



2024 版

数据存储 2030



前言

人类社会的文明史，就是一部信息存储方式和传播方式变革的历史。3500 多年前，甲骨文出现，标志着人类的文明史开启；2100 多年前，造纸术的发明，使得知识的传播更加便捷；60 多年前，以机械硬盘为代表的数字化信息记录方式的出现，使得人们可以更加高效地存储和传播信息，进一步促进了人类文明的发展和传承。

未来十年，随着以 5G/6G、AI、大数据、云计算为代表的新技术飞速发展，人类即将进入 YB 数据时代，数据存储技术的创新和发展有望开启新的文明发展时代。以数据为中心的高效、绿色和安全的数据基础设施，必将推动人类社会向更高层次的智能化进程迈进，让人类在未来的智能时代中更好地理解世界、探索世界，并勇往直前，开拓未来。

目录

01

未来数据存储场景

06

1.1 数字化技术驱动人类发展，从信息化进入数字化	08
1.1.1 医，让健康数据化，让生命有质量	08
1.1.2 食，用数据换产量，普惠绿色饮食	09
1.1.3 住，全屋智能数据交互，让空间更人性化	11
1.1.4 行，数据使能智能低碳出行，开启移动第三空间	12
1.1.5 城市：数字新基建，让城市有温度，更宜居	13
1.1.6 企业：数字化工厂重塑生产模式，增强企业韧性	15
1.1.7 能源：数据使能绿色能源，打造低碳数据中心	17
1.1.8 数字可信：数据安全应用塑造可信未来	19
1.2 数字经济的发展推动人类社会即将进入 YB 时代	20
1.2.1 数据的总量从 175ZB 到 1003ZB，将进入 YB 时代	20
1.2.2 多样化数据应用，产生多样化数据类型	21
1.2.3 AI 促进数据觉醒，开启数据黄金时代，激活多模态数据价值	22
1.2.4 云和互联网的数据激增，带来数据架构变化，产生长记忆存储新范式	22
1.2.5 端边云产生数据 70% 向数据中心集中，大规模集约化数据中心形成	22
1.2.6 2030 年数据的产生主要来自终端，未来边缘和数据中心产生的比例将增加	23

02

数据存储 2030 愿景及关键特征

24

2.1 先进介质应用	26
2.1.1 先进介质技术	27
2.1.2 介质应用创新	30
2.2 以数据为中心的体系架构	33
2.2.1 存算分离，以存强算	34
2.2.2 存算一体，以查代算	35
2.2.3 集群存储	36
2.3 数据内生安全	37
2.3.1 主动数据保护	37
2.3.2 数据零拷贝	38
2.3.3 零信任存储	39

2.4 智能数据编织	41
2.4.1 自动化数据编排	41
2.4.2 跨域数据协同	42
2.4.3 存力网络	43
2.5 数据智能	44
2.5.1 内容消费的服务型接口	45
2.5.2 数据语义提取	45
2.5.3 数据多模态分析	45
2.5.4 数据自适应建模	46
2.6 可持续存储	47
2.6.1 存储系统级节能	47
2.6.2 数据传输能效提升	49
2.6.3 芯片级节能技术	49
2.6.4 绿色集约标准	50

03

数据存储 2030 倡议

51

附录 A: 参考文献	52
附录 B: 缩略语	54
附录 C: 致谢	56



未来数据存储场景



图 1-1 未来数据存储场景



10年前，人类社会刚刚进入 ZB 数据时代，移动互联网、云计算、大数据刚刚起步；今天，这些技术已经深刻地改变人类社会，而人工智能、区块链、5G/6G、AR/VR、元宇宙等新技术，进一步推动社会进入一个新的智能世界阶段。

2030年，人类将迎来 YB 数据时代^[1]，对比 2020 年，年新增数据增长 23 倍，通用存力增长 10 倍、人工智能存力增长 500 倍^[2]。数字世界和物理世界无缝融合，人与机器实现感知、情感的双向交互；人工智能无所不及，成为科学家的显微镜与望远镜，让我们的认知跨越微小的夸克到广袤的宇宙，千行万业从数字化走向智能化；数字技术持续演进，帮助人类利用数字手段加速实现未来智能世界 2030。

未来十年，数字技术将帮助人类跨入智能世界，这是一个波澜壮阔的史诗进程，将开启一个与大航海时代、工业革命时代、宇航时代等具有同样历史地位的新时代。

■ 1.1 数字化技术驱动人类发展，从信息化进入数字化

医，让健康数据化，让生命有质量

过去的十年，是人类健康发展充满“生命力”的十年，据 WHO (world health statistic 2021) 报告显示，全球人口预期寿命从 2000 年的 66.8 岁增加到 2019 年的 73.3 岁。伴随着人口老龄化的加速，2030 年全球 60 岁以上人口占比将达到 16.5%，人类对医疗的需求也将激增^[3]。WHO 的 2019 数据显示，全球卫生费用支出占生产总值的 10%，其增长快于全球经济增长。WHO 也预测，2030 年全球护士缺口高达 570 万人，医护人员总缺口高达 1000 万人。全球的医疗资源与人口增长分布形成剪刀差。

面向未来，如何降低医疗成本，丰富医疗资源和医疗形态，创造新的预防和治疗手段，将帮助解决看病贵，看病难的问题，让人们少生病，让生命更有质量。在下一个十年，创新的解决思路正在不断涌现。通过对健康状态实时数据追踪和建模，把防病与日常生活习惯结合起来，从“治已病”转向“治未病”，包括诸如下列场景：

构建知识图谱，让健康更靠谱

得益于互联网、物联网、AI 等技术的发展，以及可穿戴设备、家用监测设备等产品的普及，对个人健康建模不再是奢望^[4]。通过实时分析用户身体指标数据、医学临床反应、健康诊疗结果等，形成健康知识图谱；通过对比分析，为用户提供定制的健康解决方案；通过对营养、运动、睡眠等维度的干预，帮助用户逐渐改善不良生活方式，促进个人形成健康的生活习惯。比如业界有公司尝试构建饮食和疾病之间关系的健康知识图谱，可帮助用户改善睡眠质量，进行有效的体重管理，全年接受健康管理的参与者平均每天睡眠增加 35 分钟，体重减轻约 1.5 公斤，从而降低因不良生活方式导致的相关疾病的发生概率。

传染病蔓延轨迹预测，让疾病预报更准确

利用自然语言处理等技术，持续收集并分析全球范围内关于重大公共卫生事件的新闻、报告和搜索引擎指数，从中提取有效数据，并进行科学建模和智能化判断分析，可以有效提升应对公共卫生事件的响应速度和决策能力。比如说业界有公司使用自然语言处理和机器学习，从官方公共卫生组织、数字媒体、全球航空公司票务数据、牲畜健康报告和人口统计声明等多种公共数据来源中，分析了数十亿个数据点，可以 24 小时不间断地分析疾病的传播与蔓延情况。

药效精确评估，从“千人一药”到“千人千药”

AI 通过学习成千上万的病理诊疗方案，并实时分析结合病人个体综合差异，可以帮助医生给出更具个性化的治疗方案。新加坡研究机构创建了以人工智能技术驱动的药效精准评估平台，该平台可以快速识别每位患者的历史临床数据，针对患者自身情况给出建议的用药剂量和联合用药方案，并在此基础上对肿瘤大小或肿瘤生物标志物水平进行修正。此外，这些数据还可用于患者疗程和后续治疗方案的制定等。

AI 精准识别靶区，减少错杀健康细胞

个性化的精准医疗的价值，也在帮助对抗人类的天敌 - 癌症上。在传统的癌症放射治疗过程中，放疗的靶区设定范围较大，消灭癌细胞的同时也误伤了大量健康细胞。自适应放疗借助 AI 技术，在放疗过程中自动识别病灶位置变化，对放疗靶区的影像进行精密地勾画，以实现精准照射，从而减少对健康组织的损害。目前 AI 精准识别靶区，已实现 CT、超声、MRI 等多种影像的靶区自动勾画，将原来 2-3 小时的勾画环节缩短至分秒级，使得放疗对健康组织的损害量降低 30%。

面向 2030 年，人类可以依托高度灵敏的生物传感器技术与智能硬件支持，实时跟踪身体各项指标，并建立个人的健康知识突破，从而实现自主驱动个人健康，减少对医生的依赖。

在 ICT 技术的驱动下，通过精密的软硬件、强大的云边端计算能力和稳定的网络覆盖，使得便携化的医疗设备可以普及到在各个基层医院、社区、甚至家庭等多种场景，可以按需实时采集医疗数据并上传至云端处理中心，在云端构

建大数据知识库，通过 AI 调度，实现远程医疗联动。构建云端知识图谱，需要数据存力的规模部署，以容纳更多知识集合。

华为预测：到 2030 年，全球通用存力总量将达到 37ZB，相比 2020 年增长 10 倍；AI 相关存力总量占比 63%，相比 2020 年增长 500 倍。



食，用数据换产量，普惠绿色饮食

民以食为天，实现“零饥饿”被联合国列入 2030 可持续发展的目标之一^[5]。据统计，至今全球仍有超过 6.9 亿人在挨饿，预计到 2030 年，受饥饿影响的人数将超过 8.4 亿。农业从业者

长期流失：根据国际劳工组织的数据，在全球范围内，从事农业工作的人的比例从 1991 年的 43.699% 下降到 2019 年的 26.757%。人均耕地面积减少：据世界银行数据显示，在 1968-

2018 这 50 年间，全球人均耕地已从 0.323 公顷下降至 0.184 公顷，下降 43%。土壤农药污染严重：据统计，目前全球 64% 的农业土地（大约 2450 万平方公里）面临着农药污染的风险，其中 31% 的土地面临着高风险。与此同时，随着消费的升级，人们对于饮食的需求越来越追求吃得健康，吃得放心。2018 年，中国获得食品行业绿色认证的产品数量达到 13,316 个，2019 年，这一数量增至 14,699 个，同比增长 10.4% 绿色认证产品的背后是对种植环境和技术更高的要求。

在迈向 2030 年的进程中，科技和数据正在为农业赋能，帮助突破种植条件的限制，全面提升粮食的产量，让绿色食品进入每个普通人的餐桌。包括诸如下列场景：

用精准的数据，让种庄稼不再只靠经验

正所谓“栽种有时，收获有时”（a time to plant and a time to pluck up that which is planted）。传统农业何时播种，何时施肥，何时除虫，仅靠经验来判断，会让农业生产有着极大不确定性和产生诸多浪费。

ICT 技术赋能农业能够通过对土壤湿度、环境温度、作物状况、地形的特征、气候预期、病虫害程度等分析，获得精准数据，通过精准控制，让土壤和作物处于最佳匹配状态。以玉米为例，仅依据数据进行的自适应播种这一改变，就能带来每公顷 300-600 公斤的增产。

农场数字工厂化，让农业生产不再受自然环境的影响

农业工厂化的一个典型案例就是在室内种植的“垂直农场”，即用数据构建突破地域限制的标准化生长环境。在垂直农场里，每个环节通过对光照、温度、用水和营养输送等的精确控制，为农作物构建起最为适宜的生长环境。垂直农场无需农药，无需土壤，减低对农业用水的浪费；不受环境气候影响，始终确保新鲜农产品的理想生长条件；创造全球可复制的智能农业模式，利用同一套 ICT 控制系统和数据模型，可在世界上任何一个地方得到几乎一致的生产效果。业界公司的尝试显示，在 7,000 平方米的空间里，可实现蔬菜每 16 天收割一次，达到每年 90 万公斤的惊人产量。

面向 2030 年，我们通过 ICT 技术将更多的农田、农具、农作物等关键农业生产要素联接起来，收集并综合利用气候、土壤、农作物生长状态等多类数据，通过精准的数据分析，利用类似“垂直农场”这样的新种植模式，实现精准的农事操作，以提升粮食产量。

华为预测：到 2030 年，全球每年产生的数据总量达 1YB，相比 2020 年，增长 23 倍。未来随着数据不断在农业中体现，我们将逐步构建一个更有弹性、更绿色的粮食系统。全球农业每年产生的数据总量达 4ZB，相比 2020 年，增长 23 倍。



住，全屋智能数据交互，让空间更人性化

随着人们对居家体验个性化追求的不断增长，基于 ICT 技术的智能家居概念正被普及。据调研报告显示，近 80% 的千禧一代和 69.2% 的婴儿潮一代都对智能家居技术抱有积极的期待^[6]。在英国，目前 80% 的消费者已经意识到智能家居技术，在消费者对技术趋势的认知度上仅次于移动支付，而互操作性已经成为他们当下最大的购买考虑因素。除此之外，对便利和安全的需求也驱动着人们对智能化空间的向往。

数字化和数据助力打造未来居家体验，包括诸如下列场景：

打造数字化的商品目录，通过自动配送，实现储住分离

随着物联网、万兆光纤等新型基础设施的触达，越多越多的新型社区理念不断涌现，为居民提供如社区虚拟团建、宠物智能管控等全局化的服务，促进居民与社区的一体化融合。其中有一些新颖的设计理念，解决了储物与居住之间的矛盾，带给人们更清爽的居家体验。

为你家里的物品建立一个数字目录，甚至进行 3D 扫描，将不常用的物品寄存在小区统一的仓库中。比如在某个周末，你需要为即将参加的派对挑选一套晚礼服的时候，可以通过全息投影的方式，虚拟选择一套合适的搭配。只需轻轻一个点击，小区自动配送系统，就会通过机器人 10 或者楼宇输送系统快速地将你所选择的衣物送上门。

全屋智能结合场景式交互，打造亲切自然的居家体验

丰富的智能家居设备和传感器，通过稳定可靠、高联接、高速全覆盖的网络，将收集到的数据传递到家庭智慧大脑。其中的 AI 引擎，通过调



节各类家居设备的运行和协同状态，以匹配用户的实时体验需求，最终给用户带来沉浸式、个性化、可成长的全场景智慧体验。多种多样的智能家居设备，通过不同的组合，形成多样化的智能场景。比如智能睡眠辅助系统，根据个体的生理健康特征和睡眠习惯，自动匹配床垫和枕头的软硬度；营造助眠的光环境，刺激褪黑素分泌；播放助眠音乐，舒缓心情；根据家庭环境中湿度、温度、氧气的浓度等指标，提供恒温、恒湿、恒净、恒氧的睡眠环境。

2030 年，人们的家中将遍布各类智能家居，生活、娱乐将被新的交互模式来重塑；楼宇将安装各种智能管控设备；社区也将拓展更丰富的智慧功能。而这一切都需要通过大带宽的联接来提供没有时延的居住体验。

华为预测：到 2030 年，全球万兆家庭宽带渗透率达 25%，全球智能家居户数达 18 亿，年数据量达 23ZB。

行，数据使能智能低碳出行，开启移动第三空间

当下私家车出行成为人类活动的重要组成环节，2020 年全美车辆行驶里程为 2.83 万亿英里；在欧洲，每辆车每年行驶平均距离超过 1 万 2 千公里。当前的交通系统面临着诸多挑战：交通变得拥堵，全球碳排放占比 26%^[7]。

ICT 技术和出行要素（车、信号灯、行人等）联接起来，通过大数据提供决策支撑，从而实现出行变得更加智能和低碳。包括下列未来场景：

自动驾驶汽车驶入“快车道”

随着自动驾驶汽车由 L2、L3 向 L4、L5 迈进，公交车、出租汽车、低速物流、垂直行业运输（物流车、矿车）或将率先实现自动驾驶商业化。

低速开放道路：自动驾驶汽车在物流配送、清洁消杀、巡逻等领域取得了积极的成果。无人物流配送具备道路场景简单、车速低、危险性小的优势，可以在公共道路提供安全的无人物流配送服务。低速无人驾驶小车在抗击疫情中为医疗物资运输配送、清洁消杀、巡逻测温等工作提供支撑。

高速半封闭道路：重卡卡车司机成本高、易超负荷运载、超工时工作，因此重卡的自动化驾驶能够迅速帮助行业降低成本，提高效率，易于形成立竿见影商业收益。据德勤《中国智慧物流发展报告》预测，无人卡车、人工智能等技术在未来十年左右逐步成熟，将广泛应用于仓储、运输、配送、末端等各个环节。

特殊封闭道路：在矿山，港口等环境中，自动驾驶提升安全与效率，创造经济价值。在自动驾驶模式下，矿卡、挖掘机、推土机等多种机械工程车辆协同作业，一旦发生故障或者危险时，指挥人员可在控制中心开启远程接管模式，将车辆移至安全区域。在中国上海洋山港，“5G+L4 级智能驾驶重卡”车速最高达到每小时 80 公里，队列行驶间距缩短至 15 米。基于

北斗系统厘米级定位，车辆在 15 秒内可实现一次误差仅为 3 厘米的精准停车，单点装卸效率提升了 10%。

日常开放全新体验：自动驾驶出租车（Robotaxi）是自动驾驶公司服务出行的必然选择，据调研报告显示，Robotaxi 可取代 63% 的网约车 / 出租车和 27% 的公共交通。未来，自动驾驶技术将推动传统车革新，打造迎合不同场景的移动第三空间，甚至会颠覆现有行业的商业模式。比如自动驾驶餐车可能是未来的标配，你和亲朋好友的聚餐可能是以全新的形式展开：预定好一顿午餐，自动驾驶餐车会准时把你们依次接上，根据需求规划好一条风景优美的行驶路线，在欣赏美景的同时，品尝美食，畅聊人生，打造真正属于你们包间。这样既避免了往返餐厅的交通，又保证了就餐期间的私密性。

城市空中交通

未来，空域是城市交通发展的重要资源，可以搭建高效的空中城市交通网络，将极大程度的释放路网资源，减少市民的出行时间，提高城市的物流效率和应急救援能力。

空中应急救援系统：在过去的十年间（2010-2020），摩天大楼如雨后春笋般涌现于全球各大城市，增添了安全隐患。高楼消防、高楼医疗救援成为未来城市的新难题之一。空中应急救援系统的出现，使得消防和医疗救援力量能够快速到达高楼层实施灭火和人员救助，保障居民生命财产安全，成为摩天都市消防、医疗隐患的新解。

空中巴士 / 空中出租车：便捷、高效的交通体验已经成为都市人的核心需求之一，eVOTL 有望成为改善市内交通体验的利器，多家公司的四座飞行器都可以达到 100 公里左右的巡航里程。目前，空中客运试点已经展开，2019 年，该领域的中国科技公司，在浙江启动了全球首

个城市空中交通客运服务，将原本需要 40 分钟的道路交通行程缩短为 5 分钟的空中之旅。为实现城市空中客运（UAM）这类的未来场景，需要高速稳定的空天地一体化网络连接和定位系统、低成本可靠的视觉传感器和激光雷达、安全稳定的自动飞行算法、以及高效实时的指挥调度平台。

未来出行是一个多维的创新系统，通过电气化、自主化、共享化、网联化打造一个智能便捷低碳的出行体验，重塑出行体验，孵化创新的出

行服务，提升交通工具的共享效率，帮助缓解交通拥堵，降低出行带来环境污染，让不断激增的出行需求和环境对低碳的追求不再是一个矛盾体。

华为预测：到 2030 年，全球电动汽车占所销售汽车总量的比例达 82%，中国自动驾驶新车渗透率达 30% 以上，整车存力超过 500PB。

城市：数字新基建，让城市有温度，更宜居

5G、云、AI、区块链、智能传感等各种新技术的快速进步，给未来智慧城市的发展带来了更多新的可能，城市场景也将成为各种新技术的最佳应用创新场所与孵化基地。2020 年，全球投入试点的智慧城市数量将近 1000 个。2020 年相关投资接近 1240 亿美金，同比增长 18.9%^[8]。

