

空天大数据存储的挑战与创新

如何构建可靠的数据底座

浪潮数据

方案架构师-张文忠

目录 | Contents



行业趋势分析



数据存储方案



项目最佳实践



行业趋势分析

商业航天业务概览

■ 行业范围：

商业航天通常包括**卫星制造、火箭发射、卫星应用、太空探索**服务等领域。

■ 市场规模：

1.中商产业研究院发布的《2024—2029年中国服务器行业需求预测及发展趋势前瞻报告》显示，中国商业航天市场规模保持高速增长，从2018年至2023年，市场规模由0.6万亿元增至1.9万亿元，**年均增长率达23%**。预计**2029年中国市场规模将达到6.6万亿元**，年复合增长率约23%。

2.2024 年前三季度发射次数已超 2023 年全年，显示出**快速增长**的趋势。

A股商业航天产业链相关上市公司共计203家。

■ 行业格局：

1.第一梯队：航天科技集团和航天科工集团占据主导地位，市场份额分别为 70.1% 和 9%。

2.第二梯队：民营火箭公司如星河动力、蓝箭航天、东方空间等也在快速崛起，2023 年民营发射次数占比约 19.4%。

3.第三梯队：中小型企业，理工导航、星际空间等以航空装备零部件系统为主。

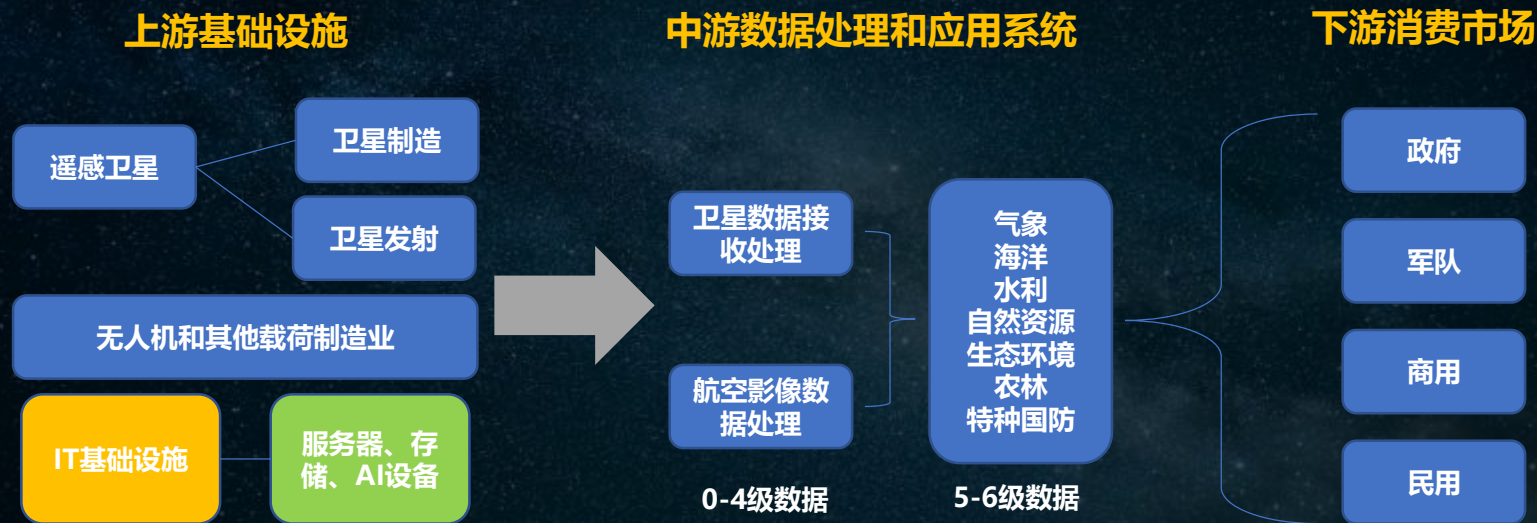
■ 发展趋势：

1.发射工业化：火箭航班化运营（如每周发射）推动成本下降 80%。

2.卫星智能化：**AI 大模型与遥感数据深度融合**，催生智能农业、智慧海洋等万亿级市场。

3.太空经济化：太空旅游、在轨制造等新场景商业化落地。

遥感行业生态



中国遥感行业生态：

遥感产业链上游为卫星制造及发射产业、航空飞机、无人机制造业及地面基础设施，包括接收设备、网络设备、IT基础设施(服务器、云存储和人工智能平台等软硬件基础设施)等。

