心目标 .....	20
1.3.1 集约化汇聚全域异构数据 .....	21
1.3.2 安全可靠绿色的数据存储 .....	23
1.3.3 高效使能多元化应用生态 .....	27
1.4 先进数据存力八大特征 .....	29
1.5 先进数据存力的核心价值及影响力 .....	33
1.5.1 产业发展升级：先进数据存力以“核心 ICT 使能产业－数据要素全产业－产业数字化赋能”三层使能机制，带动海量产业经济发展 .....	34
1.5.2 提升社会福祉：先进数据存力使能绿色发展、良政善治与民生体验优化 .....	36
1.5.3 强化区域竞争力：先进数据存力赋能新质生产力生态发展 .....	39
先进数据存力指标评测：建设指标体系保障先进数据存力高质量发展 .....	45
2.1 发展先进数据存力的六大核心考量因素 .....	46
2.1.1 容量规划 .....	46
2.1.2 资源利用率 .....	47
2.1.3 性能要求 .....	48
2.1.4 安全可靠防勒索 .....	50
2.1.5 方案级 TCO .....	51
2.1.6 原生 AI 应用赋能 .....	52
2.2 先进数据存力衡量指标体系 .....	53
2.2.1 区域发展级 .....	54
2.2.2 数据中心级 .....	55

2.2.3	存储产品级.....	57
2.3	全球各国先进数据存力发展程度对比.....	58
2.3.1	体量衡量：充足供应与持续增长支撑经济社会的可持续增长.....	58
	指标 1 单位 GDP 存储容量.....	58
	指标 2 数据存力充足性.....	59
	指标 3 数据存留率.....	60
	指标 4 存力投资增长率.....	64
2.3.2	效率衡量：10 倍性能提升与内生智能特征加速数字经济规模化增长.....	65
	指标 5 通用计算存算比.....	65
	指标 6 闪存占比.....	67
2.3.3	基本保障：满足安全稳定的基本发展要求.....	68
	指标 7 灾备覆盖率.....	68
2.3.4	前沿创新：绿色先进实现可持续发展需求.....	70
	指标 8 单位存储容量能耗.....	70
	指标 9 数据存力专利占比.....	71
2.4	全球各区域发展存力的高阶发展建议.....	73
2.4.1	全球各区域存力发展现状.....	73
2.4.2	典型国家建设先进数据存力的核心启示.....	75
<b>行动倡议：发展先进数据存力是实现新质生产力.....</b>		<b>84</b>
3.1	先进数据存力产业发展目标建议.....	85
3.1.1	将数据存储作为 ICT 产业的重要独立门类单独管理.....	85
3.1.2	确保先进数据存力产业全栈自主可控高效协同，强化先进 SSD 颗粒的国际定价权获取.....	85
3.1.3	推进以全闪化为核心的先进数据存力方案在千行百业的广泛应用.....	86
3.2	建设先进数据存力发展高阶行动建议.....	87
3.2.1	“定目标”：将区域级先进数据存力指标体系纳入政府产业发展目标.....	87
3.2.2	“巧引导”：围绕先进数据存力产业发展目标，制定监管与鼓励性政策.....	87
3.2.3	“促应用”：积极推动先进数据存力方案的在金融、政务、电信等关键行业的应用，鼓励国产开源数据库等开源软件生态发展.....	88
3.2.4	“强创新”：资助并以产学研联合实验室为核心载体，提前布局超算存储等 N+1 代先进数据存力解决方案.....	89
<b>附录 1：参考文献.....</b>		<b>91</b>



# 01

## 先进数据存力定位

迈向智能经济的“高性能引擎”







## 1.1 先进数据存力蓝图

### 1.1.1 智能时代远景展望

回顾历史，三次工业革命为人类带来了深远的变革。第一次蒸汽时代通过机械化生产提升了生产力，第二次电气时代通过电力驱动的广泛应用显著提升了生产力，前两次以硬件创新为核心，带来了物理世界的生产力跃迁。第三次信息革命则构建了一个数字化的世界，辅助并丰富了物理世界。展望未来，人类文明将从数字化时代迈入智能化时代，数字世界与物理世界深度融合，由点及面地增强了“数实融合”进程，使能数字技术与实体经济深度结合，数字智能孪生世界将应运而生。

展望未来，人类文明将从数字化时代迈入智能化时代，数字经济的新阶段—智能经济也应运而生。智能经济在吸取数字经济发展优秀经验的同时，还将兼顾到更多的社会发展需求。

6G、AIGC 等下一代技术的无缝融合，将驱动孤立、分散运作的单点智能模式向多智能系统联动的新模式发展，从而使能智能经济通过创新型的智能解决方案来提高生产力、提升社会福利和改善环境效益，促进虚实产业的融合发展，带动新一轮的经济高速增长。

图 1-1 智能时代远景展望



具体而言，我们认为智能经济有“三新”。一是数据、能源与技术的重要性提升，升级为智能时代三大核心生产要素，促进通用智能“涌现”，二是通用智能体的涌现将端到端重塑企业数据治理范式、加速数据要素市场规模化发展，三是“通用 AI 智能体”将出现，成为智能时代千行百业实现生产力跃迁的核心驱动力，丰富人类精神文明。

### 展望一：数据、能源与技术的重要性提升，升级为智能时代三大核心生产要素，促进通用智能“涌现”

首先，智能经济时的数据量与数据形态复杂度都在持续提升，非结构化数据体量的快速增长与多模态化技术带来的数据张量化 (Tensor) 成为重要趋势。

一方面，数据量正在快速增长，我们预计到 2030 年，全球产生的数据将达到 YB 级，且数据的产生场景将“云边端泛在化”。在边侧，城市治理的视频监控、工厂内高清质检摄像头产生的图片、视频等数据每年有望达到 ZB 级；在端侧，手机、PC 与智能网联电动汽车等千亿级智能终端，源源不断地收集并上传语音、文字、视频、图片等用户交互及路况导航数据；多行业、多类型的云协同发展、显著增强数据流动性，使得更多的数据被高效利用。

另一方面，数据语义、知识图谱等技术的革新与成熟将牵引海量多模态数据加速整合与普及，将使组织能够更轻松地收集和处理不同类型的数据（如图像、语音、文本、传感器数据等）整合在一起，形成张量化数据集群。正如 AMD Zen 架构设计者、前特斯拉自动驾驶硬件副总裁吉米凯勒预测，“在未来，软件和硬件将协同进化，向量计算转向张量计算，推动 AI 技术的进一步发展”。海量多模态数据将成为组织在智能经济时代赋能自身提质降本增效的核心生产要素。

其次，实现智能涌现将需要以太瓦时 (TWh) 为单位的海量能源为基础，使能大模型等原生 AI 应用。根据美国能源资讯署数据，2023 年谷歌和微软当年各消耗的电力均达 24 TWh，相当于人口 1024 万阿塞拜疆一整年的用电量，也超过了全球 100 多个国家，如冰岛、加



纳、突尼斯等。在可见的未来，AI 所带来的电力消耗也会呈现持续上升的趋势。展望智能经济时代，承担 Twh 量级能耗将成为实现智能经济的基础。

最后，以 ICT 基础设施 (6G、先进数据存力等)为核心的数字使能技术，也将有效加速多元异构数据的规整汇聚，加速“智能涌现”。无论数据以向量 (Vector) 还是张量 (Tensor) 的形式存在、也无论未来数据体量有多庞大，海量的数据都需要被集中汇聚、高效计算、安全可靠存储与有序流通，才有可能出现智能涌现。

## 展望二：通用智能体的涌现将端到端重塑企业数据治理范式、加速数据要素市场规模化发展

首先，通用 AI 智能体会在组织内端到端重塑当前“采集－汇聚－治理－分发－保障”的数据治理范式。

第一，数据的采集将不仅仅基于既定规则，通用 AI 智能体会主动要求新数据点的采集与形成以供自身主动学习。智能体从网页、边缘、终端等场景采集多元化的数据海量多模态大数据也将作为“养料”，持续提升智能体能力的通用性。

第二，源自不同种类、形式、产生场景的海量高价值多模态数据将“奔腾入海”，最终汇聚于数据中心并加速智能涌现。根据 IDC 2023 年报告，当前全球产生的数据量中，源自数据中心的占比已超 38%，预计到 2030 年达 50%以上。

第三，在数据治理环节，随着海量的数据使能数据编织 (Data Fabric) 等创新型数据管理范式在千行百业的成熟与普及，而通用 AI 智能体将成为数据编织技术的重要参与者。除了能够自动完成数据清洗、校验、确保数据质量与一致性外，智能体还可以根据业务对数据的需求、安全敏感性等因素，对数据生命周期管理与热温冷数据分级等数据治理策略类问题，提供实时洞察与建议，确保企业对数据的高效治理、以“硅基智能”使能“碳基决策”。

第四，通用 AI 智能体将深化对数据要素的挖掘，组织内的大部分数据会被更频繁地反复访问/复制，更多的数据将被“消费”。当前，在全球所有“产生数据”(Data Generated)中，90%以上的产生的数据是将组织内现有数据复制/备份产生的数据(Data Replicated)，而非新发生并被收集的“被创造数据”(Data Created)。随着通用 AI 智能体对数据要素价值的持续挖掘，其将在组织数据全局可视可管的基础上，主动为用户推送数据、提升数据分析效率与数据分析的智能化程度，实现从“人找数据”到“数据找人”的转变，进而带动组织内热数据与温数据的占比将持续提升。

最后，硅基智能体的蓬勃发展与组织间通用 AI 智能体的交互，以组织间数据交换及数据服务为核心的数据要素产业将加速发展。早在 2016 年，19 岁的美籍华裔天才少年 Alexandar Wang 与 MIT 同窗 Lucy Guo 创办 Scale AI，专注于利用人工智能和机器学习提供高质量的数据标注服务，帮助企业加速其 AI 和机器学习模型的开发、训练和部署。在短短 8 年的时间内，Scale AI 就成长为硅谷的“明日之星”、估值达 138 亿美元，包括 Meta、Uber、摩根大通等多个行业的巨头均为其客户。

在中国，政府是数据要素产业化发展的重要推动者。2019 年是数据要素市场在国家政策层面初次被提及的时间，数据被正式确立为与土地、劳动力、资本、技术等并列的生产要素。为了充分发挥数据作为新型生产要素的价值，加速数据在不同企业与组织间的流转，进而推动经济社会的高质量发展和创新，中国政府采取了举措。例如，国家数据局于 2023 年 10 月 25 日正式揭牌，旨在统筹推进数字中国、数字经济、数字社会规划和建设，协调推进数据基础制度建设（包括数据确权、流通、分配和治理等方面的制度建设），统筹数据资源整合共享和开发利用。

此外，如雨后春笋般出现并快速成长的数据交易所也是中国数据要素市场的重要组成部分，据截止 24 年 3 月，全国已有近 50 个数据交易中心。通过建立场内交易机制压缩场外交易(Over-the-Counter)空间，确保数据要素的合规、有效流通。据测算，到 2022 年中国数据



要素市场规模已达近 2000 亿元，数据资产市场总规模达 8.6 万亿，带动相关产业数字化潜在受益 34.4 万亿。如果叠加数据资产衍生市场，其潜在总规模可能超过 60 万亿。

### 展望三：“通用 AI 智能体”将出现，成为智能时代千行百业实现生产力跃迁的核心驱动力，丰富人类精神文明

数字经济时代的人工智能应用，更多是基于既定业务规则的算法，这需要开发者进行规则编写以实现智能，即所谓“在代码层面上，有多少人工，就有多大智能”，即使组织可以通过 Devops 等管理方法实现敏捷研发、借助开源社区的“群策群力”快速构建软件版本，一个特定用例的智能化程度通常也存在上限。尤其在定制化需求和小规模数据集的人工智能解决方案中，人力资源投入与智能化程度基本近似于线性关系。即使对于中小型卷积神经网络 (CNN) 等创新型 AI 应用，也常因数据体量与质量的不足而无法实现“智能涌现”，致使这些应用更多聚集于特定行业、特定场景的特定动作，如诊疗图片识别。

展望未来，“通用 AI 智能体”（智能涌现的超级 AI 应用）的到来将彻底改变这一现状。首先，通用 AI 智能体超强的技术集成能力、将自身打造为“全能型选手”。例如，通用 AI 智能体能够通过集成外挂行业知识库 RAG 等工具，主动理解行业最新动态、为用户提供最精准确切的行业信息。

其次，通用 AI 智能体具备自适应发展特征、确保“只要不断电就持续学习”。通用 AI 智能体的智能水平不会随着人力资源投入的减少而收敛，反而能够随着环境和需求的变化自我调整优化，以持续提升智能水平，乃至主动采集和利用新的数据，以丰富训练数据集和加深对环境的理解。例如，在未来智能 AI 客服有望能根据与客户的互动，自主学习和改进自己的应答方式，通过分析客户的反馈和情绪状态，不断调整话术和响应策略，从而提供更个性化和有效的服务，而非仅简单基于规则进行交互。



此外，通用 AI 智能体将具备“基于自然交互的精确执行”能力，这使得它能胜任诸多复杂且需涉及多个操作环节的工作（如客户服务、运营商网络运维等），使其成为组织内可媲美诸多“碳基员工”的“硅基专家”，有能力胜任多个岗位的工作、带来组织的生产力跃迁。

我们可以敏捷研发场景为例进行展望，对通用 AI 智能体对组织效能的提升进行管窥蠡测。

首先，通过海量集成 NPL、静态代码分析工具、自动化测试工具、可视化等技术，通用 AI 智能体可以通过阅读产品开发文档 (Product Requirement Document) 充分理解产品架构的设计。随后，智能体将自动按照项目开发计划甘特图，根据项目的当前进度，在 Backlog 中寻找相应开发任务，并相对独立地完成各子系统/模块的开发。在编码过程中，智能体会主动学习过去未能提交到主干 (Main Branch) 的错误代码以强化自适应纠错能力、使生成的代码符合公司编码规范，并在白盒乃至黑盒类测试不通过时，自动 debug 寻找原因并报错、形成 Lessons Learnt 并提交质量管理体系备份。通过多次开发任务的执行与持续迭代优化，通用 AI 智能体的编码表现将远超大多数碳基员工，并成为实现新一代敏捷开发、与时俱进革新 Devops 内涵的核心驱动力。

最后，通用 AI 智能体将凭借不断学习与进化的能力变得更为泛在，从“一个应用程序”走向能与用户实时主动交互，成为组织打造自身品牌、优化用户体验的核心抓手。正如 Open AI 的董事局主席、Meta 前 CTO 布莱特泰勒所预测，“在未来几年内，我将不会在各个保险公司的网页与 APP 间反复切换，而是选择与 AI 智能体直接对话来为家人添加保费。随着大语言模型的出现，我认为与软件直接交流可能是最符合人体工学的软件交互方式，用户不再需要反复阅读产品使用说明书，这种交互范式将重塑品牌的数字体验。”

总的来说，通用 AI 智能体将凭借自身技术高度集成化、自适应发展及基于自然交互的精确执行这三大特征，对千行百业各个形态的组织进行端到端地重塑，进而大幅度提升产品创新、客户服务与关键业务运作效率，实现生产力的跃迁。

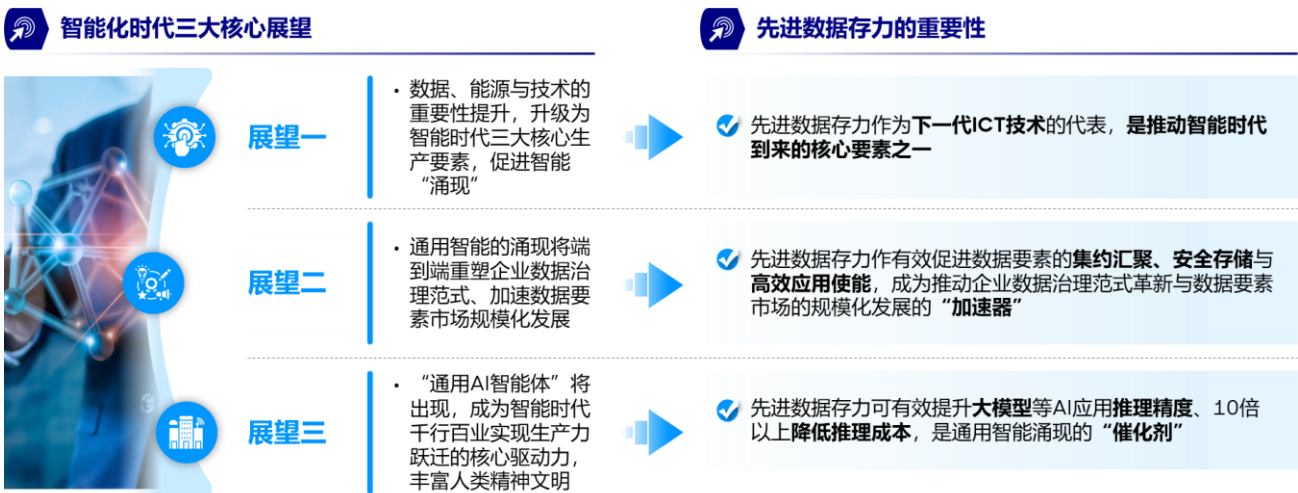


此外，通用 AI 智能体还可极大地丰富人类的精神文明。以 AI 赋能的元宇宙为例，它不仅是虚拟活动的延展，更是数字空间与现实世界深度融合的领域。在这个世界中，数据的积累和运用成为推动文明进步的关键。随着数据量的增长，“数实融合”将更加彻底，元宇宙中的虚实交互也会更加精准，个性化体验随之提升，增强人们的沉浸感。这不仅颠覆了娱乐、学习和工作的方式，也带来了文明形态的深刻变革。未来，数据将成为文化和精神传承的重要载体，推动人类文明朝着虚实融合的更高层次迈进。

### 1.1.2 先进数据存力 蓝图

从定位上看，先进数据存力是智能时代的重要组成部分与核心使能技术，将持续扮演“高性能引擎”的角色，可有力驱动数据要素规模化发展、使能社会经济智能化转型升级。

图 1-2 先进数据存力在智能化时代的重要性



为落实自身定位，我们认为先进数据存力有三大演进方向。一是实现异构数据全域采集，二是确保安全可靠降本的数据存储，三是作为高效使能多元化应用生态。更新一步，基于三大演进方向，我们认为先进数据存力具备全域泛在、性能跃迁、原生智能、集约架构、多级可靠、主动安全、绿色低碳与开放生态八大特征。



从使能场景看，先进数据存力有三大核心价值。一是牵引产业升级，不仅带动数据存储产业链的发展，还在加速数据要素全产业链发展与产业数字化赋能上起到关键作用；二是提升民生福祉、带来良政善治、环境增益与衣食住行等基本民生体验优化，三是赋能 AI 等新质生产力产业生态的发展、提升区域/国家竞争力。

最后，为更好赋能各区域国家与各行业企业评估自身在先进数据存力建设上所处的发展阶段、为国家及企业建设先进数据存力的规划提供参考，本白皮书发布了先进数据存力指标体系，以协助全球各区域政策制定者与企业管理层体系化评估自身在先进数据存力领域的建设现状，并为适配智能经济的到来提供先进数据存力的发展策略。

图 1-3 先进数据存力蓝图体系



## 1.2 先进数据存力定位

先进数据存力是智能时代的重要组成部分与核心使能技术，我们认为先进数据存力在智能经济中将持续扮演“高性能引擎”的角色，可有力驱动数据要素规模化发展、使能社会经济智能化转型升级。



一方面，先进数据存力是数据要素市场的“加速器”，“海纳百川”地汇聚多元异构（图片、语音、视频、乃至多模型向量）数据、对数据要素进行高效安全存储、使能高价值数据跨组织有序流转。

另一方面，先进数据存力是使能通用智能涌现以实现社会经济转型的“催化剂”。先进数据存力可有效强化大模型等原生 AI 应用的推理精度、显著降低推理成本，进而成为实现通用智能涌现的“温床”，加速千行百业应用迈向智能化。首先，以外挂行业知识库 (RAG) 给大模型实时推送最新、最全面的信息，使能大模型等原生 AI 应用“博闻”；其次，KV-Cache 技术多级缓存构建“HBM – DDR – SSD”三级缓存机制，实现大容量 SSD 可长期保留长序列与多轮对话，解决记忆缺失导致交互不连贯等问题，实现大模型的“强记”。

### 1.2.1 “先进数据存力 x 数据要素市场”：数据有序流通的“加速器”

总的来说，我们认为先进数据存力是数据要素市场规模化发展的“加速器”。一方面，数据存储是数据要素市场的重要组成部分、也是数据要素的核心载体；另一方面，先进数据存力将凭借自身高效汇聚海量数据、安全可靠绿色存储与高效使能北向应用三大核心能力，协助数据产业的规模化发展，保障并牵引数据要素的有效流动。

首先，数据存储本身就是数据要素市场的重要一环、也是发展数据要素市场的基础。从数据存储的功能定位看，其是数据要素流通的先导环节，也是确保数据要素市场安全规模化发展的底座。以中国数据要素市场为例，2023 年市场规模达 1200 亿元，其中数据存储环节占 20% 左右，是数据要素市场最大的细分领域之一。

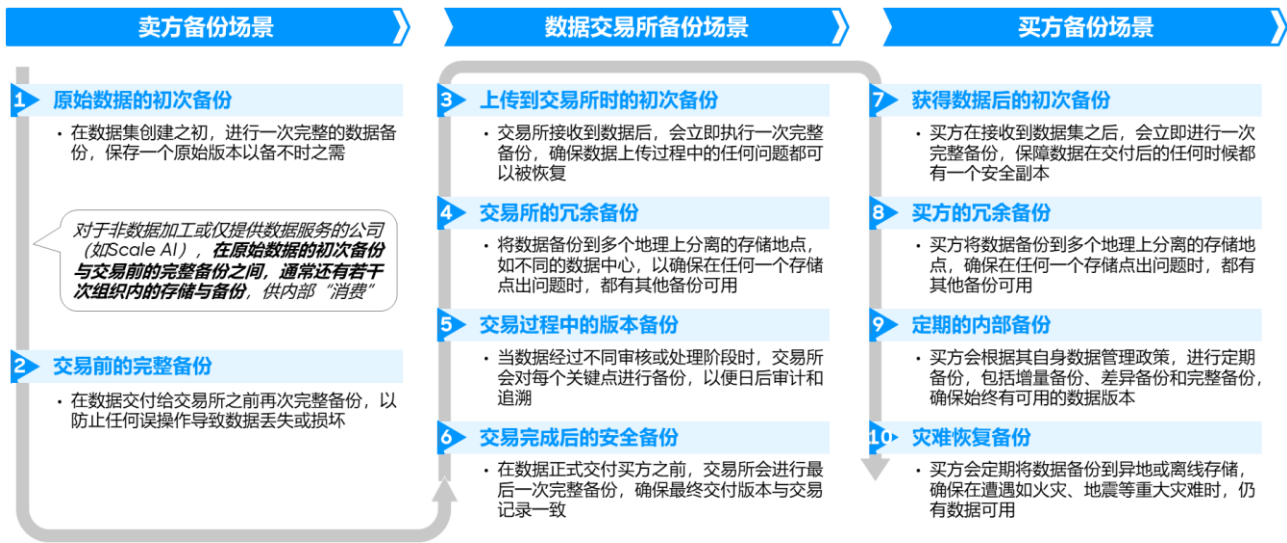
其次，先进数据存力将凭借自身多元异构数据高效汇聚、全链路各环节高效备份与使能数据安全可靠流转等核心能力，成为智能经济时代数据要素的“加速器”，进而使能数据分析、数据加工、数据交易、数据服务、生态保障及数据采集等多个相关环节稳健发展。

一是在数据采集环节，先进数据存力以多协议兼容降低采集难度、加速数据要素化进程。多元异构的数据来源将成为千行百业的数据要素化所带来的必然特征，而由于数据格式与数据类型的差异性，这些来自“千江万河”的数据往往被存储在不同的存储设备中、为数据的流通与汇聚管理造成了极大的挑战，对数据要素化的进程造成影响。面临该挑战，多协议兼容、集约化部署的先进数据存力方案，将极大加速工业数据要素化进程，使得多元异构的源数据真正实现“千江万河汇聚入海”，加速工业数据资产化进程、改善制造企业资产负债表。

二是数据资产在流通环节的脆弱性，每个在数据要素市场的各个环节，都需要先进数据存力提供高效可靠的备份，使能数据资产的全场景安全存储。以数据加工环节为例，目前千行百业对高质量数据的需求，为数据加工产业的发展带来了“源动力”，规模化发展数据加工产业所带来的就业创造等外部性效应也被各国政府所重视。例如，Scale AI 专门下设负责招募数据标注员 (Data Annotator) 的子公司 Remotasks，在全世界范围内广泛吸纳数据加工工人，用于为客户提供高质量数据集，给非洲、东南亚等地区带来了数以百万计的就业岗位。