城市的数字化，智能化已成为全球领先城市探索城市可持续发展的最关键路径，数字化和数据支撑诸如下列未来场景：

纳米传感，精准感知城市脉搏

城市数字化发展的基础是数据，而数据则来源于遍布在城市各个角落，各种各样的传感装置，在所有的传感技术之中，一种低成本、微型化的纳米传感器技术有望成为推动新一轮传感技术革命的“颠覆性”技术，美国《麻省理工学院技术评论》杂志把这种基于传感器技术的“感知城市”列为 2018 年全球十大突破性技术之一。

石墨烯纳米气敏传感器：这是一种对气味非常敏感的传感器。美国一所大学研制成功利用石

墨烯开发出新型的纳米涂层，他们将这种纳米薄膜集成到气敏传感器的电路中，与目前最好的使用碳基材料的传感器相比，对分子响应提高了 100 倍。未来传感器就能准确识别出空气中的有害气体，有毒气体，爆炸物等，从而大大提升城市对于危险物的感知能力。

纳米缝隙传感器：是一种能够识别特定频段声音的传感器，纳米裂缝传感器的表现大大优于传统传声器，能够将特定频段音源准确地识别出来。当把纳米传感器放置在小提琴的表面，它能够精确的记录乐曲中的每一个音符，并且将其“翻译”给外接设备，输出电子乐曲。当把纳米裂缝传感器佩戴在手腕处，它甚至能精确地测量人体的心跳。可以预见，这种技术的突破未来将大大加强城市对于声音的感知能力。

全光信息交换，开启万兆互联时代

城市数字化转型对海量的信息交换提出挑战，万兆互联的全光城市初步展现出了巨大发展潜力与价值。2021 年 4 月，中国上海发布了“全光智慧城市全球第一城”以 F5G 光网为底座，构建城市“1 毫秒”时延圈，实现全市光高速

枢纽布局，为后续城市智慧化发展打下了一个坚实的网络基础。未来的全光城市目标架构将包含四个组成部分：

全光接入：光联接延伸至家庭、楼宇、企业、5G 基站等城市全场景。全光传输向大型企业、楼宇、5G 基站等末端延伸，支撑各行业数字化转型，赋能 F5G+X，5G2B 等行业应用扩展。

全光锚点：家庭宽带、政企、5G、数据中心等业务的汇接点，由全光网统一传送；实现多技术协同，支持各类业务的一跳入云。

全光交换：城市光网一跳直达。通过全光交叉等技术，打造立体化的全光网络，实现一跳直达、云间高速、云光协同等。

全自动运维：实时感知网络动态，主动运维，并能够进行预测性运维，从而实现网络资源弹性化，业务自动化、资源分配自动化，运维自动化。

智慧中枢，城市从人治走向 AI 治理

随着城市全量数据的打通，融合，AI 从局部的智能走向全场景的智慧，催生新的公共治理主体，未来的城市都需要一个强大的智慧中枢平台，它一方面汇聚来自于城市各个角落的海量数据，另一方面通过平台把数据转变成一种城市治理的先进能力，普惠千行百业，极大提升城市治理效率与用户服务体验。日本丰田公司的早期探索：在丰田未来城市的规划中，每个房屋、建筑、车辆都配备有相应的传感器，这些数据会汇聚到一个城市的智慧中枢平台，由 AI 分析人们所处环境状况，通过人车分流，确保道路上车辆与行人之间的绝对安全。

基于数据的主动精准政务服务

机器识别技术的出现使得非接触服务成为可能，今天在中国大多数发达省份，政务办理已不再需要去政府的服务大厅，通过手机就能够进行



远程的自助服务，可以预见未来十年政务服务的数字化，智能化程度将会进入到更高的发展阶段。

未来随着海量数据的不断积累与汇聚，人工智能技术的不断成熟，政务服务也必将会更多地向主动服务，精准服务的方向发展，大幅提升城市治理效率与市民的服务体验。以智慧养老为例：上海的街道推行给独居的老人安装智能水表。在老人的同意下，通过实时监测独居老人的用水情况，12 小时内用水量一旦低于 0.01 立方米，街道的“一网统管”平台，就会接收到报警信息，并及时通知社区，社区志愿者就会第一时间上门查看老人的情况，通过这些智能设备的使用，使得社区对于独居老人的关怀做到细微之处，给老人的生活带来温暖。

下一个十年，将是 5G，光，AI，云，区块链，智能传感等 ICT 技术快速发展的十年，城市将会进入到万兆联接的时代：万兆的企业接入，万兆的家庭宽带接入，万兆的个人无线接入体验。

城市与 ICT 技术的结合与聚变必将会在未来产生巨大的裂变效应，大幅提升城市资源的利用率，治理的效率，用户的体验，从而真正实现城市的可持续发展目标，让城市更有温度，更宜居。

华为预测：到 2030 年，城市带来的数据占比达到 96%，和城市基础设施相关的资源监控、调度、管理产生的数据达到 42%。

企业：数字化工厂重塑生产模式，增强企业韧性

未来十年，人口老龄化导致世界出现巨大的劳动力缺口。联合国报告显示，2030年65岁以上人口比例将超过12%，25岁以下人口占比从2020年的41%，下降至2030年的39%，全球劳动力短缺超过8,520万人。以制造业为例，到2030年，全球制造业面临790万工人的短缺，影响实现产值6071.4亿美元^[9]。

同时，消费需求的多样化也在影响着生产模式的变化，倒逼企业进行生产模式的革新，如基于“一人经济”的发展，快速调整产品形态，推出一人食套餐、迷你家电，甚至迷你KTV等；还要能从情感维度主动激发消费者的购买欲望，对产品的外表、形象、含义进行快速的组合设计，如在短期内定制出各类限量款或联名款。

此外，黑天鹅事件也在对企业生产的延续性提出了新的挑战。如新冠疫情原因，2020年全球GDP损失了近3.94万亿美元的产出，供应链中断是企业增长的最大风险。为此，如何利用数据，保护数据，重塑生产模式，增强产业链韧性的也成为尤为重要问题，包括诸如下列场景：

协作机器人

越来越多的企业受制于劳动力短缺的问题，需要企业通过新生产力来迅速补位。协作机器人是工业机器人的一种，最初目的是满足中小企业的定制化和柔性制造需求，成为弥补劳动力短缺的重要补充力量。相比传统的工业机器人，协作机器人更适合干人不想干的工作，比如分类，包装，挑拣等高重复性的工作。协作机器人有几个优势：

更安全：协作机器人更加轻巧智能，携带的传感器可以确保它一触即停，可与产线上的工作人员亲密合作，共同完成任务；

更快速灵活地部署：协作机器人通过人性化的编程，如拖动画示教，自然语言和视觉指导，可

以随时投放在新的岗位上，快速完成编程和调试，迅速执行任务；

更低的TCO，更短的ROI：协作机器人的售价和每年的维修成本远低于传统工业机器人，在过去几年，协作机器人的平均售价下降了一半；

目前协作机器人在3C和汽车等制造领域应用最为广泛，同时，我们也看到它在医疗化验和检测的应用崭露头角，帮助医务人员减少重复、费时的工作流程如做尿液分析，也可以降低工作人员的传染风险如咽拭子采样。

自主移动机器人

自主移动机器人（AMR）是制造业向柔性化、智能化发展的关键使能要素，改变企业的生产流程、仓储物流等重要环节。自主移动机器人，具备丰富的环境感知能力、基于现场的动态路径规划能力、灵活避障能力、全局定位能力等。工业制造及物流领域的自主移动机器人，目前主要基于SLAM（simultaneous localization and mapping，同步定位与地图构建）技术、利用激光导航、视觉导航以及卫星定位等技术，实现自主导航。自主移动机器人可实现产线物流的自动化与无人化，货物的智能拣选、搬运以及出入库等无人化场景。

数字仿真，柔性生产

为了能够适应多变的市场需求，在激烈的竞争中取得优势地位，企业必须更为积极地拥抱新的生产模式。柔性生产、柔性制造系统等概念正越来越受更多企业的青睐。柔性生产、柔性制造系统需要通过ICT技术进行拟实生产，包括运用仿真、建模、虚拟现实等技术，对新产品的生产制造全过程进行模拟，降低新品开发和设计的成本，更精准地规划生产线的调整成本和生产能力；同时柔性生产的智能任务调度系统会根据工厂的生产能力、订单复杂度和

交付时间需求，通过分析统筹安排生产任务的发放、生产物料和工具的调配，确保充分发挥出工厂中所有设备和人员的最大生产效率；柔性生产利用视觉编程、自然语言交互、行动捕获等 ICT 技术能力，快速实现对生产设备功能的重新自动编程和定义，以满足企业柔性化生产的需求；物流管理的柔性化利用 ICT 手段来有效地进行仓储和物流管理，避免漏发、发错、发混。以家具企业为例，大规模的定制化下，所产生的每一块板，装饰条，把手等都需要有一个属于它自己的识别码或 RFID，来协助自动化的打包和装车规划，以及运输和配送环节的全流程跟踪，真正实现以消费者的需求为中心的智能定制化生产模式。

打造有韧性的智能供应系统，帮助企业应对突发性危机

越来越多的企业将打造一个有韧性，智能的供应链作为其最重要的战略布局之一。供应链可视化利用 ICT 技术，采集、传递、存储、分析供应链中的上下游订单、物流以及库存等相关指标信息，以图形化的方式展现出来。供应链可视化可以有效提高整条供应链的透明度和可控性，从而大大降低供应链风险。对于上游供货，通过对物料、设备等的追踪，实时显示其整体交付的程度，包括包装、入库、出库、质检等工序的状况，甚至可以追溯其生产流程中的各种状态。帮助企业针对物流中可能出现突

发事件，及时调整物流路线，确保物资的准时、安全地到达目的地。

对仓库运营环境信息的实时监控，建立远程监控系统，通过各类传感器，用图像化呈现仓库的温度、湿度、灰尘、烟雾浓度等运维信息，一旦发生如火灾、漏水等前期征兆，可及时介入，避免物资的损失。对货物出入库信息的实时追踪，随着货物的流通，通过 IOT、RFID、二维码等技术，自动识别并登记物品的信息，可在远端实时调取货物仓储的状态数据。

面向 2030 年，数字化转型推动企业的进一步升级。利用人工智能、传感器、物联网、云计算、5G、AR/VR 等技术来打造新生产力，弥补劳动力缺口，帮助企业把握新的业务商机，拓展企业边界。

未来，通过对产品设计、任务分配、设备功能、物流配送等环节的柔性化重塑，实现以人为中心的新生产模式。供应链也将会在数字化的助力下，变得可视化、网状化，增强企业的韧性以应对变化万千的市场环境。

华为预测：到 2030 年，企业数字化转型将进一步推动数据服务在企业中的应用，数据服务的支出占比达 87%，支撑智能制造的 AI 存力占企业 IT 投资比例达 7%。



能源：数据使能绿色能源，打造低碳数据中心

2015年《巴黎协定》在第21届联合国气候变化大会上达成全球共识：将全球平均气温相比工业化前水平的增幅限制在远低于2℃，尽力将增幅限制在1.5℃水平，在本世纪下半叶实现人为排放量与清除量的平衡^[10]。2020年9月，中国在联合国大会上提出中国双碳目标：力争于2030年前二氧化碳排放达到峰值，并争取2060年前实现碳中和。实现全球的气候控制目标，需要从能源的供应、消费和固碳等多角度入手，全方位促进全球能源结构转型。

随着能源网络复杂性的提高和行业数字化的进程的发展，ICT技术成为脱碳解决方案的重要组成部分。提高新能源的比例、适应新的能源结构、充分的发挥ICT技术和数据，让绿色能源更智能，实现经济可持续发展，支撑包括如下场景：

海上风能，潜在的主力新能源

2020年，全球可再生能源装机容量新增45%，达到280GW，其中风能新增114GW，增长率达到90%以上。目前在欧洲部分国家，正在积极利用近海发电，其中英国和德国截至2020年海上风电装机容量超过18GW，占全球海上风电的51%。即便如此，海上风能当前只提供全球电量的0.3%，还有巨大的发展空间。

海上风况优于陆上，风速通常比沿岸陆上高出25%。同时海上风湍流强度小，具有稳定的主导风向，海上风机的容量可以达到陆地风机容量的3~4倍。海上很少有静风期，其发电时间往往能达到3000小时/年，更能有效利用风电机组的容量。而伴随着技术改进，海上风电的装机成本大幅降低，预计到2040年海上发电成本将比2019年下降60%。全球风能理事会（GWEC）预测，到2030年，全球海上风电装机量将从现在的29.1GW升至234GW。未来五年海上风电的增长率将达到31.5%，海上风电迎来快速发展时期。



漂浮光伏（FPV），光伏产业新趋势

据国际能源署（IEA）发布的《2020年全球光伏报告》，截止2020年底全球光伏累计装机容量达到760.4GW。2020年，光伏约占所有新增可再生能源总发电量的42%，其中陆上大型光伏电站一直是光伏产业的建站主要模式。但陆上光伏的发展也开始面临土地获取以及成本制约的问题，同时陆上光伏在高温情况下会出现效率下降，漂浮光伏成为新的部署模式。与陆基光伏相比，漂浮光伏不但可以节省用于农业用途的土地，而且相比路基遮阳障碍物更少，灰尘数量更少，同时自然冷却潜力也会提高光伏的性能。2020年荷兰乌得勒支大学的学者基于北海实际测试及研究论文表明，海上的漂浮光伏表面温度远低于陆基光伏，两个点的光伏面板表面温度差达到9.36摄氏度。全年发电量海上漂浮光伏比陆基光伏的年均产出能高出约12.96%。

随着技术的不断成熟，漂浮光伏将迎来快速发展时期。2021年7月14日世界最大的内陆漂浮光伏系统之一——新加坡胜科登格漂浮太阳能电站正式竣工投运，覆盖水面面积45公顷

(相当于约 45 个足球场)，覆盖水面上累计安装了 12.2 万块太阳能板，产能达 60 兆瓦。据 RethinkEnergy 预计，到 2030 年全球漂浮光伏的市场容量将超过 60GW。而漂浮光伏的全球潜力达到 400GW，足以将太阳能光伏的现有装机容量翻一番，随着技术的成熟，漂浮光伏的部署速度在加速，为可再生能源的全球打开了新的领域。

打造低碳数据中心与低碳网络，加速“碳中和”进程

据 IEA 研究报告显示，自 2010 年以来，全球互联网用户数量翻了一番，全球互联网流量增长了 12 倍，数据中心和传输网络的耗电大幅上升，2019 年全球数据中心电力需求约为 200TWh，约占全球最终电力需求的 0.8%；数据网络消耗约 250TWh，约占全球用电量的 1%，其中移动网络占三分之二。中国 2030 年数据中心用电预计将达突破 4000 亿千瓦时，占全社会用电量的比重将升至 3.7%。而 PUE 每优化 0.1，可节省用电 250 亿度，减少碳排放约千万吨，若全部使用绿电，碳排放每年可以减少 3.2 亿吨。引入绿电和自然冷却降低 PUE 成为低碳数据中心的关键举措。

除了引入可再生能源、自然冷却实现数据中心

高效、节能，另外一个重要手段就是人工智能的应用。通过数据中心内的传感器收集温度、电量、泵速、耗电率、设定值等各种数据，通过 AI 算法调整数据中心的运行模式和控制阈值，从而实现降本增效。将人工智能用于数据中心冷却，实现将用于冷却的能量减少 40%。据 DCD 的报告，欧盟 Horizon2020 资助的位于瑞典 BTDC 研究项目，在自然冷却的同时，通过人工智能算法实现冷却系统、IT 负载、服务器风扇和温度协同，PUE 达到 1.01 的最高水平。随着 AI 技术的进一步成熟，配合绿电引入、自然冷却段，数据中心和通信网络将会更加省电、高效，并最终真正实现零碳目标。

到 2030 年，世界需要将排放量减少一半，以风能、光伏为代表的新能源正加速部署实现生产侧清洁替代，消费侧通过电气化实现电能替代。ICT 作为一个行业，除了自身需要节能减排以外，同时也在赋能其它行业来减少碳排放。

华为预测：到 2030 年，数据中心用电占全社会用电量的比重将升至 3.7%，其中存储系统用电占比将达 25~32%，提升数据中心绿色能源比例对碳排放改善有着重要作用。



数字可信：数据安全应用塑造可信未来

在数字化转型的加速驱动下，组织与组织之间、组织与客户之间以及组织内部的互动从物理世界迁移至数字世界，由此而产生了宝贵数字资产，建立数字信任成为组织最重要的战略目标之一。数字信任是一个复杂庞大的系统，包含隐私、安全、身份、透明、数据完整性以及治理和合规等关键领域^[11]。新的技术应用，如区块链、隐私增强技术、人工智能等，以及新的规则定义，将塑造可信的数字未来。包括诸如下列场景：

基于区块链的智能合约

数字资产为组织和个人带来了史无前例的快捷和便利，但同时也带来被窃取和盗用的高风险。基于区块链技术的智能合约以数字化的形式将合约条款写入区块链中，合约事务的保存和状态处理都在区块链上完成，由于区块链的分布式特性，保障智能合约的存储、读取、执行整个过程透明可跟踪、且不可篡改。基于区块链技术的智能合约在物流、电子商务、金融保险等多个领域有着巨大的潜在市场应用价值。据咨询公司预测，智能合约能将美国的个人房贷成本降低 480-960 美元 / 年；在美国和欧洲，将银行房贷运营成本降低 30-110 亿美金 / 年，将个人车险费降低 45-90 美金 / 年；帮助全球汽车保险公司将保险理赔成本降低 210 亿美金 / 年。

建立新的互联网个人信息调动机制

近年来，针对过度收集数据的规则制定和诉讼探索在不断推进。在公平交易的数字战略中，大数据背景下的个人信息的调动机制将会变得更加平衡，兼顾隐私权利和个人信息开发两个目的，在传统告知同意原则的基础上，强调主体对于个人信息的控制权。2021 年，《个人信息保护法》正式发布，作为中国首部关于个人信息保护的专门法律，重申了个人信息保护工

作的多项基本原则，包括公开透明、目的明确、最小必要。未来，个人信息调动机制将从规则框架上继续细化，为用户明确数据收集的场景、用途及风险。

GDPR 是目前世界上最严格的针对个人数据的隐私和安全法，由欧盟起草通过，正式生效于 2018 年 5 月 25 日。2020 年，美国发布《联邦数据战略与 2020 年行动计划》，旨在保护数据完整性、确保流通数据真实性、数据存储安全性等基本原则。2020 年 5 月 27 日，日本参议院正式通过《数字平台交易透明化法案》，该法案旨在规制特定数字平台，增加特定数字平台的公开义务。由此可见，数据反垄断趋势正在全球蔓延。未来，在反垄断法的不断完善和应用中，用户和第三方企业将从行业巨头手中获得更多的数据主权，避免大平台对个人隐私数据进行非法地获取、滥用及交易等侵犯数字安全、破坏公平竞争的行为，从而促进数字信用生态的建设。

共建数字可信的智能世界，面向 2030 年，人类可以借助区块链、人工智能等技术更好地保护个人隐私和数字资产，更精准地打击假新闻等数字造假行为，减少诈骗或数据盗用的隐患。隐私增强计算等技术为多方实现安全加密的数据共享，在不影响隐私安全的前提下，确保数据价值的流通。

华为预测：到 2030 年，50% 以上的计算场景将采用隐私增强计算技术；85% 的企业将采用区块链技术。隐私增强计算、区块链技术以及 IPFS 的应用将极大增加不可压缩的加密数据以及分布式账本数据，每年新增数据量将达到 17ZB。超过 80% 以上企业部署包括存储系统在内的多层勒索病毒防护体系。