卫星制造主要企业包括：中国航天、中国卫星、世纪空间、长光卫星、欧比特等；卫星发射企业主要包括：中国航天、零壹空间、蓝箭空间。

遥感产业链中游主要为遥感数据接收处理企业，遥感应用服务企业。

遥感数据接收处理企业主要单位包括：航天宏图(PIE)、世纪空间(PCI代理)、易智瑞(ENVI代理)、中科九度等；遥感应用服务企业主要包括航天宏图、中科星图、世纪空间、欧比特等。

卫星遥感信息化业务场景

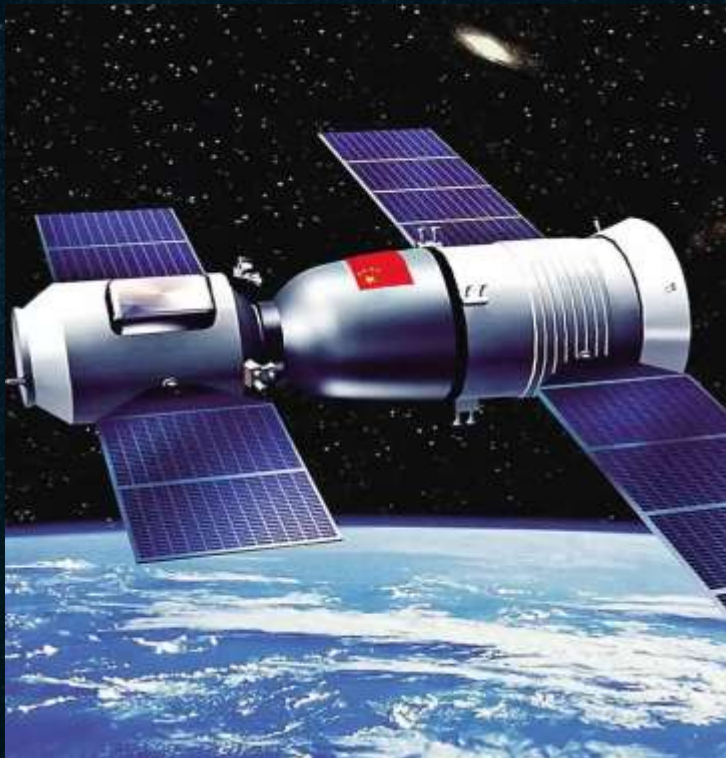
- CAD、仿真设计

- 星地协同边缘计算

- 地面测控系统自动化

- 空天地一体化观测网络

- 智慧发射场



- 遥感数据的智能解译

- 卫星智能化生产线

- 卫星数据交易平台

- 卫星数据全生命周期管理

- 在轨卫星的智能运维

场景化融合（商业航天+自动驾驶，商业航天+智慧城市）

信息化趋势-卫星遥感AI人工智能

大模型在卫星遥感技术中应用场景

智能数据处理

地物识别

动态监测

边缘计算



计算机视觉

图像识别
目标定位与检测

语义分割
三维重建



自然语言处理

遥感报告自动生成
地理空间知识图谱构建

智能检索系统



人机互动

可穿戴机器人
3D空间交互

实体沙盘+实时投影
脑机接口深度应用



智能语音

语音录入
智能导航

语音控制



知识图谱

多模态知识抽取

时空知识表示
星地协同

智能数据处理

- 深度学习自动校正图像（如NASA使用CNN消除云层干扰）
- 超分辨率重建（ESA的EnMAP卫星数据分辨率提升4倍）
- 多源数据融合（Sentinel-1雷达与Landset-8光学数据AI配准）

