

Industrial Trend of Distributed Storage



A pearl of wisdom....

Heraclitus 535 BC

„There is nothing immutable except the need to change “

Enrico* 2023 AD

„There is nothing immutable except the need to have more space for storing data “

*Enrico Venuto, Politecnico di Torino



about me...

Gaetano Pastore (Kai for whom cannot remember/speak my name 😊)...

Master Degree in Electronics Engineering @ University of Naples Federico II

20+ Years Experienced in the IT European market

SSD/Flash Storage in Europe

Software Defined Storage, Flash Accelleration SW, Object Storage

Cloud-native Distributed Storage for Kubernetes, Openstack....

Since 2020 in Huawei as Ecosystem and Industry development



New Connections and Applications Accelerate Data Production and Mobility

4 billion people connected **➤** 75 billion things connected

5G

30x
Traffic

1,000,000x
Computer graphics performance

8K/AR/VR

5G



Diverse applications

AI



Mass data

Cloud computing



Real-time processing

60 TB/day
Training data

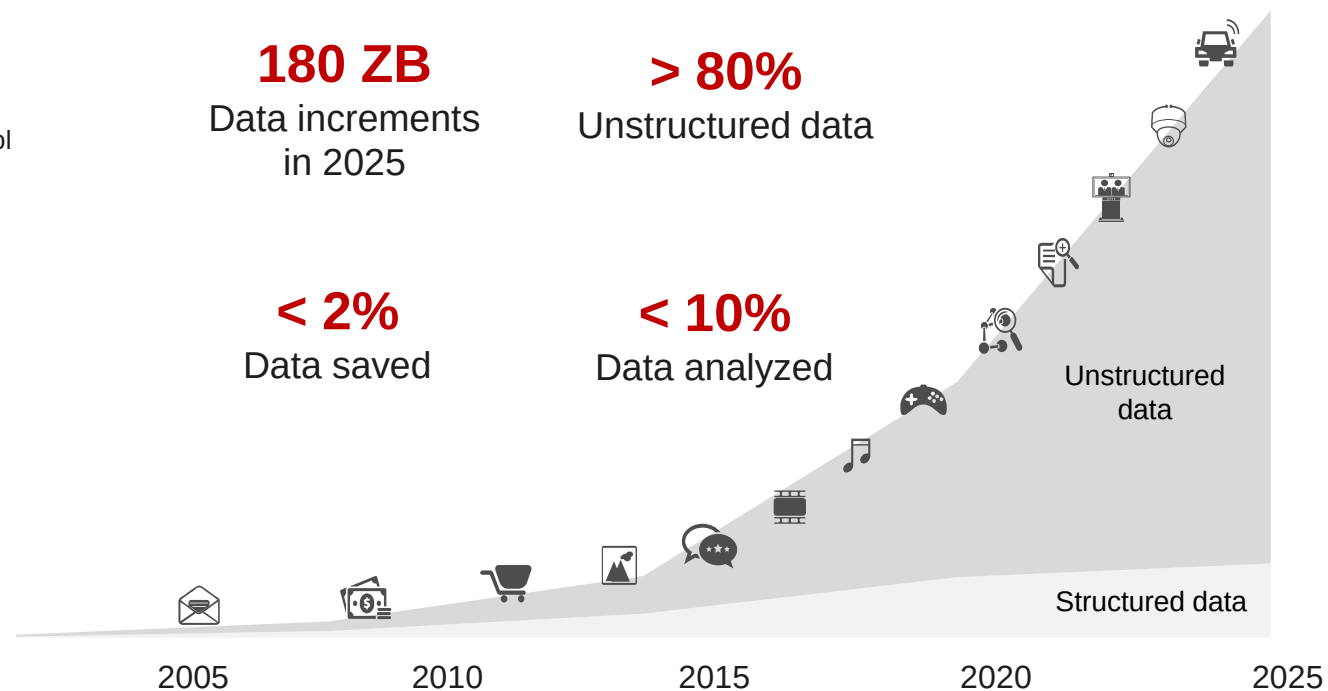
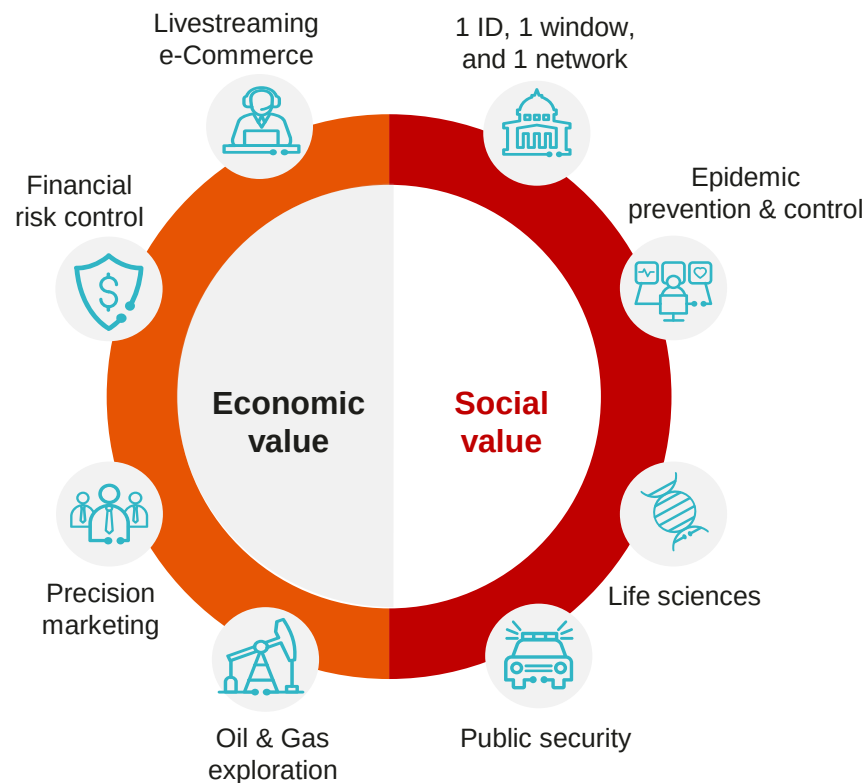
Autonomous vehicle