1.2 数字经济的发展推动人类社会即将进入 YB 时代

2030 年数字经济占比将达到 60%，数据成为数字产业发展和产业数字化的基础。

当今世界，科技革命和产业变革日新月异，数字经济蓬勃发展，深刻改变着人类生产生活方式，对各国经济社会发展、全球治理体系、人类文明进程影响深远。中国信通院《全球数字经济新图景（2020 年）》报告称，2019 年全球数字经济规模达到 31.8 万亿美元，约占全球 GDP 的 36%。数字经济保持快速增长，质量效益明显提升，数字经济增加值规模达到 35.8 万亿元，占国内生产总值（GDP）比重达到 36.2%，对 GDP 增长的贡献率为 67.7%。

全球 2030 年数字化经济的占比将达到 60%，数字产业本身的持续增长和传统产业的数字化转型在加速。数字产业到 2030 年的产值比重将达到 9%，成为经济增长的加速器；传统产业的

数字化转型，正在通过数字化的武器，指导传统产业更加互联网化、智能化、自动化，增加产业的客户渠道、减少成本、提升生产和服务效率。

2030 年产业的数字化渗透率达到 45%，数字化进程的深入让我们更加理解这个世界，推动人工智能和智能制造的发展。

在这个数据驱动的世界，人类、社会和地球资源的活动的正在越来越多的被以数据的形式观察、记录、监听、追踪和处理，这使得我们可以比以往更加精确的理解、以至于描述这个世界。数据的大量积累和机器学习的技术发展催生的人工智能的成熟，人工智能会给业务、流程、沟通带来更多的自动化，通过为客户个人偏好提供量身定制产品把效率和生产率提升到一个新的高度。

数据的总量从 175ZB 到 1003ZB，将进入 YB 时代

根据 IDC 和华为 GIV 团队预测，全球每年新产生的数据总量随着数字化的发展快速增长，从 2020 年每年产生 2ZB 到 2025 年每年产生 175ZB，2030 年将达到 1003ZB，即将进入 YB（1 Yotta Bytes = 1000 Zetta Bytes）时代。



图 1-2 全球每年新产生的数据总量预测



多样化数据应用，产生多样化数据类型

随着行业数字化深入，数据应用越来越多样化，除了传统的数据库应用，分布式数据库、大数据、HPC 高性能计算等新兴应用不断产生，当前平均一个企业的数据应用超过 100 种。

结构化数据激增促使核心系统弹性扩展：数字化、移动化技术发展使企业与客户交互渠道发生巨大变化，手机 APP 类互联网应用成为触发客户购买行为的最佳媒介。这固然引领了业务快速增长，但结构化数据的激增也给核心系统带来难以预料且波动巨大的业务浪涌。核心系统必须具备极强的资源弹性，以确保高峰期能够快速扩展以保障业务正常运行，而平时能够释放闲置资源避免浪费。同时多读多写会成为核心应用的标准配置，以保证系统极高的可靠性。

非结构化数据进入生产决策系统：非结构化数据，包括文本、图像、视频和音频等，因其丰富的信息内容和形式的多样性，成为企业数据资产的重要组成部分。**到 2030 年，全球每年产生的数据总量达 1YB，其中 80% 以上都是非结构化数据。**非结构化数据已经在企业得到广泛应用，以 AI 为例，56% 的企业至少把 AI 用于一个企业功能，其中大量场景是基于非结构化数据进行分析处理的。而随着企业数据治理能力提升，数据驱动业务增长成为可能，非结构化数据开始走向生产决策系统，比如金融行业的在线实时授信、医疗行业的病理分析等。**预计 2030 年 80% 的非结构数据将成为生产决策数据。**

AI 促进数据觉醒，开启数据黄金时代，激活多模态数据价值

AI 大模型已超出人类想象的速度，将我们带入智能世界。数据作为人工智能三要素之一，决定了 AI 智能的高度，其价值得到了前所未有的提升，全球步入一个数据觉醒的新时代。

首先，热数据会急速增多。据统计，2023 年我国新增数据存储量仅占全年数据生产总量的 2.9%，海量的数据在源头就被放弃，没有被存储。随着人工智能的能力不断增强，热数据的规模和重要性随之增长，大量数据不再丢弃而是存储下来为人工智能提供即时的、有价值的输入。**预计 2030 年这些热数据将 100% 由 SSD 闪存所承载。**

其次，温冷数据价值得到重新利用，温冷数据逐步变热。温冷数据是那些不经常访问的数据，如备份和归档数据，传统上被认为是价值较低的。但人工智能的全量数据训练需求，使得这些温冷数据重新获得了关注。通过将温冷数据纳入训练过程，AI 不仅能够提高模型的准确性和泛化能力，还能够发现之前未被注意到的数据价值。这些需要快速访问的温冷数据也被称之为活跃归档数据。**预计 2030 年 60% 以上的企业每日访问活跃归档数据至少一次以上。**

云和互联网的数据激增，带来数据架构变化，产生长记忆存储新范式

近年来，云计算和互联网技术的迅猛发展，为各行业带来了前所未有的变革。当前，云和互联网已成为数据存储需求增长最快的领域之一。据统计，企业级 SSD 盘约三分之二发往云和互联网厂商。

为了应对数据量的爆炸性增长和业务需求的快速变化，云和互联网厂商正在推动 Diskless 架构改造，如 Google Cloud 等。Diskless 架构通过

将服务器本地盘拉远，构建起由 Diskless 服务器和远端存储池组成的全新架构，实现了计算资源和存储资源的彻底解耦与灵活共享。这一变革不仅大幅提升了资源利用率与扩展性，还显著降低了运维复杂度和能耗成本。Diskless 架构凭借其灵活的架构设计和高效的存储资源管理能力，有望在 2030 成为主流架构，为云和互联网行业的持续发展提供强有力的支撑。**预计 2030 年超过 80% 的云和互联网企业采用 Diskless 架构。**

端边云产生数据 70% 向数据中心集中，大规模集约化数据中心形成

数据的来源于端边云，但数据的产生和数据的存储偏好并不相同。

端（Endpoint）：终端包含所有的在网络边缘的终端设备，包括 PC、手机、工业传感器、汽车、可穿戴设备等等。**预计 2030 年超过 75% 以上的端侧产生数据将被人工智能实时处理。**

边（Edge）：边缘是指用来处理企业级负载的服务器和设备，但其位置没有放在核心数据中

心，而是放在分支机构的 Server Room、工作场地、或是无线基站，以便靠近数据处理、减少网络延迟。**预计 2030 年超过 80% 以上的边侧数据将被人工智能实时处理。**

云（Core Data Center）：核心数据中心是指大规模的数据中心、包括企业数据中心、IDC、以及公有云厂商的云数据中心。**预计 2030 年超过 90% 以上的数据中心侧数据将被人工智能实时处理。**

2030 年数据的产生主要来自终端，未来边缘和数据中心产生的比例将增加

随着端设备的增加，2030 年数据产生依然主要来自终端设备。根据预测 2030 年智能汽车增长、可穿戴设备、工业 IoT 等大幅增长，端设备产生的数据将增长 14 倍，占比为 52%。

2030 年边缘设备也将大幅增长，其中 5G MEC、CDN、Robo、高新媒体处理器都将得到大幅应用；家庭数字化处理中心也在形成雏形，未来每个家庭有一个数字化处理中心，连接所有家庭数字或智能化终端，如手机、可穿戴设备、

智能冰箱等设备，存储数据、处理数据、帮助管理生活。到 2030 年边缘产生的数据增长 22 倍，占比达到 21%。

云(Core Data Center)是数据汇聚、处理、备份、复制、转移的关键节点，每一种操作都将产生新的数据，数据中心的对数据的操作有放大效应；未来随着数据中心数据汇聚的越多，放大效应越大。到2030年，数据中心产生的数据增长 18 倍、占比为 27%。

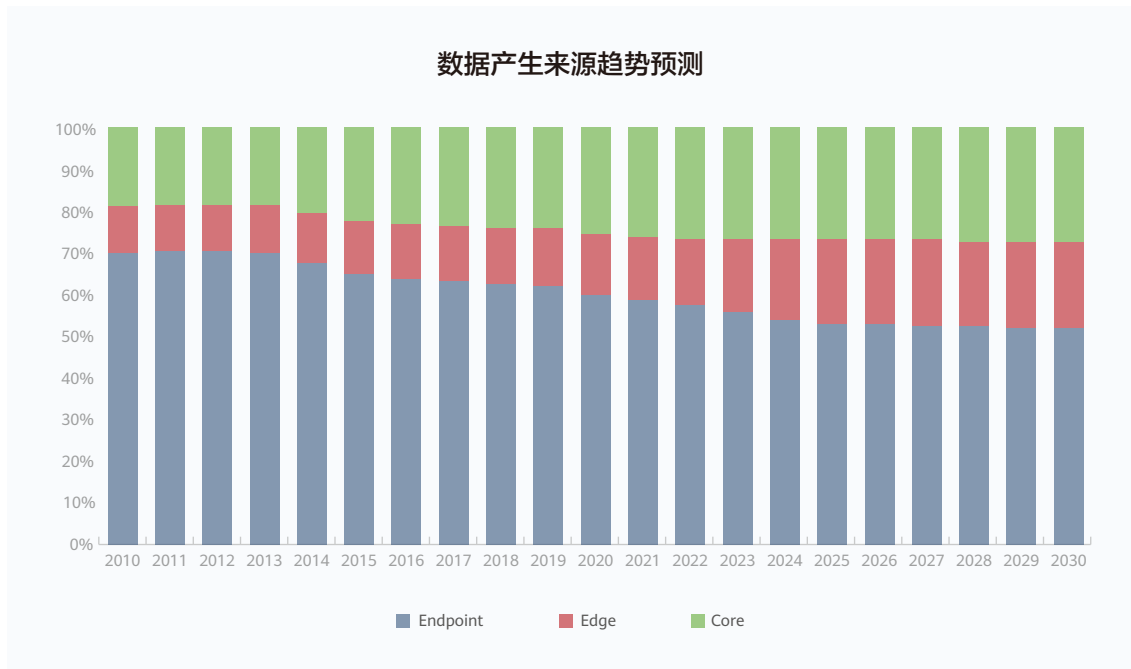


图 1-3 数据产生来源趋势预测

2030 年数据的存储位置进一步向数据中心集中，终端上产生的数据通过应用系统、备份系统很大一部分被保存在数据中心。随着网络建设和带宽的增加，数据保存在数据中心更加方便和安全，网盘、照片、账号、应用数据都在数据中心保存。以各种应用账号为例，不同客户端使用同一账户登录，通过数据中心保留的账号和状态数据，你可以体会到统一的视图服务。

到 2030 年，被存储下来的数据约有 65% 被存

放到数据中心，端设备无需长期保存数据，定期备份到数据中心是未来趋势。随着对实时处理、低时延要求的应用增长，边缘存储数据的场景也在增长，如智能驾驶训练终端、实时边缘流处理、5GMEC、VR/AR 边缘中心等，边缘处理数据比例在 2030 年达到 10%。

各种分散的数据向数据中心汇聚，让我们更加有可能对数据价值进行深度挖掘，成为数字化、智能化坚实的数据基础。



数据存储 2030 愿景及关键特征

未来 10 年，数据的年复合增长率接近 40%，数据类型呈现多样化，单一的存储介质难以满足多样化的数据存储需求，需要多元化的介质来应对存储成本高、功耗大、持久性差等挑战。海量多样化数据促进多元化先进介质和介质应用发展，结合智能化数据缩减和联合数据编码技术，存储容量密度将实现数倍提升。

高速增长的数据量与缓慢增长的数据处理能力已成为数据产业的基本矛盾，数据存力与数据发展严重失衡。经典的以 CPU 为核心的架构理念将难以支撑海量数据存储和处理的需求，需要以数据为中心来实现整个架构的重构^[12]。新体系架构宏观上存算分离，微观上存算一体，通过高通量、超低时延和高扩展的互联总线，打破资源的边界，形成处理器、内存和存储资源池化，以存补算，实现数据处理效率的数倍提升。

日益增长的数据流转需求与日趋严重的数据重力，已成为影响数据价值发挥的基本矛盾。智能

数据编织支持数据的跨域智能高效流动，打破空间的约束，实现数据的所见即所得，支撑数据流动效率百倍提升；数据内生安全使能数据使用权、管理权和归属权分离，促进数据的可信流转^[13]，通过主动防御构建安全可信的数据应用环境，保障数据隐私，实现数据可信流转效率千倍提升。

日益复杂的存储系统已无法满足新兴多云应用的智能化数据业务需求，需要数据业务逻辑与数据智能解耦。在未来的数据基础设施中，存储将具备数据感知、数据理解的新能力，支撑数据服务走向千行百业，数据业务百倍增增长。

持续增长的存储能耗，与全球低碳发展要求仍存在差距，对存储的绿色低碳能力提出了新的要求。新节能材料、以光代电和动态节能技术促进芯片节能，新型液冷散热和智能化的整机调控技术促进整机节能，系统级多维度、智能化资源调控技术从数据全生命周期上实现减排，未来将实现能效效率数倍提升，支撑未来数据产业的可持续发展。

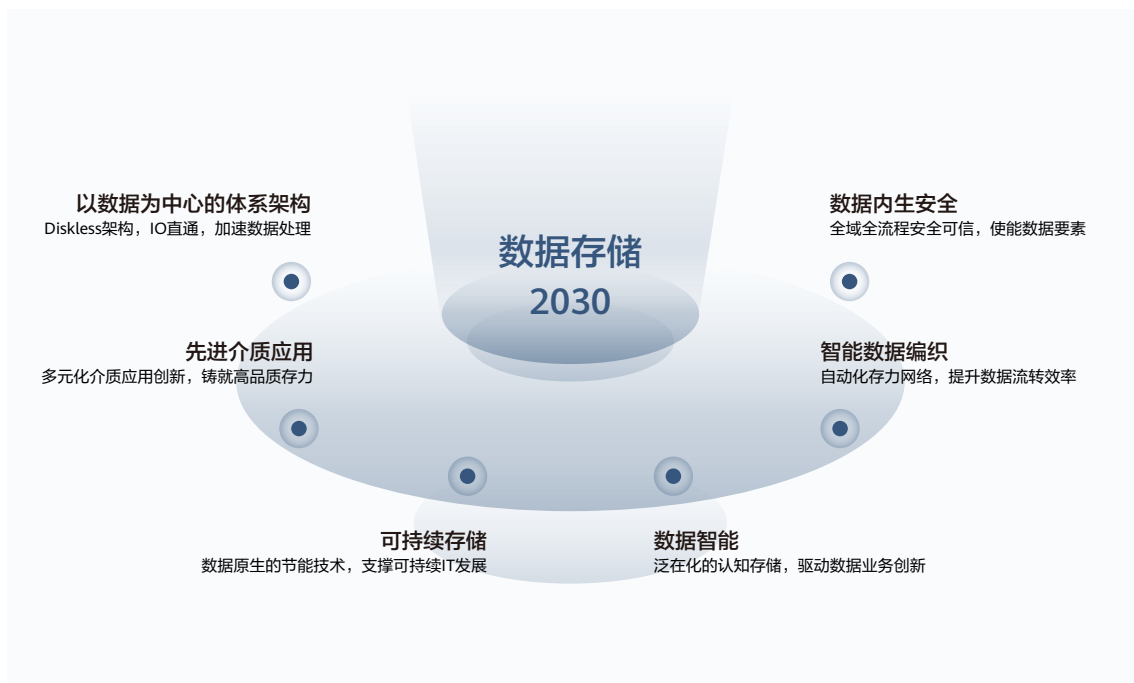


图 2-1 数据存储 2030 六大关键特征

综上所述，未来存储需要具备先进介质应用、以数据为中心的体系架构、数据内生安全、智能数据编织、数据智能、可持续存储共六大特征。

2.1 先进介质应用

随着 AI 大模型逐步走向多模态，数据将迎来觉醒。越来越多的视频、图像数据将被保存下来用于训练，预计到 2030 年，全球每年新增 1YB 的数据，用于大模型训练的数据量有望增长 1000 倍以上，达到 400EB，其中有接近 50ZB 的价值数据需要存储，相比 2020 年增长 23 倍，要求存储介质必须具备大容量、高性价比、低能耗，以及高可靠、高扩展、长寿耐用和高安全性，同时要求存储具有数据计算和分析能力，以便更快的获取数据。

鱼与熊掌不可兼得，不同的介质具有各自的优劣势，需要通过多种介质组合来应对挑战。根据不同介质的演进趋势，预计到 2030 年，介质容量密度有望提升 10 倍，但相比存储数据量 23 倍的增长仍存在较大差距，需要介质应用创新来填补这个差距。

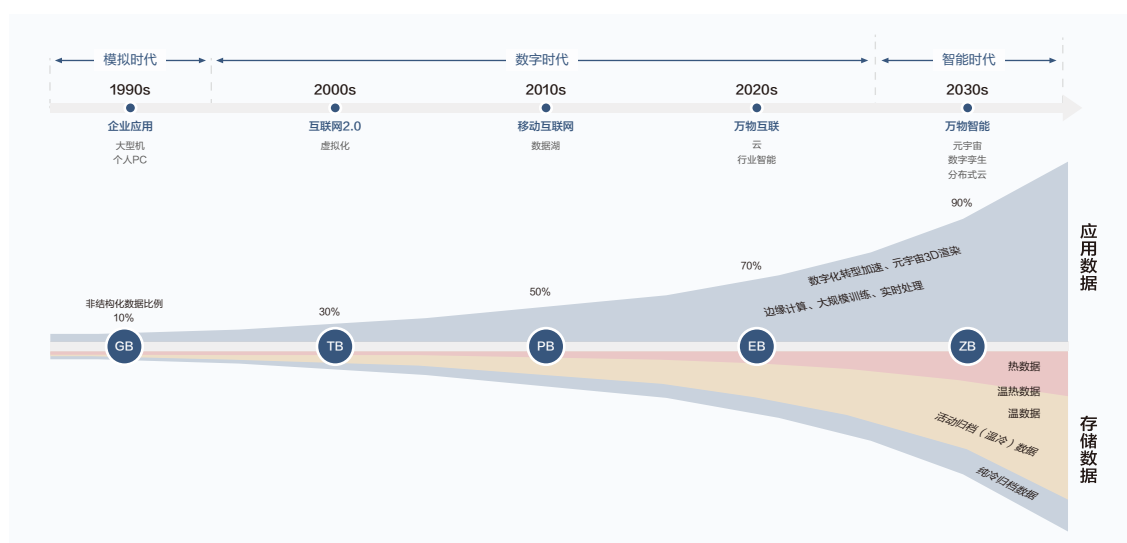


图 2-2 数据量增长趋势

不同的数据对存储介质的要求不同，按访问频率可将存储数据大致分为热、温、冷三类。

热数据：共占比约 30%。其中，AIoT、边缘计算、机器人和自动驾驶实时处理需要纳秒级数据访问能力，属于极热数据，占总存储容量的 1.5% 左右，需要极高性能的内存型介质；而银行、电子商务等在线交易类业务，以及 EDA 等工业制造类业务也需要频繁实时访问能力，属于一般热数据，容量将增长超过 35 倍，需要高性能的存储介质。

温数据：HPDA 等数据密集型业务需要对大量数据做分析，没有很高的访问频率和实时性要求，但数据量占比达到 60%，相比 2020 年预计增长超过 25 倍。这部分数据存储除了要求大容量介

质外，对成本、功耗也及其敏感，需要高性价比的存储介质。

冷数据：历史文献、国家档案及其他一些法律规定的需要长期存储的数据，平时极少访问，占比大约 10%，容量增长预计接近 20 倍。由于需要长期存储，可靠性要求变得更高，需要高可靠、长寿命介质。