地物识别与分类

- 建筑物提取：Maxar的AI模型在迪拜实现95%自动识别精度
- 作物分类：PlantVillage系统结合Sentinel-2数据识别14类作物（准确率89%）
- 灾害评估：日本ALOS卫星+AI实现台风后损毁评估提速300%

动态监测系统

- 非法捕捞监测：Global Fishing Watch实时分析70万艘船只
- 森林砍伐预警：Amazon Conservation的RAISG系统提前14天预测风险
- 城市扩张分析：中国高分卫星+AI跟踪雄安新区建设进度

信息化趋势-行业可信数据空间

- 10月18日, 国家数据局对外公布《可信数据空间发展行动计划（2024—2028年）》（征求意见稿）（下称《计划》），《计划》提出到2028年，可信数据空间标准体系、技术体系、生态体系、安全体系等取得突破，建成**100个以上可信数据空间**

地方发展

- 2024年，上海浦东新区从健康医疗、金融、交通、气象等数据价值高、需求密集领域入手推进公共数据场景授权落地。到2025年，至少实现20个公共数据赋能产业发展和民生服务高价值案例
- 2023年，深圳数据交易所达成国内首笔基于数据空间技术的场内数据业务合作，打造我国首个智能制造领域数据空间应用案例
- 2022年，江苏省率先开展了可信数据空间行业试点工作，在家纺行业与电子行业部署可信数据空间测试床，取得良好成效