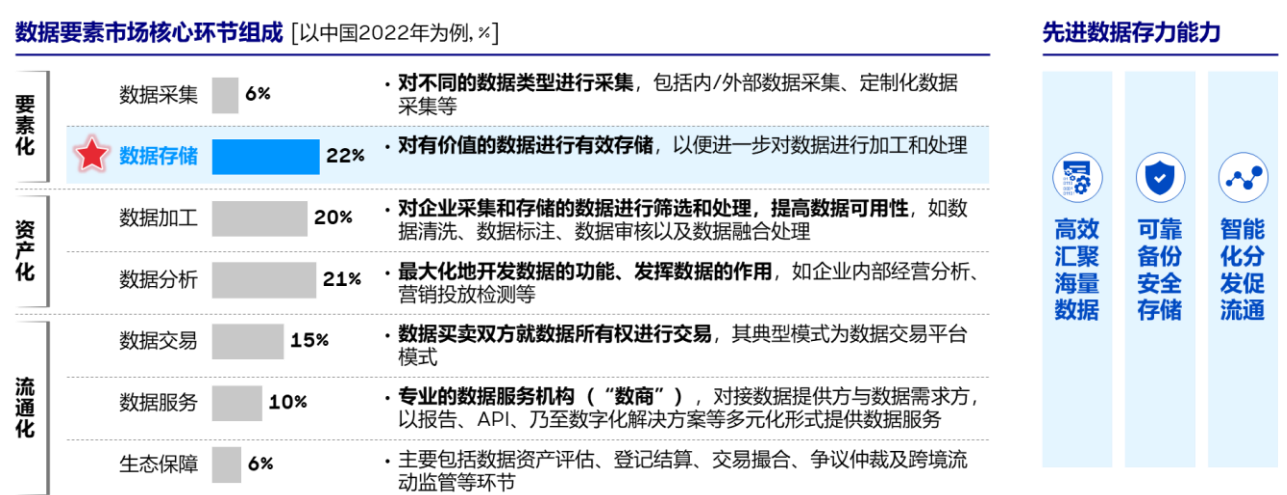
随着全球数据要素市场的发展与数据交易所等中介的普及化，我们认为数据在流转过程中需经过至少 10 次备份，方可确保端到端安全可靠。首先，数据卖方需要进行原始数据的初次备份与交易前的完整备份；对于非数据加工或仅提供数据服务的公司，在原始数据的初次备份与交易前的完整备份之间，通常还有若干次组织内的存储与备份，供组织内部“消费”。其次，数据交易所等中介一般有高标准的安全要求，会配置严密的备份和恢复方案，以应对各种潜在的风险和威胁，在未来可能至少有 4 个刚性备份场景，分别为数据上传到交易所时的初次备份、交易所的冗余备份、交易过程中的版本备份与交易完成后的安全备份。最后，数据买方有至少 4 个备份场景，分别为获得数据后的初次备份、买方的冗余备份、定期的内部备份与灾难恢复备份。

图 1-4 数据要素市场数据备份链路图



三是先进数据存力可有效使能数据资产的安全流转与分发。在数据分析环节，先进数据存力赋能数据编织，实现全局可视可管，为数据服务提供安全底座，进而主动为用户送数据、提升数据分析效率与智能化程度；在数据交易环节，先进数据存力与区块链等前沿存储技术协同，保障赋能，先进数据存力以全栈可靠方案，确保数据交易可信可控；在数据服务环节，在数据服务商的产业模式下，拥有极致安全的数据存储系统，可确保安全交易服务、成为赋能数商产业发展的重要基石；在生态保障环节，先进数据存力在存储层具备的数据加密与隐私保护等特征，为数据要素市场的可持续规模化发展保驾护航。

图 1-5 先进数据存力对数据要素市场的使能



事实上，为使能数据要素的高效汇聚、可靠备份、安全存储与有序流转，打造专业的数据存储基础设施集群已成为智能经济时代的大势所趋。例如，中国政府及产业界联合打造存力中心，积极数据基础底座，赋能新质生产力，就是以存力加速数据要素发展的业界领先实践。

**存力中心推动数实融合，打造端到端数据管理服务中心。**首先，其以区块链技术赋能数据交易、确保数据交易安全可靠、全网共知；其次，支持多源异构数据的存储协议，如块存储、文件存储、对象存储等均可支持。以对海量异构数据高效的汇聚能力，有效推动数据加工的发展，使数据要素从无序低价值到有序高价值；最后，存力中心通过端到端数据要素治理，打造多元化数据服务赋能千行百业、强化社会福祉。例如，存力中心在政务领域打造快速企业资质快速获取能力并应用于成都等城市，以前需要跑 8-9 个单位才能获得资质，如今可大幅精简资质审批流程，实现“上一张网，办所有事，最多跑一次，一次能办成”，资质获取时间缩短至 15-20 天。

### 1.2.2 “社会经济转型 x 先进数据存力”：通用智能涌现的“催化剂”

首先，先进数据存力可有效强化大模型等原生 AI 应用的推理表现、10 倍以上地降低单 Token 的推理成本，进而加速“智能涌现”。先进数据存力不是静态的“仓库”，而是动态且高度智能的“催化剂”，以优异的读写性能与软件算法等技术，催化通用智能涌现。

第一，先进数据存力的“以存强算、以存代算”特征，可有效强化大模型等原生 AI 应用的推理表现。“以存强算”特征通过外挂行业知识库有效提升了推理精度、增强了 AI 在处理长文本和复杂对话中的准确性和连续性，使得大模型在与用户交互时有能力进行“超长上下文”对话，而不会突然“断片”；“以存代算”特征则使大模型具备“记忆”能力，通过记录用户与模型的交互记住常见问题的答案，并在类似问题出现时，直接调取存储系统中“现成记忆”，以确保用户快速得到准确回答、避免再次用显卡推理。





第二，**10 倍以上地降低单 Token 的推理成本，降低大模型训练门槛**。我们以万卡智算场景为例，经测算当前运营一座能用于大模型训练与推理的万卡规模智算中心需要每年投入高可达 10 亿人民币，如果其中 10% 的 Token 通过“以存代算”的模式推理出，那么**综合推理成本的节降空间在未来可达亿元级**。事实上，诸多热门应用场景的 Token 重复率远超 10%，如客户服务中心大约 60%-80% 的查询是重复性问题，在企业的在线自助服务平台和用户论坛中 75%-85% 的问题可能是重复性查询。这意味着“以存强算、以存代算”有着广阔的应用前景、也将成为“智能涌现”的有效助推器。

其次，智能经济相较数字经济，还伴随组织、流程与管理机制的全面革新，而先进数据存力在其中将有效加速组织流程与管理机制的创新。例如，在过去几年中随着数据量的激增，全球有越来越多的组织强化“**首席数据官**”岗位的设置，以强化组织对数据价值的挖掘与安全合规管理。据某全球性企业调研显示，任命了首席数据官的企业从 2011 年至 2022 年月跃迁了 28.5%，其中北美地区渗透率达到 38.5%，亚太地区也提升至 10.2%。

然而，积极以组织变革拥抱数据要素价值的企业在数据管理上面临着诸多挑战。一是各部门可能由于**数据标准不统一**，导致数据的转换与治理的成本随数据量的提升而快速增长。例如，不同的事业部在如何将管理成本分摊到各产品的问题上可能采取不同的会计处理办法，使得将这些数据统一并提供给数据需求方时，需要投入额外的成本；二是**数据服务的自助化与智能化程度有进一步提升空间**，限制了业务部门对数据的挖掘潜力。例如，诸多部门可能有高频读取数据的需求，而这些需求可能因数据提供方是“碳基员工”而无法实现“24 小时自助式数据服务”。

面临种种数据治理与组织管理上的问题，先进数据存力可作为核心使能技术，牵引组织管理与流程体系的创新。例如，该企业可以将公司内部所有系统的数据都汇聚在**集约化存储资源池**中，以湖仓一体等创新解决方案为载体，对财务职能可能需要用到的所有数据进行统一管理与及时更新；二是建立所有业务职能均可统一访问交互的**AI 智能体**，实现数据服务的

“自助化”；三是由财务职能确立统一的会计处理准则并输入智能体，使其能够理解这些会计处理准则(如成本均摊规则)，如有更新需求可统一输入至大模型，确保无论谁来“取数”结果都一样，财务职能仅需按需更新维护会计处理准则即可。总的来说，数据治理与构建良性的数据服务，无论对什么行业的组织来说，都是一个“长坡厚雪”的系统工程、需持续投入方有成效。而**先进数据存力**，为组织管理范式与流程机制的变革提供了新的可能性。

最后，先进数据存力还可在智能经济时代促进社会的可持续与良性发展、强化各国在未来关键产业中的竞争力。一方面，智能经济相较数字经济更为普惠，除赋能千行百业提质增效外，还可促进绿色发展、良政善治与民生福祉。如将这些成效换算成经济效益，由先进数据存力牵引的正向社会经济效益可达万亿级。

另一方面，先进数据存力也与诸多中长期应用前景广阔的产业联系紧密，布局先进数据存力就有助于强化各国在未来关键产业中的竞争力。以中国政府近年提出的“新质生产力”中的9大未来产业中的量子信息为例，通过先进的数据存储技术，量子通信网络能够更有效地存储与管理大量的量子密钥和纠缠态信息，确保这些数据在需要时能够快速、准确地访问和使用。例如，当银行在全球范围内进行跨国金融交易时，量子存储器可以安全地存储和传输加密密钥，确保交易过程中不受黑客攻击和数据泄露的威胁。

## 1.3 先进数据存力核心目标

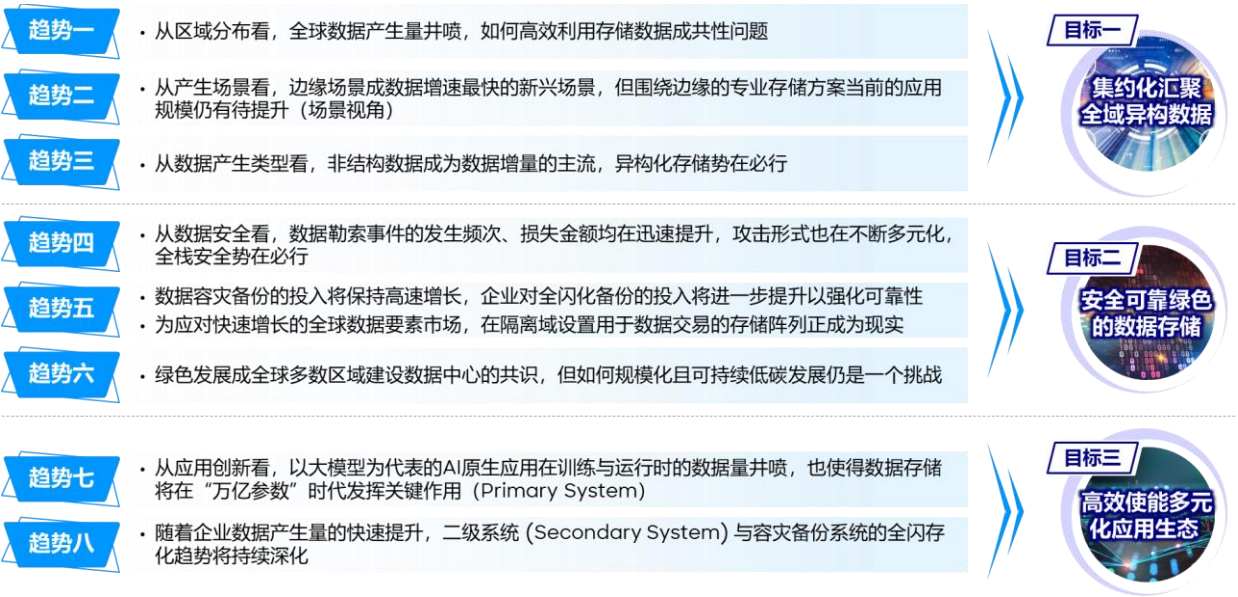
全面审视先进数据存力行业的变化趋势，我们发现全球数据产业呈现**8大趋势**，对数据存储行业提出了新的要求。基于全球数据产业的趋势与先进数据存力在智能经济中的定位，我们认为先进数据存力的目标有三。一是**集约化汇聚全域异构数据**，对多种类型的数据“海纳百川”；二是**安全可靠绿色的数据存储**、确保数据要素在组织内与跨组织流转时的安全可靠与





绿色低碳；三是**高效使能多元化应用生态**，先进数据存力以优异的性能，为组织核心系统及二级系统等多元化北向应用（如 AI 大模型）的高质量运行提供有力支撑。

图 1-6 先进数据存力三大发展目标



### 1.3.1 集约化汇聚全域异构数据

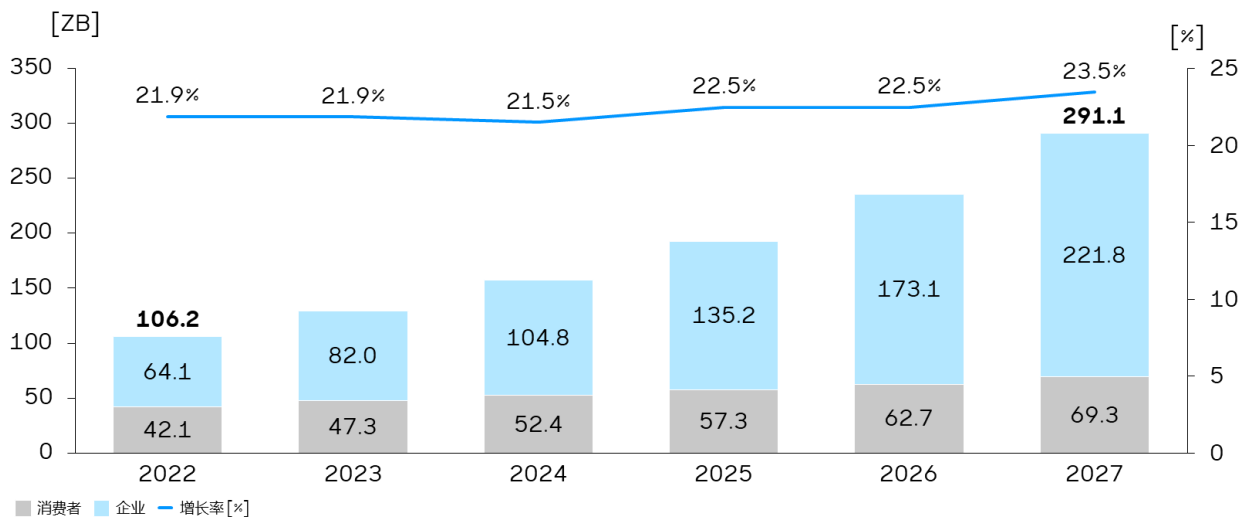
#### 趋势一：从区域分布看，全球数据产生量井喷，如何高效利用存储数据成共性问题

参考 IDC 报告，虽然 21-23 年全球各区域增量保持 **10%以上** 增长，但与此同时中美集中度进一步提升（两国占比从 20 年仅左右 50% 到 23 年接近 60%）。以美国、中国为代表的大国存储方案与产业发展范式，将不仅仅与本国存储产业的发展相关，也将有条件引领全球数据存储产业发展范式、为其他区域提供“范例”。

图 1-7 全球数据产生量



全球数据产生量 [ZB, 2022 - 2027]



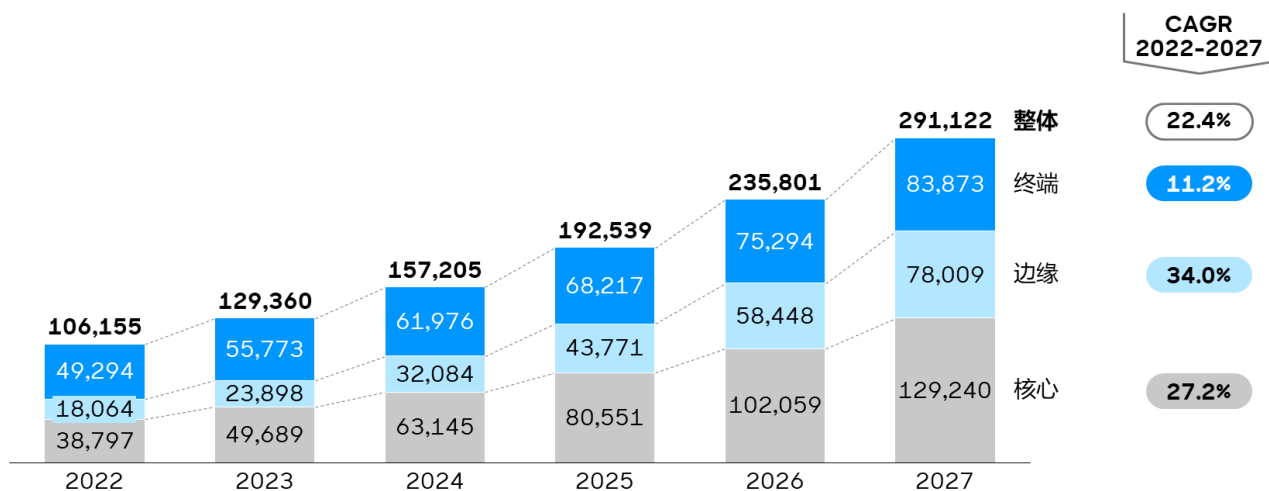
**趋势二：从产生场景看，边缘场景成数据增速最快的新兴场景，但围绕边缘的专业存储方案当前的应用规模仍有待提升**

根据 IDC 数据圈报告，在“端”(Endpoint)、“核”(Core)、“边”(Edge)三侧的数据增长中，边缘数据以 30% 以上速度快增、超过“核”(数据中心)与“端”，以边缘为核心场景的数据存储需求与重要性正快速提升。

图 1-8 全球数据产生场景



全球数据产生场景 [EB, 2022 - 2027]

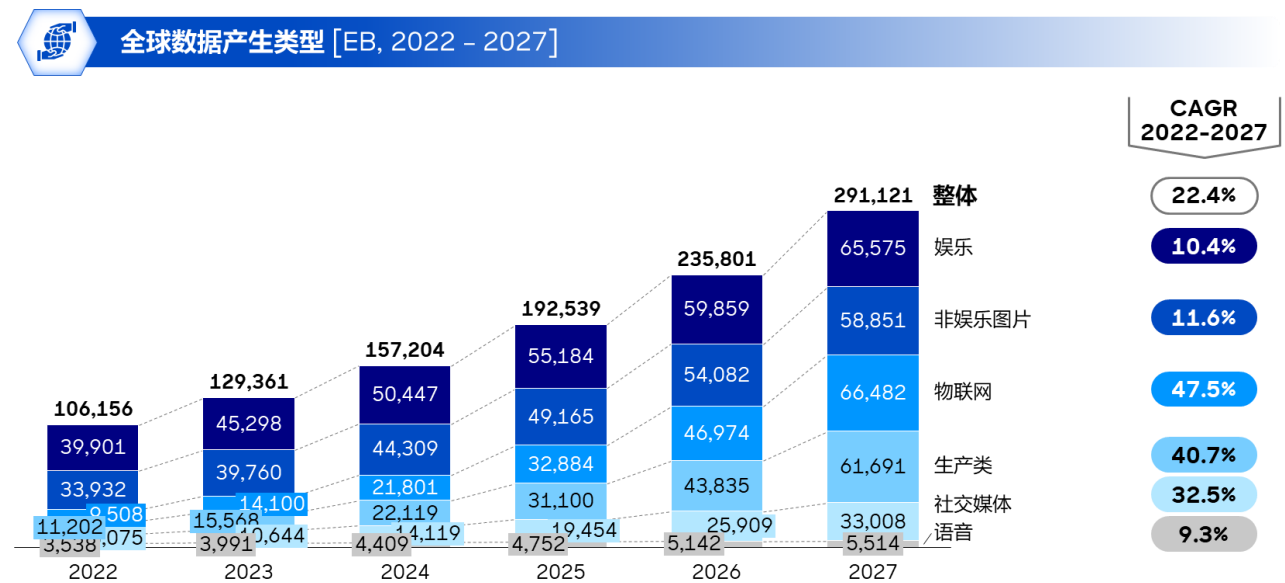




**趋势三：从数据产生类型看，非结构数据成为数据增量的主流，异构化存储势在必行**

根据 IDC 数据圈报告，23 年全球数据增量中 80% 以上增量来自图片、视频等非结构化数据，存储的场景。如何高质效地存好、用好这些非结构性数据，成为企业与公共部门等数据要素所有方在数据要素加速发展的产业背景下，一个必须回答的问题

图 1-9 全球数据产生类型



### 1.3.2 安全可靠绿色的数据存储

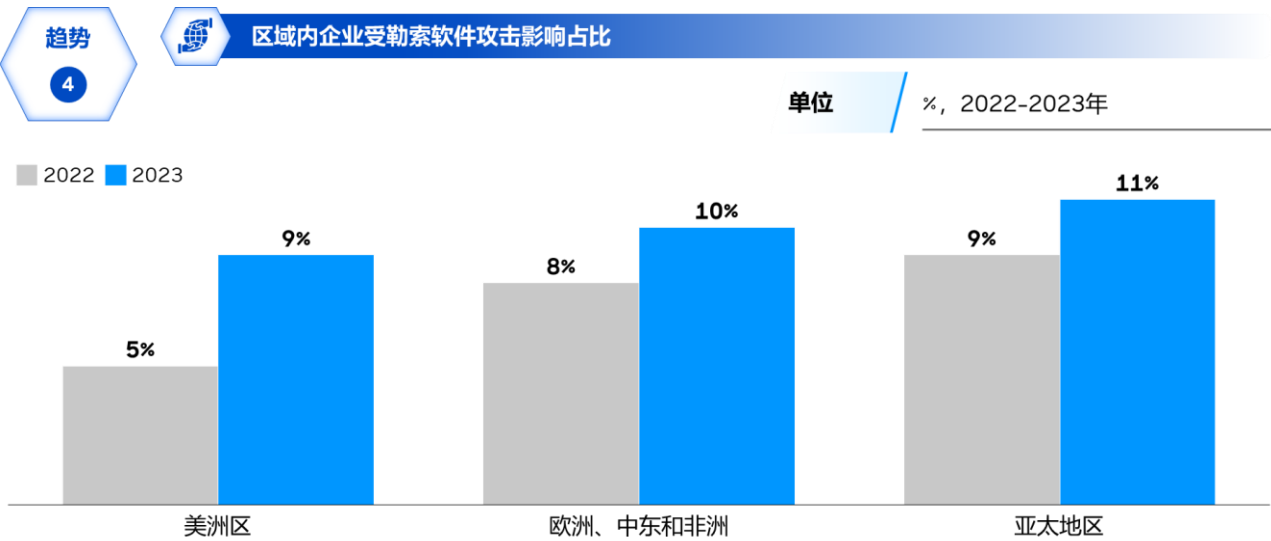
**趋势四：从数据安全看，数据勒索事件的发生频次、损失金额均在迅速提升，攻击形式也在不断多元化，全栈安全势在必行**

当前数据安全形势正变得更加严峻，每年数据勒索事件的发生频次、损失金额都在快速提升。以美国 FBI 的 IC3 年度报告为例，在 2023 年一年中，美国就遭遇了多达 2825 次数据勒索事件<sup>错误!未找到引用源。</sup>。数据勒索软件攻击，直接损失高达 5 千 9 百万美元，与 2022 年相比，勒索攻击次数与损失金额分别上升了 18% 和 74%。在 2825 次的勒索攻击中，有 42% 的事件都是针对美国的关键行业，其中公共医疗、关键制造业、政府机关、科技行业和金融服务业分列第一至

五位。例如，金融服务行业在 2023 的年度累计统计下平均每起攻击事件损失 9 百万美元（含业务中断和停工损失、数据恢复和修复成本损失、品牌名誉损害以及律师费等）。

除更加频繁的勒索频次外，损失金额也在不断提升。Sophos 的 <The State of Ransomware> 表明，如不单单考虑赎金支付的损失，平均总成本则达到了支付赎金的 9 倍，而在仅仅两年之后，平均恢复总成本则达到了 15 倍<sup>2</sup>。以中国为例，2023 年，采用双重/多重勒索模式的勒索组织相比 2022 年新增了 36 个，勒索组织在要求更多赎金的同时，还会窃取企业数据以数据泄露风险威胁企业以及企业合作伙伴，在 2750 起事件中，有超两成案例窃取了超 500 GB 以上的企业及个人数据。

图 1-10 全球部分区域内的企业受勒索软件攻击影响占比

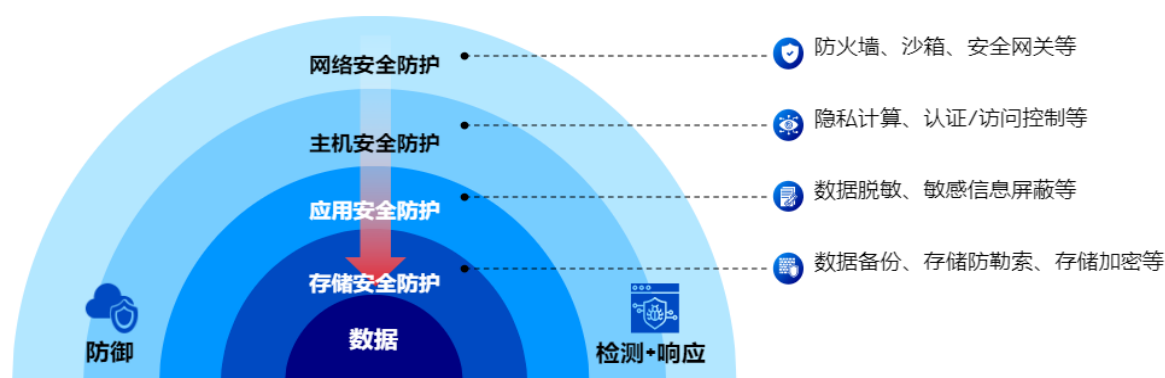


此外，网络攻击方式与技术也变得更为多元，仅凭传统的网络层防护（防火墙等）已不足以应对当前的攻击形式。例如，数据勒索组织开始更多采用如 BYOVD 技术和离地攻击技术（LOTL）复杂尖端技术进行勒索攻击<sup>3</sup>。攻击者利用这两项技术可以利用已知的漏洞驱动程序和合法系统工具绕过系统防火墙，“合理”地收集信息、执行代码，且黑客可以在不引起警报的情况下恶意篡改、删除以及窃取数据。



在此背景下，企业亟待端到端强化自身安全防护能力，构建“网络-主机-应用-存储”四级安全防护体系势在必行。而在四级安全防护体系中，存储防护是数据安全的“最后一道防线”，其重要性不容忽视。