当前，为了训练大模型，越来越多的冷数据将被激活变成温数据，温数据的占比有望超过 60%，传统的热、温、冷三层数据将变为热、温冷两层数据，数据比例有望从 20%:30%:50% 到 30%:70%，冷数据的价值被激活，我们将迎来“数据觉醒”。

先进介质技术

多样化的数据驱动存储介质多样化发展，在不同的应用领域构建竞争力。其中，极热数据内存型介质将以 DRAM 为主、SCM 为辅，内存的分级将成为新生态；热数据介质将全部使用 NAND Flash，Flash 闪存技术则会继续向高密度、低时延演进；在温、冷数据介质技术中，磁带有望继续向高密度、高并发方向演进，光盘则会走向更大容量、更高并发、更长寿命。

1、热数据介质技术

内存是计算机体系架构中非常重要的组成部分，用于程序与数据的缓存。随着数据密集型应用的发展，需要处理的数据量将从当前的 GB 级增长至 TB 级，驱动内存型介质向更大容量、更低功耗、更高并发方向发展。

1) 内存架构将走向多层次化

当前，DRAM 占据内存型介质的主流地位，由于 20nm 以下制程工艺的容量密度提升空间有限，未来 10 年将在 10nm 级别持续发展。随着大数据集对大内存需求的增加，SCM 等新介质技术的发展推动内存架构走向多层次化，逐渐跟 DRAM 形成互补。

2) SCM 将持续开拓新场景

基于新型材料和结构的 SCM 能够拥有与 DRAM 相媲美的性能，并具有持久化的新特征。在某些特定领域上，使用 SCM 实现的 CIM 内存作为 DRAM 的补充，起到了很好的加速效果，并且未来围绕 SCM 的新生态将不断丰富。各种具有持久化能力的 SCM 介质可以使高性能热数据得到快速访问。现有存储系统中，处理器经常花费大量时间在 IO 等待上。未来，创新的内存型持久化存储子系统有机会改善这种状况。

3) NAND Flash 在 3D 堆叠的方向持续演进，加速替代 HDD

未相比于 HDD，SSD 在性能、功耗、容量上具有明显优势，在 To C 场景下的 HDD 已经被 SSD 替代，而在 To B 的场景下的 HDD 也有望被加速替代。

业界采用增加堆叠层数的方式来实现 SSD 的代次演进，从而提升单位硅片面积的存储容量，降低单位存储空间的成本。但随着堆叠层数的增加，堆叠 memory hole 的深宽比（孔的深度与孔径的比例）增大，对蚀刻以及沉积的工艺带来了更高的挑战，从而限制了层数的持续增加。为了进一步提升存储密度，提升 NAND 阵列的有效面积占比，未来 CMOS 外围电路与 NAND 阵列堆叠的立体架构将成为主流。

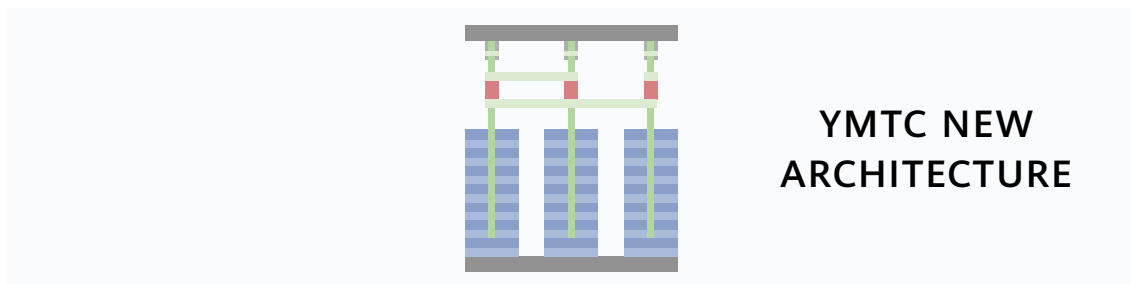


图 2-3 3D NAND 原理

通过堆叠和 3D 立体架构，预计 2030 年单位芯片面积的容量密度相比 2021 年有望增加 10 倍。但受技术的复杂性、工艺良率等因素影响，SSD 成本却难以下降 10 倍。反而由于工艺的影响、内部布线的干扰和密度的增加，SSD 的底层误码率可能进一步恶化，对低误码率和低延时高吞吐的纠错算法提出了新的挑战。

2、温数据介质技术

按照 SSD 和 HDD 的技术演进趋势，预计到 2030 年，HDD 的成本优势依然存在，这使得 HDD 在注重性价比的温数据存储场景中依然是主流介质^[14]。

HDD 的技术主要向提升密度的方向演进，由于 HDD 的磁记录只能附着在基板的表面，密度提升只能通过增加碟片数和提升磁密度来构建，受 HDD 形态和超顺磁性限制，当前 HDD 的容量密度已经接近极限，短期 HDD 密度的提升将向着突破形态限制和突破超顺磁性限制的方向演进，比如超厚 HDD，能量辅助磁记录技术（HAMR、MAMR）等^[15]，长期技术演进则包括磁记录技术的突破和材料的突破，如斯格明子，磁光、磁电结合的技术和材料等。

3、冷数据介质技术

到 2030 年，冷数据介质将仍然以磁带和光盘为主。光盘具备高可靠性、长寿命、对存储环境要求低的特点，更适合做为超长周期冷数据的存储，而磁带则主要用于中长期的冷数据存储。在数据驱动的智能世界时代，数据变得更热，相应的，对冷数据介质也提出了新的要求：低成本、可快速读取。

1) 磁介质技术

磁带记录是通过磁带介质顺序卷带的方式实现数据存储的一类技术。磁带通常使用空间折叠方法实现容量扩展。以 LTO-9 为例，其介质记录面积是同期 HDD 的 100 倍。当前磁带容量密度仅为 HDD 的 1/100，未来有望通过突破磁畴微缩、高精度伺服控制和超低误码率的磁信道编码技术，实现其容量超过 HDD 约 100 倍以上。

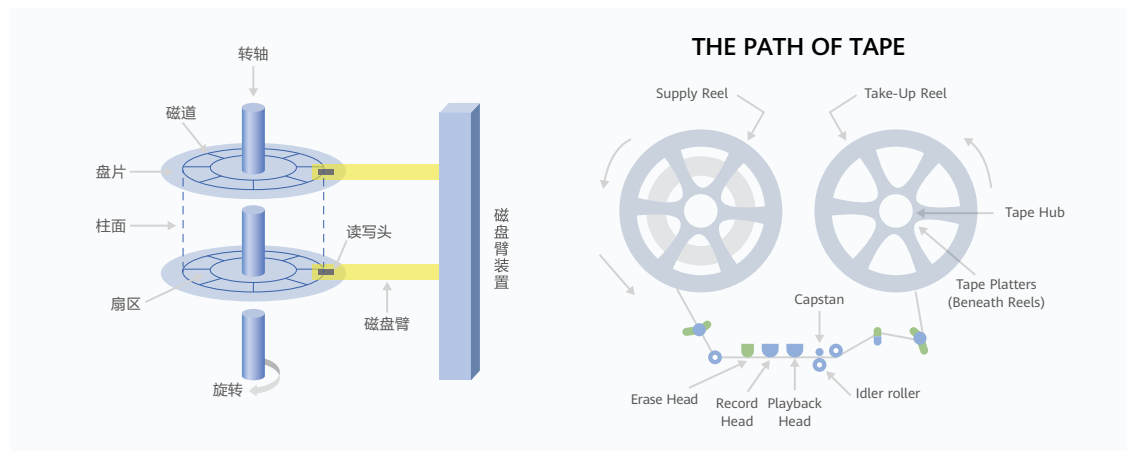


图 2-4 磁存储原理（磁盘、磁带）

磁带的线性运动可以使用更多的磁头并发读写，当前 LTO-9 32 个磁头并发带宽已超越 HDD 1 倍以上，未来有望实现超越 HDD 10 倍以上。从工作原理上看，磁带在顺序读写方面有优势，但随机读写时磁头的定位时间随容量而增加，影响数据实时性访问^[16]。未来，一方面可通过带宽优势换取时间，另一方面可通过数据的布局 and 调度算法进一步提升数据实时访问性能。从材料上看，磁带寿命受存储环境温度影响明显，当温度处于 35~40 度时，磁带存储寿命会下降数倍，增加数据丢失风险。未来需进一步探索新型材料、制作工艺和环境控制技术，延长磁带存储寿命。

2) 光存储介质技术

光存储介质技术未来将朝着大容量、低成本的方向发展。当前光存储的主流技术是蓝光存储，它最初用于消费领域，但容量只有 500GB/disc，单个光头吞吐率只有 40+MB/s^[17]。未来光存储将在超分辨、多阶、多维、镜面超多层以及体材料等技术上实现突破，将光存储容量提升到 300~700TB/disc，将吞吐率提升到百 MB/s。在 20 年内，单盘容量有望达到百 TB。

由于冷存储长寿命的要求，光存储未来面临的另一大挑战是如何实现光存储介质中的数据在几十数百年后可以被安全准确的读出^[18]。

超分辨光存储技术：光存储通过激光照射记录

材料，使其发生物理化学变化来记录信息。缩减波长和提升数值孔径可实现激光光斑变小，提升光存储记录密度，而波长和孔径的大小受衍射极限的限制，未来有望通过多束光叠加干涉超越衍射极限，进一步提升记录密度，提高光存储容量密度。

多维 / 多阶记录光存储：多维度光可突破单维度光只能记录单 bit 的限制，实现多 bits 信息记录。目前正在研究中的技术是由存储介质三维空间、偏振和光强度的五维光存储，未来有望解决光信号的空间干扰问题，向六维及以上维度发展，实现容量密度的进一步提升。

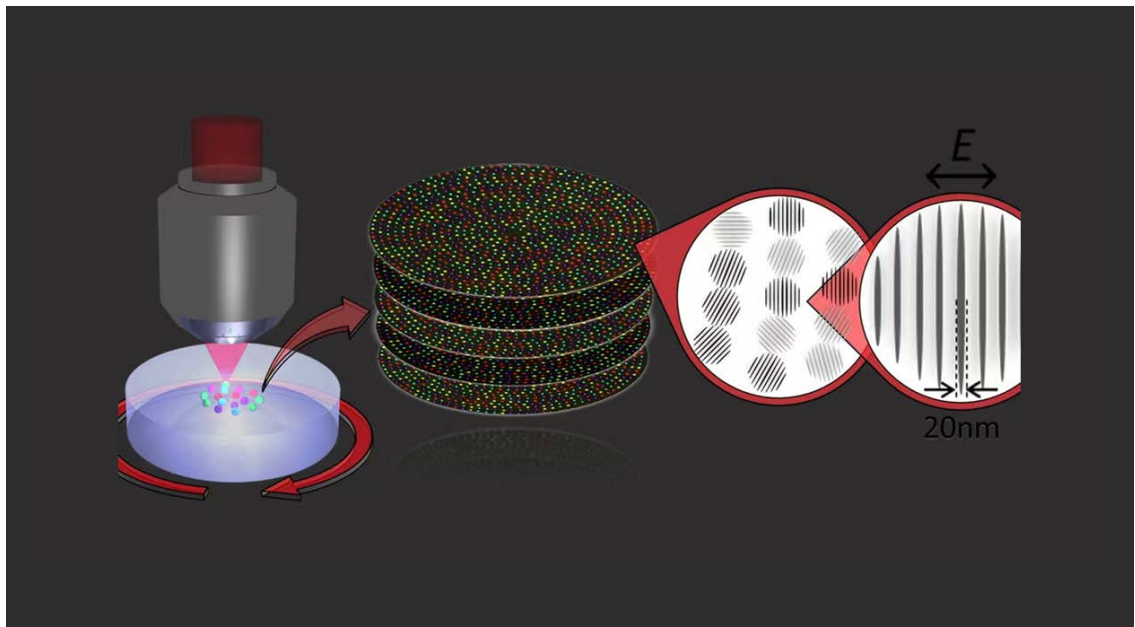


图 2-5 光存储原理

多层 / 体记录光存储：通过单光盘层数的叠加可实现光存储密度提升，如蓝光存储已实现六层商用，未来有望解决层间光干扰问题，向数十上百层方向发展。全息光记录采用相变体材料，可在存储介质体内部实现不同层和不同角度的信息记录，通过多层和体记录技术叠加，光存储可向更高密度演进，有望突破百 TB/disc 的容量密度。

伺服驱动技术：光驱包含激光器和光电调制设备，目前在多维光存储中使用的飞秒激光器和光电调制设备成本较高。随着飞秒激光产业的发展，未来有望进一步突破光高频高压电路技术，降低宝石级晶体的成本，实现光存储产业的大规模商用。受限于光存储的写入原理，单路激光的读写带宽只有几十 MB/s，未来有望通过高精度的伺服控制技术，实现多路光并行读写，提升吞吐率。

介质应用创新

1、介质工艺技术

受半导体制造工艺和介质结构物理极限的限制，SSD、DRAM 等介质的集成度无法持续提升，未来可通过 Wafer 级创新、Chiplet 级创新、接口和协议创新来进一步提升介质密度和寿命，降低介质功耗，增强介质的可靠性。

Wafer 级创新：Die-On-Board (DOB) 技术可以将存储颗粒或芯片集成到电路板上，提供更高的密度和更好的性能。Wafer-Scale 技术直接使用多个 NAND Die 的晶圆而无需对晶圆进行切割和封装，实现更高密度、更快速度和更高可靠性。当前 Wafer-Scale 技术还不成熟，需要解决超大芯片的制造、芯片的功能管理和监控、跨芯片连接、芯片散热、可靠性管理等问题。未来，有望采用先进的工艺技术、创新的芯片设计方法、智能测试手段等，在保持高密度和低能耗优势的同时，实现更高容量和更佳耐久性。

Chiplet 级创新：Chiplet 可以将不同功能模块集成在单独封装的芯片中，实现更好的灵活性和扩展性、更优秀的性能和功率效率。当前 Chiplet 技术仍然面临着芯片间通信和同步、缓

存一致性、传输速率匹配等多方面的技术挑战。未来，有望通过智能化控制算法、高效的芯片缓存一致性协议、存储介质内部封装处理器、异构处理器和加速器等技术，将计算芯片和介质芯片封装在一起，构建存算一体的 Chiplet 介质，实现高性能、低功耗、易扩展。

接口和协议创新：随着介质走向多元化，多种介质接口间数据传输存在较大的协议转换开销，在性能、安全、通用性上有较大的改进空间。ZNS (Zone Namespace) 都是用于闪存设备的高速存储协议，支持基于更小数据块的高效空间管理，缓解了 SSD 设备性能的不平衡问题，提高了 SSD 的垃圾回收和数据迁移等方面的性能，目前需要解决兼容性、应用迁移等问题。Plog 用于数据持久化的存储管理，可跨越多种存储介质，在不同存储系统之间的传输和处理数据。Plog 协议通过自动重传和自我修复机制，确保数据的一致性、可靠性和完整性，提高数据传输和访问效率。未来，随着多元化介质技术的不断发展，需要定义新型高性能接口和协议，进一步提高兼容性和数据的访问效率。

2、新型数据编码

数据编码技术包含缩减数据量的压缩编码 (Sayood, 2017)、抗数据错误的纠错编码以及抗数据丢失的纠删编码 (Peterson, Peterson, Weldon, & Weldon, 1972)，是支撑存的下 (空间) 以及存的久 (时间) 的核心技术之一。未来面对海量多元的存储数据以及介质融合的存储系统，通过智能化数据压缩、联合编码、智能化数据分类，有望突破数据编码技术，实现存储有效容量提升、集约节能、长期可靠。

$$H(p) + D(p \parallel q) \leq E_p l(X) < H(p) + D(p \parallel q) + 1$$

该定理表明数据压缩后的平均码长 $E_p l(X)$ 在最好情况的下等于数据本身信息熵 $H(p)$ 加上模型不准确导致的额外长度 $D(p \parallel q)$ ，最坏情况比最好情况大 1。

图 2-6 无损数据压缩的理论

智能数据压缩：数据压缩是按照特定的编码机制用短比特数据表示信息的过程。在数据存储中，有损压缩编码和无损压缩编码并存。当前的有损编码还无法打破经典的率失真率理论，未来需要探索语义提取和语义压缩技术，扩展率失真函数，建立新的理论体系，实现有损压缩的技术突破；业界主流的无损压缩方法以LZ和熵编码为核心，在面对非结构化数据压缩方面压缩效果欠佳。基于统计和动态预测模型的压缩方法可以有效提升

非结构化数据的缩减率，但存在模型依赖于数据和专家经验，发展缓慢；基于AI的预测模型通过对数据特征的自动提取和模型的自学习，可以超越专家设计的预测器。现有的基于AI的压缩算法面临泛化能力差和算力消耗大的问题，未来有望通过迁移学习、元学习、大模型等技术提升模型泛化能力和算法效率，实现存储系统中缩减率数倍提升。

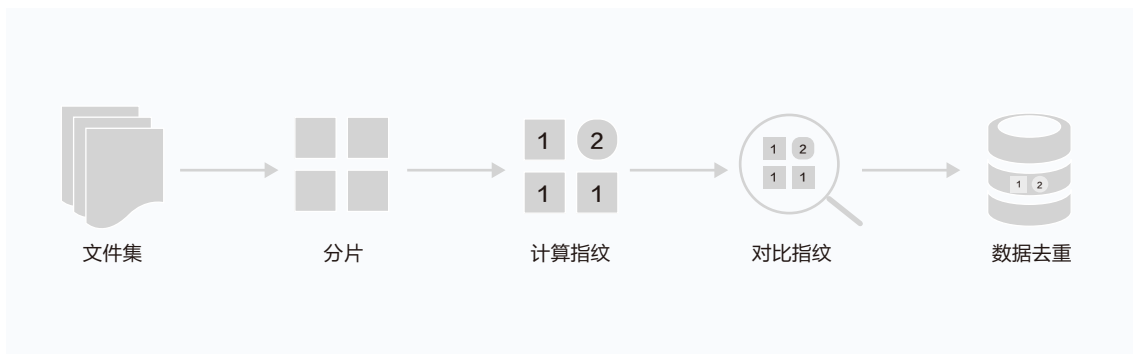


图 2-7 数据重删的基本原理

数据重删：重复数据删除技术（简称“重删”）是通过数据块级别的内容识别，实现重复数据块删除的一类技术。随着处理器技术和新型存储介质的出现，重删技术正逐渐从离线走向在线处理，数据重删的粒度也在不断缩小，从早期的文件级重删发展为近期的字节级相似重删，对于系统的

算力和 IO 吞吐率提出了更高的挑战。面向海量多元化数据重删，在高维数据场景相比结构化数据场景重删率还有数量级的差距，未来随着应用语义重删技术的发展，有望从根本上解决非结构化数据的存储效率问题。