五大层次

积极推广企业可信数据空间

重点培育行业可信数据空间

鼓励创建城市可信数据空间

审慎探索个人可信数据空间

探索构建跨境可信数据空间

重点行业



浪潮观点：AI战略建设路径





数据存储方案

数据要素加速数据中心向“数据型”中心转变

卫星遥感数据已成为国家战略资源和新型生产要素，其价值创造模式正从传统的观测服务向智能化决策支持跃迁。

数据是核心驱动力

卫星遥感
观测服务决策效率提升



自动驾驶
事故发生率降低



智能制造
生产效率提升



模型、智能体、数据空间

放大

叠加

倍增

劳动

资本

土地

数据

知识

技术

管理

数据中心向数据型中心演进

东数西算
10大国家数据中心集
群



跨国公司
国内数据中心建设



物理中心

2001-2006年

联网中心

2006-2012年

云中心

2012-2020年

算力型中心

2020-2025年

数据型中心

2025年以后

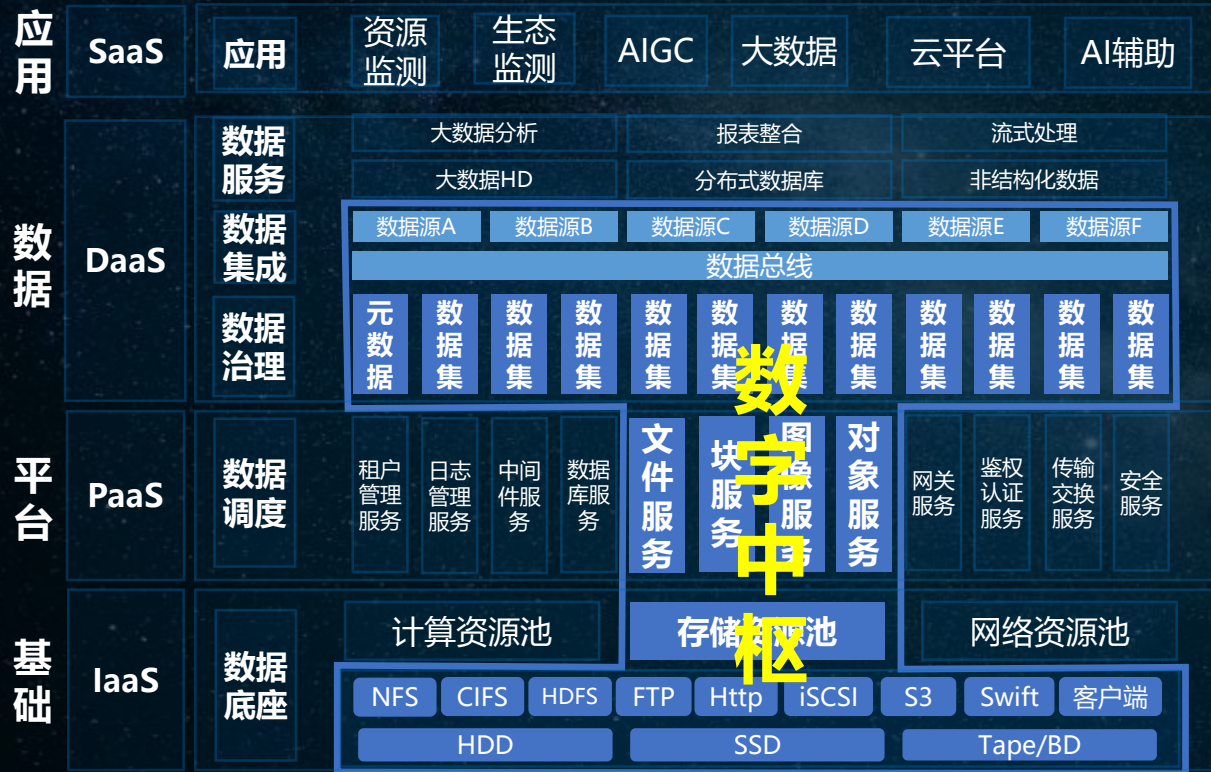
以“人”为中心

以“设备”为中心

以“数据”为中心

以数驭物，“人机物”三元融合

规划建设智慧空天大数据中心



- **空间经济计量:** 卫星数据GDP贡献率测算 (预估2030年达0.8%)
- **认知智能服务:** 知识图谱使数据价值密度提升10倍
- **星际资源勘探:** 月球/火星遥感数据商业应用