图 1-11 四级安全防护体系

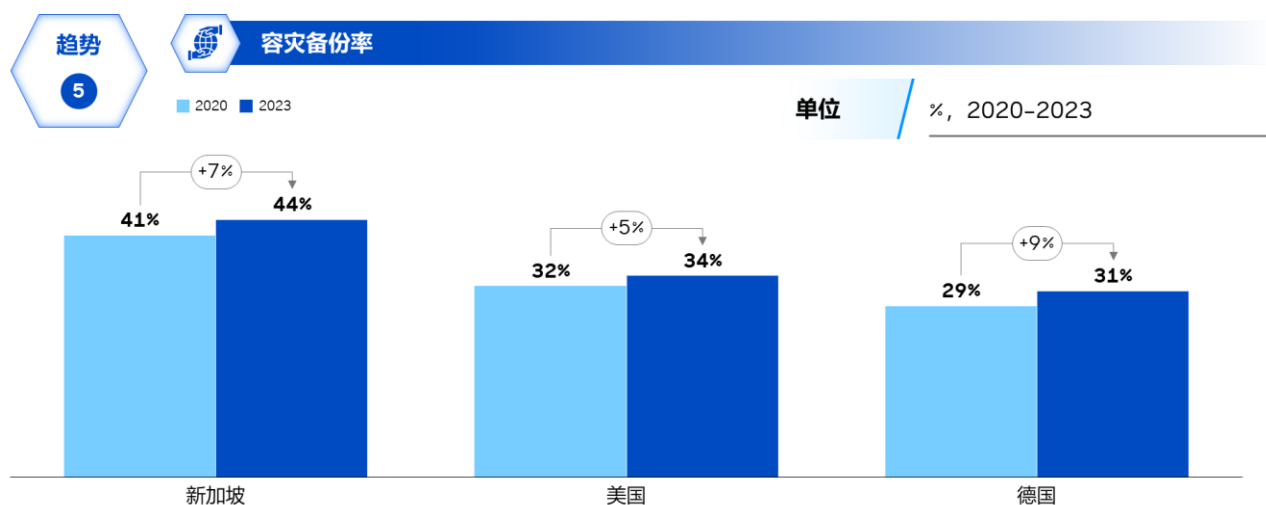


### 趋势五：数据容灾备份的投入稳步增长，企业对全闪化备份的投入将进一步提升

Gartner 报告显示，全球企业对容灾备份（Backup & Recovery）的投入稳步增长。例如新加坡、美国、德国在 2023 年的容灾备份率相较于 2020 年已增长 7%、5%、9%<sup>4</sup>。全球的数据要素正在资产化，而且网络环境越来越复杂，正如趋势四内容所涉及的数据安全内容，企业对备份产品安全性的重视程度也在进一步提升，高效强劲的备份产品也可以极大程度上保证企业的数据安全或挽救企业的损失。

以中国金融行业为例，数据要素的快速发展带来了海量机遇，但也带来了潜在数据安全风险。尽管银行等金融机构当前可以通过身份鉴权、报文加密、沙箱检验等技术提升数据安全性，乃至建设专线对接政府客户进行数据交换，但数据安全风险仍然存在。因此，部分中国国有银行正在考虑在隔离域 (Airgap) 增设专为跨组织数据交易使用的先进数据备份设备，以从根本上解决潜在的数据安全隐患。

图 1-12 全球主要区域容灾备份增长

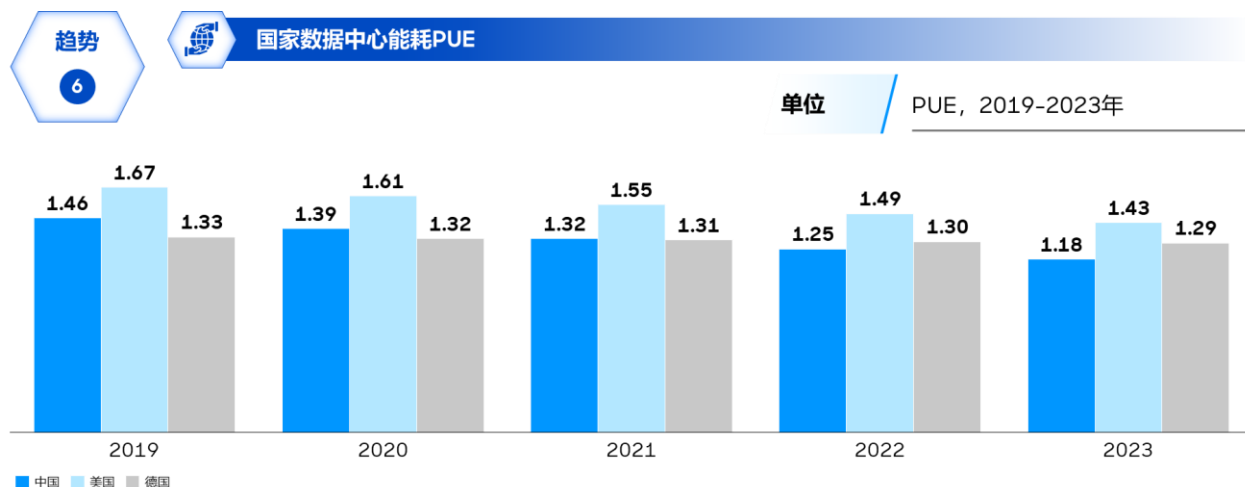


**趋势六：绿色发展成为全球多数区域建设数据中心的共识，但如何规模化且可持续低碳发展仍是一个挑战**

一方面，随着全球主要国家数据中心 PUE 数据在持续下降，IT 设备本身的能耗节降成为重要的绿色化发展趋势。例如，美国 PUE 预计从 2019 年的 1.67 降低至 23 年 1.43，中国 PUE 从 19 年的 1.46 降低至 22 年 1.25，德国 PUE 从 21 年 1.31 降低至 23 年 1.29<sup>5</sup>。

另一方面，俄乌战争带来能源成本上升，使得欧洲区域与中亚等相关对数据中心节能与集约化部署诉求进一步提升，绿色低碳化发展成为更加刚性的需求，例如欧盟的《欧盟能源效率指令》、德国的《能源效率法》和《德国数字化战略》都对数据中心绿色化发展有要求，对全球其他国家降低数据中心能耗的指标与规范设定有借鉴意义。

图 1-13 全球部分国家数据中心 PUE



### 1.3.3 高效使能多元化应用生态

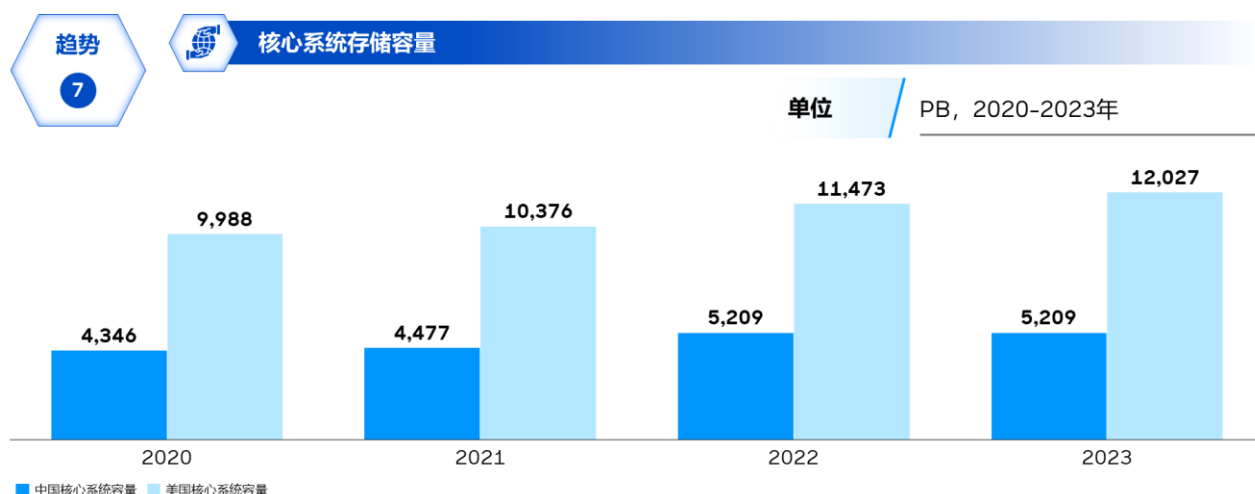
**趋势七：从应用创新看，以大模型为代表的 AI 原生应用在训练与运行时的数据量井喷，也使得数据存储将在“万亿参数”时代发挥关键作用**

随着参数从千亿级向万亿级、乃至十万亿级跃迁，加上数据膨胀倍数带来的数据体量跃迁，针对大模型的数据存储方案也面临革新，更高性能的读写带宽、更快地 IOPS、外挂型知识库等创新性数据存储方案，将深度赋能大模型训练。

2021 年至 2023 年间，在全球外置数据存储设备市场，一级系统容量已经提升了 13.82%，一级存储系统容量从 22082 PB 增长至 25622 PB；同时中国在 AI 大模型成为发展数字经济的重要节点也在提升一级存储系统的容量与投资，投资额从 2021 年 24 亿美元增至 2023 年 29 亿美元，而容量也从 4477 PB 增长至 5209 PB，与全球趋势相同。尽管中国增长喜人，与数据强国美国相比，依然差距明显，2023 年的中国一级系统容量依然不及美国的一半，但通过统计预测结果，中国的一级存储系统将在未来保持稳定提升的态势。

图 1-14 全球部分国家核心系统存储容量





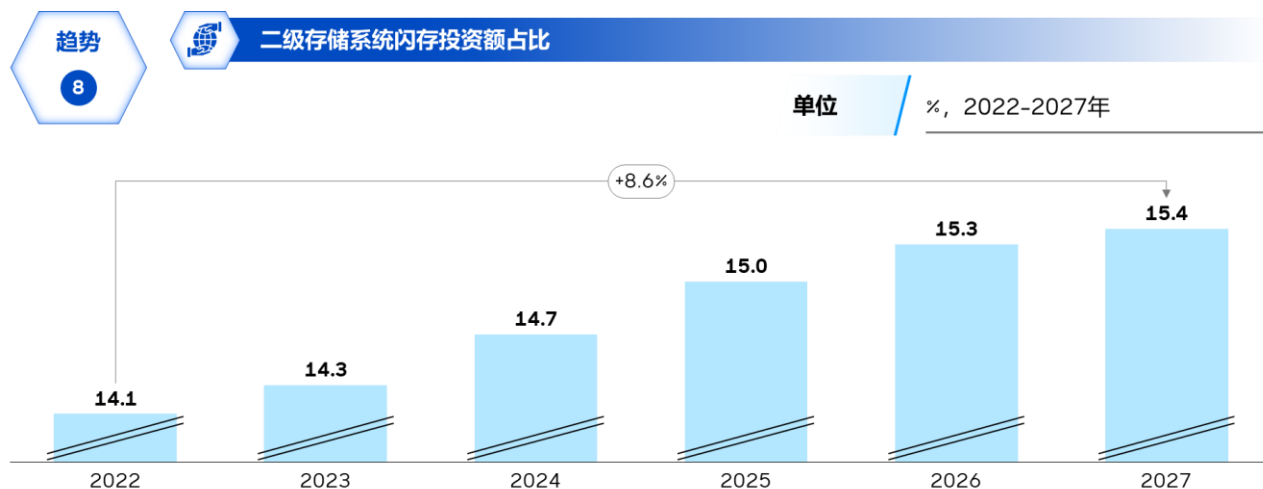
#### 趋势八：随着企业数据产生量的快速提升，二级系统的全闪存化趋势将持续深化

根据 Gartner 最新报告，全球二级系统的闪存占比在逐年提升，对外置数据存储设备市场而言，全球二级系统的闪存投资占比将从 2022 年 14.1% 提升至 2027 年 15.4%。整体而言，我们认为全球二级系统的全闪存化趋势将持续深化，其中核心原因有三。

首先，随着企业数据量的持续提升，企业越来越难以满足在规定的备份时间窗内完成备份，及 RTO 与 RPO 等系统可靠性指标要求的难度在持续提升，使得选型 SSD 的全闪存化方案形成一种趋势。其次，随着大数据智能分析、湖仓一体等创新性应用的逐渐丰富，也使得企业二级系统内的“热温”数据的比重持续提升，这也为性能更为优异的 SSD 方案逐渐将在二级存储中成为“趋势性刚需”。

例如在今年中国“618”打折促销期间，中国某头部运营商在连续三天内产生的移动数据所需要的备份时间每天均超过寻常预留备份窗口时间内 HDD 方案备份上限的 10%，因此传统 HDD 方案已明显不能满足需求。数据备份设备的全闪存化对该运营商而言势在必行。

图 1-15 全球部分国家二级系统闪存投资对比



## 1.4 先进数据存力八大特征

为更好地应对行业发展趋势，数据存储产业应从“存力”迈向“先进数据存力”。我们认为，先进数据存力应具备全域泛在、性能跃迁、原生智能、集约架构、多级可靠、主动安全、绿色低碳与开放生态八大核心使能特征。

### 全域泛在：从聚焦数据中心，到多场景泛在化

边缘数据呈现高度异构化、流通体量大、对集约化要求程度较高等特征，专用化边缘存储方案势在必行。例如，某汽车电子 EMS 工厂每日工厂产线数据在 TB 级，其中质量检测产生的图片数据为重要来源，如何有效管理这些数据对该厂是一个挑战。如果将数据全部上云并非优选；但倘若没有有效的数据存储与管理机制，也会使得这些数据难以价值化。此时，边缘存储成为了一个很好的解决方案：通过将主要的数据存储在边缘场景，仅将重要数据上云（如 IPQC 质检过程中的不合格件），该厂极大缓释了边缘数据体量膨胀带来的管理难题，同时也有效利用了这些数据。

### 性能跃迁：从传统机械硬盘，到端到端全闪化

大模型等 AI 原生应用的普及对存储产品带来性能跃迁需求，千亿级大模型在 10 倍数据膨胀后，对存储系统提出了更高的性能要求，使得全闪化势在必行，进而带来从 GB 级到百 GB 级带宽、从百万级到千万级 IOPS 的性能跃迁，有效赋能 AI 时代业务需求。此外，端到端全闪化也意味着 **SSD 从聚焦核心系统，到向二级系统和容灾备份系统迁移**。随着 SSD 与 HDD 的价格比不断下行，最终 TCO 与性能更优的 SSD 方案也将成为传统意义上“非核心系统”的主流方案。

## 原生智能：从基于既定规则，到以存强算代算

我们认为先进数据存力的原生智能有三重内含，一是对大模型的训练加速赋能、实现存力对算力的强化乃至部分替代，二是使能业务处理提效、缩短断点续训间隔期，三是具备智能涌现特征的设备运维自治。

首先，原生存力智能的核心赋能场景是以 **LLM 大模型为代表的 AI 原生应用**。通过外挂行业知识库、并将最新的行业知识（新闻、图片、音视频等）以“向量形式”存在知识库内，用以强化大模型的预测精度。外挂向量知识库有效地解决了大模型的时效性问题（如 GPT 4.0 当前仍无法获得 2024 年之后发生的事件或数据的最新信息），以“向量存储”为载体、把海量非结构数据处理成多维向量，给大模型输送最新、最全面的信息，有效强化推理准确度。

“以存代算”也是原生存力智能使能大模型的另一特征。当前的 LLM 大模型的本质是“连续预测”，基于用户提出的问题，结合上一个字预测下一个字、形成与用户的持续交互。而以 KV-Cache 技术赋能构建“**HBM – DDR – SSD**”三层缓存机制，保存“工作记忆”（如历史问答记录），解决多轮对话和长序列“场景记忆”问题（比如用户就一个较复杂的问题多轮询问、澄清、再询问），实现高性能记忆存储。相比重新计算，直接调取 KV-Cache 缓存中 SSD 层数据的成本，相比重新推理一遍的成本降低 **10 倍以上**。

第二，以先进性能加速业务处理、缩短断点续训间隔期。随着数据并行、模型并行和流水线并行成为大模型算力基础设施部署运行的范式，使得 LLM 的训练效率得到指数级提升；与



此同时，这种并行架构也使得**高性能专业存储**在**AI 架构**中的重要性持续提升，也对存储 IOPS 与读写带宽等性能提出了更高的性能要求。

**第三，智能涌现式的设备运维自治。**包含基础设施资源规划、设备预测性维护、智能故障诊断、全局可视化等特征，成为企业 IT 基础设施的“智能顾问”，从反应式运维管理到走向智能涌现与“自闭环式”运行。

### 集约架构：从存算分离阵列，到集约化资源池

先进数据存力集约架构的核心内涵有二，分别是存算解耦与以数据为中心。首先，存算解耦的架构将服务器本地盘拉远，构成无硬盘服务器和远端存储池，实现了真正意义的存算解耦，极大提升存储资源利用率。在传统的数据中心架构下，“存力”与“算力”通常作为两个独立的阵列，与网络运力设备组合到一起形成“端到端全栈解决方案”。尽管目前也可以通过存算运三力的各自优化实现一定程度上的集约化部署，但并未对存算独立的阵列架构进行根本性变革。而存算分离的创新型架构，极大**减少**了服务器因折旧周期（通常 3~5 年）低于部分业务数据生命周期（有时可多达 10~15 年）而导致的数据迁移成本。其次，先进数据存力的架构以数据为中心，以 CXL、RDMA 等技术赋能，构建高通量超融合网络、避免了用户数据和控制数据（元数据等）的低效交织，缩短 IO 处理路径，最终实现高吞吐、低时延的极致性能体验。

先进数据存力的创新型架构将带来两大核心影响。一是由存算解耦带来的北向兼容与开放 ICT 基础设施生态，先进数据存力将能够**兼容多元化北向算力 (X86/ARM 等)**，促进更多元化解决方案的形成；二是以集约化部署能力使能存力走向全场景，**加速先进数据存力在边缘智能等场景的加速普及渗透**，提升存力的泛在性。

### 多级可靠：从结果导向管理，到四级可靠体系

先进的存储系统解决方案的可靠性应包含“数据级可靠－部件级可靠－解决方案级可靠－云级可靠”四个层级。例如，数据层级应以异构算力（如 DPU）等技术手段进行 EC 技术处理，实现对数据的合理冗余保护；再如，部件层面确保 SSD 可靠且业务均衡负载，而非简单“集成硬盘”；又如，解决方案层面确保极致高韧、即使控制器只剩 1 个也能确保业务连续性、在分布式集群存储下能够满足双活与业务连续性需求、且在全闪化方案中可以不牺牲快照等增值功能的可靠性；此外，云级场景有效支持对象存储场景下的文件级颗粒度恢复。

### 主动安全：从被动应对冲击，到主动全栈防护

在网络环境愈发复杂、勒索成本逐年升高、攻击方式持续多元化的时代，企业在安全管理上应“被动为主动”，形成“网络－主机－应用－存储”全栈数据安全防护体系，而数据存储的主动安全特征，也将成为先进数据存力的重要组成部分。先进数据存力的主动安全有四大核心内涵。一是“主动”，构建事前、事中、事后端到端防勒索体系，二是“高韧”，能够支持数据快速恢复（如分钟级 CKPT 数据恢复）；三是“可视”，清晰可视的掌控系统容灾业务的运行情况，快速方便的完成数据恢复和测试演练，四是“全域自治”，能够为多类型存储设备（全闪存、分布式存储、安全备份一体机等）提供安全策略统一配置管理、全局侦测分析及主动防御能力，也能在边缘、中心等多场景提供安全可靠的数据保护。

### 绿色低碳：从政策要求驱动，到先进方案牵引

在全闪化介质与集约化架构的赋能下，先进数据存力从存储介质、软件算法与数据处理架构等多方面齐创新，可有效降低数据存储系统的能耗水平，为数据中心的可持续绿色化发展提供了一个可规模化复制的解决方案。首先，数据存储系统大部分能耗来自于存储介质，而 SSD 的选用与单盘容量的持续提升，将带来显著的能耗节降。

一方面，相同容量下 SSD 相比机械硬盘的能耗降低 70%，空间占用节省 50%；另一方面，SSD 单盘容量持续提升，未来主流 SSD 单盘容量有望达 HDD 的 2 倍以上，使能数据中心



节约化部署。除先进介质外，软件算法与数据处理架构等创新也将进一步使能存储设备的集约化部署、加速绿色转型，如数据压缩技术使得更小的存储空间可容量更大体量的数据；近存计算与专用数据处理器 (DPU) 带来的架构创新可有效减少数据搬迁成本；高密盘框的设计使得相同空间内可存储更多的数据，从而减少了每 TB 存储的能耗和相应的碳排放。

## 开放生态：从数据存取 ATM，到自助数据消费

当前的存储设备一般提供块、文件、对象等基础数据接口，进一步可以提供诸如 Table 格式对接数据库应用，DataSet 向量接口对接训练推理型应用，资产类接口对接数据交易型应用等。在未来，数据服务和 API 通过提供更先进的功能、更好的性能和更安全的数据访问，可以超越传统的数据源或表界面，通过自主创建更复杂的应用程序对外提供创新的方式来利用数据的力量。例如，在未来数据接口可以与 NLP 技术整合，提供 ChatGPT 型服务接口，使用户能够使用自然语言与数据互动。

## 1.5 先进数据存力的核心价值及影响力

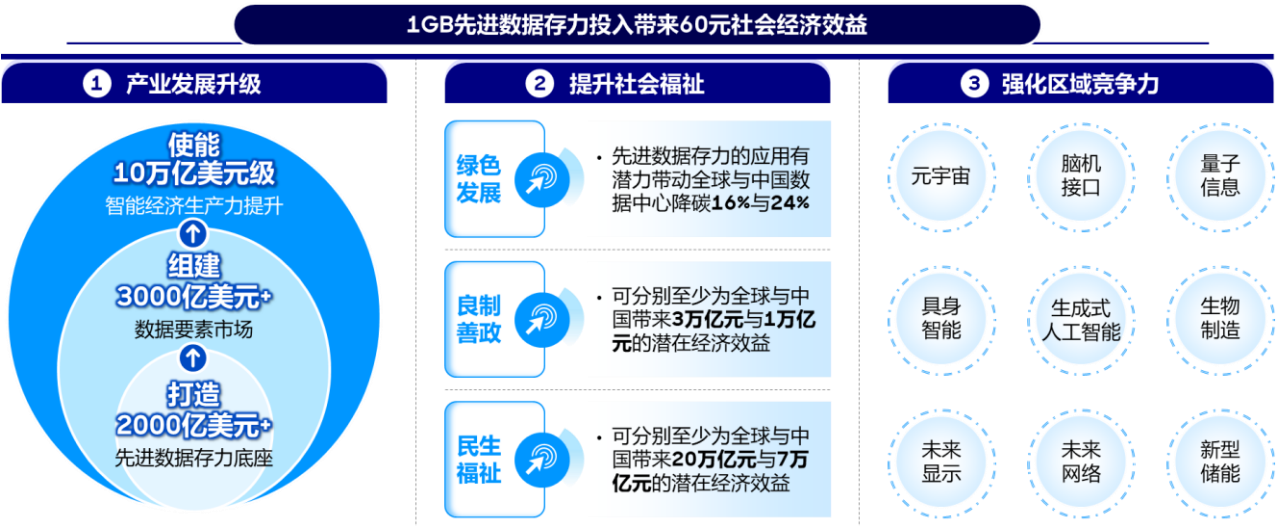
海量的数据使能千行百业走向智能经济、增加智慧化政府治理能力、提升人民生活幸福度。从城市治理、绿色发展，到民生的衣食住行医疗保障，再到企业的数智化升级打造新质生产力，AI 应用将更为普惠且泛在。总的来说，我们认为先进数据存力在赋能产业升级发展、提升社会民生福祉与强化区域竞争力三大价值与影响。

据测算，1 GB 先进数据存力投入预计能带来至少 60 元 GDP 的增长。总的来说，我们认为先进数据存力技术的广泛使用能够提升价值回报原因有三。一是性能 10 倍提升的同时不丢失可靠性，可有效使能全场景应用；二是存算分离架构的集约化资源池带来更高的单位容量



利用，有潜力将存储利用率头绳 30% 以上；三是有效使能 AIGC 等具备“智能涌现”特征的新质生产力行业，显著提升推理精度、10 倍以上降低推理成本，带来巨大经济价值。

图 1-16 先进数据存力的核心价值与影响力



### 1.5.1 产业发展升级：先进数据存力以“核心 ICT 使能产业 - 数据要素全产业 - 产业数字化赋能”三层使能机制，带动海量产业经济发展

首先，先进数据存力作为智能经济时代的核心 ICT 使能技术，未来有潜力成长为千亿级 ICT 使能产业。第一，数据中心的存算网架构将在先进数据存力的牵引下发生结构性变化，专业存储设备的市场规模将快速提升，并将在数据中心建设成长为一个独立且核心的模块；第二，先进数据存力将与 AI 大模型等创新型数字化技术深度耦合，共同成为智能经济时代的“原生智能解决方案”，使得先进数据存力底座成为智能经济时代的“势在必行”；第三，当前全球主流区域的存力与算力配比并不均衡，在多数地区存力的投资增长显著低于算力，先进数据存力产业的可持续增长潜力巨大。

据测算，到 2030 年全球先进数据存力产业的发展潜力可至少达 2500 亿美元。作为全球算力第二大国，预计中国到 2030 年的算力规模在全球占比将至少达 30%，这意味着届时中国