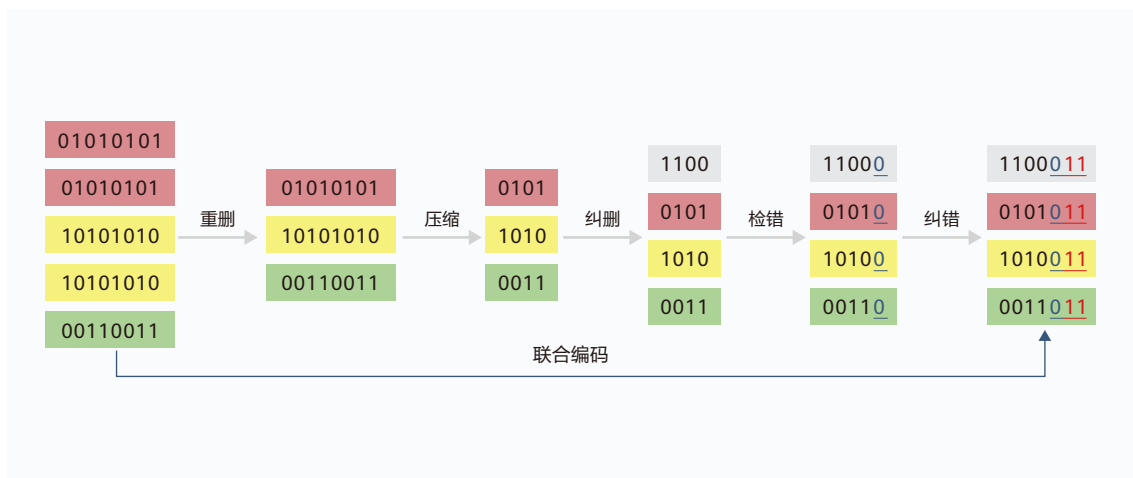


图 2-8 数据联合编码

数据联合编码：香农的分离理论 (Shannon, 1948) 证明了在码长趋于无穷的前提下信源编码与信道编码分开设计可达到整体系统最优，在有限码长的场景下，联合信源编码与信道编码可能

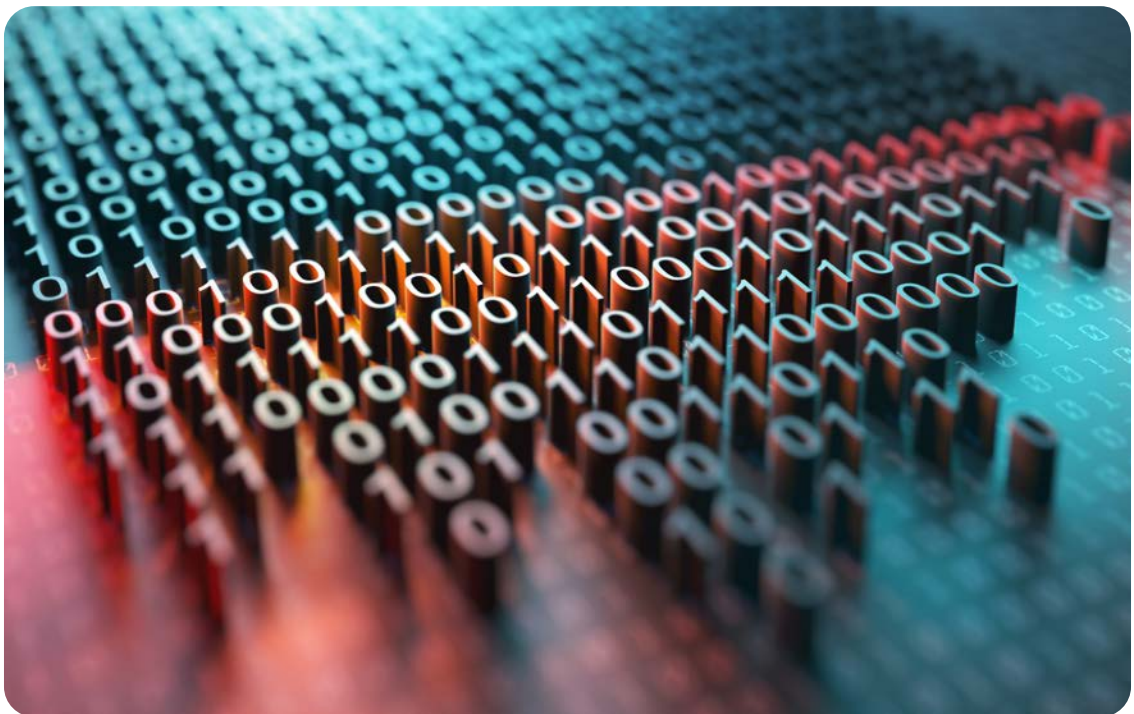
取得增益 (Jiang & Bruck, 2008)。未来通过设计联合编码可实现更高密度的存储，同时可简化系统，实现更低能耗。



图 2-9 智能分类

智能化数据分级分类：存储是一个多元化和层次化介质的系统，不同介质的可靠性、延迟、带宽和成本差异较大，需要选择与之匹配的数据编码算法 (Kim, Gupta, Urgaonkar, Berman, &

Sivasubramaniam, 2011) (压缩、纠错、以及纠删)。未来需要突破智能化的数据分类技术实现不同的数据编码与介质的最优匹配，提升数据的密度与可靠性，同时降低延迟。



2.2 以数据为中心的体系架构

在大数据、人工智能、HPC、IOT 等新型数据密集型应用的推动下，数据量爆炸增长，年复合增长率近 40%，其中热数据占比将超过 30%；另一方面，摩尔定律、Dennard 缩放定律的放缓，CPU 性能年化增长降低至 3.5%。高速增长的数据与缓慢增长的数据处理能力成数据产业的基本矛盾，数据存力与数据发展严重失衡。

在传统的以 CPU 为中心的数据中心架构中，

业务在空间、时间的不均匀性导致本地存储资源利用率低，本地内存、存储闲置率超过 50%^[19]。数据的移动、数据格式的反复转换消耗了大量 CPU 时间，使得数据处理效率低下。此外，在一些拥有十万卡 GPU 的大型智算中心中，CheckPoint 数据需要秒级快速恢复能力，而存储节点和集群的可扩展性差，读写带宽无法持续提升，导致宝贵的 GPU 训练时间被浪费，GPU 的有效利用率难以提升。

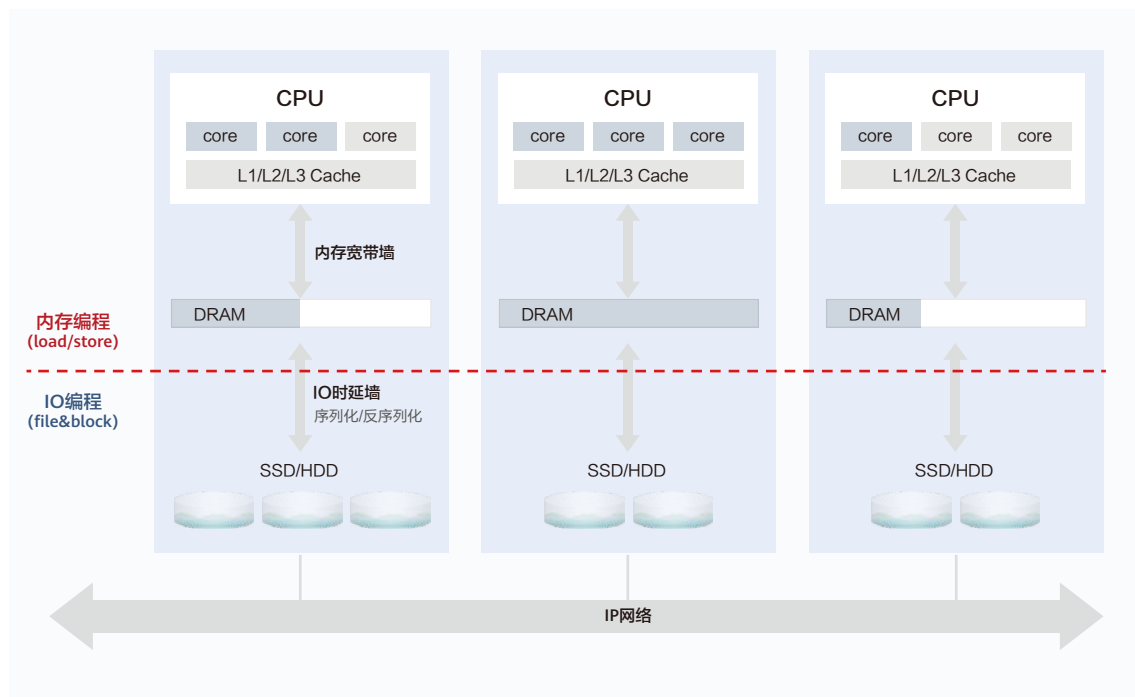


图 2-10 以 CPU 为中心的架构

现有数据中心架构存储、计算资源利用效率低下，为了提升数据处理效率和存储资源利用率，未来数据中心架构需要从“以 CPU 为中心”走向“以数据为中心”，包括三个方面：

1) 在宏观上存算分离，计算、存储资源独立部署，通过高通量^[20]数据总线互联，统一内存语义访问数据，实现计算、存储资源解耦灵活调度，资源利用率最大化。

2) 在微观上存算一体，围绕数据，近数据处理，减少数据非必要移动，在数据产生的边缘、数据流动的网络中、数据存储系统中布置专用数据处理算力，网存算融合提升数据处理效率。

3) 高可扩展的集群存储，在横向上从数十个扩展到数百个控制器，扩展能力提升数十倍，实现 EB 级容量，在纵向上从数百个扩展数千个 XPU，扩展能力提升数十倍，实现近存加速。

存算分离，以存强算

存算分离不再局限于CPU与SSD、HDD外部存储解耦，而是彻底打破各类计算存储硬件资源的边界，将其组建为彼此独立的硬件资源池（例如CPU池、DPU池、内存池、闪存池等），实现各类硬件的弹性扩展及灵活共享。存算分离架构具备三个特征：存储资源池化、全内存语义访问的内存型存储、NPU直通存储。

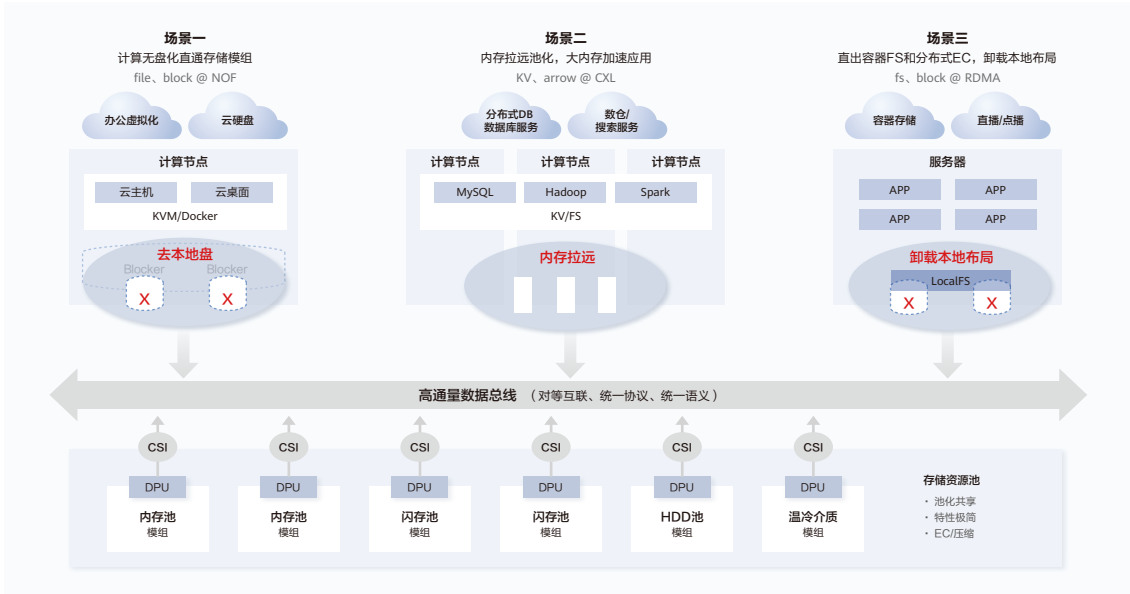


图 2-11 存算分离架构

1、存储资源池化

新型存算分离架构将服务器本地盘拉远构成无盘化（diskless）服务器和远端存储池，同时还通过远程内存池扩展本地内存，实现了真正意义上的存算解耦，可极大提升存储资源利用率。业务使用时，可根据应用需求选择不同性能、容量的虚拟盘及池化内存空间。首先，存储资源池化可以避免本地存储空间超配造成的空间浪费；其次，资源池化可避免数据跨总线和跨设备流动，减少数据移动，提升性能，降低功耗；最后，当服务器出现故障或者更新换代时，数据免迁移。

通过 NVMe over RDMA 网络技术，可以实现外存 SSD 池化，为远端访问 SSD 提供本地一致的访问性能。未来有望通过新型内存型网络（如 CXL、Unified Bus）、内存介质智能分级和内存统一编址等技术，实现内存池化，十倍扩展内存容量，降低应用获取大内存的成本。

2、全内存语义访问的内存型存储

传统应用通过文件、对象、块接口访问数据，IO 栈协议厚重，IO 开销超过 30%。采用内存语义和内存数据格式访问接口，可实现 IO 零开销、格式零转换、数据零流动。当前，内存语义访问仍面临应用数据访问接口生态、内存语义网络标准化的挑战，未来有望形成统一的内存语义标准协议，实现内存语义的数据互通，进一步提高数据访问效率。

3、NPU直通存储

传统互联总线以CPU为中心，CPU成为系统瓶颈，系统无法大规模扩展；协议类型七国八制，协议反复转换，影响系统效率；不同设备不同通信语义，数据格式反复转换，造成额外开销。需要定义高通量数据总线，支持设备对等互访，消除协议转换，简化数据访问，高通量数据总线具备如下四个特征：

1) 对等互联: 打破以 CPU 为中心的主从结构, CPU、DPU、存储对等互联, 数据访问不再经过 CPU, 异构多样数据处理设备对等直访数据, 提升数据搬移效率。

2) 统一协议: 抽象设备内、机柜内、数据中心不同通信需求, 制定统一基础协议功能, 实现在处理器与存储、不同存储设备之间采用一致访问协议。

3) 统一语义: 把不同访问需求抽象成统一的访问语义, 支持实现跨系统、跨不同类型设备数据的共享和访问机制。

4) 高通量: 单盘 SSD 带宽将演进到 25GB/s, 内存支持 100GB/s 带宽, <50ns 时延, 新型总线需要实现 SSD、内存、处理器互联, 以及扩展到机架间互联, 同时满足大块数据传输高带宽和小块数据传输低时延需求, 未来总线需要支持 TB/s 级带宽, 10ns 级时延。

存算一体, 以查代算

在以数据为中心处理范式中, 数据处理的通用计算走向数据处理专业化, 由数据搬移到处理器走向近数据布置算力, 在靠近数据的地方, 以最合适的算力来处理数据, 在数据产生的边缘、在数据移动中、在数据存储中就近完成数据处理。数据存储作为数据载体, 不仅提供数据存取服务, 还提供近数据处理加速服务, 数据就近处理有三种主要方式: 多样化存算融合、数据存储与网络融合、数据处理与网络融合。

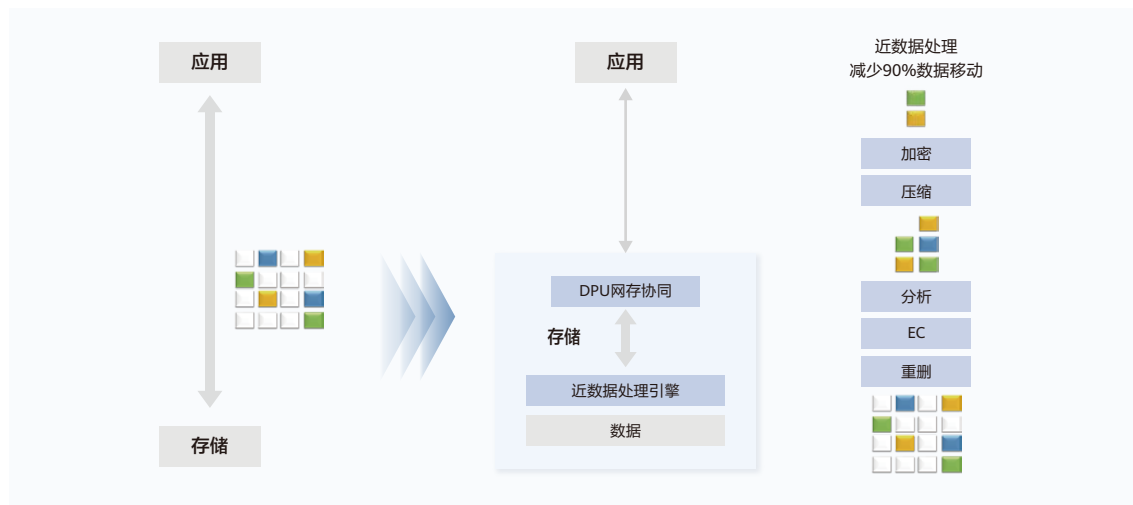


图 2-12 存算融合原理

1、多样化存算融合

存算融合是通过算子下推到存储器内部或存储模块上移到处理器内部来减少数据搬移, 解决网络时延和带宽瓶颈, 提升数据处理效率的一类技术。存算融合包括存算集成和存算一体^[21]。

存算集成 (SCI) 是在存储部件上, 集成指令运算

单元和算子单元, 实现数据预处理。例如, 在 SSD 内、内存上增加固化的数据预处理单元 (如压缩、编码引擎), 实现数据处理加速功能。或在处理器内部集成较大容量存储器来减少数据访问, 最终提升数据处理效率。未来, 面向前一种场景, 如何定义高效的前向兼容指令集、新的算子抽象, 仍面临巨大挑战, 有望通过共性指令集研究和自定义算子, 实现通用场景下数据的高效处理。

存算一体（CIM）是采用非冯诺依曼架构，基于存储单元与计算逻辑合一，打破计算和存储的边界，实现数据处理过程中极少数据搬运，相比传统冯诺依曼架构提升十倍以上能效。由于当前承载介质的局限性，在数模转化的效率、计算的精度和规模上仍面临巨大挑战，未来有望通过介质改良和发现新的介质材料来实现突破。

2、数据存储与网络协同

通过网络感知存储的语义，实现数据存储服务的卸载和数据流的调度，提升数据访问性能，加速数据应用服务。当前已经在存储访问协议卸载（文件协议、对象协议、KV 键值卸载等）、加速存储 IO（数据直通、IO 零拷贝）、卸载数据布局（索引卸载等）等方向上展示出巨大的应用潜力。通

过智能网卡可以实现灵活的存储业务卸载，但仍面临编程友好性和运行效率挑战，未来有望通过定义高效的存储算子，实现灵活性和高性能兼得。

3、数据处理与网络协同

通过与网络协同，卸载主机数据处理的开销，比如安全类数据处理（比如 SHA256、格密码）、数据压缩类数据处理（ZSTD、LZ、CDC）、数据保护类（EC）、数据分析类（Scan、Filter、Merge 等）等操作不再由通用处理器负责，随数据流卸载到专用数据处理器。以 DPU 为代表的专业数据处理器具备成本更低、功耗更低、即插即用、即换即用等独特优势，数据流动中加速数据处理，释放通用处理器算力，倍数级提升大数据、HPC、数据库等应用性能。

集群存储

在大型的智算中心中，单个存储节点或一百以内的节点扩展性能，无法满足计算集群对百 PB 数据、千亿文件、百 TB/s 带宽的能力要求。预计到 2030 年，存储集群的横向扩展能力有望达到数千节点；同时，为了近数据处理提升效率，越来越多数据读写任务被卸载到 XPU 上，存储支持的 XPU 数量也与日俱增，未来有望支持百万 XPU 同时工作，且弹性扩缩。集群存储容量有望提升 100 倍，达到数百 PB 级，同时解决高性能和大容量的难题。