数据应用服务

存储即平台

数据资产管理

数据存储管理

数据存储底座

场景共同体

卫星遥感数据面临的挑战与问题

存储传输效率低

作业时间长



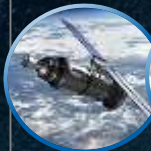
数据吞吐



- 存储设备吞吐量低
- 作业传输时间长，计算作业时间长

多介质、多平台、多传感器数据

遥感数据



监控视频



多传感器



- 海量、多源、异构的海洋数据
- 各个存储平台支撑的业务系统间无法进行数据流转

运维管理复杂

成本



维护



- 运维人员需要具备多个厂商多套设备同时运维的能力
- 运维成本高

灵活性差

数据扩展



- 无法根据新业务需求，灵活调整对外接口协议能力

提高性能

- ✓ 提高TB级带宽
- ✓ 解决数据传输时间过长，计算时间短的问题

多协议互通互访

- ✓ 实现多协议互访互通，灵活调整多协议对外服务能力
- ✓ 多业务平台间按需数据高效流转，消除数据孤岛

简化运维+降低成本

- ✓ 一套管理软件infinistor完成分布式集群存储的统一监控
- ✓ 提高运维效率，提供一体化智能运维综合解决方案

弹性扩展+灵活扩容

- ✓ EB级容量扩容空间
- ✓ 弹性扩容，在线增删节点，客户无感知，集群性能随着节点的增加而线性增长

卫星遥感数据处理整体方案设计



HPC应用



大数据 云计算平台



人工智能

高性能计算

图形图像渲染 建模仿真分析

环境科学分析

分析模拟软件

Ansys

中尺度大气模式WRF

FEPGA

地质建模EVS

分子模拟软件

模型服务+云平台服务

因子分析

决策树分析

统一报告平台

资源库平台

推理预测

资源监测

气象预报

环境监测

星际资源勘探

备份归档

数据备份

数据归档

AUS流处理

计算结果导入

SDH批处理

预处理 >> 模型训练 >> 仿真验证 >> 推理 >> 数据归档

自动建模工具

AI框架

性能分析工具

数据管理与标注平台SEED

模型开发与管理平台AI-Station

高并发仿真平台

采集处理高效存储池

融合共享存储池

训练/推理高性能存储池

备份归档存储池



存得多

EB级扩展，满足未来极致容量需求



算得快

AS13000，满足极致性能需求



用得好

多协议支持，支持文件、对象S3、
大数据HDFS、POSIX

面向HABC的存储解决方案



数值模拟



AI预测



大数据分析



数据建模

智能融合存储系统

EB级超大规模智能存储编排



解耦可组合存储
软硬件协同设计

AI使能全生命
周期数据编织

多模态数据层次化
异质存储融合架构

方案架构：解耦、融合、扩展

数据汇聚

实测数据

遥感数据

社会科学数据

数据编排

工作流驱动

主动特征析取

索引检索

软硬协同优化，高效数据流动

数控分离架构，单节点带宽大于**100GB/s**

硬件加速设计，**30TB/s**集群带宽，**10w+**高并发需求

存储空间压缩，降低客户成本