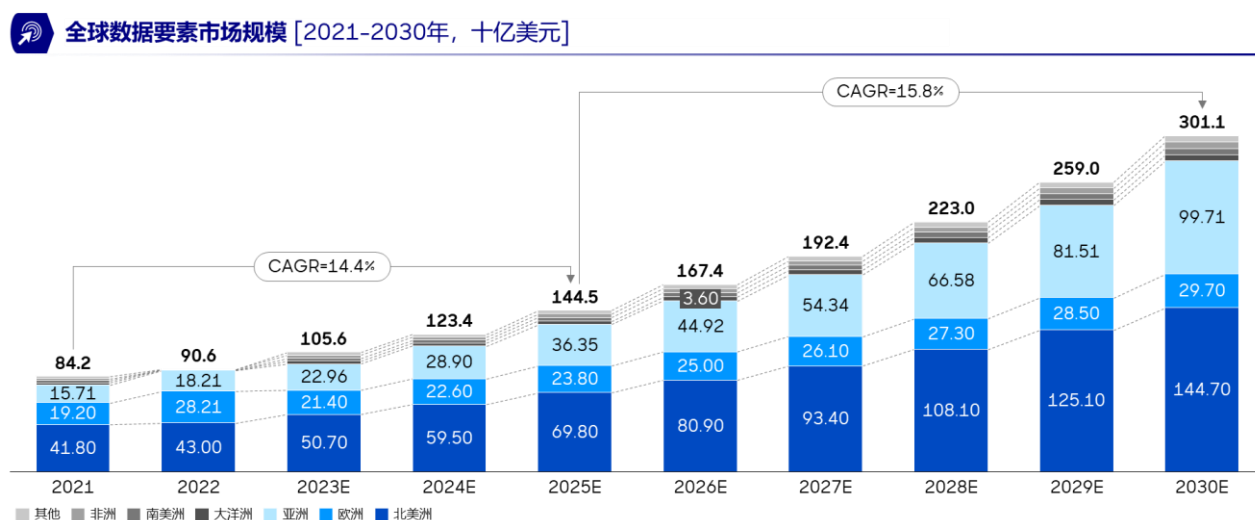




算力将有望超过 35 ZFlops，如果数据存储产业在通用计算及智能计算场景下，均按较为理想的存算比进行配置，中国将有条件打造**千亿美元级**先进数据存力底座产业（以当前汇率约合 5300 亿元）。更进一步，NAND Flash 颗粒等关键部件占存储设备成本达**60~70%**，如国将这些零部件的产值纳入经济效益考量，则**随着存储芯片等关键部件国产化进程的加速**，到 2030 年预计中国的先进数据存力**端到端全产业链规模，将有潜力达到万亿元级**。

其次，先进数据存力产业将与数据要素市场深度耦合，并驱动数据要素市场的规模化发展。到 2030 年，全球数据要素市场规模将达全球数据交易市场有望达**3011 亿美元**，年均复合增长率**超 15%**；中国数据要素市场将达到**5156 亿元**，年增长率**超 40%**。在数据要素市场发展，数据存储环节**直接贡献 20%**，并作为“压舱石”充分使能数据要素的安全流转、间接贡献**80%**。

图 1-17 先进数据存力的核心价值与影响力



最后，先进数据存力可作为智能经济时代的核心使能技术，赋能千行百业提质增效、带来生产力的提升。据测算，以先进数据存力为代表性技术的数智化解决方案，到 2030 年对生产力提高所带来的潜在价值预计可达**十万亿美元级**。对中国而言，到 2030 年智能经济带来对中国各行业带来的生产力提升经济效益可达**1.5 万亿美金**，其中制造业、贸易、建筑、健康

服务业与交通运输业均将获得巨大收益。例如，在制造业，通用 AI 智能体可以实时监控和调整生产线的各种参数，通过自我学习和优化进一步提高生产效率。它能够自主识别和解决瓶颈、协调不同机器的工作速度和任务分配，从而实现生产线的最高效率和最优配置。

### 1.5.2 提升社会福祉：先进数据存力使能绿色发展、良政善治与民生体验优化

先进数据存力除使能产业发展提质增效外，还可通过使能绿色发展、提升政府治理能力与优化民生体验，带来社会福祉的提升。绿色发展方面，先进数据存力可带来全球数据中心带来显著的降耗减排；提升政府治理能力方面，以先进数据存力赋能的数智化系统可在防灾护生、预防犯罪、电子政务、环保监测等场景中可展现显著的价值；优化民生体验方面，先进数据存力凭借其广域泛在的特征，在衣、食、住、行、医疗、教育、养老等各个方面作为“使能底座”，提升人民生活幸福度。

#### 1) 绿色发展：量化存力产业端到端降耗节能效应

虽然大模型等 AI 原生应用持续快速增长，数据中心的能耗于近年快速提升、并逐渐成为全球的“能耗大户”，而先进数据存力可有效确保数据中心的有效降碳。经测算，先进数据存力方案从 OEM 制造到在数据中心投入使用的生命周期能带来可观的能耗节降，可为全球与中国的数据中心节降 16% 与 24% 能耗。首先，应用先进数据存力(SSD 替换 HDD) 可直接为全球及中国数据中心能耗降低碳排放；其次，先进数据存力的应用可有效赋能集约化部署、降低数据中心的整体投资与运营成本，进而创造良好的条件牵引更先进绿色的冷却系统(如液冷系统) 的加速落地。

图 1-18 先进数据存力对绿色经济的影响



国家类型	范围 1	范围 2	范围 3
<b>释义</b> 	<ul style="list-style-type: none"> <li>由公司控制的设施设备直接产生的温室气体排放，例如企业现场的柴油发电机、锅炉、燃气炉等</li> </ul>	<ul style="list-style-type: none"> <li>由公司购买的能源的间接温室气体排放，包括电力、制冷、供热和蒸汽</li> </ul>	<ul style="list-style-type: none"> <li>涵盖企业价值链中除范围1&amp;2的所有间接排放，如员工出行、供应链、废弃物回收、采购</li> </ul>
<b>数据中心</b> 	<ul style="list-style-type: none"> <li>先进数据存力凭借其存算分离架构、SSD先进介质、近存计算、数据压缩等技术，有效使能数据中心的存算资源利用率的提升，可显著降低能耗</li> <li>先进数据存力使能集约化部署，使得数据中心有条件在不额外加大冷却成本使用更先进的冷却系统（如聚α烯烃油冷、液冷），提高冷却效率，降低冷却能耗</li> </ul>		<ul style="list-style-type: none"> <li>先进数据存力具备智能化运维特征，有效降低运维环节排放</li> <li>闪存单位存储密度远高于HDD，有效降低物流环节的温室气体排放</li> </ul>
<b>OEM整机制造商</b> 	<ul style="list-style-type: none"> <li>先进数据存储设备通过产品设计及工艺创新，如零波峰焊接、绿色PCB设计等，降低温室气体排放</li> <li>使用更多再生能源，例如光伏发电或风能发电</li> </ul>		<ul style="list-style-type: none"> <li>闪存盘的外壳通常用铝等可回收且散热性强的材料，可带来显著的降耗减排</li> </ul>
<p>💡 如果全球数据中心均使用先进数据存力方案（如以SSD替换HDD介质、使存储设备具备智能化运维特征），则全球数据中心及中国数据中心能耗有望降低9%和15%；通过先进数据存力方案使能集约化部署，并在数据中心成本可接受范围升级冷却系统，至少可降低数据中心能耗10%，最高可能降低50%数据中心能耗</p> <p>🔍 尽管从制造能耗来看，单TB的SSD相比单TB的HDD的制造能耗更高，但是如果考虑5年生命周期，使用全闪存存储方案的综合能耗亦具备显著优势</p> <p>🔄 此外，OEM厂商还可通过创新工艺及回收等措施创新降低制造环节的排放（例如零波峰焊接）</p> <p>XX：直接影响    XX：间接影响</p>			

## 2) 良政善治：先进数据存力推动社会治理现代化，使能社会经济可持续发展

先进数据存力是城市建设治理的底座，不仅在防灾护生、预防犯罪、电子政务、环保监测等场景中可展现显著的价值。通过整合与分析海量数据加速决策过程、提升响应效率、降低运营成本并保障数据的安全性与稳定性。预计到2030年，先进数据存力在全球范围内为社会治理带来至少**3万亿元**的经济效益，而在中国也可以带来至少**1万亿元**的经济效益。

例如，先进数据存力技术可有效使能灾害预警系统，深度挖掘历史灾害数据价值，在灾害发生前迅速识别潜在风险、提供预警方案，避免或减少经济损失。经测算，全国每年以先进数据存力赋能的先进灾害应急响应方案<sup>7</sup>，可带来**3000亿元**经济效益，包括减少的直接经济损失、紧急安置人口及贵重物品转移<sup>8</sup>、电梯救援成本减少等项目。

再如，先进数据存力可深度分析犯罪历史数据，融合实时监控数据，实现对犯罪行为的精准预测和实时预警，帮助预警和防范安全隐患。以先进数据存力赋能的智能犯罪预警系统，可实现对潜在犯罪行为的精准预测和实时预警。经测算，全国每年通过先进数据存力支持的犯罪预警系统可节省成本达**4100亿元**，包括案件减少的数量及经济损失<sup>9</sup>，重大经济犯罪案件追回金额，非法集资案件追赃挽损等项目<sup>10</sup>。

### 3) 民生福祉：先进数据存力全方位提高人民生活幸福感

随着经济的高速发展，人民对生活幸福感的要求也随之逐步提高。在衣、食、住、行、医、教育、养老等各方面，满足人民对于高质、高效幸福生活的期许应是各行业发展的应有之义。如果量化为经济效益，由先进数据存力技术赋能的千行百业，能为全球和中国分别带来**20 万亿元**和**6.9 万亿元**的社会经济福祉。

例如，在出行领域，先进数据存力通过赋能多端响应的数据系统，不仅能够为用户和城市管理者解决拥堵问题，还能提升用户通勤途中的舒适度和幸福感。以智慧交通控制系统、智能导航和自动驾驶的协同作用的应用场景为例，先进数据存力赋能的智能解决方案能够通过智慧交通控制系统解决拥堵问题，通过智能导航和自动驾驶减少用户拥堵时间，大幅提高通勤过程的舒适度。在快速记录、同步城市全路段车流量及实景图后，系统能够实时调节各路口交通信号灯最佳工作节奏，并将信息同步推荐至用户端实现无延迟智能导航。此外，先进数据存力赋能的智能驾驶能让用户的通勤时间变成休息、学习或娱乐时间<sup>19</sup>，这将大幅降低通勤过程的时间浪费并彻底消除诸如通勤时间过长会诱发抑郁之类的隐患，一定程度上提高居民生活幸福度。系列应用预计能为全球和中国带来经济社会福祉**2.4 万亿元**和**7700 亿元**。

又如，在医疗领域，先进数据存力用通过大数量级动态存储平衡区域间用户所享受的医疗资源并实现共同效率最大化。以共享医疗的应用场景为例，先进数据存力赋能的智能解决方案可以实现对区域内患者全生命周期内健康信息的存储与实时更新<sup>20</sup>。例如面对急诊响应，数据库能够在极短时间内调取患者健康信息（省去检查和病例书写环节），最大限度为急诊病人争取治疗时间，提高急救效率；再如，面对疑难杂症场景，有了先进数据存力赋能的全样本数据库，在偏远地区的患者不必再舟车劳顿来到大城市寻医问药，本地医院可以用数据库快速匹配类似病例，同曾经处理过类似病症的科室取得联系并进行即时会诊<sup>21</sup>。据测算，共享医疗的广泛应用预计能为全球和中国带来经济社会福祉**6.6 万亿元**和**2 万亿元**。



### 1.5.3 强化区域竞争力：先进数据存力赋能新质生产力生态发展

九大战略性新兴产业的发展都依赖于先进数据存力，以推动技术进步和应用优化。数据存力通过其汇聚、存储和分发数据的功能，支持了这些领域的高效运作。比特化使得所有领域能够通过数字化技术进行信息分析和处理，提升了数据的快速处理能力。无论是生成式人工智能的模型训练，还是量子信息的密钥管理、元宇宙的虚拟数据存储，数据存力通过提供快速、高效的存储，提供了必要的技术支撑，使这些领域能够处理复杂的数据任务，推动国家技术创新和应用发展<sup>22</sup>。

#### 1) 生成式人工智能

生成式人工智能通过推动科技创新、产业发展和人才培养，全面增强国家在全球竞争中的战略优势。人工智能通过突破关键核心技术，推动产业链与创新链的深度融合，加快传统产业的升级改造，形成现代化产业体系。同时，生成式 AI 还通过优化学科建设和人才培养机制，培养适应未来发展的高端科技人才，推动产学研用的紧密合作<sup>23</sup>。

此外先进数据存力使得 AI 能够“以存代算、以存强算”，带来千行百业生产力的跃迁。首先，“以存代算”的特征赋予了大模型“记忆”功能，通过记录用户与模型的交互过程，记住常见问题的答案。当类似问题再次出现时，模型可以直接从存储系统中提取已保存的“记忆”，确保用户能够快速得到准确的回答，从而避免再次使用显卡进行推理。其次，先进数据存力通过高效存储、数据处理、对仿真模型的支持和虚拟调试的应用，优化生产线的效率。比如某汽车 Tier-1 依托数据存储技术，实现了对印刷电路板的全面高效检测，提升了产能。

#### 2) 量子信息

量子通信技术通过量子密钥分发 (QKD) 等先进加密方法，提供了信息传输的绝对保密性，保卫国家信息安全。在全球金融、军事和政府等关键领域，这种高度安全的加密手段可以有



效抵御黑客攻击和数据泄露，确保敏感信息的安全，从而增强国家在国际科技和安全领域的战略优势。

先进的数据存力支持大规模部署和高效管理量子密钥及纠缠态信息，因此对量子通信行业至关重要。例如，在银行跨国金融交易中，量子存储器能够安全地存储和传输加密密钥，确保交易过程的安全性。强大的数据存储能力使得量子通信网络能够稳定运行，提升系统的可靠性和安全性，为量子通信的广泛应用提供了必要支持。

### 3) 元宇宙

元宇宙通过推动网络通信技术、促进技术进步、吸引顶尖人才和创造就业机会，提升国家的竞争力。元宇宙是一个融合虚拟现实、增强现实、区块链和互联网技术的数字生态系统，允许用户通过数字身份在多个互联的“数字孪生”世界中进行社交、娱乐、购物和教育等活动。元宇宙驱动了网络通信的发展，6G、VR 和 AR 等技术的耦合应用将成为元宇宙的杀手级应用场景，为核心技术的进步提供了良好的培育土壤，吸引高质量的人才和就业岗位，从而推动了相关技术的创新和应用。

先进的数据存力负责存储、备份和分发元宇宙这个巨型数字孪生世界中所有的虚拟形象及活动的数据。数据存储不仅支撑着用户在元宇宙中的持续体验，还确保了虚拟环境的稳定性和可靠性。通过高效的数据存储解决方案，元宇宙能够有效管理和处理大量的动态数据，从而提升整体系统的性能和用户体验。

### 4) 脑机接口

脑机接口不仅推动高科技医疗设备的创新，还促进了社会福祉的发展。脑机接口是一种通过将人类大脑与计算机直接连接，让大脑与外部设备交流的技术。脑机接口的发展鼓励了智能科技在医疗领域的显著进展，改善了残疾人以及社会整体的生活质量，从而增强国家在全球竞争中的地位。





先进数据存力显著提高了信号处理的速度和精度，使得脑机接口能够更快速、更准确地解读脑电波。脑机接口领头羊 Neuralink 公司将会通过脑机接口帮助解决脑损伤疾病，为患者创造协调的未来。这一愿景离不开高效的信号采集和处理和深度学习算法，而先进数据存力的进步大大提升了信号处理的速度和精度，使得脑机接口能够更快、更准确地解读脑电波，促进了医疗、广告等领域的应用，例如准确解读脑信号以有效控制假肢、进行神经康复，并辅助诊断和治疗神经系统疾病。

## 5) 具身智能

具身智能在工业生产、家庭服务等领域代替人类实现高效和高精准度的服务，显著增强国家竞争力。具身智能是指具有物理实体和感知能力的人工智能系统，而人形机器人是最具代表性的具身智能形式，它们通过类似人类的身体结构和运动能力，能够在复杂环境中执行任务，与人类互动，模仿人类行为。在工厂中，人形机器人替代人工搬运货物，提高生产效率和工作安全性，推动制造业自动化，降低运营成本。在码头操作中，人形机器人可以高效处理装卸和运输任务，加快物流周转，提升国家在国际贸易中的地位。家庭服务领域的机器人能够减轻人们的日常负担，提高生活质量，同时推动服务业技术革新，创造新的就业机会和经济增长点<sup>24</sup>。

先进数据存力不仅在性能上推动了机器人技术的发展，还在商业上促进新的“以数据为资产”的商业模式。技术上，高速的读写带宽、优秀的 IOPS 速度和先进的闪存技术使具身智能技术能够迅速处理和存储大量数据，从而提高其操作精度和效率。商业上，埃隆·马斯克指出，人形机器人在工作时存储的数据将会是企业的重要资产，这进一步突显了数据存储对提升机器人功能及其商业价值的至关重要性。

## 6) 生物制造

未来生物制造技术通过资源循环利用、环保效益和创新推动，显著提升国家的经济竞争力和可持续发展能力。首先，生物技术能够有效利用工业废弃物和农林废弃物，转化为高价值产

品，如用钢铁尾气生产的高蛋白饲料和燃料乙醇、用秸秆生产的生物乙醇。这种资源循环利用不仅减少了环境污染，还降低了对粮食的依赖，提升了资源利用效率。其次，生物制造的可再生性支持了国家的“双碳”目标，有助于实现可持续发展。最后，通过技术创新，我国能够在全球市场中占据先机，形成新的产业优势，推动经济向绿色、智能、高效方向发展，并加快工业转型升级进程<sup>25</sup>。

**先进数据存力通过高效管理和处理生物数据，推动了生物制造领域的精准研究和自动化生产。**例如，在合成生物学中，先进数据存力存储海量 DNA 序列数据，使科学家能够快速分析数以亿计的基因信息，从而优化微生物菌株设计，并用于生产医药或生物燃料等领域。高效的数据存储与读取能力不仅提高了研究效率，还推动了生物制造工艺的自动化，使复杂的生产过程更加稳定和可控，助力行业快速发展。

## **7) 未来显示**

**未来显示领域的技术创新将推动高科技行业的升级、提升国家文化软实力，并增强国家安全和战略防御能力。**未来显示的创新将引领新一轮科技进步，推动电子产品、医疗设备和汽车显示等行业的升级。同时，先进的显示技术将推动媒体、教育和文化产业的发展，提升国家的文化软实力和信息传播能力，并在军事和国防领域提升指挥系统和信息化作战平台的效率，增强国家的战略防御能力。

**先进数据存力技术可以提升未来显示的数据处理能力，支持创新应用，并且增强设备性能。**随着显示技术的进步，数据量大幅增加，先进存储系统能够高效管理和处理这些数据，确保实时、准确地展示复杂图像和视频。此外，这些存储技术支持虚拟现实、增强现实等应用，推动了沉浸式体验和动态内容的展示。同时，通过提升存储速度和容量，存储技术改善了显示设备的响应时间和图像质量，增强了视觉效果和用户体验。

## **8) 未来网络**



未来网络技术通过加速数据传输和提升带宽，优化智能城市、自动驾驶和物联网应用，增强国家科技创新。5G 提升自动驾驶的安全性和效率，而 6G 技术则通过极高的传输速率和低延迟，支持智能城市升级、实时数据处理和无人机应用，从而进一步提升国家的技术实力和全球竞争力。先进数据存力通过优化数据管理和分配，为未来网络的发展提供了关键支持。随着网络架构向“以应用服务为中心”转型，数据存储技术可实现资源的高效调度和管理。通过融合计算和存储资源，先进数据存力技术将进一步推动未来网络的发展，助力人工智能等前沿技术的应用落地。

## 9) 新型储能

新型储能技术的发展显著增强了国家的电力系统稳定性和能源利用效率。通过灵活布局和短建设周期，这些储能技术，如电化学储能、压缩空气储能和重力储能，能够在电力供需不平衡时高效调节电力，支持新能源的开发和消纳，减少对传统化石能源的依赖。此外，新型储能的规模化应用带动了相关产业链上下游的协同发展，加速技术创新与产业进步，从而提升国家在全球能源竞争中的地位。

先进数据存力为新型储能提供了数据管理和分析支持，尤其是在国网荷储一体化发展中发挥了关键作用。高效的数据存储系统能够实时监控和优化储能设备的运行，管理大规模储能数据，从而提高储能系统的响应速度和稳定性。例如，在湖北应城的 300 兆瓦压缩空气储能电站中，数据存储技术用于实时监测和管理地下储气库的运行状态，确保了储能系统的高效和安全运行。

图 1-19 先进数据存力对新质生产力 9 大未来产业的赋能

## 九大未来产业



## 产业对提振区域竞争力的影响



## 先进数据存力的赋能

生成式人工智能	<ul style="list-style-type: none"> <li>生成式人工智能通过推动科技创新、产业发展和人才培养，全面增强国家在全球竞争中的战略优势</li> </ul>	<ul style="list-style-type: none"> <li>先进数据存力使得AI能够“以存代算、以存强算”，带来千行百业生产力的跃迁</li> </ul>
量子信息	<ul style="list-style-type: none"> <li>通过量子密钥分发（QKD）等先进加密方法，提供了信息传输的绝对保密性，保卫国家信息安全</li> </ul>	<ul style="list-style-type: none"> <li>支持大规模部署和高效管理量子密钥及纠缠态信息，对量子通信行业至关重要</li> </ul>
元宇宙	<ul style="list-style-type: none"> <li>通过推动网络通信技术、促进技术进步、吸引顶尖人才和创造就业机会，提升国家的竞争力</li> </ul>	<ul style="list-style-type: none"> <li>负责存储、备份和分发元宇宙这个巨型数字孪生世界中所有的虚拟形象及活动的数据</li> </ul>
脑机接口	<ul style="list-style-type: none"> <li>推动高科技医疗设备的创新，促进社会福祉的发展，改善残障人士等弱势群体的生活质量</li> </ul>	<ul style="list-style-type: none"> <li>支持生物学信号到电信号的转换，推动存储器体积的缩小和性能的提升</li> </ul>
具身智能	<ul style="list-style-type: none"> <li>人形机器人在工业生产、家庭服务等领域代替人类实现高效和高精度的服务，显著增强国家竞争力</li> </ul>	<ul style="list-style-type: none"> <li>不仅在性能上推动了机器人技术的发展，还在商业上促进新的“以数据为资产”的商业模式</li> </ul>
生物制造	<ul style="list-style-type: none"> <li>推动技术创新、优化关键产业链、并提升公共健康和食品安全水平</li> </ul>	<ul style="list-style-type: none"> <li>先进数据存力通过高效管理和处理生物数据，推动了生物制造领域的精准研究和自动化生产</li> </ul>
未来显示	<ul style="list-style-type: none"> <li>推动高科技行业的升级、提升国家文化软实力，并增强国家安全和战略防御能力</li> </ul>	<ul style="list-style-type: none"> <li>提升未来显示的数据处理能力，支持创新应用，并且增强设备性能</li> </ul>
未来网络	<ul style="list-style-type: none"> <li>加速数据传输和提升带宽，优化智能城市、自动驾驶和物联网应用，增强了国家的科技创新和全球竞争力</li> </ul>	<ul style="list-style-type: none"> <li>优化数据管理和分配，为未来网络的发展提供了关键的技术支持</li> </ul>
新型储能	<ul style="list-style-type: none"> <li>比特与瓦特齐头并进，显著增强了国家的电力系统稳定性和能源利用效率</li> </ul>	<ul style="list-style-type: none"> <li>提供了数据管理和分析支持，尤其是在电网荷储一体化发展中发挥了关键作用</li> </ul>



# 02

## 先进数据存力指标评测

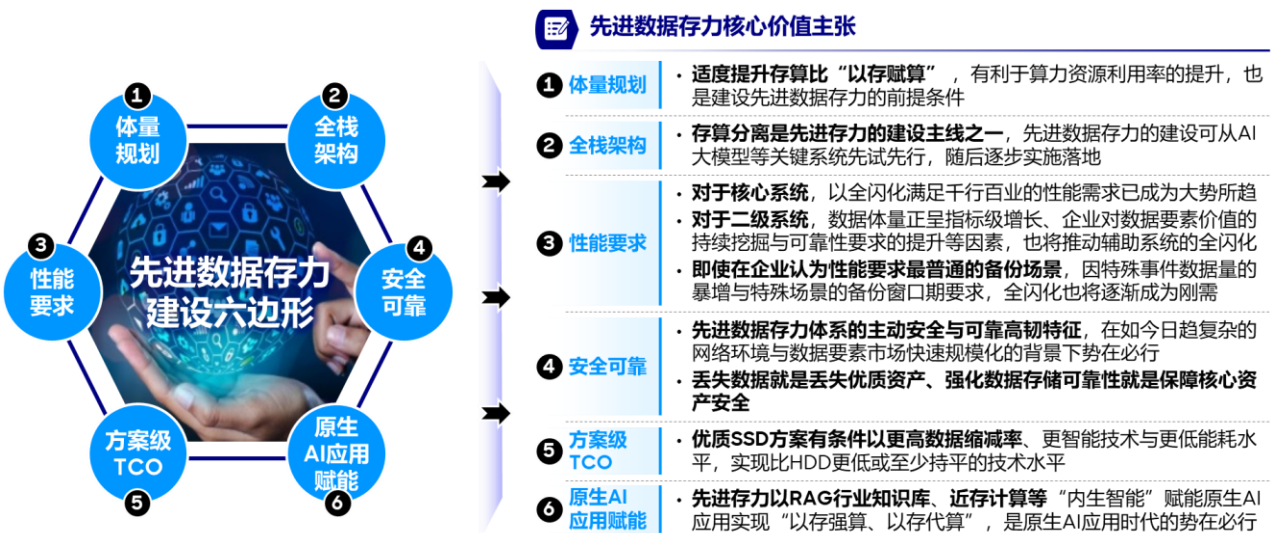
建设指标体系保障先进数据存力高质量发展



## 2.1 发展先进数据存力的六大核心考量因素

虽然先进数据存力可有效赋能千行百业加速数智化转型，但不同行业企业所面临的情景与业务需求不尽相同。因此，企业应体系化规划并推进先进数据存力的建设、以正确且全面的认知对先进数据存力进行适度超前的持续投资、以应对智能经济时代的业务挑战。