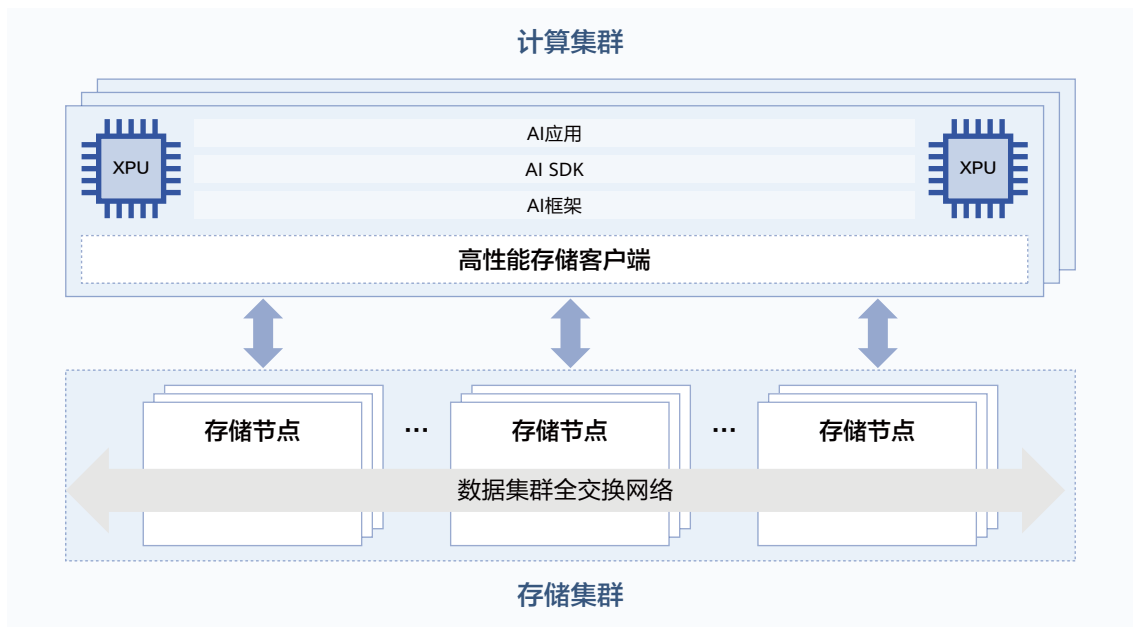


图 2-13 集群存储示意

2.3 数据内生安全

数据作为新型生产要素价值日益凸显，其作为高价值目标所面临的攻击面和攻击强度越来越大，当前基于边界的被动防御体系无法满足未来数据安全的需求^[22]。在数据价值释放过程中，针对数据的隐私保护需求日益旺盛，围绕数据“可用不可见、可见不可得”的隐私计算在充分保护数据和隐私安全的前提下，实现了数据价值的转化与释放。数据流转是释放数据价值

的必要途径和手段，由于数据可复制性、可共享、可无限供给，因此在流转过程中如何保证数据产权、使用权和控制权得到有效保护，是当前数据基础设施需要解决的首要问题。

未来，数据内生安全将成为数据基础设施的基础能力，需要在主动数据保护、数据零拷贝、零信任存储和 AI 安全等技术方向上持续突破。

主动数据保护

数据安全攻防态势研究表明，当前的被动防御安全体系无法有效抵御勒索等病毒攻击，需要从数据安全态势感知、数据时间线旅行、原生防篡改、多维联动响应等多个技术方向，构建主动数据保护安全体系。

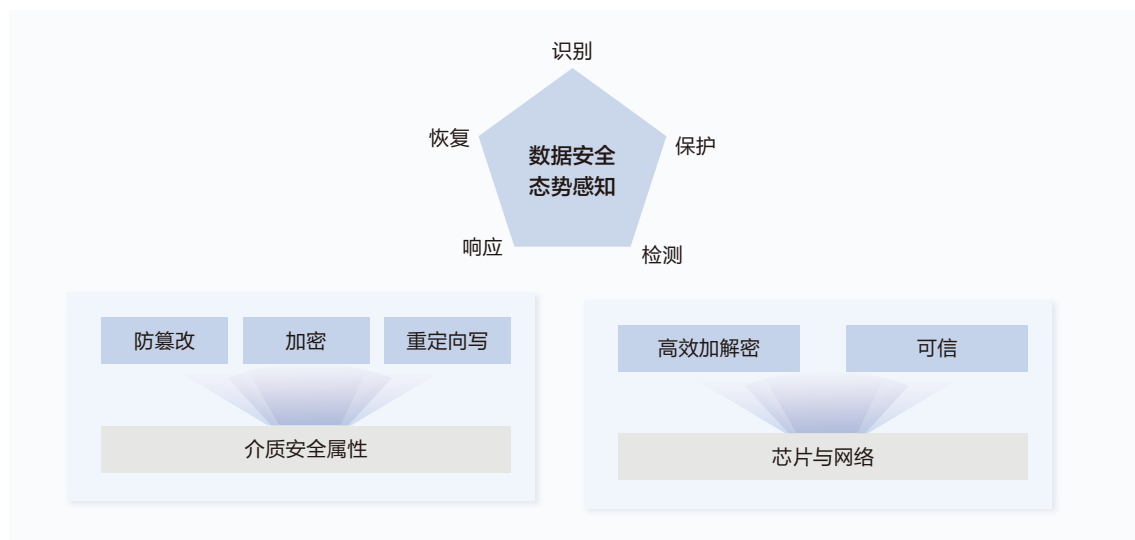


图 2-14 主动数据保护

数据安全态势感知：数据安全态势感知技术是在一定的时间范围内采集数据访问行为、数据信息熵、数据内在关联、数据分布等，结合大数据分析技术动态度量与评估数据安全风险和威胁，支撑后续自主防御决策和行动。当前业界的主要痛点是如何做到高效精准的威胁检测与态势感知能力、对威胁的动态评估能力不足等问题，未来有望通过海量数据的采样理论、异构数据的融合处理、不完全信息条件下的活动辨识等方向的研究，逐步提高检测的准确度和性能，增强未知数据威胁检测能力。

数据时间线旅行：数据在遭到内外部攻击损坏后，数据基础设施需要具备在最短时间内将受损数据恢复到任意历史时间点，实现数据零丢失的能力，同时为了实现攻击溯源，必须具备最细粒度的数据重放能力，支撑数据安全策略的调整与优化。当前业界的主要挑战是快速精确定位受损数据的时间点、自动化行为溯源等问题，未来有望通过 IO 级数据恢复、因果根因分析等技术实现数据在时间线上的旅行。

原生防篡改：当前数据防篡改的能力主要是依

靠系统级的数据访问控制技术实现，由于系统攻击面较大，很难有效保证数据的防篡改，未来有望通过系统级数据访问控制技术结合介质物理防篡改属性，实现物理级数据原生防篡改能力。

多维联动响应技术：多维联动响应技术需要通过网络设备、安全设备、终端 EDR 设备、存储

设备的跨设备联动协同，实现多维度威胁处置闭环，防止威胁范围扩散。当前业界的主要痛点在自主决策和响应技术上，即如何制定智能化的响应策略，给客户便捷有效的处置备选方案。未来有望通过 AI 安全分析、因果分析与推理等技术突破，有效提升自主决策和响应的智能化程度，实现真正的快速准确响应。

数据零拷贝

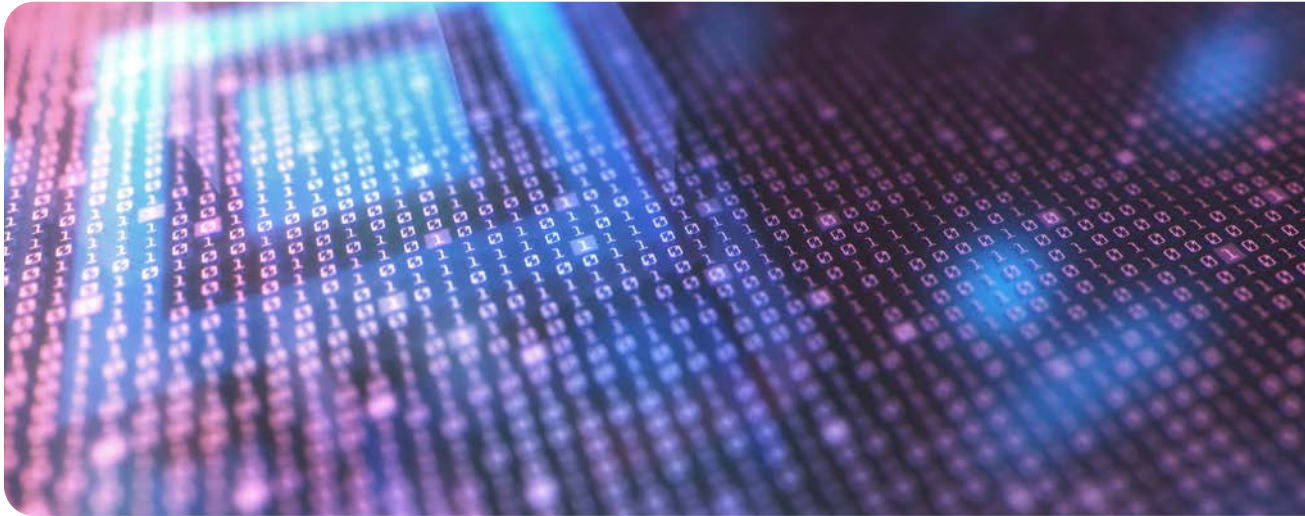
数据要素价值释放过程概括为三个阶段，第一个阶段是数据支撑业务系统运转，推动业务数字化转型与智能决策，第二个阶段是数据流通对外赋能，让不同来源的优质数据在新业务和场景中汇聚融合，实现双赢、多赢的价值利用，在该阶段需要解决数据共享与数据访问控制之间的效率问题，通过基于密码学的访问控制、数据自保护技术、高效透明审计技术、高效网络加密传输等技术能在保证数据主权安全的前提下，实现数据高效流动与使用；第三个阶段是无边界零拷贝，最大限度地消除数据孤岛，通过零数据拷贝访问技术打破数据边界，实现数据共享。



图 2-15 数据价值释放模型

基于密码学的访问控制：主要通过密码学的方式保护数据机密性，对于不符合访问控制策略的用户，无法解密。当前基于属性加密 (Attribute-Based Encryption, ABE) 方案已可支持任意逻辑的完备访问控制策略。相比于传统的一对一公钥加密，ABE 是一对多的，极大降

低网络通信开销和关键节点加解密计算开销。未来需要研究控制加密的技术，对离开信任域的密文进行策略判断和随机化处理，确保不符合预定访问控制策略的数据，无法离开信任域进入流通阶段。



数据自保护技术：近年来，数据安全正逐渐从以系统为中心的数据强管控，演变成为以数据为中心的全生命周期安全防护。数据自保护技术就是在隐私层面满足“可用不可见、可见不可得”要求的一类技术。当前主要采用密态计算技术来实现数据隐私保护，但仍存在关联信息隐私泄露的潜在问题，数据的使用范围、方式、有效期和访问权限难以约束，未来有望通过数据胶囊等技术，将访问策略、使用控制策略和密态数据封装在一起，确保数据拥有者对数据的自主可控，实现数据的安全流转。

高效透明审计技术：当前的数据可信审计的主流技术是区块链技术，但存在开销相对较大、

共识算法效率低、数据冗余存储等问题，未来有望通过高效透明审计技术来构建数据防篡改的审计方案，实现更加高效的可信数据存储，更加贴近实际生产过程中对数据读写时间的要求。

零数据拷贝访问：当前由于各个应用数据模型的差异，导致大部分的应用基于独立数据副本结合自身数据模型形成了大量的烟囱式应用，未来有望将应用数据模型下沉至数据存储层，基于同一份数据自动生成应用数据模型，从而消除数据孤岛，同时结合细粒度访问控制、基于芯片认证的可信网络传输等技术实现跨信任域高效数据访问。

零信任存储

零信任存储是基于零信任模型的扩展，旨在解决当前存储面临的数据泄漏、完整性被破坏、数据可用性破坏等诸多安全问题，在零信任存储中，所有的数据访问与操作都被视为未被验证的，访问主体、数据以及数据操作动作三者基于最小授权原则，通过持续验证、动态授权等方式实现最小粒度数据访问控制。同时要实现零信任存储需要从数据存储与使用环境安全、数据全路径安全加密等几个方向突破。

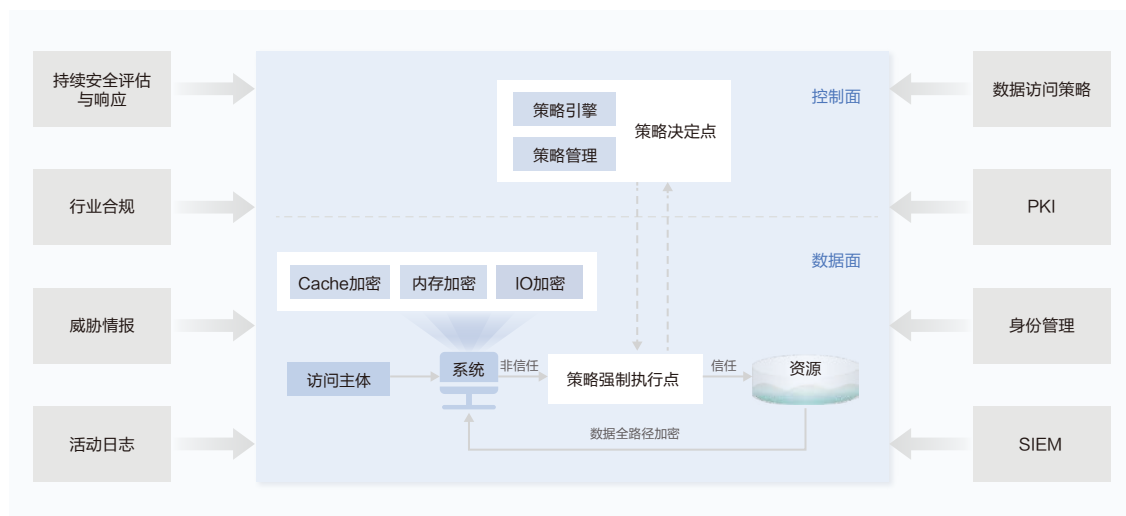


图 2-16 零信任存储

强制数据访问控制：细粒度数据访问控制是基于最小授权原则，利用数据访问主体特征、数据属性、细粒度数据处理动作三者之间的映射，以确保最小粒度数据集只能被主体在特定的条件下访问与使用。未来，由于被授权实体和数据的海量性、数据处理的复杂性、控制条件的不确定性，访问控制策略设计日趋复杂，访问控制策略配置错误将带来重大安全隐患。为了应对这个挑战，未来有望通过形式化验证、自动策略生成、合规审计等技术来保证复杂策略的一致性和正确性，解决大规模形式化验证的性能、自动化策略生成的机制、复杂规则匹配的问题。

数据全路径加密：当前基于边界的数据安全体系，对数据全路径的安全假设存在数据泄漏的风险，我们需要考虑从内存、存储 IO、网络 IO、Cache 等数据处理的全路径进行加密，并通过统一密钥管理实现原生数据安全能力的共享。

隐私计算：为了保证数据在计算过程中的隐私安全，数据安全计算应运而生。数据安全计算的主要技术流派包括：针对 AI 场景的联邦学习^[23]（Federated Learning）、基于硬件安全基础的可信执行环境（Trusted Execution Environment）、基于密码学算法的多方安全计算（Secure Multi-Party Computation，

MPC）、针对向验证者证明某陈述正确性场景的零知识证明（Zero Knowledge Proof）等。

1) 可信执行环境：实现敏感数据处理的硬件隔离技术，主要挑战在于硬件安全隔离机制实现的完备性无法用数据证明，难以自证清白，存在安全漏洞风险，但和密码学技术相比，TEE 对性能影响小，未来基于 TEE 的隐私计算将成为业界普遍需求，预计 2030 年 50% 以上的数据处理场景将使用该技术。

2) 基于密码学的同态加密、安全多方计算技术因其安全性在数学上可证明，从而成为业界公认最理想的隐私计算技术。但主要挑战在于其性能比常规计算降低一万倍以上，需要大幅提升才能满足应用需求。随着近似计算的成熟，同态加密、安全多方计算在人脸识别、健康数据分享等特定领域以获得应用。未来，突破基于硬件加速的同态加密、安全多方计算技术，将在金融、医疗等行业的高安全应用场景获得广泛商用。

3) 多方计算的基础是多方之间共享秘密，如果通过零知识证明等密码学方法实现，性能开销非常大，利用 TEE 实现多方之间的秘密共享，不但可以大幅度提升多方计算性能，而且在信任 TEE 基础上安全性可数学证明，未来有广泛的应用前景。

2.4 智能数据编织

数字技术的不断发展催生了大量的跨域数据流动的需求，对数据的可用性和质量提出了更高的要求。但地域的阻隔和数据治理的困难限制了数据的自由流动，最终形成了数据重力。数据编织是以一种自动化的方式，动态地协调分布式的数据源，跨数据平台地提供集成和可信赖的数据，支持广泛的不同应用的使用^[24]。智能数据编织可基于人工智能和知识图谱等技术，不断识别和连接来自不同应用的数据，以发现

可用数据点之间独特的业务相关关系。在数据网络中，边缘、数据中心、云端频繁的数据交换，智能数据编织可通过对现有的、可发现和可推断的元数据资产进行持续分析，完成跨平台的数据整合，为应用提供高效数据流动和处理。为了更好地实现智能数据编织，需要在跨域数据协同、自动化数据编排和高效快速存力网络等技术方向上持续突破，以解决数据重力问题。

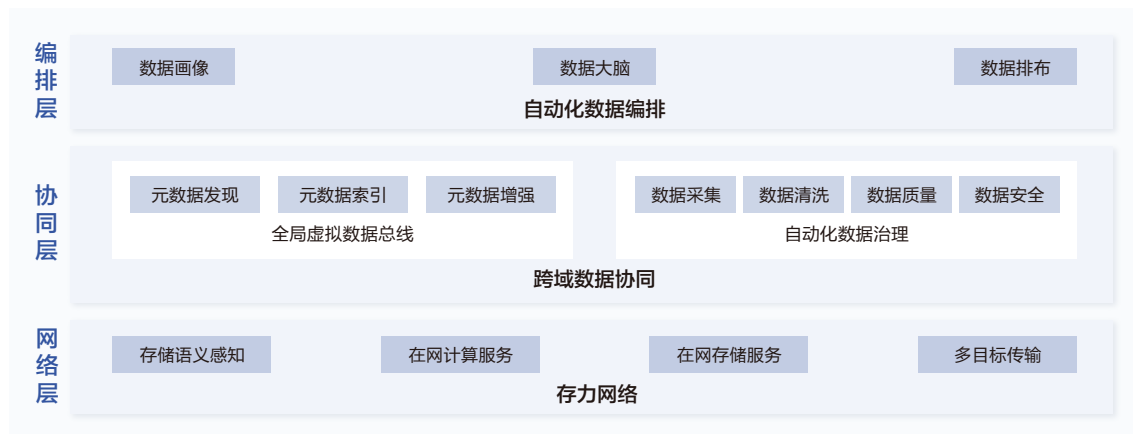


图 2-17 智能数据编织框架

自动化数据编排

当前，由于数据内容无法感知网络状态，应用的意图也无法有效传递给网络，导致数据的放置跟网络匹配失衡，数据存取延迟大、网络利

用率低下。未来需要通过构建数据画像和数据大脑，实现业务无感、业务性能无损的数据最优排布。

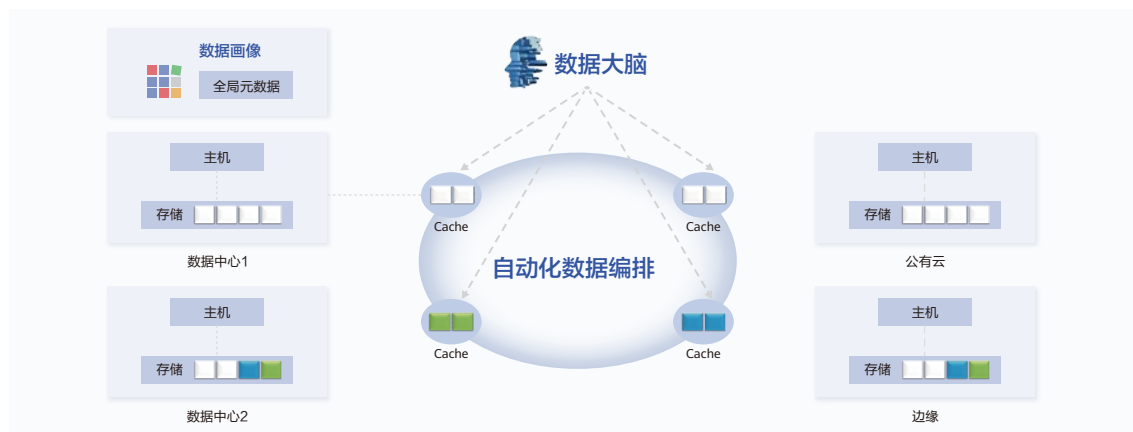


图 2-18 自动化数据编排框架

数据画像：数据画像是通过获取存储网络状态、数据块的时空信息、应用的标签等来感知应用特征。当前业务感知的粒度和精度较差，未来有望通过深度图神经网络、因果学习等技术，构建包含数据重力、数据量、数据活跃度、网络带宽和时延等多个因素在内的海量数据多维度画像，实现精准的业务感知。