高密硬件存储，最高105盘位，单节点**2.5PB**

空间压缩算法，最高可压缩**40%**

智能数据管理，更加便捷高效

数据热度识别分级，自动迁移、删除

热温冷数据全生命周期管理，**保证成本和性能**

面向人工智能应用的存储解决方案



KV存储加持, 提升GPU利用率

计算过程和结果写入缓存, Prefill时延降低**80%**
GPU多轮利用, 可处理每秒请求数量提升**3-8倍**

横向数据流动, 提升存储效率

多种协议实时互通, 实现数据即时共享
数据冷热分层自动流转, TCO降低**20%**以上

纵向数据访问, 加快训练速度

智能缓存优化, 小文件性能提升**5倍**
GPU直通存储, 大模型训练时间缩短**50%**

面向可信数据空间的存储解决方案

分布式智能融合存储

数据归集

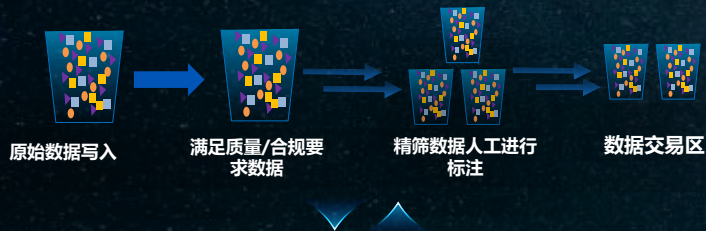
原料区

加工区

粗加工语料区

精加工语料区

运营区



块

文件

KV

对象

大数据

协议融合

SSD

HDD

磁带

蓝光

公有云

介质融合

统一存储池，数据高效归集

统一元数据管理、全局统一视图
数据加密、可信流通

存储资源优化，加速数据处理

数据编织服务，**提升标注效率30%**
AI智能调度算法与数据流转策略，**快速完成数据流转**

智能数据管理，更加安全高效

数据中心级全生命周期存储，**管理效率提升5倍**
多模态混合应用负载，**TCO降低20%以上**

面向数据全生命周期管理的解决方案



面向遥感卫星行业的融合存储系统

多协议融合互通架构 多元异构集群纳管

NFS/CIFS/FTP

S3/Swift

HDFS

数据免迁移 即时共享 语义无损

文件

对象

大数据



大模型存储



通用型

性能型

容量型

KV型

AS13000G7系列

TB级带宽
千万级IOPS
EB级容量
推理加速

新一代数据加速引擎 端到端性能优化



IO聚合压缩

智能预读

分布式
并行客户端

Data Turbo架构平台

字节级数据锁

全局
一致性缓存

分布式哈希表

数控分离

盘控一体

元数据索引



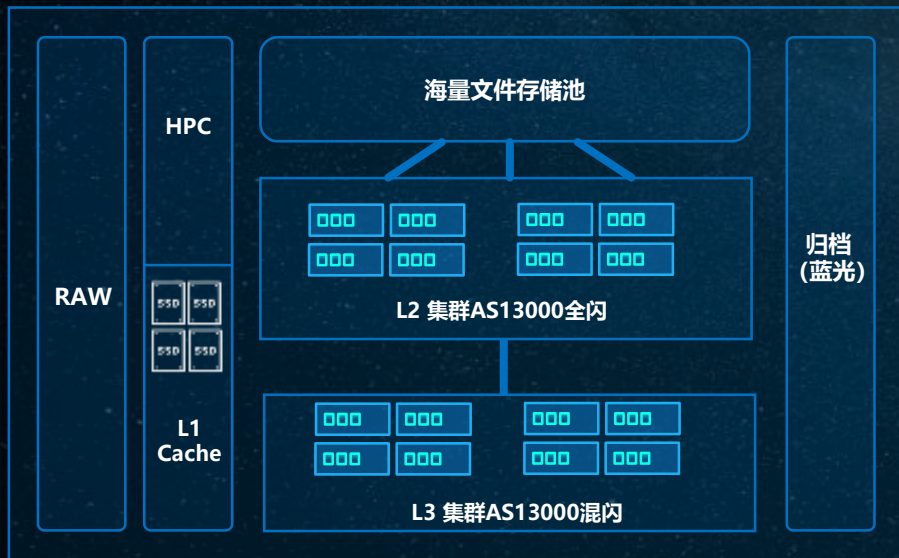


项目最佳实践

最佳实践：助力FAST探索宇宙起源

200+节点，为FAST大科学装备提供超**百PB**大容量全生命周期存储解决方案

国家天文台 FAST数据存储与计算



客户需求

- 世界上口径最大、灵敏度最高的射电望远镜FAST，具有强大的脉冲星搜寻能力，年观测时长超过5000机时
- 数据接受率最高可达38GB/s，提供足够的访问带宽和较低的延迟
- 年观测数据量10PB-12PB，海量观测数据长期保存

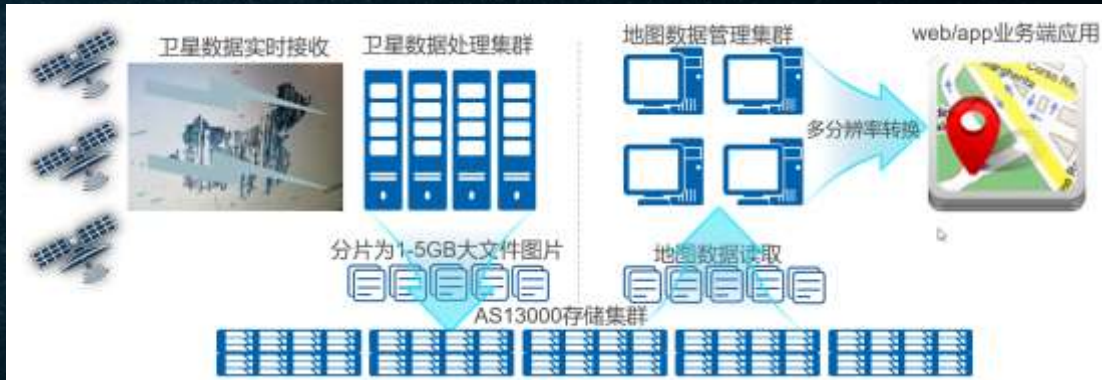
解决方案

- 提供全闪文件系统，80GB/s读写带宽
- 提供分级存储方案，AS13000+ 蓝光，实现数据的全生命周期管理
- 采用4U60高密节点，节省TCO和机房空间