图 2-1 建设先进数据存力的六大核心考量维度



### 2.1.1 容量规划

当前，中国在数据存储能力方面面临显著挑战，尤其是与全球领先国家如美国、德国相比，存力充足性明显不足。2023年，中国存力充足性为6.4%，而美国与德国这一数据分别为10.7%与13.4%。这一差距在企业级应用中尤为突出，众多行业普遍面临存储资源不足的问题，且存储规模与业务逻辑的匹配度存在偏差。例如，某汽车行业零部件供应商基于其业务特征，为其EMS工厂设定了依据营收额配置相应容量数据存储容量的标准，这种过于简单的配置逻辑，非常容易导致“数未能尽其用”、数据价值未得到充分挖掘和利用。总的来说，我们认为提升存力配置有三大好处。





首先，更多的存力资源可减少服务器等待 I/O（输入/输出）操作完成的时间，从而可以更高效地利用 CPU 和内存资源。例如，本地的高速存储（如 SSD 和 NVMe 设备）可以显著加快数据访问速度，极大减少 CPU 等待时间，更合理的存算比使得服务器的存储资源与计算资源更匹配，数据可以在靠近计算的地方存放，提升数据的“本地化”程度，加速访问速度进而提升整体算力资源利用率。

其次，增加存储容量和优化存算比让更多的虚拟机或容器在服务器上运行，这些技术需要大量的存储资源用于映像、日志和其他运行时数据。更大的存储容量允许数据以更高效的方式进行分区和调度，减少数据在不同存储设备之间传输的时间，提高数据处理效率。

最后，更大的存力赋予更高效的备份和恢复策略，应对各种突发事件的能力就更强。正是因为更可靠的数据存储，使得计算可以更加持续和高效地进行。在备份或数据恢复过程中的停机时间减少，服务器可以更连续地利用计算资源。

### 2.1.2 资源利用率

当前，主流数据中心的数据存储设备以“服务器加本地盘”的模式为主，造成数据存储资源的利用率不高、副本备份数过多等挑战。由于每台服务器独立管理本地存储，数据副本的数量非常多，增加了存储管理的复杂性和成本。在此背景下，如何提升存力资源利用率成为一大核心挑战。而先进数据存力所提供的解决方案，就是存力资源池化为核心的“存算分离”架构。在存力资源池化的架构下，海量的专业存力设备之间通过高速互联、形成了一个巨大的“共享资源池”，这个巨大的资源池能够兼容各种存储与通信协议，与北向的算力资源池形成高速互联，有效使能存算运三大 ICT 底座资源利用率的提升。

存算分离是先进数据存力的建设主线之一，也是数据中心架构变革的关键演化趋势之一。对计算与大数据领先的国家及企业而言，存算分离的架构已然成为大势所趋。例如，在中国，多家领先 ICT 解决方案供应商也已广泛采用存算分离架构，并提供相关解决方案；在美国，

硅谷领头羊（如谷歌、亚马逊、微软和 Meta）已在其数据中心广泛采用存算分离架构、美国的电商、文娱、物流、生活服务等多个行业的诸多企业也已广泛应用存算解耦架构。

**推动存算分离有六大优势。**一是**资源独立扩展**，存储和计算资源可以独立扩展。企业可以根据需求增加存储容量或计算能力，而不需要同时增加另一种资源；二是**维护和升级的灵活性**，升级或维护某一部分时不会影响到另一个部分，提高系统完成任务的可靠性、减少停机时间；三是**优化性能和更高的可用性**，比如使用更快的存储介质（如 SSD）来提高数据读写速度，同时使用更高性能的计算节点来提高处理能力；四是**数据持久性和安全性**，数据存储独立的存储系统中，计算节点的故障不会影响数据的完整性和安全性；五是**负载的优化和分布**，可以在网络上分布计算负载和存储负载，优化网络资源的使用，进而实现更高效的网络流量管理，减少网络瓶颈，提高总体系统性能。六是以**数据为中心的架构与近存计算等技术**，可有效降低网络等资源损耗，让数据“就近处理/计算”。

当前数据中心在推进存算分离架构时，通常不会一次性对所有软件系统进行转换，而是采用**渐进式的策略**。例如，先从大模型等核心系统入手，在试点验证成功后，逐步推广到更多的业务系统和支持系统，通过不断的优化和调整，最终实现全面的存算分离架构。这种方式确保了风险的最小化、资源的合理利用和业务的连续性。

因此，全球各国政府及企业应**加速推动存算分离架构、并持续投资推动替换**。通过强化对数据中心先进架构变革的持续推动、积极培养相关产业人才（如了解新型架构的运维人才）等举措，确保持续投入与平稳转型。

### 2.1.3 性能要求

对于核心系统 (Primary System)，以全闪化满足千行百业的性能需求已成为大势所趋。全球范围看，核心系统的全闪化率在 2023 年已达 27%，预计 2027 年达 35%。以中国为例，主流电信运营商的 ERP 与 OSS 等核心业务系统基本实现全闪化改造；在金融行业，领先保



险公司的核心保险系统，财务系统等对业务时效性要求高的系统，都采用全闪化解决方案。再如，上海市青浦区建设城建管理系统时，通过实时动态查看街边影像，以 **AI 影像识别判定违章搭建**，系统的读写带宽达到 3 GB/S，IOPS 达 30W，延时要求 50 微秒，MTBF 要求 80 万小时，因此选型全闪化 **SSD 解决方案** 以满足上述性能要求。

对于二级系统 (Secondary System)，数据体量正呈指标级增长、企业对数据要素价值的持续挖掘与可靠性要求的提升等因素，也将推动二级系统的全闪化。以某中国电信运营商为例，其规定节点读写带宽当量达 500 MB/s 时，就必须以全闪化 SSD 方案满足时效性要求，因此其 BSS 与 OA 系统基本已完全实现全闪。在未来，为满足集团提出的二级系统 RTO 达 2 小时的可靠性要求，该运营商将推动 BSS 与 OA 系统的全闪存化。

即使在企业认为性能要求最普通的备份场景 (Backup)，因特殊事件数据量的暴增与特殊场景的备份窗口期要求，全闪化也将逐渐成为刚需。例如，在 2024 年巴黎奥运会的期间，在长达 2 周的时间内，某电信运营商 OSS 系统的 HDD 备份方案无法满足保证的访问量暴增带来的数据备份需求，导致备份用时超 20-25%；再如，某运营商的洲际 SD-WAN 业务有数据加密需求，为避免 HDD 方案的备份时间过长带来安全隐患、其采取全闪化备份方案以适配该特殊业务场景的“极速备份”需求。

最后，在追求性能的同时，组织也不应牺牲系统的可靠性，除采取全闪存 (SSD) 先进介质外，还应关注快照 (Snapshot) 等用于提升可靠性的高级功能。例如，快照功能可以用于保护数据、防止数据丢失。企业可以定期创建数据快照，以便在遇到数据丢失、损坏或安全事件（如勒索软件攻击）时，迅速恢复到快照时间点的数据状态，从而将损失降到最低，进而有效提升数据存储系统的“韧性”。而先进数据存力方案通过优化设计与各种技术手段，可很好地实现高性能与可靠性的“鱼与熊掌兼得”，即数据存储系统提供高性能数据访问时，提升可靠性的高级功能处于开启状态，这值得千行百业的组织与企业重点关注。

## 2.1.4 安全可靠防勒索

构建主动安全与可靠高韧的数据存储系统，在如今日趋复杂的网络环境与数据要素市场快速规模化的背景下势在必行。

一方面，网络攻击与数据勒索问题已成企业“难以承受之痛”，强化对端到端防勒索的投入势在必行。例如，某电信运营商在建设数据中心时，7~8%的投资流入构建端到端数据安全上，其中2~2.5%为数据存储相关安全投入。另一方面，随着数据的要素化与资产化，数据存储系统的可靠性与韧性也愈发重要。在2030年数据要素市场达万亿、数据资产超十万亿的背景下，企业以数据资产入表为核心抓手，提升净资产额与高质量资产比重、进而优化资产负债表、提升资本市场估值，正成为强化自身竞争力的核心手段。

当前，勒索软件对电力、电信、医疗等包含关键基础设施在内的多个行业的攻击愈发频繁，勒索赎金的规模也越来越大，使得组织对数据存储的防勒索解决方案势在必行、而并非一个“可选项”。

例如，2023年1月，英国皇家邮政遭受LockBit勒索攻击，并索要8000万美元赎金。由于皇家邮政最终未满足攻击者的赎金要求，该攻击导致其国际邮件投递服务瘫痪，数百万封信件和包裹滞留在该公司的系统中。1月12日，皇家邮政声称网络攻击事件迫使他们停止了国际邮政服务。尽管公司聘请了专业机构帮助恢复业务，但遗憾的是，勒索攻击的影响一直在持续。皇家邮政所在母公司财报显示，截至2023年9月，公司收入同比下降6.5%，给出的原因正是工业行动和勒索软件入侵。

再如，2023年8月，Rhysida勒索软件组织对美国多个州的医院和诊所发动了一系列勒索攻击。该组织声称已窃取了1TB的机密文件和1.3TB的SQL数据库，其中包含了约50GB的病人隐私数据。Rhysida勒索软件组织要求医疗信息系统服务商Prospect Medical Holdings支付赎金以恢复数据，并威胁将泄露受窃数据。然而，Prospect Medical Holdings



公司拒绝支付赎金，并与执法机构合作进行调查。为了应对此次攻击，他们采取了紧急措施，包括隔离受感染的系统和从备份恢复数据，恢复过程非常耗时和复杂。

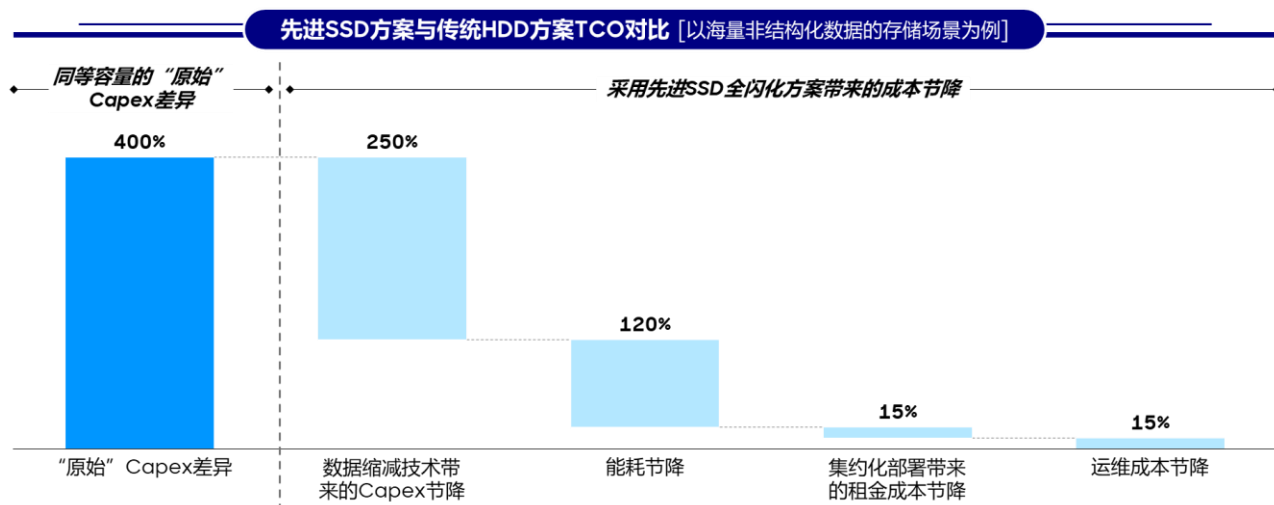
在此背景下，丢失数据就是丢失优质资产、强化数据存储可靠性就是保障核心资产安全。因此，企业需强化对数据存储系统可靠性的重视，强化对 RTO、MTBF 等关键可靠性指标的管理、重视对容灾备份 (Backup & Recovery) 的投资。例如，某电信行业企业要求 OSS 等核心系统每年的故障时间不超过 26 分钟（即可靠运行时间超过 99.995%），对极关键节点要求每年的故障时间不超过 2 分钟；即使对 OA 等非核心系统，每年的故障时间亦不能超过 8 小时（即可靠运行时间超过 99.9%）。

### 2.1.5 方案级 TCO

当前，诸多企业在二级系统与备份容灾系统选型时，因侧重 Capex 成本考量而选择 HDD 方案。然而，随着 NAND Flash 芯片的持续更新、数据压缩技术的持续创新与先进 SSD 方案对设备运维能力的持续演进提升，优质 SSD 方案有条件以更高数据缩减率、更集约化的部署、更自治化的运维与更低的能耗水平，实现比 HDD 更低或至少持平的技术水平。

例如，在海量非结构化数据为核心且存储设备生命周期 5 年的场景下，当前先进 SSD 全闪存方案的单位容量 TCO 已可与传统 HDD 方案基本持平。当前 SSD: HDD 价差大概在 5:1 左右。首先，更优的数据缩减技术使得可实现数据缩减率领先传统 HDD 方案至少 50%；其次，先进 SSD 方案的年度 Opex 相比传统 HDD 方案可下降至少 80%，一是先进 SSD 介质带来更低的能耗成本，二是集约化部署带来的租金成本节降，三是高度自治化运维带来的维护检修等费用降低。

图 2-2 先进 SSD 方案与传统 HDD 方案的 TCO 对比



此外，企业也应明确以方案级 TCO 为单位，结合性能与安全可靠性要求等因素综合考量后，再进行全闪化与传统 HDD 方案的选择，而非在二级系统及容灾备份系统的选型时单纯以成本为考量，以最大化先进数据存力对业务运行的保障与赋能。

### 2.1.6 原生 AI 应用赋能

展望 2030，原生化 AI 应用将成为大势所趋。“地基不稳，地动山摇”，如何打牢 ICT 基础设施地基以适配原生 AI 应用时代，成为企业必须思考的一个问题。一方面，以 RAG 向量存储强化推理精确度提升用户体验、进而“以存强算”，成为诸多部署大模型企业的共性需求。例如，中国电信发布的星辰大模型，在 2024 年的版本中，就将加入核心网/无线接入/宽带接入等系统的最新数据、强化模型推理能力、带动集团运维能力提升。另一方面，以 KV-Cache 多级缓存（建立 HBM – DDR – SSD 三级数据存储体系）可解决多轮对话和长序列场景记忆问题、有效环节大模型“幻觉”，实现“以存代算”，可很好地缓释算力供应链紧缺挑战，实现 10 倍以上的推理成本节降。

“万卡时代来临，AI 大模型是实践存算分离架构应用潜力最大的核心系统之一，万亿级参数使得服务器外挂硬盘的模式不再适用，集约化全闪 SSD 存储资源池成为刚需。”

——中国电信解决方案专家





## 2.2 先进数据存力衡量指标体系

为适配 AI 新时代的业务需求、我们对提出了先进数据存力指标体系，发布“区域发展－数据中心－存储产品”三级指标体系，作为各区域与千行百业的企业组织评估自身竞争力、建设存力体系的参考，并对区域发展层级的关键指标进行了详尽梳理分析、对各国的关键指标进行追踪与更新。

我们在测算过程中共选取了 19 个不同区域和发展水平的国家进行深入研究，综合比较样本国家在数据存力各个维度的表现进行国际横向对比。全面的先进数据存力指标体系，将更好帮助各个国家厘清现有的发展优势和尚待提升的空间，助力其加快建设全面的数据存力支撑经济社会高质量发展。

图 2-3 先进数据存力指标排名汇总

一级维度		体量				效率		基础保障		前沿保障	
二级维度		充足			增长	平衡	敏捷	可靠	绿色	先进	
三级维度		单位GDP存力容量	数据存力充足性	数据存留率	存力投资增长率	通用计算存算比	闪存占比	灾备覆盖率	单位存储容量能耗	数据存力专利占比	
亚太	中国	9	11	7	11	13	19	19	14	2	
	日本	10	10	6	18	3	18	17	8	6	
	新加坡	1	3	1	17	1	11	1	9	7	
	泰国	12	13	19	15	11	17	6	18	16	
	印度	13	15	13	3	9	9	11	19	3	
	沙特阿拉伯	16	17	16	14	11	5	15	12	15	
	土耳其	14	16	17	19	14	8	18	15	17	
美洲	南非	6	7	18	2	7	4	14	16	5	
	美国	5	4	3	16	3	3	8	2	1	
	加拿大	7	6	5	12	2	16	2	7	4	
	墨西哥	19	19	12	4	10	13	4	17	10	
	哥伦比亚	17	14	15	8	12	1	16	11	19	
	智利	15	12	11	7	9	14	12	13	18	
	巴西	18	18	14	1	8	2	5	10	13	
欧洲	德国	3	2	2	9	3	12	9	4	9	
	法国	11	9	10	6	5	15	3	6	11	
	英国	8	5	8	13	6	7	7	1	8	
	瑞典	2	1	4	10	3	6	10	5	14	
	捷克	4	8	9	5	4	10	13	3	12	

### 2.2.1 区域发展级

先进数据存力的区域发展级指标可用于衡量全球各经济体的先进数据存力发展情况，整体而言沿用了**体量、效率、基础保障与前沿保障**四大维度，并结合发展先进数据存力的六大核心考量因素进行了补充更新。

充足的体量仍是任何区域实现先进数据存力产业的前提基础，而在智能经济时代，全球数据要素市场快速发展，数据将逐渐演化为全球各经济体的核心资产。因此，我们在原有数据存力指标体系的基础上，补充了“数据存留率”作为核心指标，建议各区域强化对数据要素的挖掘与存留。从统计结果看，中国当前的数据存留率仅 2.8%，而美国达到 7.3%，这说明中国对数据要素的挖掘相较美国等国而言仍有较大的提升空间。

图 2-4 先进数据存力指标体系（区域发展级）

指标维度	衡量范围	指标名称	计算公式	单位
体量	充足	单位GDP存力容量	· 数据存储设备容量/当年区域GDP产值	GB/万美元
		数据存力充足性	· 区域内存储设备容量/当年区域产生的数据总量	%
		数据存留率	· 企业数据存留量/企业当年数据产生量	%
	增长	存力投资增长率	· 当年数据存力投资/去年数据存力投资	%
效率	平衡	存算力比	· 已有数据算力总量/已有数据存力的容量	GB:GFlops
	敏捷	闪存占比	· 当年用于闪存投资额/当年总体数据存力投资额	%
基础保障	可靠	灾备覆盖率	· (当年容灾投资 + 当年备份投资) / 当年数据存力投资	%
前沿保障	绿色	单位存储容量能耗	· 当年数据中心存储设备能耗/已有数据存力的存储容量	Kwh/TB
	先进	数据存力专利	· 当年本土企业申请的数据存力专利数量/全球数据存力相关专利数量	%



## 2.2.2 数据中心级

在数据中心层面的先进数据存力建设上，我们建议企业在原有指标体系上，额外关注**数据覆盖率**（以数据地图已映射数据资产数量/全部数据资产数量的比例表示）、**年均设备汰换率**（以当年淘汰存储设备容量/区域内存存储设备容量比例表示）、**存算解耦架构比**（以存算分离式架构容量当量/存储总容量表示）以及**智能计算存算比**（智能计算场景中有效数据存力容量/有效数据算力总量）。

首先，**高数据覆盖率**表示企业能够全面了解和管理其所有数据资产、能够对本企业的单个/多个数据中心的数据进行统一管理，这对于企业的数据治理、合规性和安全性管理等方面至关重要。其次，关注**维保期内设备占比**可使得企业很好地掌握自身数据存储设备的资产健康度，通常而言数据存储设备使用至多**5年**就应换新、在使用期也应定期关注资产健康度，而不应超过这个周期，乃至“一直用到坏为止”。在数据要素市场快速发展、数据成为新生产要素与可交易高价值资产的背景下，企业在建设先进数据存力时不应过于关注成本，因为**丢失数据就是丢失企业在智能经济时代的核心资产，也将损害企业的核心竞争力**。

其次，**存算解耦**作为现代数据中心架构设计的一种关键趋势，也值得组织重点关注。存算解耦架构比的提升，首先能够显著提升各区域 ICT（信息与通信技术）基础设施的利用率。存算解耦通过提升资源利用率和赋能算力多元化发展，有效缓解了算力产能紧张的问题，为数据要素市场的快速发展提供了坚实的支撑。同时，它也为企业在智能经济时代保持核心竞争力和持续创新提供了有力保障。

最后，**智能计算存算比**作为评估智能场景下数据存储与数据处理算力匹配度与效率的关键指标，也应得到企业的重视。我们建议企业综合考量硬件架构（如是否采取存算分离）、数据类型与规模，以及应用场景等多因素的影响，在不同的场景下做出合理的存算比配置。例如，图像和视频数据通常需要大量的存储空间和高速的数据处理能力。

以多模态大模型为例，当前美国典型多模态 LLM 在部署时，存算比**通常在 1:10 内**、智算场景下的存算比相对合理，而中国**普遍高于 1:40**，这一现象**值得关注**。提升对先进数据存力的建设以优化存算比、利用其“博闻强记”的优势实现大模型推理精度的**显著提升**与推理成本的显著降低，是中国强化自身 AI 产业实力、打造具备国际竞争力 AI 产品的一大机会点。

图 2-5 先进数据存力指标体系（数据中心级）

指标维度	衡量范围	指标名称	计算公式	单位
体量	存量	数据中心存储容量	· 各类存储设备容量加总	EB
		数据存力充足性	· 存储容量/企业高峰数据量预计	%
	增量	存力投资增长率	· 当年数据存力投资/去年数据存力投资	%
效率	平衡	通用存算比	· 已有数据存力的容量/已有数据算力总量	GB:GFlops
		智能计算存算比	· 智能计算场景中有效数据存力容量/有效数据算力总量	%
		存力使用率	· 已占用存储容量/总可用存储容量闪存	%
	敏捷	闪存占比	· 介质投资/整体存储投资	%
		数据覆盖率(数据管理)	· 数据地图已映射数据资产数量/全部数据资产数量的比例	%
		维保期内设备占比	· 维保期内存储设备容量/数据中心内总存储容量	%
基础保障	可靠	灾备覆盖率	· (容灾备份软件投资+备份存储投资)/全部存储投资	%
		事后 RTO时间恢复目标	· 从故障发生导致业务停顿之起到IT系统恢复至可支持各部门运作、恢复运营所需的时间	小时
		RPO时间恢复目标	· 故障发生后数据可以恢复到的时间点	小时
	经济	单位存储拥有成本	· 总体拥有成本TCO/容量	\$/GB
前沿保障	绿色	存储设备能耗水平	· 数据中心存储设备总能耗/有效容量	Kwh/GB
	智能	智能存储功能丰富度	· 数据加密、审计日志、快照、异步复制、元数据检索等，根据客户具体需求进行满足	
		存储API开放程度	· 已开放API接口/所有可用API接口	%
	架构	存算解耦架构比	· 存算分离式架构容量当量/存储总容量	%



2.2.3 存储产品级

存储产品级，我们建议将原有的**单位容量价格指标** (存储费用/容量) 改为**单位容量 TCO** (存储系统 TCO/容量)，以综合考量先进数据存力解决方案的软硬件特征 (先进介质、高密盘框等) 对**单位容量 TCO** 的影响，加速先进数据存力产业的**规模化发展**。

图 2-6 先进数据存力指标体系（存储产品级）

指标维度	衡量范围	指标名称	计算公式	单位
体量	可用容量	原始容量	• 存储系统配置的物理容量	PB
		可用容量	• 原始容量减去用于RAID保护和元数据的容量	PB
		有效容量	• 可用容量×(写入系统的数据量/占用的容量)	PB
		存力体量效率	• (有效容量+可用空闲容量)/原始容量	%
	可扩展性	扩容设计预留量	• 预留节点×单个节点硬盘数量×各硬盘容量	PB
效率	性能	IOPS	• 1000ms/(寻道:磁头移动至正确磁道上所需时间+旋转延迟:盘片旋转将请求数据所在扇区移至读写磁头下方所需时间)	
		读写带宽	• Throughput MB/s = I/O Per Second *Size of I/O = I/O per Second * KB per I/O : 1024	MB/s
		响应时间	• 响应时间=等待时间+服务时间(读写耗时)=(队列长度+1)*平均响应时间	s
基础保障	可靠	MTBF	• 每两次相邻故障之间的工作时间的平均值	天
		MTTR	• 将有缺陷的部件或系统恢复工作秩序所需要的平均时间	小时
		可用度	• MTBF /(MTBF+MTTR)	%
	经济	单位容量TCO	• 存储系统5年 TCO/容量	\$/GB
前沿保障	异构兼容性	支持的存储协议	• CIFS、NFS、iSCSI、FC.....	
		连接方式	• SAN、NAS、DAS	