1 PB/day
Production data

Digital, connected factories

Data infrastructure

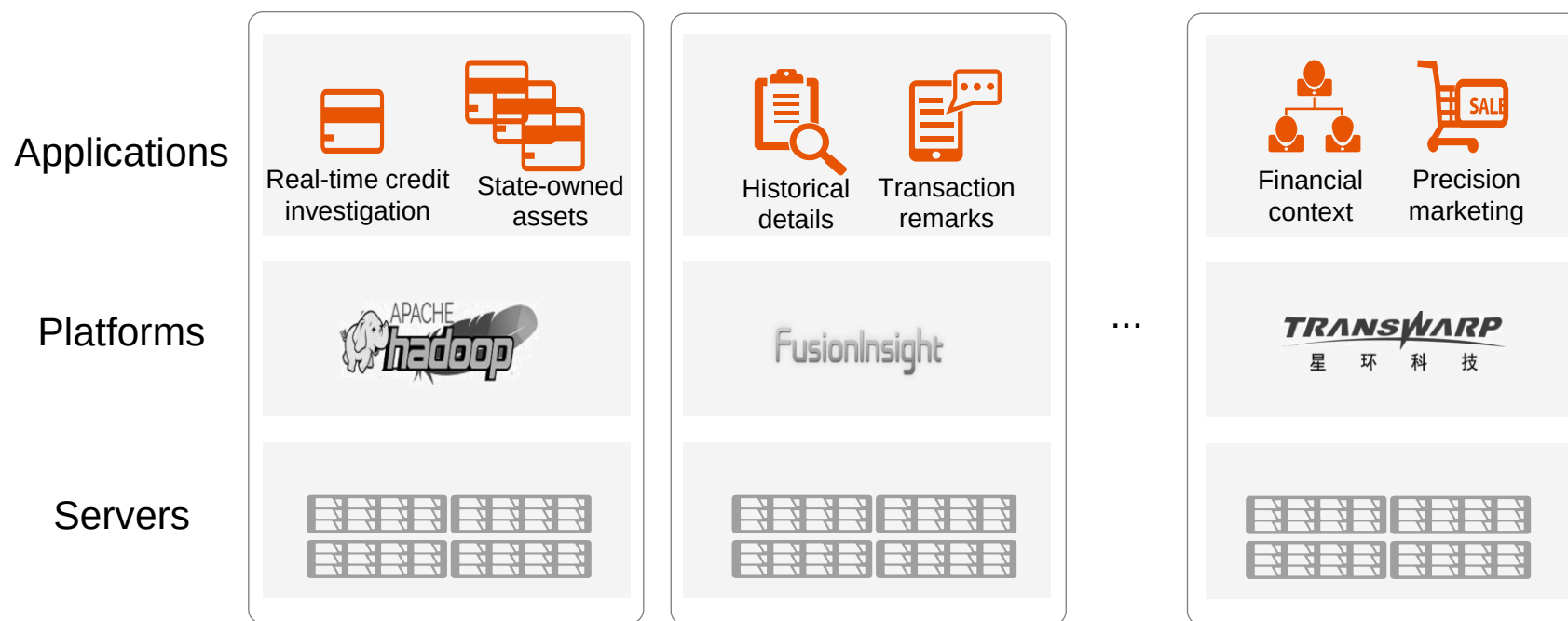
In the Era of Mass Diversified Data, the Value of Data Is Widely Recognized