最佳实践：ZY卫星中心存储平台

300+节点，百PB+，远程数据容灾，数据全生命周期存储解决方案

卫星	轨道	原始数据 日增长	0级数据 日增长	产品生产 日增长	归档数据 日增长	每卫星 日增长
ZY302	6条	150GB (25*6)	1.46TB	1.5TB	1.8TB	4.9TB
GF6	15条	450GB (30*15)	5.8TB	1 TB	5.13TB	12.3TB
GF1B	8条	240GB (30*8)	2.4TB	1 TB	1.5TB	5.13TB
GF1C	8条	240GB (30*8)	3.1TB	1 TB	1.5TB	5.13TB
GF1D	8条	240GB (30*8)	3.1TB	1 TB	1.5TB	5.13TB
日均总增长量						32.59 TB
年均总增长量						11.89 PB



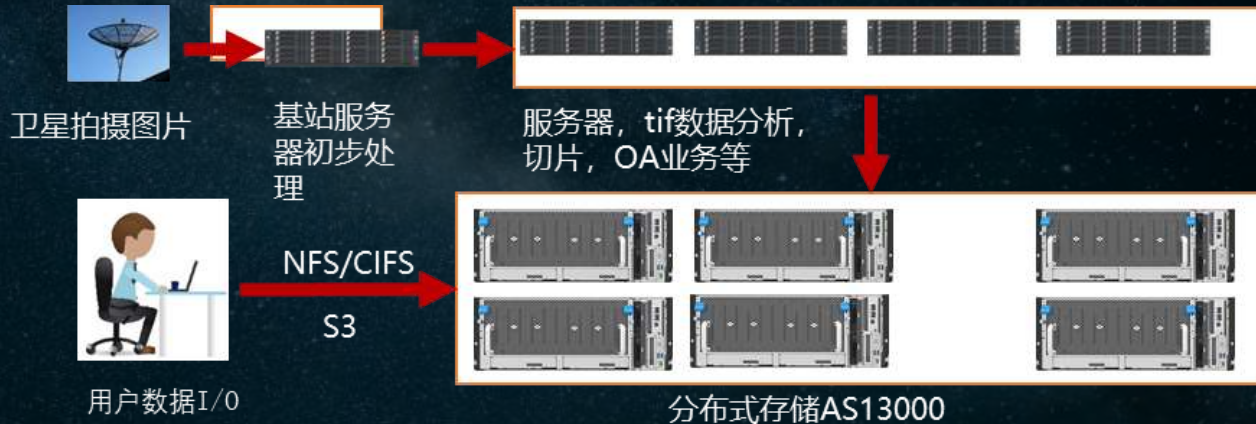
- 支撑“资源一号02C”、“资源三号”共三颗资源卫星处理能力，后续会增加到10颗卫星。
- 日处理3万张卫星资源监测照片，覆盖250万平方公里；
- 接收到的卫星信息经过数据处理集群后形成多个1G~5G大小的tiff文件，写入到存储集群中
- 要求单个节点读写带宽900MB，总带宽随节点数线性增加，不低于40GB/s，7*24小时不间断运行
- 卫星数据图片的存储、高速访问，大文件小块IO
- 单存储池存储容量要求大于70PB
- 将遥感卫星产生的原始数据，通过hadoop计算集群和GPU异构计算集群计算，产生1级到4级的不同的处理数据，并将数据存入分布式，对外售卖或供各类机构使用。