## 2.3 全球各国先进数据存力发展程度对比

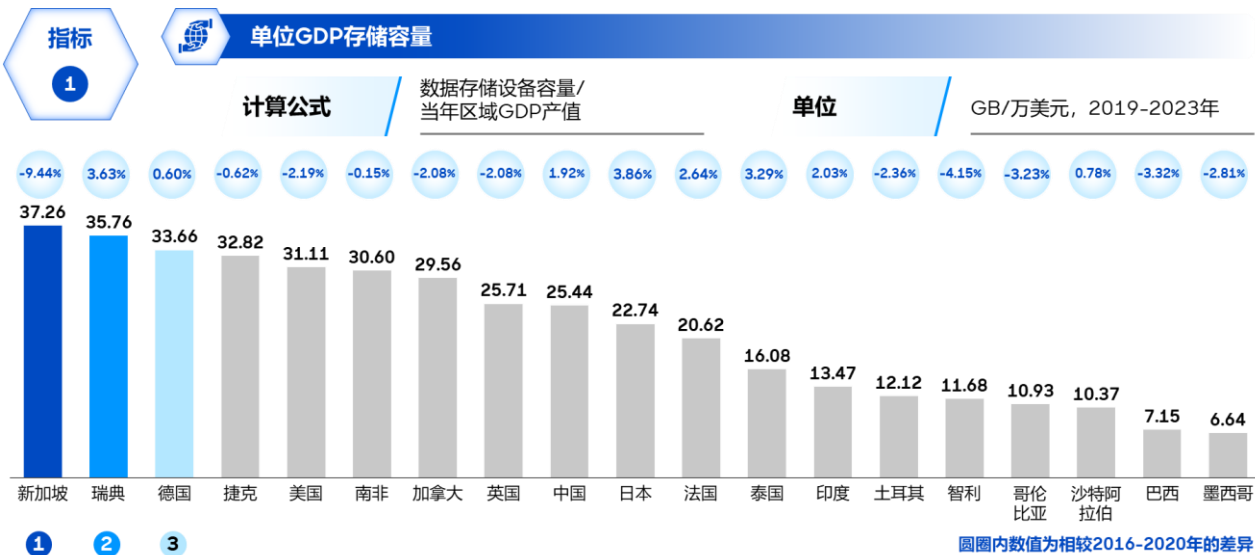
### 2.3.1 体量衡量：充足供应与持续增长支撑经济社会的可持续发展

#### 指标 1 单位 GDP 存储容量

##### 1) 指标解释与结果展示

存储空间的容量，是人们对数据存储能力的最基础的认知。考虑到经济体量较大的国家往往会拥有较大体量的存储空间，本文在计算存储能力的容量时采用单位 GDP 存储容量来确保不同国家的可比性，即用数据存储空间容量除以 GDP 规模。每一万美国 GDP 对应的存力容量越高，表明该国的数据存力容量水平越高，该国数字经济在 GDP 中的渗透程度越高，能更好支撑经济社会高质量发展。

图 2-7 各国单位 GDP 存储容量表现



##### 2) 指标结果分析与典型案例

2023 年的测算结果显示，新加坡、瑞典、德国、捷克等国家的单位 GDP 存储容量较高，每一万美国的 GDP 对应的存储容量分别达 37.26 GB、35.76 GB、33.66 GB、32.82 GB，存





储在经济发展中的支撑能力较为明显。日本、中国、法国等国家处于中间水平，每一万美元的 GDP 对应的存储容量约在样本国家中位数 20-25 GB 左右，墨西哥、巴西等发展中国家的排名较为靠后。

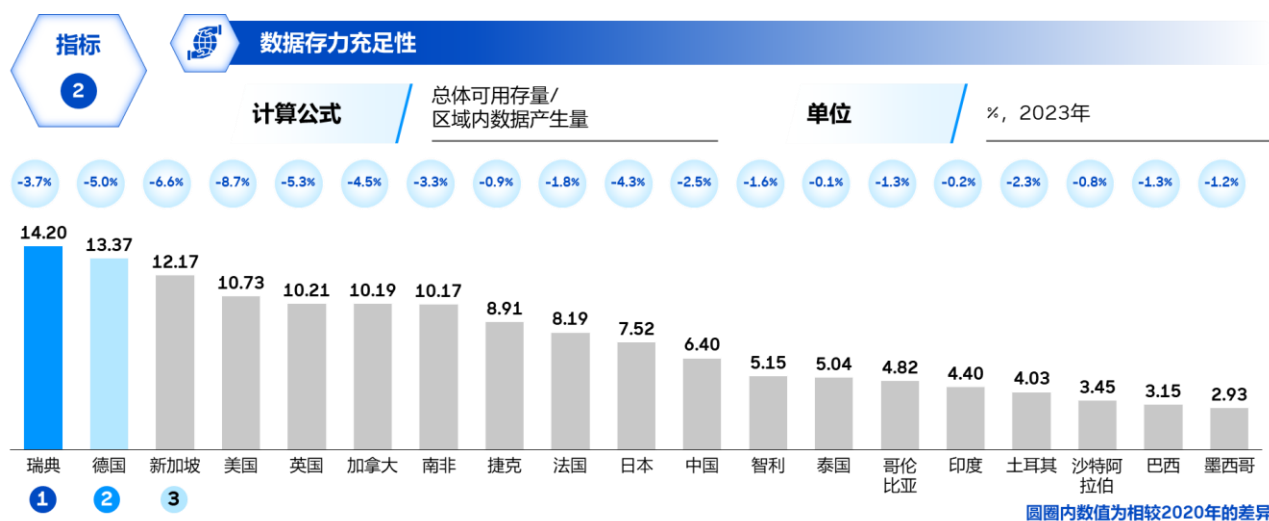
相较于 2020 年，不同国家的单位 GDP 存储容量有不同趋势，但是全球主流国家的存力投资容量 2021-2023 年这两年间至少增长了 15%<sup>26</sup>，这表明全球主要区域及国家对存力容量高度重视且持续通入。另一方面，部分国家受到通货膨胀等因素影响，导致名义 GDP 增速快于实际存储容量投入。例如，巴西名义 GDP 在 21-23 年增长了约 24.8%，但通货膨胀率也在同期显著增长，如巴西在 2021-2023 年的通货膨胀率分别是 10.6%、5.8%、4.6%，存力投资快速实际经济增长<sup>27</sup>。

## 指标 2 数据存力充足性

### 1) 指标解释与结果展示

数据存力充足性衡量的是数据量需求和存储容量供给之间的匹配关系。通过计算一国数据存储容量除以一国数据总产量可以反映当年一个国家或地区所生产的所有数据中，有多大占比可以被存储，以反映存力的充足性可以衡量一个国家的数据存力是否足够充足以支撑其经济社会高质量的发展需求。

图 2-8 各国数据存力充足性表现



## 2) 指标结果分析与典型案例

2023 年，当前水平世界排名前三的国家还包括瑞典、德国和新加坡，数据存力充足性分别为 14.2%，13.4%和 12.2%。全球数据产生量在快速增长，存力供给增速相对较慢，因此数据存力充足性出现下行趋势。随着语音、视频、图片等非结构数据体量的快速增长，对安全可靠高性能的数据存储方案的需求将会持续提升。

例如，德国自 2021 年开始就通过了一系列联邦政府数据战略（如欧洲“Gaia-X”项目），强调服务数据安全存储、处理和共享的基础设施建设，增强欧洲数字主权，这有效刺激了数据存储产业的建设需求。

2023 年，中国的数据存力充足性数据为 6.4%（2020 年为 8.9%），当前数据充足水平尚不充足。对比同期美国数据能发现，中国存力供应起步晚、基数小，充足性绝对水平不足、增长速度仍相对较慢。

## 指标 3 数据存留率

### 1) 指标解释与结果展示



数据存留率测算在单位时间内，一个区域内产生的总数据量中，被有效保存至硬盘存储（“盘化”）的数据总量占比。如果说数据存力充足性表示一个区域“承接数据产生量的潜力上限”，那么数据存留率则表示“有多少数据被该区域保存至存储系统并被妥善保管”，更高的数据存留率直接反映了一个区域/国家对数据要素的重视程度及实际利用情况。

具体而言，数据存留率其中的计算过程涉及了存力充足性、新增数据占用率以及新增数据周转率等因素，其中“区域内存力供应 \* 被保存数据存力资源占用率”表示每年新增被盘化保存的数据当量，受到数据保存策略、冷热温数据生命周期管理策略等因素影响，是一个应被重点关注的综合性指标。总的来说，我们认为数据存留率的重要性来自两方面。

一方面是深化对历史数据价值的挖掘，从而促进社会福祉、保障民生民权、使能产业发展。例如，在德国博物馆、图书馆和国家档案馆保存的具有历史和文化价值的文件，需要永久电子化保存，以方便公众查询和历史研究；再如，德国公司的股东会议记录、董事会决议等数据，需电子化保留至少 10 年，以确保公司治理和法律合规。

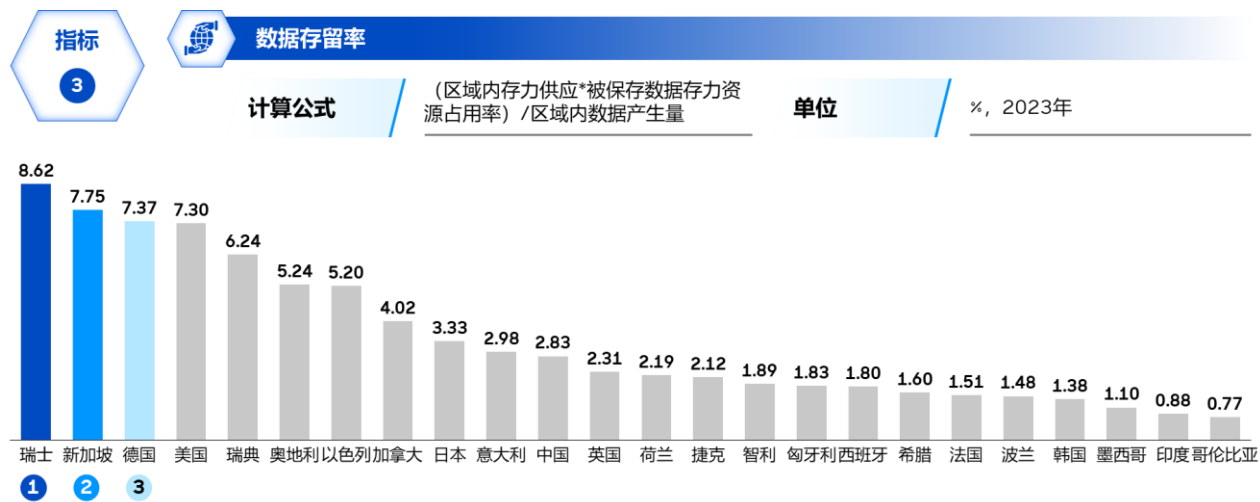
另一方面为加速智能时代的到来，国家和企业要挖掘并留存更多的数据供 AI 发展使用。当前以大模型为代表的原生 AI 应用，更多是学习正确的结果（即“经验”）。在未来，以数据为“燃料”的 AI 原生应用，有条件通过学习错误的结果（“教训”），以及从“过程数据”中进行学习，加速通用人工智能 (AGI) 的到来。

正如图灵奖得主、谷歌前副总裁杰弗里辛顿曾做训练过一个尝试识别手写数字的小型神经网络，在训练数据集编写时，辛顿故意让训练数据集的错误率达到 50% 且不更改这些被错误标注的数据集。但当他基于该数据集，使用反向传播训练法 (BP 算法) 进行持续梯队优化时，神奇的事情发生了：在识别新的手写数字时，模型给出答案的误差降到了 5% 以下。这意味着，这个小型神经网络“意识”到训练数据集本身是有问题的，已具备了一定的“分辨”能力，成功避免了过拟合 (Overfitting) 等问题。辛顿后来感叹道：“使用错误的训练数

据，可能反正能让 AI 获得更好的结论，正如聪明学生可以知道老师讲的知识中，有一半可能并不那么准确、因此能够 (通过学习与反思) 比其导师更聪明，AI 也是如此。”

大胆展望未来，除了“错误的数据”外，大模型等 AI 原生应用也可能通过对“过程数据”的学习，在执行任务时将不仅能“知其然”，还能“知其所以然”。例如，随着敏捷开发成为趋势，开发者的代码也被分为了成功提交上主干 (Main Branch) 且符合公司编码要求的“正确代码”，及未成功提交主干的“过程代码”。通过全盘学习“正确代码”与“过程代码”，辅以海量数据输入与先进算法架构，AI 原生应用有条件真正理解敏捷开发全过程、从而具备检查错误代码乃至生成正确且符合公司编码规范代码的能力。

图 2-9 各国数据存留率表现



## 2) 指标结果分析与典型案例

中国的数据存留率仅为 2.83%，不及美国的一半，也落后于瑞士、新加坡、德国、美国等国，几国分别为 8.62%、7.75%、7.37%、7.30%。数据存留率其中的计算过程涉及了对国家的存储容量供应、企业级数据存力充足性、生产数据周转率等领域，存留率应当被视作为



一个综合性指标。因此通过此指标可反映出中国的综合实力较孱弱，进步空间大，中国目前的存力水平依旧与世界数据存储强国有较大差距。

中国与欧美等国数据存留率有差距，核心在于潜在价值数据目前还没有在存储系统中存留。而数据存不下来、不被组织保存下来的核心原因，是因为数据没法很好地赋能决策、数据的价值也未得到充分挖掘。在智能经济时代，随着数据要素市场的规模化发展与大模型等 AI 原生应用对海量高质数据的需求，势必会强化千行百业对数据价值的深化挖掘，使得更多行业的数据将需要被很好地保存、以形成高质量数据集。

例如，在气象预测领域，早在 2018 年，欧洲中期数值天气预报中心 (ECMWF) 就启动了“欧洲气候数据存储 (CDS)”服务，提供过去、现在和未来气候信息的一站式服务。具体而言，CDS 包含一个工具箱，可以让用户建立自己的基于网络的应用程序，分析、监测和预测气候驱动因素的变化及其对商业部门的影响，如分析地表温度和土壤湿度的变化及其对能源、水资源管理或旅游业的影响。该数据存储服务可开放获取并免费供所有人使用，极大地改善了获取气候数据和工具的途径，为社会经济与气象科学发展带来了海量价值。

总的来说，中国可在四个维度进行考量并且高质量优化数据存留率。一是以深入数据要素的价值挖掘、将更多数据保留以直接提升数据存留率；二是在重点行业围绕核心应用（如金融、电信、政务等行业的核心系统）加速推行存算分离架构的加速落地，强化构建集约化共享存力资源池、以间接提升被保存数据的有效占有率、降低“服务器 + 本地盘”方案带来不必要的冗余数据备份、排掉“被保存数据的存力资源占用率”中的“水份”、使能高质量发展；三是持续以大模型等创新型应用为牵引、强化对数据价值的挖掘、提升热温数据占比进行，进而带动数据周转率的提升；四是通盘考量热温冷数据生命周期管理策略（即冷、热、温数据的存储周期），确保较高安全等级的数据得到足够长时间的存留，不可一味追求高存留率，致使关键或高度敏感的业务数据的存储周期过低。



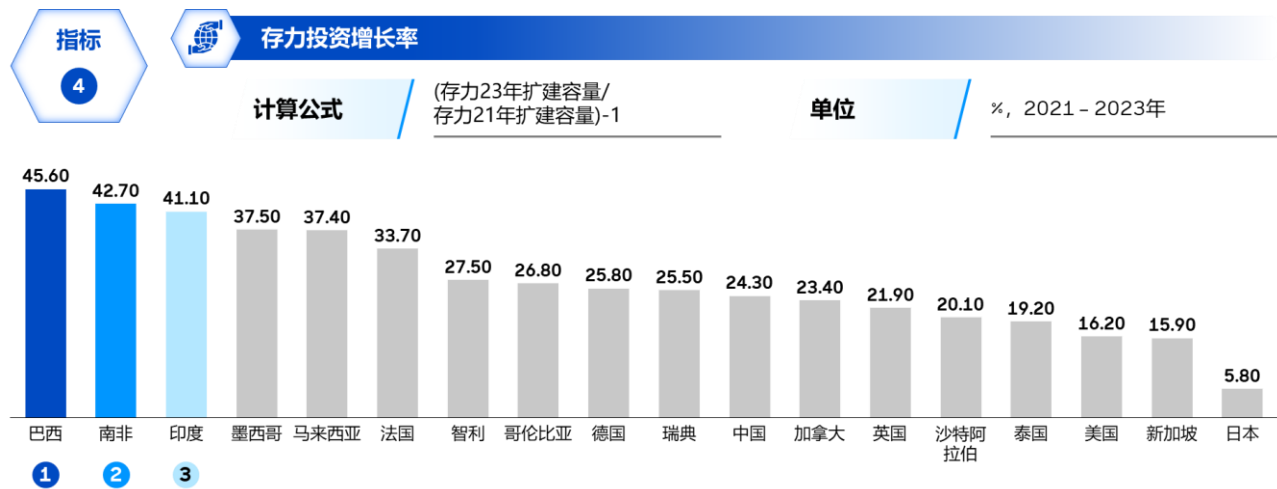
在引导数据存留率的提升上，政府有不应简单地数据存留率视作企业的内部管理行为，而应积极发挥引导作，“有为而治、主动引导”。第一，政府应积极发展数据要素市场发展，加速完善数据要素的确权、定价机制，使能数据在不同组织间的有序流通与交易，从源头上提升企业挖掘、存留数据的积极性，促使企业将数据视作智能经济时代的“核心资产”。第二，政府可强化以大模型等原生 AI 应用在政务、金融、电信等行业的部署、以对智能涌现的促进、强化对数据要素价值的深度挖掘，实现从“人力头脑风暴式数据挖掘”到“硅基智能体自适应数据挖掘”范式的转变。第三，政府应针对各个行业出台数据分级管理机制、推动更多组织日常数据的保存。例如，英国国家卫生服务体系 (NHS) 也实施了电子健康记录系统，患者的病例保存期限受到一系列法律法规和指导方针的规定（如 X-ray 及其他放射影像通常与病历保存时间一致，多为最低 8 年）；再如，荷兰的医院与诊所遵循严格的法律规定，确保 X-ray 影像的长期存储（可长达百年）。第四，政府可主动规定敏感重要数据的存储周期要求、避免企业一味追求过高的数据存留率而忽视对重要敏感数据的存储周期要求。例如，德国要求所有本国企业保留商业信函与员工记录至少 6 年，相比之下中国在规范性与要求上不如德国设定严格，用人单位保存终止的劳动合同信息仅 2 年。

## 指标 4 存力投资增长率

### 1) 指标解释与结果展示

存力投资增长率指标使用 2021 年至 2023 年三年间各国的存储投资容量的年复合增长率来衡量各国增长动能。高增长率体现的是在各国已有存力规模上的追加投资增速。

图 2-10 各国存力投资增长率表现



2) 指标结果分析与典型案例

巴西、南非等发展中国家的存储市场增长迅速，2021-2023 年复合存力投资增长率（按容量）均超过 40%。巴西政府重视数据中心的行业发展，将其视为鼓励包括银行业等多行业发展的重要推力，宽松的环境也吸引了针对性的投资。2017 年来，谷歌已经在巴西持续投入建设 ICT 基础设施、强化存力投资建设，并将在圣保罗等地进一步扩大投资水平。

中国 2021-2023 存力容量扩建投资复合增长率仅为 24.4%，与 17-19 年 44.7%增幅相比，出现较大滑落。此外，中国存力容量扩建投资仍有相当一部分来源于二级系统的 HDD 规模增长，相比之下同为发展中国家的南非和巴西，在 19-23 年期间，全闪化投资超过 55%。

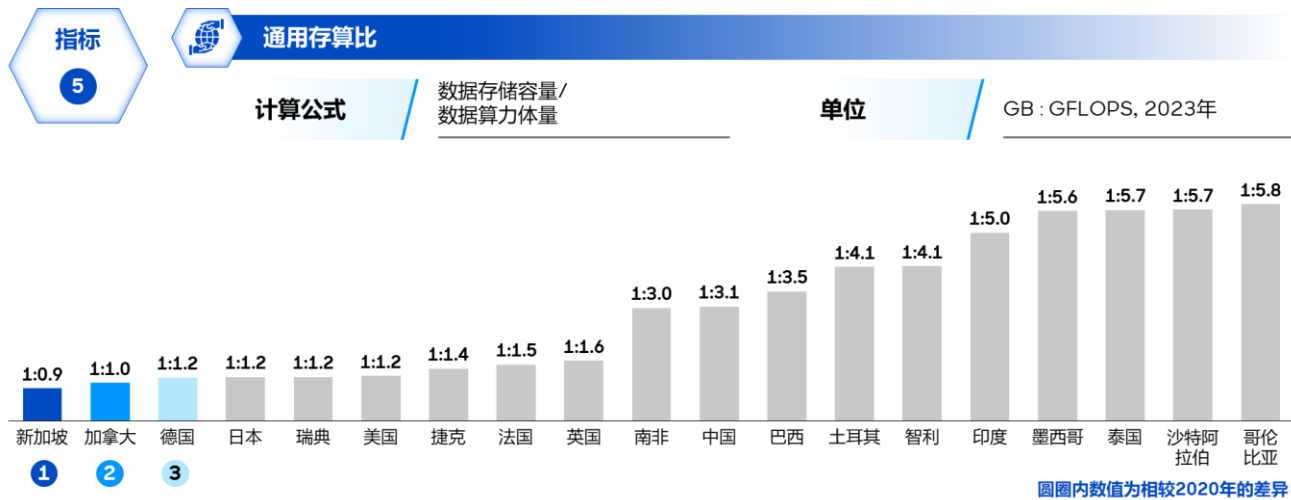
2.3.2 效率衡量：10 倍性能提升与内生智能特征加速数字经济规模化增长

指标 5 通用计算存算比

1) 指标解释与结果展示

存储和计算均为数字基础设施中的重要组成部分，通过计算数据算力的体量和数据存力的容量可以评估数据中心建设的平衡性，避免因存力的短缺造成算力的冗余而影响数据处理中的效率。随着数据为王的时代到来，数字应用的关注重点也从追求计算机的计算速度转变为追求大数据处理能力，单纯的数据算力无法单方面支撑整个数据中心的运营，还需要相匹配的数据存储容量。

图 2-11 各国通用计算存算比



## 2) 指标结果分析与典型案例

存算比表现相对较好的国家中，新加坡作为亚太地区的数据枢纽，不仅政府推出一系列如“智慧国计划”的政策激励强调 ICT 基础设施，还是多个全球 ICT 解决方案供应商的区域总部，如谷歌、微软等。

据预测，全球通用计算总量将在 2030 年达到 3.3 ZFLOPS，与比 2020 年相比增长十倍，这意味着全球年均算力将在这十年间呈 25% 的年复合增长率。中国作为算力全球第二，且将人工智能产业作为核心未来产业且 AI 产业应用潜力最大的国家之一，其对算力（尤其 AI 算力）的需求量必定不低于 25%。当前，中国在 2021-2023 年算力年复合增长率约为 25%



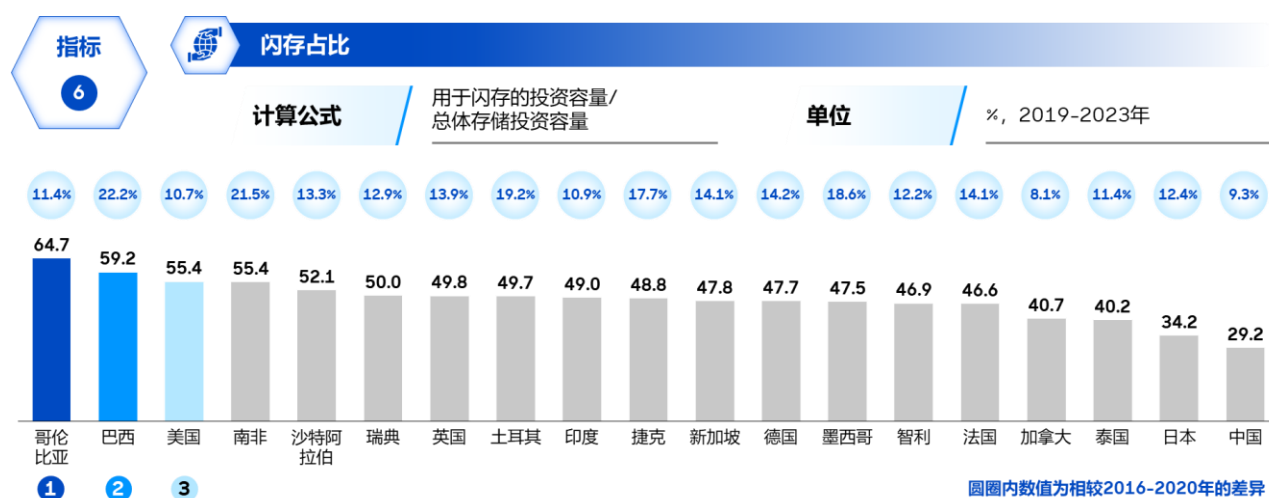
左右，这意味着算力配置仍有一定增长空间；21 到 23 年中国目前的同期存力年复合增长率仅为 11.5%，存力有进一步增长的空间。

## 指标 6 闪存占比

### 1) 指标解释与结果展示

闪存是业内升级演进的重要趋势，通过计算各个国家当年数据存力总投资额中闪存相关产品的投资占比可以进一步分析不同国家对于敏捷高效生产力的使用情况，占比越高代表该国家进行快速的数据调取和挖掘时，闪存能够提供强大的底层支撑能力，存力敏捷性更强。

图 2-12 各国闪存占比表现



### 2) 指标结果分析与典型案例

闪存技术相对于传统 HDD 响应快、容量更高，且近几年由于竞争逐步完全、价格正在逐步下行。中国对于 2019 年至 2023 年 5 年的累计闪存投资额仅为总体 5 年投资额的 29.2%，在世界内均处较低水平。排除小基数的影响，美国仍然是当前世界范围内闪存占比指标表现最好的经济体。如此看来，中国企业需更积极地挖掘数据要素价值、提升热温数据比例、从

源头提升对先进介质的需求；同时，积极关注 SSD 介质在方案层级 TCO（如低碳带来的运营成本节降）而非单纯聚焦 SSD 与 HDD 在 Capex 上的对比。

相较于 2020 年的累计统计数值，哥伦比亚依然维持世界领先的地位。哥伦比亚的数据中心产业持续蓬勃发展，尤其在电信、电子政务、智慧城市、金融等关键行业持续加码先进闪存介质的应用。例如，Claro Columbia、Movistar Columbia 等电信运营商在其数据中心内积极应用全闪化数据存储解决方案。再如，金融等行业同样需要快速交易、实时分析、保障敏感数据安全也加大了国家对于先进介质的拥抱，因此哥伦比亚的闪存化率保持世界领先。