数据大脑：当前的数据编排存在数据分散、维度爆炸、无法标准化、对开发人员能力要求高、无法感知客户应用等挑战，同时面向多云场景，衍生出对数据流动可信的管控诉求。通用的数据编排平台无法做到多方兼顾，未来有望通过意图 API、机器学习、大数据分析等技术，生成结合行业应用的最优数据排布策略，同时可针

对数据编排的整个流程提供强大的安全管理和审计能力，最终实现数据编排的自动驾驶。

数据排布：数据排布是根据业务策略将数据放置到最优位置的一类技术，帮助用户通过内容名字访问就近的数据，期望以最小的代价获得最佳的体验。如对于冷数据，可以通过东数西存的方式将数据放置在西部，降低运营成本。当前的数据排布存在不同业务间数据共享差、数据访问长尾、数据缓存命中率低、网络带宽占用大等问题，未来有望通过业务逻辑和数据逻辑分离、数据网络编码、数据预取淘汰算法等技术突破，实现自适应数据缓存和就近读写缓存加速，提供应用无感、成本最优的数据访问体验，使得数据更容易被发现和利用。

跨域数据协同

企业使用多个地域的数据中心或多个异构云供应商，提供统一的计算 / 存储服务，以提升基础设施能力、控制成本；资产、软件、应用的分布跨多个数据中心或多个云环境，催生了跨域的数据协同和数据整合。跨域数据协同在如下两个技术方向上存在机遇和挑战：

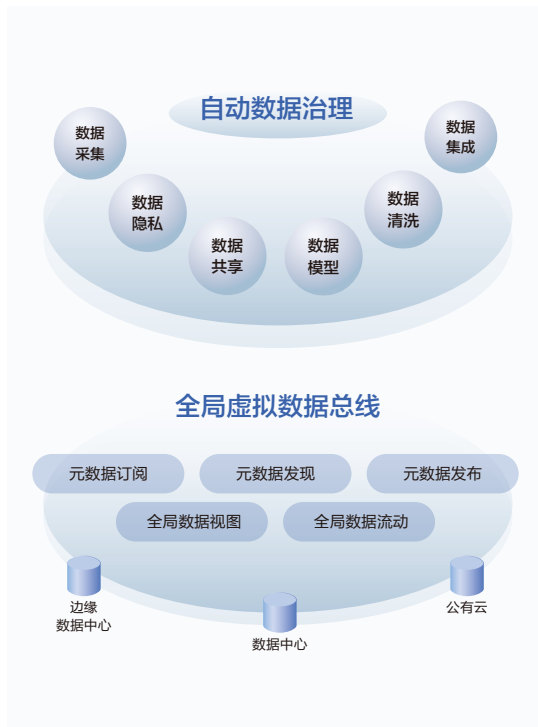


图 2-19 跨域数据协同框架

全局虚拟数据总线：在公有云和企业数据中心的，当前主要采用分区管理的方式来管理数据，产生了大量的“数据孤岛”。未来有望通过对元数据进行发布、发现、订阅来实现高效按需互联，构建全局虚拟数据总线。全局虚拟数据总线需要具备统一的数据命名空间和透明的数据流动能力，为客户提供跨云的全局数据空间，以及安全、高效、易用的数据网络。

自动化数据治理：不同来源和不同类型的数据需要互联互通和高效协作，通过统一规范数据模型和数据体系，将数据采集、数据集成、数据清洗、数据质量提升与数据安全保证等基础功能集成，并提供自动化处理的能力，可以提升从数据采集到数据价值发现的效率。当前数据治理技术尚不成熟，需要在异构数据集成、数据血缘管理和数据分类分级上取得突破，构建“统一、高效、智能”的数据中台服务，有效提高数据的质量和可用性。

存力网络

未来，数据的产生天然形成数据孤岛，跨域数据流动存在广泛需求。当前数据访问的网络延迟大、系统效率低，严重阻碍了数据应用的发展，需要打破地域和区域的限制，构建高效快速的存力网络，实现应用无感、地域无感的数据访问。从存储业务的原始需求出发，未来存力网络应提供如下 4 种能力：

存储语义感知：传统网络仅感知网络语义，如 IP 地址、TCP/UDP 端口号等，对所有网络报文一视同仁。未来智能数据网络能够进一步感知存储语义，如根据存储语义分辨报文的重要性和优先级实现策略转发、识别报文之间的关联性实现 Co-Flow 调度、基于存储 IO 语义进行路由等，从而实现对存储报文的差异化处理，充分利用有限网络资源，支撑数据在不同节点间的频繁交互。

在网计算服务：传统网络仅具有报文转发和路由能力，未来智能数据网络将进一步赋予网络计算能力。通过抽象运算算子，设计图灵完备的指令集，实现高效的数据处理引擎。一方面，数据处理引擎可由网络转发设备承载，能够实现数据的随路处理，在数据搬移的必经之路上对数据进行加解密、压缩、去冗、校验等计算处理，实现数据计算和数据传输的实时并行；

另一方面，数据处理引擎可由端侧网卡设备承载，能够实现数据不动计算动的近数据计算，节省数据搬移带宽和提供低延时服务。此外，SDXI (Smart Data Accelerator Interface) / NVME 等接口协议转换也可由端侧网卡设备承载，提供硬化的数据流动能力。

在网存储服务：当前网络的主要功能是数据包的搬运。未来，有望利用网络自身大量数据包的转发处理能力，对外提供多样化的随路存储服务，比如分布式锁、元数据缓存、事务并发控制等服务，实现 sub-RTT 的服务响应时间，大幅提升数据访问效率。

多目标传输：在网络控制协议上，传统的 TCP/IP 网络是基于网络生存性设计的，存在高通量和低时延不可兼得的问题，未来有望通过 RDMA over WAN、F6G、全光网络等技术，同时实现超低时延和高吞吐；在网络路由协议上，传统网络是满足单目标来设计的，无法同时满足路径最短、网络利用率最大、负载均衡等多目标的需求。而在数据存储网络中，既存在低延时网络的实时交互的数据库查询需求，又存在高通量网络的大文件传输需求，未来有望建立多目标的网络协议，实现多样化的数据服务。



2.5 数据智能

预计到 2030 年，以数字孪生、元宇宙、ChatGPT 等技术为代表的智能数据基础设施无处不在，和人的生活紧密结合。当前，日益复杂的存储系统已无法满足新兴多云应用的智能化数据业务需求，需要数据业务逻辑与数据智能解耦。数据基础设施面临着三大挑战：1) 各个应用的数据分散在各个角落，形成一个个数据烟囱，应用间数据无法共享；2) 对数据价值的挖掘消耗了海量的资源，反复对数据进行建模、训练、推理，不可持续发展；3) 针对海量应用的数据管理的复杂度不断上升，数据预处理的效率成为核心瓶颈，严重制约了应用的发展。

数据智能，意味着数据存储将具备数据感知、数据理解、新型数据服务等能力，支撑数据服务走向千行百业，数据业务百倍增长。数据存储将向泛在化、多样化内涵、认知存储等三个趋势发展。

泛在化：数据存储将走向小型化、便携化、绿色化、智能化，呈现出低功耗、可移动、生物性、量子性等特征，将催生便携式存储 (Portable Storage)、计算型存储、类脑脑机存储、生物 DNA 存储等一系列新形态。其中，便携式存储将会是最早大规模商用的产品，短期内以数据

存储、数据移动为主，实现数据在端侧、边缘和数据中心或云中的快速流转；中长期来看，便携式存储通过可组合乐高式设计，形成一个高可靠、高安全、免运维的智能移动存储设备，集数据存储、数据移动、数据交互、数据处理为一体，实现数据的实时共享、实时交互、实时处理。

多样化内涵：从数据产生的源头来看，层出不穷的新型应用使得数据的产生主体从传统服务器向多样化的数据产生源发展；从数据格式来看，传统应用以图形图像数据格式为主，新型脑机接口、生物仿生、AI 等应用将推动数据格式走向多样化，并催生出 Vector、Tensor、RAG 和长记忆存储等新的数据范式；从数据语义来看，自动驾驶、无人机、机器人等端侧设备，将产生大量的复合语义数据。

认知存储：当前的存储设备仅提供数据存放功能，访问层次多，无法满足极致的应用体验。未来的存储设备将成为具有认知能力的智能设备，能够自动处理和分析数据，依托数据进行自适应建模，形成领域知识，并通过“学习”不断提高自身的处理能力^[25]。

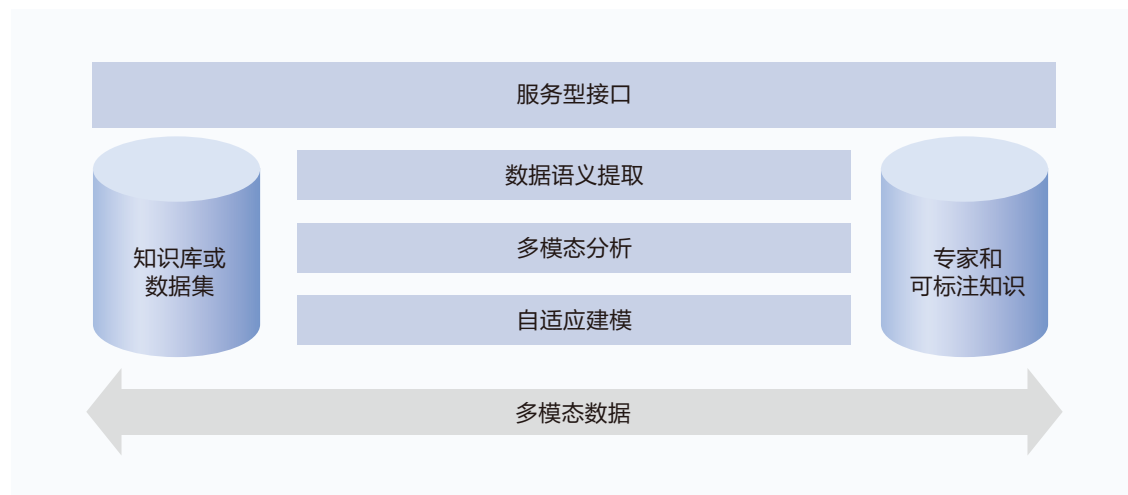


图 2-20 数据智能框架

未来，数据智能技术将向以下 4 个方向演进：

内容消费的服务型接口

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。未来的数据服务和 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。它们能够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。

对数据工程师来说，未来的数据接口允许自定义查询、过滤和以编程方式操纵数据。它们可以提供允许分页或过滤结果、实时数据流和事件驱动的数据处理的端点。这有助于提供更高级的功能，为特定的使用案例量身定做。

对普通用户来说，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。这意味着用户可以用简单的语言提问，而界面会理解问题的意图并提供相关的信息。

对商业决策用户来说，数据接口可以使用预测性分析来提供基于历史数据的洞察力和预测。这可以帮助用户识别数据中的模式和趋势，这些模式和趋势可能无法通过简单的分析立即看到，而需要先进的可视化技术，为用户提供对数据更全面的理解。

总之，未来的数据存储将从数据存取 ATM 机进化为内容消费的超级市场。

数据语义提取

数据语义是面向数据业务运用智能化的技术从数据中抽取与业务目标相关信息的一种技术。通过数据语义的提取可以压缩原始数据，提高系统效能。

当前，语义提取技术仍面临诸多挑战。首先，

语义提取技术主要基于自然语言处理、知识图谱和深度神经网络，受到深度神经网络理论可解释性和规模的限制，语义提取技术距离达到期望准确性还存在一定距离，且基于语义的推理准确性难以达到原始数据的效果。其次，语义提取的泛化性较差，可部署性差，需要多次训练。最后，提供能够独立于软件 / 硬件支持甚至平台之间差异的完整的数据语义服务是其可用性的基础。数据语义的可移植性，需要定义完整的数据定义和描述的语义方案，不断推动数据服务的标准化和产业化。未来有望通过 NLP、预训练大模型技术来实现数据语义提取技术的突破，实现语义推理无损。

数据多模态分析

未来，多模态数据的整合将更加普及和成熟。随着传感器技术的不断发展，人们将可以更轻松地收集和不同处理不同类型的数据，包括图像、语音、文本、传感器数据等，并将它们整合在一起^[26]。例如，一个自动驾驶汽车可以同时获得和处理多个数据源，如道路数据、交通数据、车载传感器数据、以及周围环境及座舱的图像和语音数据，并将它们整合在一起，以更准确地态势感知和做出智能决策。

同时，多模态数据的整合是一项复杂的任务，需要将来自不同来源的数据进行标准化和统一处理，以便在不同应用程序之间进行交换和共享。未来可能的融合方式包括：

- 1) 多模态数据融合：将多种类型的数据融合在一起，例如结合语音、图像、传感器等多种数据源，进行多模态数据分析。这种方式可以提高数据的信息量和可靠性，同时也可以解决单一数据源无法解决问题的情况。
- 2) 多层次融合：将不同层次的数据融合在一起，例如结合底层的传感器数据和高层次的语义信息，进行多层次的数据分析。这种方式可以提高分析的精度和深度。

3) 多源数据融合：将来自不同来源的数据融合在一起，例如结合社交媒体数据、物联网数据、企业内部数据等多种数据源，进行多源数据分析。这种方式可以提高数据的完整性和覆盖范围，同时也可以发现不同数据源之间的关联和联系。

当前，数据多模态融合分析技术主要基于规则、特征和语义的融合算法，涵盖了机器学习、深度学习、计算机视觉、自然语言处理、传感器技术等多种技术。未来，数据多模态融合分析技术会解决强依赖数据同分布和封闭领域的问题，通过空间变换、自监督学习技术和 AIGC（人工智能技术生成的内容），实现跨模态学习，提升学习能力，自动学习各模态间的语义对齐关系，提高模态融合的理解精度。

数据自适应建模

数据自适应建模（Data Adaptive Modeling）可以从输入数据中自动识别和学习潜在的模式和结构，并生成相应的预测模型。当前，数据自适应建模还面临诸多挑战。首先，由于样本采样的偏差，应用环境与模型训练场景的差异造成的数据漂移问题普遍存在。当数据的分布产生漂移时，旧模型难以适应新的环境，需要重新训练。其次，为应对新环境和新场景，数据自适应模型需要快速适应新数据，解决快速

响应和高效预测的问题。

当前自适应建模主要依靠神经网络和机器学习技术，神经网络技术需要寻找合适的网络结构，机器学习需要选择合适的特征，未来有望通过增量学习、迁移学习、领域适应方法、生成对抗网络等方向的技术突破，实现应对复杂多变场景的自适应建模，促进自适应建模的大规模实用化^[27]。



2.6 可持续存储

预计到 2030 年，每月读取一次全球数据所需电量，为全球年产量的 4%~6%，产生的二氧化碳需要全球树木耗时 7 天吸收。如何降低每 bit 数据读写能耗，对构建可持续发展的数据基础设施至关重要。

基于经典的冯·诺伊曼架构，数据在存储和计算单元间传输所需能耗，占 IT 系统总能耗的 60%~90%，数据密集型应用的能耗问题尤为突出。以数据为中心的体系架构，将解决数据传输能耗大的问题。

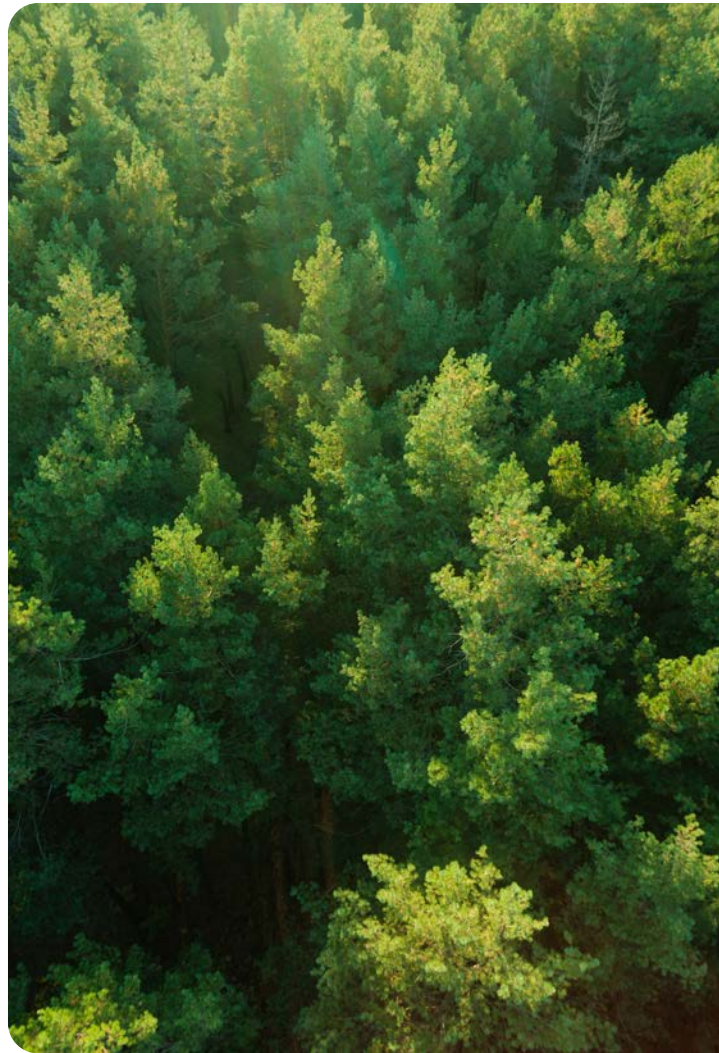
未来，低功耗介质、以光代电等技术将减少能耗产生，存储系统、整机和环境等节能技术，将进一步减少二氧化碳产生并提高能耗效率，从芯片、介质及网络全方面减少能耗，达成每比特最优能效和最少碳排放。

存储系统级节能

存储系统级节能是通过感知计算、存储、网络设备的运行状态，识别数据冷热特征，并结合业务负载规律，构建系统调优模型。存储系统根据模型，调整软硬件工作状态，实现整系统能耗最优。存储系统级节能有以下四类技术：