最佳实践：CG卫星数据融合存储平台

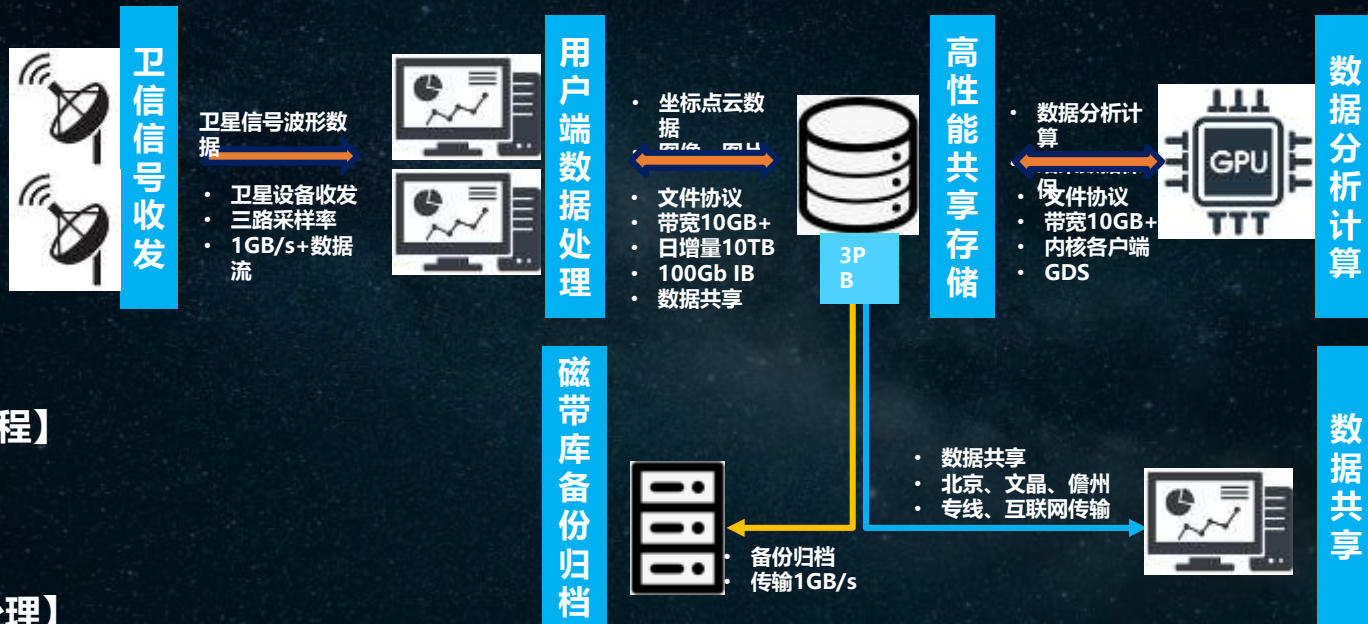
45PB+卫星数据，协议融合满足数据处理、加工和共享需求

- CG卫星是我国第一家商业遥感卫星公司
大容量归档存储，数据归档业务（遥感图像10GB大文件tif，小IO顺序读写图像分析），**45PB**，**融合互通**

- **容量：45PB**
- **目录结构：4级**，按卫星名-编号-年月时间
目录下文件数量不等，基本上都是上G的大文件，也有小量KB级描述文件，文档等
- **文件**：几兆到几G不等，大多数为大文件



最佳实践-DZ所雷达数据生命周期管理



【数据流程】

【数据处理】

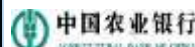
- 数据采集：非相干散射雷达信号发送/接收
- 数据处理：处理、标记、转换为数据格式
- 数据存储：生成数据进行保存
- 分析计算：GPU/CPU集群进行数据分析计算
- 数据共享：科研数据多站多用户访问，数据共享
- 备份归档：配置数据生命周期管理策略进行备份归档

【存储优势】

- 提供高性能并和存储方案，支持全闪、混闪数据分层，可按需扩展。
- 支持IB RDMA高速网络，支持GDS访问
- 支持纠删码冗余技术
- 支持数据生命周期管理方案

行业用户11000+，累计出货量15万台

金融



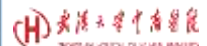
通信



政府



医疗



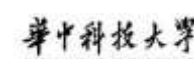
能源



交通



教育



制造



场景共同体，场景化定制



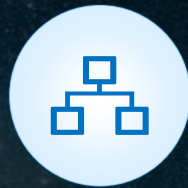
应用创新



方案共创



产品定制



市场协同



服务外包



技术整合

$$1 + C_6^n$$

THANKS



更多资讯，请关注浪潮数据存储公众号