领先国家闪存率高的核心原因，主要在于更高自动化程度的应用系统对 SSD 全闪化方案提出的刚性需求，以及在选项数据存储方案时对方案级 TCO 的重视。一方面，高度自动化的系统对数据处理速度和效率有极高要求，推动了对 IOPS 和读写带宽等高性能指标的需求；另一方面，全闪化方案的 TCO 当前在诸多场景下已有条件与 HDD 方案相当乃至更优，对方案级 TCO 的重视也使能领先国家更倾向于使用全闪存化解决方案。

### 2.3.3 基本保障：满足安全稳定的基本发展要求

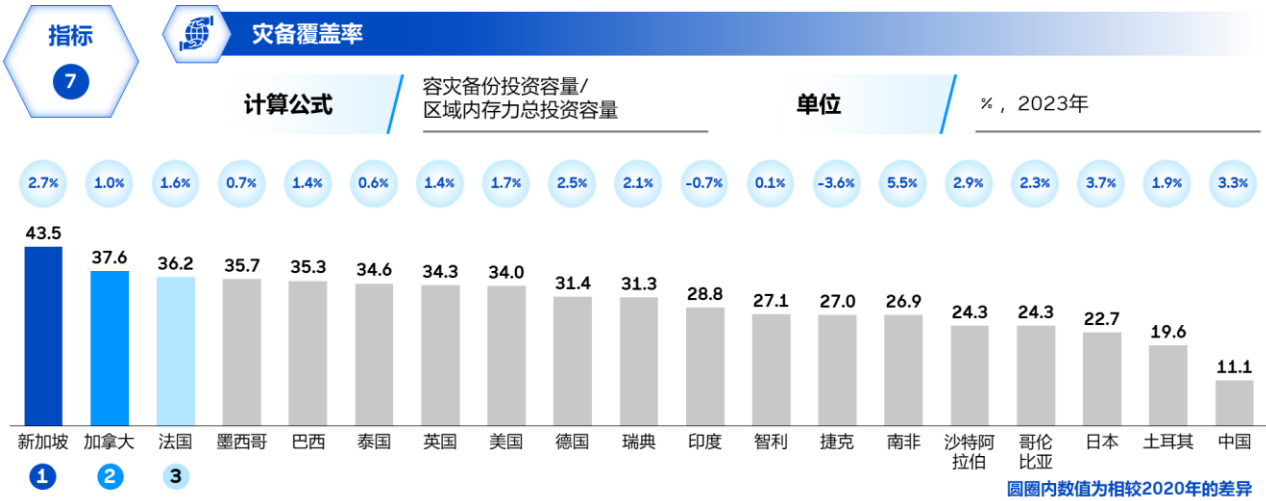
#### 指标 7 灾备覆盖率

##### 1) 指标解释与结果展示

容灾备份投资是可靠性重要表征，充足的容灾备份能有效应对各类网络安全攻击和意外。灾备覆盖率衡量的是数据存力投资中用于备份和容灾恢复的投资占比。更高灾备覆盖率意味着企业能将自身更大的业务数据存储更高的年限。

图 2-13 各国灾备覆盖率表现





2) 指标结果分析与典型案例

从样本数据来看，数据存储的灾备覆盖率仍低于预期。新加坡 (43.5%)、加拿大 (37.6%) 和法国 (36.2%) 的存储投资中用于灾备的占比最高，表明其数据应对各类意外的韧性较强。而中国 (11.1%) 的灾备投资占比相对较低，可以看出中国在存储灾备行业投入严重不足，一旦发生应急事件或会造成较大损失。

2024 年 4 月，中国工信部发布《关于做好 2024 年信息通信业安全生产和网络运行安全工作的通知》已经提及，现阶段应当“加强投入保障，各企业加强头筹规划能力，以增强容灾备份能力”。目前，中国容灾备份等级最高的金融业仅实现 4 级（美国已拥有多个 6 级存储区域），大部分行业备份水平仍处于 1-2 级，可见中国容灾备份行业未来的增长空间巨大，政府需要以更大的力度采取措施鼓励、敦促容灾备份在国计民生各个行业发挥应有的风险防控作用。

新加坡为数据中心产业、金融服务与金融科技产业等国家关键战略产业提供了安全稳定的数据存储方案，新加坡对数据容灾备份的重视和显著投资是出于对自身重点产业的保护，在新加坡，有许多数据中心直接采用了 Tier 3 和 Tier 4 级别的灾备需求。

2020-2023 年，日本推出了一系列政策以提升数据灾备覆盖率和加强灾害管理，提高灾害管理的效率；法国通过经济复苏计划（France Relance）加强了数据灾备的覆盖率，特别是通过对数字化和绿色技术的投资来提高数据中心的安全性和可持续性。

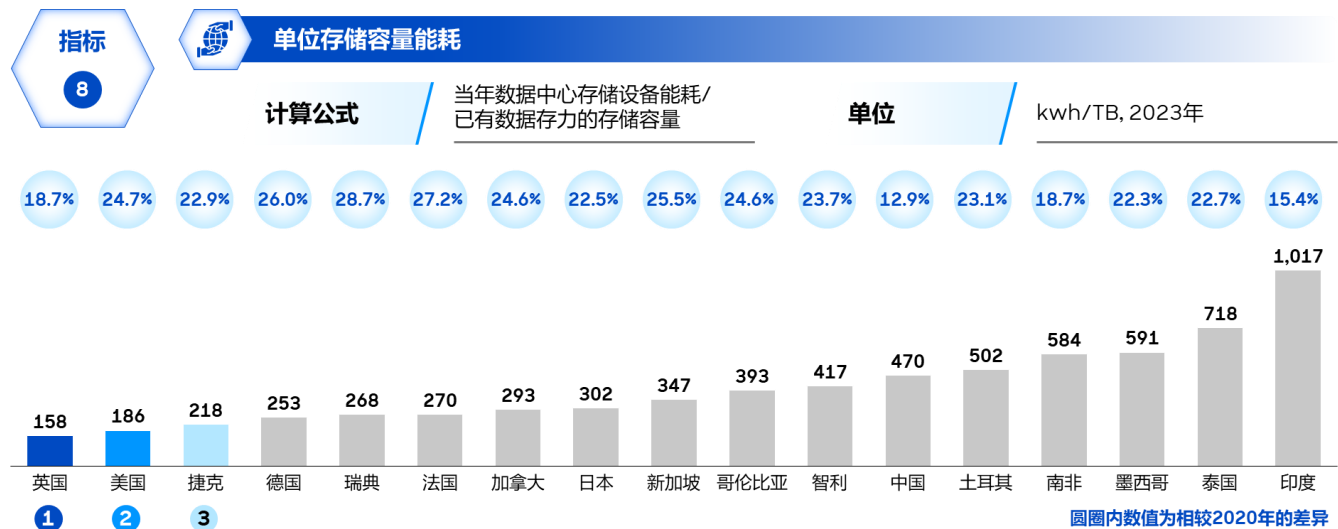
2.3.4 前沿创新：绿色先进实现可持续发展需求

指标 8 单位存储容量能耗

1) 指标解释与结果展示

存储设备导致的碳排放与日常使用过程中的能耗直接相关，在衡量绿色指标中，通过将存储设备一年耗电量除以总存储容量来进行不同国家存储设备绿色性的比较，单位存储容量能耗越低，则代表数据存力碳排放低，更好支持社会实现节能减排。

图 2-14 各国单位存储容量能耗表现



2) 指标结果分析与典型案例

2020 年到 2023 年，全球主要国家的单位容量能耗都有下降的趋势，核心原因是各国闪存化率的提高与对数据中心绿色化发展解决方案的投入。总体而言，当前数据存储行业正见证



着向低碳、高效两个维度的转型。从排名来看，英国、美国、捷克、德国、瑞典和法国单位容量能耗最低，而南非、墨西哥、泰国和印度最高。

中国相较 2020 年有一定程度下降，但整体而言降幅相对美国、德国等仍有一定差距，主要的原因之一是全闪化率的提升程度较低。在 2023 年，中国对数据存储的投资额中，仅 32.6% 为闪存投资，相比之下德国、新加坡、英国、美国与印度等国的闪存投资额占比均超 50%。尽管国内部分数据中心在积极尝试创新型解决方案（如创新型冷却系统），但提升闪存占比对节省存储设备能耗而言，是更为本质且效果更为显著的降碳方案。

为持续降低单位容量能耗水平，中国一方面需强化全闪化进程、从源头降低存储设备的能耗水平，另一方面在数据中心层面探索降低能耗的解决方案，如创新型冷却系统、数据中心智慧运维方案、引入可再生能源等，双管齐下以实现可持续降碳。

## 指标 9 数据存力专利占比

### 1) 指标解释与结果展示

数据存力相关专利体现了科技创新水平，也帮助企业增强市场竞争力。本文统计了 2023 年当年全球“数据存储”和“存储器”相关的专利申请量，并按照申请人注册地址区分国别。

图 2-15 各国数据存力专利占比表现

指标

9



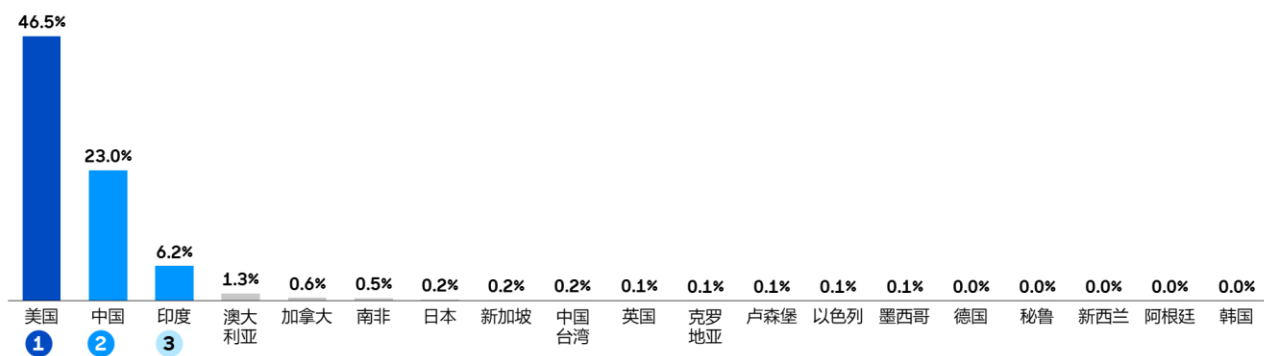
数据存力专利占比

计算公式

$$\frac{\text{当年本土企业申请的数据存力专利数量}}{\text{全球数据存力相关专利数量}}$$

单位

%, 2023年



## 2) 指标结果分析与典型案例

2023 年，全球与数据存储和存储器有关的专利申请数超过 28 万。从专利数量上来看，美国始终保持世界范围内的领先优势，中国已在 2023 当年已显现追赶劲头，但总量上仍与美国仍有差距。

尤其在高价值专利领域，中国和美国差距十分显著。在 2023 年，美国超过 30 万美元以上的高价值专利申请数超过 2.7 万份，超本国申请专利数的 20%，与此同时，中国基本没有高价值专利申请。这意味着，在尖端存储领域的创新上，中国与美国仍存在较大差距。美国高价值由一众高科技公司贡献，主要是谷歌、英特尔、Sonos、苹果等在为专利研究做贡献，高价值专利内容的关键词有“服务器架构”、“存储网络”、“存储池”等。另一方面，美国的专利申请环境玩家多元、内容均衡，美国前十专利申请人仅贡献了整体专利申请数的 10%。而中国的前十专利申请人贡献了整体专利申请数的近 20%。

美国的领先优势源于其强大的研发基础、充足的资金支持、丰富的人才储备和完善的知识产权保护体系。尤其需要注意的是，美国达成了产学研一体化的统一性，通过政府和企业、高校之间的密切支持合作，促进技术研发和创新。比如，Manufacturing USA 计划通过联邦和企业共同资助，促进工业界与学术界合作，推动先进制造业发展。加州大学伯克利分校



的实验室 AMPLab 与戴尔公司展开长期的技术合作，专注于大数据处理和存储技术的研究，共同推动了技术的应用和优化。

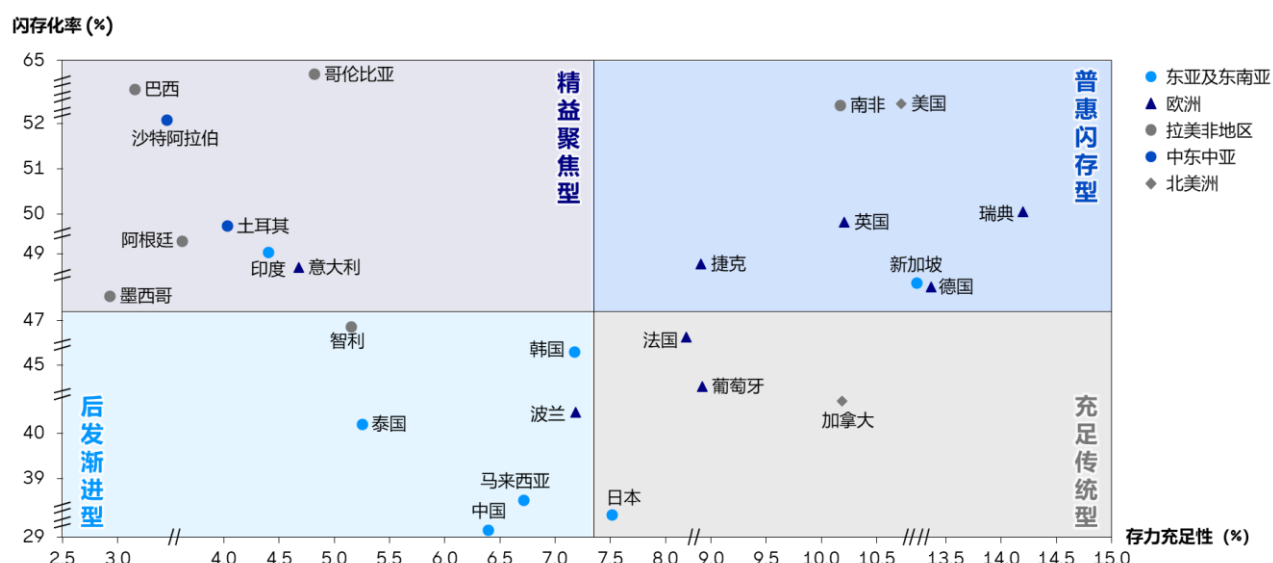
亚洲方面，印度的表现非常突出，主要是因为近两年有诸多跨国公司在印度申请专利。例如，2023 年印度前十专利申请人有 7 家是大型跨国公司，例如 Qualcomm、小米、三星等，且这 7 家公司贡献了印度整体专利申请数的 25%<sup>29</sup>。

## 2.4 全球各区域发展存力的高阶发展建议

### 2.4.1 全球各区域存力发展现状

不同国家区域的发展阶段不尽相同，我们按照**存力充足性**（衡量充分性）与**全闪化率**（衡量先进性），将全球区域细分为 4 大类型，分别为“普惠闪存型”、“充足传统型”、“精益聚焦型”与“后发渐进型”。

图 2-16 全球各区域存力产业发展现状





**“普惠闪存型”**国家存力充足性以及闪存化率均较高，通过定位全球存储产业地位、出台激励与规范政策、推动联合创新以及鼓励头部企业采纳先进方案，全面发展国家先进数据存力产业。例如，美国的数据存储充足性和高闪存化率得益于政府的高度重视和财政支持，如《芯片法案》等政策提供了大量资金支持。通过产学研一体化的合作，企业与国家实验室共同推动先进存储技术研发。此外，产业界对先进存储方案的积极应用，如英伟达在 AI 基础设施中的全栈化解决方案，进一步提升了美国的存力水平。未来，美国将继续在存储技术创新、数据安全与可持续性以及国际合作方面发力，巩固其在全球存储产业中的领导地位。

**“充足传统型”**国家的存力充足性处于世界领先地位、但闪存化率低于全球平均。这些国家通常已具备较好的数据中心产业基础，但国家内的存力升级的速度仍有待进一步提升。以加拿大为例，加拿大的存力充足性得益于广泛分布的灾备覆盖的发展。目前加拿大的 HDD 和 SSD 容量分别占总容量的 86%和 14%，其闪存化率较低与其资源型的经济结构和数据密集型产业的需求相对较少有一定关系。未来，以加拿大为代表的充足传统型国家可以通过政策支持 and 激励措施，鼓励企业和数据中心投资于闪存技术；此外，这类国家还可以加强对绿色 ICT 产业的打造，进一步推动闪存化进程。

**“精益聚焦型”**国家闪存渗透率高但存力充足性水平较低、“小范围先进、普惠性不足”。巴西的高闪存率与低存力水平之间的差异，主要源于其数据中心服务的集中性。数据中心主要服务于政府、金融、油气等高需求行业，这些行业对数据存储的安全性、可靠性和性能要求极高，因此推动了闪存技术的广泛应用。然而，这些设施并未普及到全社会，普通用户和小型企业的存储需求相对较低。此外，巴西正积极发展电商和智能制造等新兴产业，这些领域对先进数据存力的需求逐步增加，但整体仍处于初期发展阶段，导致人均存力水平仍然较低。展望未来，以巴西为代表的精益聚焦型国家可考虑加大对电商、智能制造等新兴产业的支持，推动数据存力的广泛应用；与此同时，加强 ICT 基础设施建设，普及先进存储技术，使更多中小企业和个人能够受益。



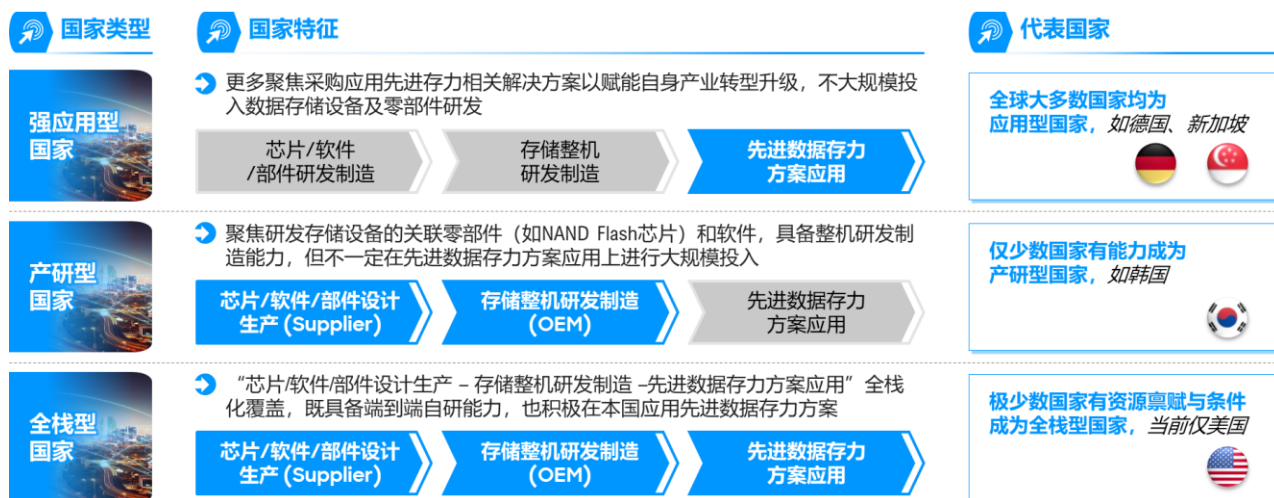
“后发渐进型”的国家的存力充足性以及闪存化率均较低。以中国为例，多数应用系统的自动化水平较低导致对存储的性能要求有限、对方案级 TCO 重视程度不足等原因，使得闪存化率偏低，当前中国全闪化率仅为 6.22%，远低于世界平均水平 17.01%。但辩证地看，

“后发渐进型”国家也有更大的“弯道超车”机遇，这些国家并不存在需大规模替换现有传统存储方案，而可直接加速全闪化方案的投资建设，实现存力产业“跃迁”。

## 2.4.2 典型国家建设先进数据存力的核心启示

成功从发展先进数据存力产业获益的国家可分为三类，一类是强应用型国家，一类是注重芯片和整机研发制造的产研型国家，还有一类是全栈型国家。强应用型国家更多聚焦采购应用先进数据存力相关解决方案以赋能自身产业转型升级，不大规模投入数据存储设备及零部件研发，全球大多数国家均为应用型国家，其中德国、新加坡为强应用型国家的代表；产研型国家聚焦研发存储设备的关联零部件（如 NAND Flash 芯片）和软件，具备整机研发制造能力，但不一定在先进数据存力方案应用上进行大规模投入，仅少数国家有能力成为产研型国家（如韩国）；全栈型国家实现了“芯片/软件/部件设计生产 - 存储整机研发制造 - 先进数据存力方案应用”的全栈化覆盖，这样的国家既具备端到端自研能力，也积极在本国应用先进数据存力方案，如中国、美国。本白皮书以德国、新加坡、韩国与美国为例，对三类国家的发展特征及关键成功要素进行梳理总结，以期形成可借鉴的经验供其它区域参考。

图 2-17 全球各国建设先进数据存力产业的领先实践



## 德国：政策与产业应用双元并举，打造全球数据存储强国

德国作为欧盟的经济领头羊，在欧洲数据存储产业中扮演了积极拥抱者的角色，同时以严格的政策要求促进数据存储等 ICT 产业的安全可持续发展，具体而言可总结为三个阶段。第一阶段始于 1990 年代中期，德国经历互联网服务快速增长，企业开始建设和租用数据中心；第二阶段发生在 2000 年代初期，谷歌、微软等科技巨头开始进入市场并设立研发中心，德意志电信等开始大规模建设数据中心以提供多样化服务；第三阶段从 2010 年开始大数据时代来临，带来了数据中心的建设量井喷，并于 2016 年起强化数据保护法与绿色发展法案，强化安全可持续发展。如今，德国数据存储产业在存力充足性、全闪化率、可靠性（如灾备覆盖率）、绿色可持续发展实现全球领先（如 PUE），在政策与企业的双重驱动下成为欧洲存储强国。

总结德国存储产业发展的成功要素，我们发现其具备三大特征。一是政策监管与德资企业对合规性的忠实，带来更高的刚性容灾备份需求，二是德国企业对数据要素的挖掘充分；三是高度重视可持续发展，持续产业政策与标准要求牵引。



首先，德国对多种类型数据的存储要求设定了更高标准，使其存力投资与灾备覆盖率充足。例如，在 2020 年《商法典》及《税法》的修订版中，要求德国企业保留商业信函与员工记录至少 6 年，账簿、发票等会计记录至少 10 年。相比之下，尽管中国也有如《劳动合同法》法规，但是中国在规范性与要求上不如德国设定严格，如用人单位保存终止的劳动合同信息仅 2 年，对诸多企业日常经营类数据缺乏相关硬性规范和严格要求。此外，德国全球性企业通常也会对数据存储年限做出严格且具体的要求。例如，某德资领先汽车行业 Tier 1 知名供应商要求员工档案至少保存 10 年<sup>30</sup>、年度财报与 BP 报告等敏感信息至少保存 15 年，且在全球范围内贯彻该要求（包括中国区）。

其次，德国政府高度重视可持续发展，持续产业政策与标准要求进行牵引。在 2014 年，德国推出了 DIN EN 50600 数据中心标准，为谷歌、微软等一众巨头确立了数据中心的建设参考规范。在 2021 年，德国政府《联邦政府数据战略》提出建设可持续的数字基础设施，而德国复兴信贷银行也根据德国的《可再生能源法修订版》与《德国气候行动计划 2050》推出了针对绿色数据中心的最高百万级的贷款计划，以低息率、长期还款期等政策优惠，根据企业项目规模等因素对在德数据中心进行改造与建设。例如，德意志银行在 2017 年与 Pure Storage 合作，采用了 Pure Storage 的全闪存阵列替代了传统机械硬盘改造了数据中心，极大程度提高了德意志银行的数据处理效率，比如提升了 CRM 和 ERP 系统性能，使得德意志银行处理大量客户数据和实时信息抓取管理的效率有效加强。

此外，德国企业对数据要素的挖掘充分。例如，某外资汽车零部件供应商设立专门的数据管控组织，按需协助业务部门抓取并导出所需分析的数据，从组织架构上为研产供销服等各职能提供数据服务，促进各职能提升对自身日常运营及成本项目等数据的掌控与实时分析能力。基于企业与政府的共同推动，德国当前具备超 350 个数据中心，一级存储系统在全存储系统中占比高达 76%、全闪存占整体存储容量的 51%，成为数据存储强国。

## 新加坡：以政策优惠招商引资，持续强化本土运营商创新能力形成互补

首先，新加坡在过去半个世纪以来持续重视创新产业，并积极以税收优惠及补贴政策招商引资并吸引外资高科技企业的持续投资、提升经济水平、创造就业机会、培养高水平人才。例如，1999 年发布的《全球贸易计划》<sup>32</sup>与 2010 年发布的《创新发展计划》<sup>33</sup>等政策，提供高达 10%的企业所得税等优惠，并成功吸引多家 ICT 企业在新开设分支机构与持续投资，其中就不乏数据存储产业领先企业，西部数据于 81 年入驻、希捷于 82 年入驻。

再如，2016 年发布的 **<TechSkills Accelerator>**<sup>34</sup>等补贴政策，通过提供新员工 90%的培训补贴、应届生 50%-70%的培训与实习费用补贴以及定制化培训计划激励谷歌等企业培养本地新加坡人才，向数字化未来更进一步。例如，谷歌 2022 年在新招聘约 200~300 名员工，而其中来自新国高校的候选人就达到 65 人，占当年招聘员工量的至少 20%。