1) 硬件功耗智能调优

采用大数据和 AI 分析历史数据，探索影响能耗的关键因素，获得 PUE 预测和节能收益模型。利用寻优算法获取调优参数组，预测出设备硬件（CPU、磁盘、网络、风机、冷泵等）的调优策略，达成整系统最优能耗控制。当前解决方案存在模型泛化性差、少样本及实时性差等问题，节能效果不理想，需要大量人工干预。同时，AI 模型解释性差，安全运营风险较大^[28]。未来，有望通过融合专家经验的模型模块化、少样本概率建模、高效在线训练 / 推理、域自适应等技术，减少人工干预，提高模型可解释性，大幅降低硬件能耗。



2) 数据分级节能技术

当前大部分电能用于维持服务器、存储、网络等设备状态，在非工作时间段，存储设备仍需消耗大量电能，如何减少非工作状态的设备能耗至关重要。数据冷热分级就是在数据中心中，根据数据使用频率，把数据存放在磁光电混合介质中，从而有效降低能耗，平衡性能和成本。当前基于人工经验的数据分级策略和容量规划，存在资源浪费大的问题，未来需解决 IO 访问建模、数据布局，高效预取等问题，有望通过构建精细化的数据使用模型，在保证性能的同时最小化数据存取能耗。

3) 存储设备散热技术



图 2-21 存储设备散热技术

数据中心的冷却系统用电量，占总用电量的 30%-60%，如何将热量更高效的带出 IT 设备，增加整机散热能力，减少制冷系统功耗，成为绿色低碳的关键因素^[29]。当前数据中心主要采用风冷散热技术，COP(Coefficient of Performance)约为 5~10 水平，存在噪音大，无法满足数据中心散热能耗的需求。未来片上的免导热材料、闭式射流冲击冷却散热技术、新型冷却技术、整机余热回收等技术的突破，可有效提高散热效率，从硬件各层面减少碳排放。需要解决零热阻的焊接技术、高比热容且耐腐蚀的非水工质液冷材料和电机转换效率等问题。

4) 资源感知的统筹调度技术

清洁能源是在生产过程中不产生任何温室气体的能源，包括太阳能、风能、水能、地热能、生物能源和核能。大型数据中心可集约布局到清洁能源丰富的西部，同时为实现就近计算，计算也放在西部，但会存在计算延时大的问题。东部保留小规模、极低时延的边缘数据中心，

以满足业务低时延要求并减少需要搬移的数据。需开发数据放置策略及跨多 DC 调度引擎，动态实时感知计算、网络、存储资源位置、状态以及可用性、异构性，地域资源定价、碳排放标准等关键因素，实现跨 DC 统筹分配。结合数据智能协同调度，实现跨 DC 的全局统一数据抽取、分析、聚合框架，获得计算效率、数据移动效率与能耗效率最优。



数据传输能效提升

当前数据中心网络通信设备约占总能耗的 15%，在 AI、大数据分析等新应用驱动下，数据中心对数据传输带宽需求越来越高。随着 400G、800G 时代来临，网络带宽想进一步提升，功耗将成为瓶颈。预计 2030 年电费支出约占数据中心年运营成本的 95%，网络设备能耗占数据中心总能耗的 20%，亟需优化数据传输能效。

当前数据中心网络方案中，“光 - 电 - 光”的转换过程及电信号的处理能耗最大，减少网络中光电转换次数，可有效减少整体能耗。光交换基于光信号直接映射到出端口，无需额外光电转换，且具备 10TB 级带宽，ns 级时延，每瓦 TB 级能效优势。当前光交换基于时间交换技术，光路切换时延需数十毫秒，通过光电混合技术可构建高通量网络。未来，需突破纳秒级切换光器件技术和高速交换算法，实现低功耗的全光数据中心网络。

芯片级节能技术

当前存储系统中绝大部分的能耗都由芯片产生，如何降低芯片能耗至关重要。随着芯片元件集成度提高，单位体积内散热增加，但由于芯片

材料散热速度有限，“热耗效应”限制芯片性能发挥。如何增加芯片算力的同时控制芯片能耗成为一大挑战。异构多样化算力集成、片上动态能效智能管理等技术，可有效解决高算力与低功耗的矛盾。

芯片节能技术主要有以下几个研究方向：

1) 低功耗原材料

随着冷源结构、氧化物材料、碳基纳米材料等新兴芯片材料的发展，3D 封装、Wafer Level Chip 等封装技术的进步，互补场效应晶体管 (CFET) 等低功耗技术突破，芯片集成密度有望持续提升。

2) 高密低功耗工艺

经典物理规律为随着芯片元器件尺寸减少，能耗随之减少，但在纳米、埃米尺度下，会出现“尺寸效应”，此规律不再适用。未来有望通过 DTCO\STCO 技术，寻找最佳芯片设计和光刻工艺方案，使晶体管数量增加至千亿级。

3) 芯片能耗管理

片上能耗管理技术，是通过控制芯片电压与时

钟频率，达成降低能耗的一类技术。当前芯片电压和时钟频率都是由芯片整体控制，按模块最大需求设置，能量浪费大。片上能耗管理技术有望根据业务负载，按分核级别控制电压和核内子模块的时钟频率开关，使芯片能耗 / 算力成比例增长，达到所耗皆所用。未来，可结合 AI 和传感器技术实现功率预测、功率封顶和部件功耗控制，达成各部件能效比最优。

4) 数据处理专业化

随着摩尔定律放缓，单 CPU 性能提升面临瓶颈，算力年增长量小于 50%，供需差距越来越大。随着登纳德缩放定律的终结，采用多核技术提

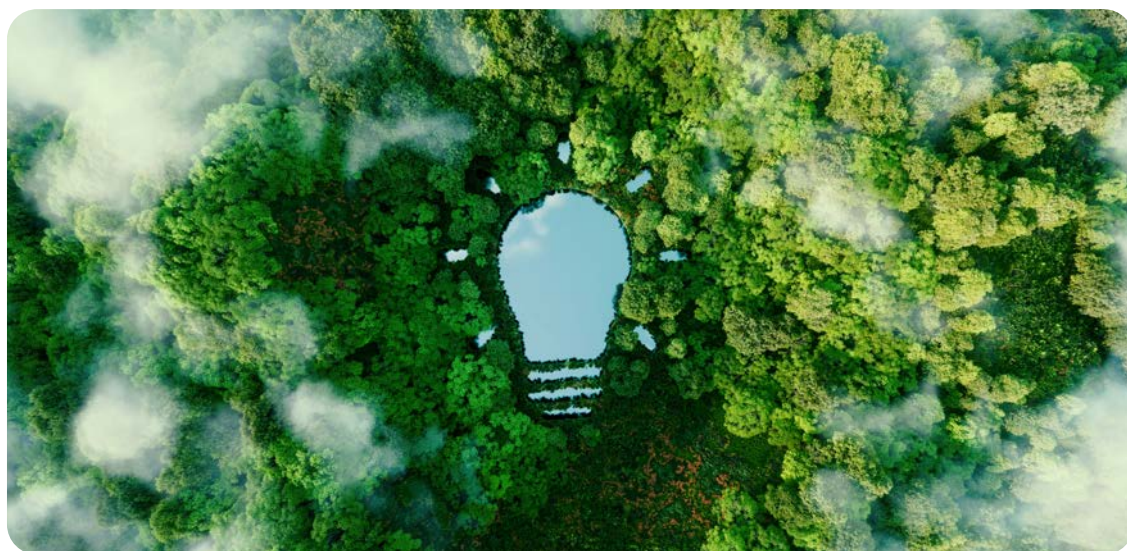
升算力的方法，将导致能耗大幅增加。经典通用处理器架构难以满足多元化应用发展。通过特定领域的体系结构设计，可以满足多样化算力需求，大幅降低系统功耗^[30]。

当前特定领域体系架构通过高效的并行形式、层次化内存结构、混合精度和特定领域编程语言实现多样化算力。由于系统架构、指令集和编程模型的差异大，使多样化算力的发展，面临程序跨平台运行难和编程复杂性高的挑战。未来，有望通过统一指令集、异构资源抽象、高效资源调度和异构编程模型的技术突破，实现大规模多系统的编译器、编程语言、加速库、开发工具等异构软件平台。

绿色集约标准

2022 年全国数据中心耗电量约为 2700 亿千瓦时，较 2021 年增加 25%，占全社会用电量的 3.1%，预计到 2030 年数据中心能源消耗翻一番。数据中心用能存在高碳排、重污染的情况，存储产业亟需逐步完善绿色集约相关法规标准，助力国家“双碳”战略。

当前面向数据中心的绿色标准，如能效仿真模型、节能技术、LCA 碳排放测评、减碳低碳、循环回收等标准已有初步发展，但存储产业未形成统一绿色集约标准。未来，存储产业有望建立统一绿色集约标准，覆盖数据全生命周期碳足迹、芯控接口、数据传输功耗、存储设备能效、存储设备碳排放强度和存储设备可再生能源利用率等关键指标，制定能耗基准，形成存储绿色低碳综合评价体系。





数据存储 2030 倡议

随着数字经济的全球化，数据存储产业作为数字化基础设施的基石，具有至关重要的作用。为了应对 2030 年 YB 级数据带来的挑战，不断提升数据存储的整体竞争力和产业协同性，实现产业级重大创新和突破，我们建议在以下几个方面持续开展协同技术创新：

1、应对多样化数据的挑战，发展多元化介质，重点推进介质应用创新，大幅度提升介质容量密度和每比特能效；

2、突破传统的冯诺依曼体系架构，推进以数据为中心体系架构的建设，构建高通量对等互联总线，建立统一的互联互通标准和协议，重构数据基础设施。

3、重视存力建设，以存补算，从数据处理的全流程来建立存力量纲，提升数字化基础设施的整体效率；

4、推动零信任存储体系建设，促进数据的产权、使用权和控制权分离，建立数据重力相关指数的行业标准，大幅度提升数据可信流动效率；

5、推进每比特最优能效和碳排放的绿色集约标准体系建设，发展可持续 IT，从以环境能效为中心的节能要走向以 IT 系统效率为中心的可持续节能。

让我们共同努力，一起开创数字化基础设施的大航海时代！

附录 A：参考文献

- 【 1 】 Seagate and IDC, Data Age 2025, May 2020
- 【 2 】 Gartner, “Forecast: Hard-Disk Drives, Worldwide, 2020-2026” , 2022. <https://www.gartner.com/document/4014430>.
- 【 3 】 World Health Organization. World health statistics 2021: monitoring health for the SDGs, sustainable development goals. 2021. <https://apps.who.int/iris/handle/10665/342703>.
- 【 4 】 德勤中国, 数字化健康白皮书, 2021
- 【 5 】 联合国, 2030 年可持续发展目标在中国, <http://sdgcn.org/sdg2.html>
- 【 6 】 realtor.com, Smart Home Technologies Reshape Real Estate Preferences in 2020, <https://www.realtor.com/research/smart-home-tech-2020/>
- 【 7 】 World Economic Forum, Raising Ambitions: A new roadmap for the automotive circular economy, 2022, https://www3.weforum.org/docs/WEF_Raising_Ambitions_2020.pdf
- 【 8 】 IDC, 全球智慧城市支出指南, 2020
- 【 9 】 Korn Ferry, Future of Work---The Global Talent Crunch, 2018, <https://www.kornferry.com/content/dam/kornferry/docs/pdfs/KF-Future-of-Work-Talent-CrunchReport.pdf>
- 【 10 】 联合国环境规划署, Emissions Gap Report 2020, 2020, <https://www.unep.org/emissions-gap-report-2020>
- 【 11 】 Abbosh O., Bissell K., Reinventing the Internet to Secure the Digital Economy, 2019, https://www.accenture.com/_acnmedia/thought-leadership-assets/pdf/accenture-securing-the-digital-economy-reinventing-the-internet-for-trust.pdf
- 【 12 】 Hennessy, John L. and Patterson, David A., Computer Architecture, Fifth Edition: A Quantitative Approach, Morgan Kaufmann Publishers Inc., 2011
- 【 13 】 中国信息通信研究院, 数据要素白皮书, 2022
- 【 14 】 Gartner, HDD and SSD market forecast, 2021
- 【 15 】 Yang S, Zhang J. Current Progress of Magnetoresistance Sensors. Chemosensors, 2021
- 【 16 】 Takeshi H., Hitoshi N. A study on high-density recording with particulate tape media for data storage systems, Synthesiology, 2017

- 【 17 】 SONY, Optical disc archive generation 2 white paper, 2016
- 【 18 】 Yuan X., Zhao M., Guo X., Li Y., Gan Z. and Ruan H., Optical tape for high capacity three-dimensional optical data storage, Chinese Optics Letters, 2020
- 【 19 】 舒继武, 新型存算分离架构技术展望, 中国计算机学会通信, 2022
- 【 20 】 范东睿, 叶笑春, 包云岗, 孙凝晖, 中国高通量计算机的自主研发之路, 中国高性能计算发展战略专题, 2019
- 【 21 】 Conte T. M., DeBenedictis E. P., Gargini P. A. and Track E., Rebooting Computing: The Road Ahead, IEEE Computer Society Press, 2017
- 【 22 】 邬江兴, 网络空间内生安全发展范式, 中国科学: 信息科学, 2022
- 【 23 】 Yin X. F., Zhu Y. M., Hu J. K., A Comprehensive Survey of Privacy-preserving Federated Learning: A Taxonomy, Review, and Future Directions, ACM Computing Surveys, 2022
- 【 24 】 Gupta A., Key Pillars of a Comprehensive Data Fabric, Gartner, 2021
- 【 25 】 Microsoft Azure, Azure 认知搜索中的知识存储, 2023, <https://learn.microsoft.com/zh-cn/azure/search/knowledge-store-concept-intro?tabs=portal>
- 【 26 】 Baltrušaitis T., Ahuja C., Morency L. P., Multimodal machine learning: a survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019
- 【 27 】 Wilson G., Cook D. J., A Survey of Unsupervised Deep Domain Adaptation, Association for Computing Machinery, 2020
- 【 28 】 Yu Y., Wu C., Zhao T., OPU: An FPGA-based overlay processor for convolutional neural networks, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2020
- 【 29 】 国家信息中心, 数据中心绿色高质量发展研究报告, 2022
- 【 30 】 Hennessy J. L., Patterson D. A., A New Golden Age for Computer Architecture, Communications of the ACM, 2019

附录 B：缩略语

缩略语	英文全称	中文全称
ABE	Attribute-Based Encryption	基于属性加密
AIoT	Artificial intelligence of things	人工智能物联网
CBA	CMOS Banded Array	CMOS 与 Array 键合
CFET	Complementary Field Effect Transistor	互补场效应晶体管
CIM	Compute In Memory	存内计算
CMOS	Complementary Metal–Oxide–Semiconductor	互补式金属氧化物半导体
CNA	CMOS Next to Array、CMOS Near Array	CMOS 在 Array 旁边
COP	Coefficient of Performance	性能系数
CUA	CMOS Under Array	CMOS 在 Array 下边
CUE	Carbon Use Efficiency	碳利用效率指标
DNA	Deoxyribonucleic acid	脱氧核糖核酸
DPU	Data Process Unit	数据处理单元
DRAM	Dynamic Random Access Memory	动态随机存取存储器
DTCO	Design-Technology Co-Optimization	设计工艺协同优化
EDA	Electronic design automation	电子设计自动化
HAMR	Heat Assisted Magnetic Recording	热辅助磁记录
HDD	Hard Disk Drive	硬盘
HPDA	High Performance Data Analytics	高性能数据分析
IoT	Internet of things	物联网

缩略语	英文全称	中文全称
LCA	Life Cycle Assessment	生命周期评估
LTFS	Linear Tape File System	线性磁带文件系统
LTO	Linear Tape Open	开放线性磁带
MAMR	Microwave Assisted Magnetic Recording	微波辅助磁记录
MRAM	Magneto resistive Random-Access Memory	磁阻式随机存取存储器
PB	Petabyte	拍字节, 千万亿字节
PLC	Penta-Level Cell	五层式存储单元
PUE	Power Usage Effectiveness	能源利用效率
QLC	Quad-Level Cell	四层式存储单元
SCI	Storage Compute Integrated	存储计算集成
SCM	Storage Class Memory	存储级存储器
SSD	Solid-State Drive	固态硬盘
STCO	System Technology Co-Optimization	系统技术协同优化
STT-MRAM	Spin-transfer torque MRAM	自旋转移转矩磁性存储器
TEE	Trusted Execution Environment	可信执行环境
UB	Unified Bus	统一系统总线
Wafer Level	Wafer Level	晶圆级
YB	Yottabyte	尧字节, 一亿亿字节
ZB	Zettabyte	泽字节, 十万亿字节

附录 C：致谢

数据存储 2030 编写过程中得到了来自华为内外部多方的大力支持，100 多位来自华为的专家和社会各界知名学者参与了材料的讨论、交流，贡献思想、共同畅想了 2030 年数据存储产业的发展方向和技术特征，在此对所有参与技术交流和讨论的学者们致以诚挚谢意！

（学者名单按照姓名字母排序，不分前后）

包云岗（中科院计算所，研究员）

崔鹤鸣（香港大学，副教授）

陈明宇（中科院计算所，研究员）

冯 丹（华中科技大学，长江学者特聘教授）

顾 荣（南京大学，特聘研究员）

过敏意（上海交通大学，教授，IEEE Fellow，欧洲科学院院士）

黄 勤（北京航空航天大学，教授）

蒋德钧（中科院计算所，副研究员）

金 海（华中科技大学，长江学者特聘教授，IEEE Fellow）

李 祎（华中科技大学，副教授）

刘贤明（哈尔滨工业大学，教授）

陆游游（清华大学，副教授）

缪向水（华中科技大学，教授）

任 奎（浙江大学，教授，ACM Fellow，IEEE Fellow）

舒继武（清华大学，长江学者特聘教授，IEEE Fellow）

唐 卓（湖南大学，教授）

王 聪（香港城市大学，教授）

王则可（浙江大学，百人计划研究员）

王肇国（上海交通大学，副教授）

邬贺铨（中国工程院院士）

谢长生（华中科技大学，教授）

赵世振（上海交通大学，副教授）

周 可（华中科技大学，长江学者特聘教授）

2024 年版本刷新说明

华为积极与业界知名学者、客户、伙伴深入交流，投入对智能世界的持续探索。我们看到智能世界的进程明显加速，新技术和新场景不断涌现，产业相关参数指数级变化。为此，华为对 2021 年发布的《数据存储 2030》进行系统刷新，展望面向 2030 年的场景、趋势，并对相关预测数据进行了调整。

华为技术有限公司



深圳龙岗区坂田华为基地

电话: +86 755 28780808

邮编: 518129

www.huawei.com

商标声明

 HUAWEI, HUAWEI,  是华为技术有限公司商标或者注册商标, 在本手册中以及本手册描述的产品中, 出现的其它商标, 产品名称, 服务名称以及公司名称, 由其各自的所有人拥有。

免责声明

本文档可能含有预测信息, 包括但不限于有关未来的财务、运营、产品系列、新技术等信息。由于实践中存在很多不确定因素, 可能导致实际结果与预测信息有很大的差别。因此, 本文档信息仅供参考, 不构成任何要约或承诺, 华为不对您在本文档基础上做出的任何行为承担责任。华为可能不经通知修改上述信息, 恕不另行通知。

版权所有 © 华为技术有限公司 2024。保留一切权利。

非经华为技术有限公司书面同意, 任何单位和个人不得擅自摘抄、复制本手册内容的部分或全部, 并不得以任何形式传播。