近年来，为适配数智化时代的到来，新加坡在 2014 年提出了智慧国度 2025 计划，以数据中心建设为核心，发展先进数字化产业。一方面，满足当地产业生态的数字化升级要求，如新加坡本土企业 Singtel、Keppel Data Centres 运营的数据中心为有数据存储需要的企业提供托管服务；另一方面，将算力与存力等数据服务发展为一个独立的战略性产业，如 Singtel 利用自己的数据中心为其他企业提供**数据存储服务**。

更进一步，为强化本土企业创新能力、强化自主创新能力，新加坡本土运营商也积极与在新 ICT 厂商合作、联合研发数据中心管理工具方案等创新型业务。如此，新加坡运营商与外企 ICT 基础设施厂商就形成了互补关系，共同为用户企业提供数据存储托管服务套餐。例如，Keppel 与思科一同开发数据中心管理工具，Singtel 与微软 Azure 联合创新数据存储技术等。基于这些创新型解决方案，运营商提升了对自身数据中心的运营管理能力、也实现了第二增长曲线，如 Keppel Data Centres 的数据中心服务可占 Keppel 集团总收入的 7%。





此外，新加坡高校也积极与数据存储企业合作，形成产学研一体化生态。如希捷与南洋理工大学合作开展了数据存储课程和项目，共同研发高密度数据存储解决方案，包括热辅助磁记录技术等先进存储方案。经过多年的发展，以数据中心为核心的产业生态已为新加坡直接贡献了约 15 亿美元 GDP 增长（占新加坡年 GDP 增量超 10.5%）与大量就业岗位，带动提升了新加坡的数据存储产业生态发展。

总的来说，新加坡发展数据存储产业的成功要素有三。一是设立清晰的政策目标，到 2025 年数据存力产业投资规模达到 10 亿新加坡元，每年培养数以千计的人才完成高级技能培训以支持国内外存储领域企业，并新建或改建 20 个绿色标准数据中心；二是利用优质外商投资与补贴政策吸引优质 ICT 企业入驻；三是产学研深度协同、持续保障头部企业的高素质人才供给。

### 美国：多级政府齐发力，打造产学研一体化的先进数据存力生态

在政府层面，美国联邦及州政府高度重视数据存储产业，在过去十年积极出台多项政策，以直接财政支持先进数据存力产业发展。例如，美国联邦政府于 2022 年推出了芯片法案，以超 500 亿美元的费用资助美国本土半导体产业用于半导体和存储技术的研发、生产以及人才培养，并给予了 240 亿美元的投资税抵免。再如，纽约州的 Strategic Plan 中提到的 Tech Innovation for All Program 希望将政府、企业与科研组织相结合以达成科技技术创新研发，其中包括数据存储技术，为此政府提供了 5 千万美元的赞助费用。

除直接对企业进行财政支持以外，以政府机关鼓励并资助国家实验室推动产学研一体化，也是美国实现存力产业领先的核心抓手之一。企业通过国家科学基金会与能源部的 CRADAs 政策（合作研究和开发协议政策）达成与国家实验室的直接合作，开发 N+1 代产品的架构与特征，如 Seagate 与劳伦斯伯克利国家实验室（美国能源部所有，加州大学运营管理）一同

开发了应用于大数据和快速数据处理需求的创新存储介质与架构；另一方面，也可以为下一代产品做预研，如 HP 与橡树岭国家实验室一同开发了应用于超算环境下的数据管理和存储解决方案。

针对超算场景，美国政府资助的实验室亦多有布局，进而带动了先进数据存力产业的进一步发展。例如，美国橡树岭国家实验室(Oak Ridge National Laboratory) 的 Frontier 超级计算机在第 62 届 TOP500 排名为第一，是世界上第一台百亿亿次级别的超级计算机。Frontier 实现了 1.194 Exaflop/s 的性能，存储环境则具有近 700 PB 的容量，这使得其存算比也达到 **11 PFlops : 7 PB** 的标准。具体而言，Frontier 系统使用了超 5000 个 NVMe 设备，总读/写速度达到 10TB/s，并能够进行超过 200 万次随机读取操作，辅以额外的 480 个 NVMe 设备构成的元数据层，用于处理混合工作负载和维护文件系统的结构。

再如，23 年 6 月 22 日，英特尔（Intel）官方宣布美国能源部阿贡国家实验室已经完成基于英特尔 CPU 及 GPU 的新一代超算“Aurora”的安装工作，上线后将提供超过 2000 PFlops 的浮点性能；存储方面，Aurora 集成了超过 1024 个存储节点以每秒 31TB 的总带宽提供 220PB 的总的存储容量，这使得其存算比达 **9.1 PFlops : 1 PB**。

此外，美国产业界也对先进存储方案有着开放的心态与积极的应用。例如，英伟达优先采购专业外置存储、为其 AI 算力基础设施架构提供有效补充，与其算力生态形成高效协同的全栈化解决方案。正如黄仁勋在 2024 年 GTC 大会上强调：“在 AI 大模型时代，专业存储伙伴对我们非常重要！”，并强调 Milvus 等开源向量数据库在大模型训练场景下值得关注，认为向量数据库乃是解锁大型自然语言模型潜力的至关重要的基石<sup>37</sup>。”

## 韩国：选定高价值环节多年持续投入，政府与头部企业紧密协同



总的来说，韩国并不是一个“强应用型国家”，但却由于其在存储领域的商业优势拥有行业内的定价权。三星通过其在存储领域的市场主导地位，依靠庞大的产业链控制市场定价，不仅为自身带来了可观的经济效益，同时创造了大量的就业岗位，显著提升了整个韩国存储行业的竞争力。2023 年，三星存储业务的收入达到了 674 亿美元，占集团总营收约 36%；海力士在 2023 年的存储业务收入为 298 亿美元，占其集团营收的约 70%，先进的存储芯片产业为韩国创造了海量经济价值与可观的高薪工作机会，有效促进社会发展。

回顾历史，韩国将存储产业作为国家半导体领域的第一优先级，自 1970 年代开始，举国之力持续研发，最终在存储领域底层技术领先世界，在全球拥有存储芯片的定价话语权。回顾韩国半导体发展历程，在 1970 年代间，以三星为代表的韩企开始为日企代工电子产品，并逐渐完成了技术转移；同时韩国政府修改了外资投资法，确保对本土企业的强力支持。在 1980 年代，韩国政府启动了《半导体产业培育计划》以推动集成电路产业的发展，同时期三星也开始研发 DRAM，并在 1990 年代末市场份额成功反超日本，时至今日，三星和海力士的 DRAM 市场份额达到的 43%和 28%。

随后，三星在 1990 年代初开始研究 NAND 芯片及 SSD，经过多年发展，最终在 2016 年前后确立并夯实了自身在存储芯片领域的技术领先地位。除三星等企业自身的大举研发投入，韩国政府也在持续更新自身政策、对以存储为代表的半导体产业出台多项专项政策。例如，在 2010 年推出《国家半导体战略》，确立了在 2030 年成为半导体领域世界第一的目标，并在 2021 年实施《K-半导体战略》拟投资超 4000 亿美元持续推动半导体产业<sup>35</sup>，又在 2023 年提出《K-芯片法案》<sup>36</sup>，对 2023 年起在国家战略产业投资可商业化设施的企业，在两年内抵免大、小公司最高达 15%、25%的投资税抵免，且基于 23 年是否有额外投资，可额外增加 10%的 2023 年本年的税务额外优惠。在多重政策的加持下，三星与海力士等巨头最终确立自身在终端市场的领先地位，更进一步，三星与海力士分别于 2010 年与 2013 年前后，推出数据中心级推出企业级 SSD 产品。

2024 年 6 月，韩国政府公布了新一期半导体行业发展规划，拟投入 190 亿美元以增强韩国半导体基础设施建设，由韩国开发银行赞助韩国芯片制造商和供应商，并延长了税收优惠政策，继续推动韩国的半导体和存储器的发展。

除了韩国政府和企业的齐心协力，双方也积极带动国家高校及研究中心，以求技术突破、克服人才短缺，继续维持韩国的存储技术领域领先地位。例如，韩国政府每年以 2.7 亿美元的资金直接资助韩国人工智能和大数据研究院以求 AI 和大数据计算、存储的能力突破。再如，三星与海力士作为龙头企业也积极与韩国各大学合作推动项目以培养人才克服短缺，两家公司于延世大学、成均馆大学以及韩国科技技术研究院等高校推出与半导体相关的课程，提供培训项目以及奖学金支持，希望在五年内通过联合课程培养 1160 名存储产业人才。

### 启示：先进数据存力产业的稳健可持续发展，需政府与企业深度协同，在政策引导、技术创新与市场应用上齐发力

首先，一个国家需明确认知自身在全球产业格局中的位置，为自身在全球存储全栈产业生态中进行“战略定位”（如需全栈化发展、还是仅核心零部件的突破）。例如，美韩通过产业链协同，分别实现了应用生态的繁荣与存储芯片产业的领先。对中国而言，积极发展并端到端覆盖“芯片/软件－存储设备整机－存算网全栈解决方案”全产业链环节已成为大势所驱。

其次，基于产业的“战略定位”制定发展目标、提供政策支持、鼓励技术创新。一是提出清晰可量化的存力产业发展目标（如新加坡智慧国度 2025 配套计划提出数据存力产业投资规模达到 10 亿新加坡元，且培养 2500 高级存储产业人才，并新建或改建 20 个绿色标准数据中心），二是为支持这些目标的实现提供监管与鼓励性政策支持，监管性政策对数据保存与可持续发展有严格规范，如德国《商法典》与《税法》的修订版以及《DIN EN 50600》数据中心建设规范系列等，鼓励性政策提供资助与税收优惠，如新加坡《知识产权发展激励计



划》、《创新发展计划》以及国家研究基金等，三是以产学研合作等机制为抓手，鼓励企业提前布局下一代存力技术、强化头部企业的创新能力。

随后，各国在市场应用上应在公共部门（如政府及银行等国有部门）积极开展先进数据存力方案试点。如果公共部门不进行大规模应用、打造可复制的先进产业案例，则国内产业将始终难以形成先进且成熟的解决方案组合、实现真正的引领。此外，也应积极鼓励头部 AI 原生应用厂商（大模型厂商等）强化对先进数据存力方案的应用。

尤其对中国这种旨在成为全球先进数据存力全栈技术创新引领者的国家而言，以政策支持与产学研协同等机制鼓励全栈技术创新十分重要。一方面，端到端全栈创新所需的研发投入巨大，国家可定向扶持重点实验室、鼓励企业与高校联合起来，共同进行下一代存力技术创新（如在超算环境下的数据存储技术）；另一方面，存储行业规模效应显著，从美韩等国的发展历程看，各国最终能引领存力整个行业的仅为极少数厂商，因此应在先进数据存力全栈环节（存储芯片、解决方案供应商等）选择少数厂商组成产业联盟、以头部企业牵引行业标准制定，共同定义并持续赋能“先进数据存力”产业的发展。

图 2-18 对全球各国政府建设先进数据存力产业的建议

	强应用型国家	产研型国家	全栈型国家
“定目标”	明确本国产业在“芯片/软件 - 存储设备整机 -先进数据存力方案应用”中的聚焦环节及演化路径，选择成为 <b>强应用型、产研型或全栈型国家</b> ，并 <b>确立产业核心发展目标</b>		
“巧引导”	重视对 <b>各类数据生命周期的管理</b> ，如出台数据敏感度分级与保存周期政策	政府资助 <b>产学研联合创新项目、强化N+1代技术预研</b> ，如三星与成均馆大学实验室每年得到韩国政府大量资助	在强应用与产研型国家政策基础上， <b>强化本国数据存储端到端产业链的构建</b> ，鼓励 <b>本土全栈化解决方案的形成与应用</b>
	强化对 <b>存储整机厂商的招商引资优惠</b> ，如新加坡政府的税收优惠政策等	定向补贴 <b>NAND Flash晶圆厂</b> ，确保先进SSD颗粒 <b>量产节奏</b> 持续领跑全球	
“促应用”	对所有类型国家，在金融、政务、电信等对 <b>数据可靠性与性能要求高的行业或公共部门</b> 优先打造 <b>先进数据存力的试点项目</b>		
“强创新”	如有余力，可考虑以 <b>产学研一体化方式强化存力或存算网一体化解决方案层面的创新</b> ，如Keppel与思科一同开发基于云的数据中心管理工具	以 <b>产学研联合实验室(如美国橡树实验室)</b> 为核心 <b>创新创新载体强化预研、持续注资投入</b>	在金融、政务、电信等关键行业优先 <b>全栈化应用本国数据存储方案</b>
		成立区域化 <b>数据存储产业联盟</b> (如美国Storage Networking Industry Association)，推动 <b>技术标准的制定和统一、分享产业最佳实践、带动全产业链良性发展</b>	



# 03



## 行动倡议

发展先进数据存力是实现新质生产力的必由之路





## 3.1 先进数据存力产业发展目标建议

### 3.1.1 将数据存储作为 ICT 产业的重要独立门类单独管理

在数字经济时代，算力、存力与运力的概念被提出，分别对应计算、存储与网络产业。尽管三力并驾齐驱的格局已然形成，但存力本身并未被作为一个重要独立的门类进行单独管理。究其本源，一是因为存力产业缺乏具体的品类映射（如服务器/GPU之于“算力”、交换机/路由器/防火墙等网络设备之于“运力”），变成了服务器硬盘与专业存储设备的统称，而专业外置存储的重要性并未得到重视；二是数据存储在数字经济时代并未与人工智能等战略性新兴产业建立起直接的关联关系，而是更多作为ICT底座的一部分提供赋能支撑。

然而，在智能经济时代，事情正在起变化。首先，存算分离的架构创新使得抽象的“存力”有了具体的实体映射（以专业外置高性能数据存储设备为代表），先进数据存力以集约化存力资源池、高通量网络与开放算力生态组成的创新型架构，使得专业的外置数据存储设备 (External Controller Based Storage) 成为智能经济时代先进数据存力的核心价值载体。其次，先进数据存力“以存强算、以存代算”的特征，使其能够加速 AI 产业的发展、与算力形成补充共同促进“智能涌现”，因此智能经济时代，建设先进数据存力也需要作为发展 AI 产业的重要组成部分，在制定 AI 产业的发展战略时一同纳入考量。

### 3.1.2 确保先进数据存力产业全栈自主可控高效协同，强化先进 SSD 颗粒的国际定价权获取

先进数据存力不仅仅是一个硬盘，其能力更不局限于一个机架，而是可以推动数据中心架构端到端变革、加速通用AI智能涌现的革命性力量，其全栈生态涉及“零部件－存储整机－存算网一体化解决方案”三个层次。

近年来，全球自由贸易流通的格局与高科技领域的全球化产业分工格局于近年面临越来越多

的不确定性，中国当前如希望在先进数据存力领域取得可持续的全球领先地位，需构建自主可控的先进数据存力供应链，将产业全栈端到端自主可控作为产业实现可持续发展的基础。

关键零部件层面，尽管国内涌现了长江存储等 NAND Flash 厂商且在颗粒创新上取得了一定突破，但从全球视角看，当前 NAND Flash 颗粒的定价权集中在极少数国际品牌。如果不加速国产关键零部件厂商的规模化应用并快速取得定价权，则中国在 SSD 全闪化方案的牵引上将始终“受制于人”。因此，建议政府定向对头部国产 NAND Flash 厂商的规模化应用进行支持，强化中国在全球 SSD 领域的定价权，为先进数据存力解决方案在千行百业的应用落地提供有力支撑。

存储整机层面，国内虽已有一批较为成熟的 OEM 厂商，但在产业协同上仍有进一步强化的空间。先进数据存力产业尚在发展初期，产业标准尚需完全统一（如接口标准、分布式标准、运维标准、安全标准等），此时正需要各类产业生态厂商（如 OEM）积极协作强化沟通、共同制定先进数据存力相关产业标准，推动先进数据存力方案走向千行百业。

存算网全栈化解决方案层面，端到端国产化是打造本土安全可靠生态的刚需。在存算分离的架构下，集约化存力资源池、高通量网络与北向算力生态深度耦合，若其中任意一方非国产，则全栈化方案整体亦无法实现真正意义上的安全可靠。

### 3.1.3 推进以全闪化为核心的先进数据存力方案在千行百业的广泛应用

为实现先进数据存力产业的全球领先，高闪存化率是必要条件。一是性能跃迁，SSD的优异特征可以很好地满足业务性能要求，相较HDD，SSD能实现读写带宽与IOPS等核心指标10倍的提升，以及更高的存储可靠性，确保企业能够适应当前数据量快速提升、应对数据治理日益复杂等挑战；二是全场景普适，当前优异的SSD全闪化方案，已可在解决方案级与HDD方案在TCO上打平，这意味着全闪化方案事实上已可全场景普适（如AI大模型等核心系统、数据备份），未来随着数据压缩等技术的进一步发展，全闪化方案的优势将进一步扩大。



总的来说，全球领先的闪存化率是发展先进数据存力产业的基础。然而，当前中国的闪存化率仅**29%**，相比全球主要经济体而言显著更低，不及世界平均；与此同时，美国的闪存化率高达**55%**、德国亦达**48%**，与中国拉开显著差距。

因此，我们建议将推进以全闪化为核心的先进数据存力方案在千行百业的广泛应用，作为中国先进数据存力产业发展的一个核心目标。

## 3.2 建设先进数据存力发展高阶行动建议

### 3.2.1 “定目标”：将区域级先进数据存力指标体系纳入政府产业发展目标

建议将闪存化率、数据存留率区域级先进数据存力指标体系纳入政府产业发展目标，确保各级政府充分认知先进数据存力产业的重要性、投入必要资源。一方面，建议各级政府以数据存留率等指标确保各行各业的企业与组织对数据要素的充分利用；另一方面，各级政府可以闪存化率等指标进行牵引，加速全闪化方案在千行百业的落地，鼓励地方政府及国有企业优选全闪化方案，加速先进数据存力产业生态的规模化发展。

### 3.2.2 “巧引导”：围绕先进数据存力产业发展目标，制定监管与鼓励性政策

总的来说，建议政府从监管性与鼓励性政策出发，结合当前实际情况与先进数据存力产业发展目标制定相应政策。监管性政策的核心目标是确保数据要素在千行百业的企业与组织间安全有序流通、对不同敏感度的数据制定差异化且符合国情的合规性要求，而鼓励性政策的核心目标，则是确保先进数据存力产业全栈生态的良性发展、推动先进数据存力应用普及。

监管性政策上，建议对标德国等强应用型国家的数据分级管理政策，完善数据敏感度分级与保存周期政策、强化企业对数据备份的重视。例如，德国要求所有本国企业保留商业信函与员工记录至少**6**年，账簿与发票等会计记录至少保存**10**年。相比之下，中国在规范性与要

求上不如德国设定严格，如用人单位保存终止的劳动合同信息仅 2 年，且对诸多企业日常经营类数据缺乏相关硬性规范和严格要求。

鼓励性政策上，建议强化对长江存储等 NAND Flash 晶圆厂定向补贴。当前 NAND Flash 颗粒的定价权集中在极少数国际品牌、国产品牌不强势的竞争格局，使得企业级 SSD 价格时常出现波动、不利于以全闪化先进方案的规模化应用。以政府支持实现国产存储芯片晶圆厂在先进 SSD 颗粒的研发、量产与规模化应用上“领跑全球”，可有效控制国内先进 SSD 颗粒价格水平、确保以全闪化方案为核心的先进数据存力方案在各行业的落地。

### 3.2.3 “促应用”：积极推动先进数据存力方案的在金融、政务、电信等关键行业的应用，鼓励国产开源数据库等开源软件生态发展

首先，建议积极促进以全闪化为核心先进数据存力在千行百业落地，在金融、电信、政务等关键行业先试先行，核心原因有三。一是这些行业对数据存储的性能有更高的要求，先进数据存力的优异特征可以很好地满足读写带宽、IOPS 等业务性能要求；二是这些行业有着更高的数据合规要求，无论是端到端防勒索保障数据安全、分钟级 RTO 确保数据可靠性、还是长周期安全稳定的数据备份要求，先进数据存力方案都可以很好地满足；三是“树标杆促转型”，全闪化方案与存算分离的架构等先进数据存力方案的应用，可以很好地提升存算运三者的资源利用率，带动数据中心整体的成本节降，并且有助于形成标杆项目规模化复制。

其次，建议政府组建包含存储整机厂商在内的先进数据存力产业联盟，共同制定先进数据存力相关产业标准。通过产业联盟的建立与持续发展，先进数据存力有望加速落地。

最后，建议政府鼓励国产开源数据库等开源软件生态发展、拓展先进数据存力解决方案的应用空间。培养国产开源软件生态是具有良好的外部性效应（如鼓励国产数据库等软件的创新）、可持续累计（因生态伙伴持续参与贡献代码）且终极可靠（零风险）的举措，值得持续投入强化。尽管国内企业可以选择多元化的手段缓释采取国外开源软件的潜在安全风险，



如持续贡献社区提升话语权、主动向社区提供代码修改建议确保软件安全性、使用开源社区中成熟的版本确保稳定性与兼容性，但当面临重大国际地缘风险时，他国的开源生态依旧不是完全中立且可靠的。

### 3.2.4 “强创新”：资助并以产学研联合实验室为核心载体，提前布局超算存储等 N+1 代先进数据存力解决方案

通过对韩国、美国等在先进数据存力领域创新能力领先国家的对标研究发现，由政府大力资助的产学研实验室，是两国实现数据存储创新能力持续领先的重要抓手，也是这些国家提前布局 N+1 代先进数据存力解决方案（如超算存储）的核心载体。

一方面，当今时代技术革新的速度持续加快，技术演化方向的不确定性正在持续提升，这使得完全依靠自研，以“内生方式”覆盖所有潜在创新点模式的可行性受到挑战。在此背景下，以产学研联合实验室为载体，使能企业与学界的合作集思广益、共研共创，可为业的当代产品设计创新带来新的动能。例如，三星结合自身技术规划，与韩国科学技术院联合研发存内计算 (Process-in-memory) 技术，并将其作为通用技术模块 (Common Building Block)，赋能自身存储芯片产品线研发的提质增效。

另一方面，对于技术路线不那么清晰的下一代产品 (N+1 代)，国家资助的产学研实验室则是理想的预研载体。一是因为预研话题过于超前、企业参与的积极性可能相对有限，二是由于部分研发所需的运行环节与测试设备的组建成本高昂、通常需国家出资负担。例如，希捷与劳伦斯伯克利国家实验室（美国能源部所有）一同开发了应用于大数据和快速数据处理需求的创新存储介质与架构；再如，HP 与橡树岭国家实验室一同开发了应用于超算环境下的数据管理和存储解决方案，这些实验室均不离美国政府的资助。

图 3-1 对中国政府建设先进数据存力产业的建议





## 1 “定目标”

- 将闪存化率、数据存留率区域级先进数据存力指标体系纳入政府产业发展目标，确保各级政府  
政府对先进数据存力产业的重要性得到充分认知、投入必要资源

## 2 “巧引导”

- 监管性政策上，建议对标德国等强应用型国家的数据分级管理政策，完善数据敏感度分级与  
保存周期政策、强化企业对数据备份的重视
- 鼓励性政策上，建议强化对长江存储等NAND Flash晶圆厂定向补贴，早日实现国产存储芯  
片晶圆厂在先进SSD颗粒的研发、量产与规模化应用上“领跑全球”

## 3 “促应用”

- 建议积极促进以全闪化为核心先进数据存力在千行百业落地，在金融、电信、政务等关键行  
业“先试先行”
- 组建包含存储整机厂商在内的先进数据存力产业联盟，共同制定先进数据存力相关产业标准
- 鼓励国产开源数据库等开源软件生态发展、拓展先进数据存力解决方案的应用空间

## 4 “强创新”

- 资助并以产学研联合实验室为核心载体，提前布局超算存储等N+1代先进数据存力解决方案





## 附录 1：参考文献

1. 美国联邦调查局发布 IC3 2023 年年度报告
2. Sophos 网络集中安全管理发布 The State of Ransomware
3. 360 安全中心发布 2023 年勒索软件流行态势报告
4. Gartner 存储报告
5. Germany | Data Center Market Overview | Cloudscene
6. TechTarget Data Center
7. 应急管理部发布全国自然灾害基本情况
8. 应急管理部发布 2022 年全国自然灾害基本情况\_中国政府网
9. 中华人民共和国最高人民检察院 (spp.gov.cn)
10. Reports on world crime trends (unodc.org)
11. 国家发展和改革委员会 (ndrc.gov.cn)
12. UN E-Government Survey 2022
13. 华语医疗智库 (cn-healthcare.com)
14. 中共中央国务院印发《“健康中国 2030”规划纲要》
15. 中国大陆食源性疾病暴发监测资料分析 (zgspws.com)
16. 世界经济论坛 (weforum.org)
17. 国际劳工组织数据库
18. 世卫组织数据库
19. 网易数读
20. 国际货币基金组织：人工智能的治愈能力 (imf.org)
21. 医政医管局完善院前医疗急救服务指导意见
22. 北京采编中心
23. 中国科技网经济参考报
24. 上海科学技术委员会

25. 央视网：发展新质生产力 生物制造 制造万物
26. Gartner 存储报告
27. 世界银行
28. IDC & Gartner
29. Patsnap Database
30. Commercial Code (Handelsgesetzbuch – HGB) (gesetze-im-internet.de)
31. Industry 4.0: Innovative Flash Solutions for Worldwide Competition | Pure Storage
32. 新加坡全球贸易商计划指南 | 文章 - 汇丰机汇 (hsbc.com)
33. 新加坡经济发展局 (edb.gov.sg)
34. SSG | TechSkills Accelerator (TeSA) (skillsfuture.gov.sg)
35. 韩国政府发布政策
36. 全球芯片赛跑：韩国出台《K-芯片法案》
37. Nvidia GTC 2024