Some industry challenges in different verticals

Finance: Numerous Big Data Application Silos Hinder Data Sharing

A bank with 14 business departments and 66 big data clusters



Inefficient deployment

Mobile strategy acceleration

Physical deployment with days of capacity expansion

Low resource utilization

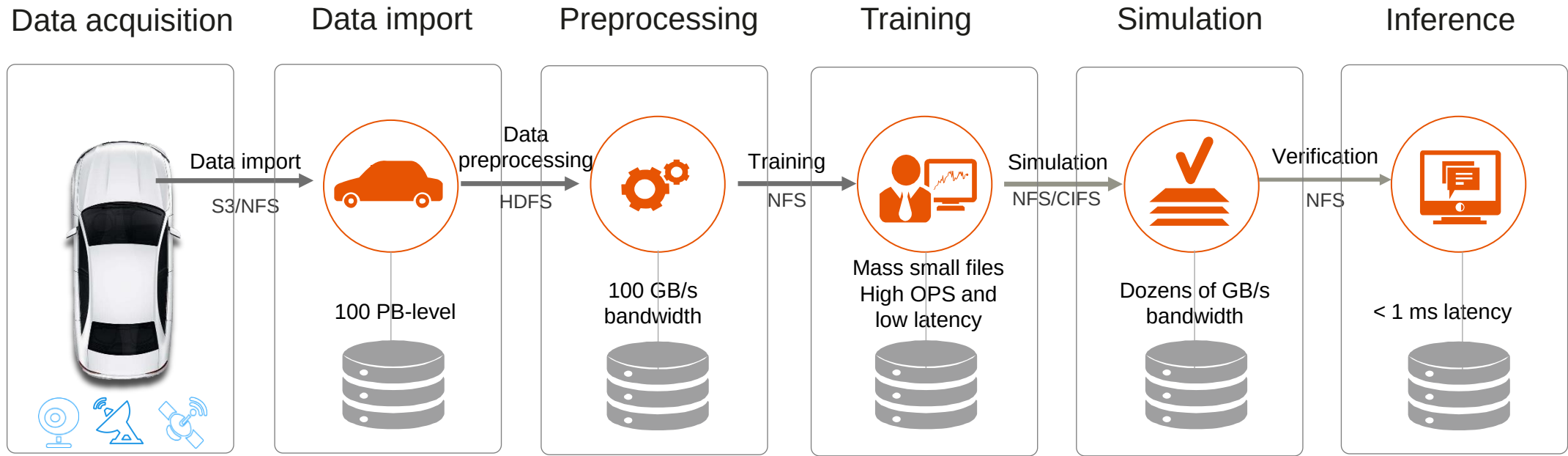
Independent expansion impossible due to storage-computing coupling

11% of memory utilization and 35% of disk utilization

Inefficient analytics

Difficult data sharing due to multiple vendors and clusters

Various Protocols and Workloads in Different Phases Slow Down Data Mobility



Insufficient storage space

60 TB/vehicle per day, hundreds of PB in total

Cross-protocol data migration

S3/NFS/HDFS multi-interface access

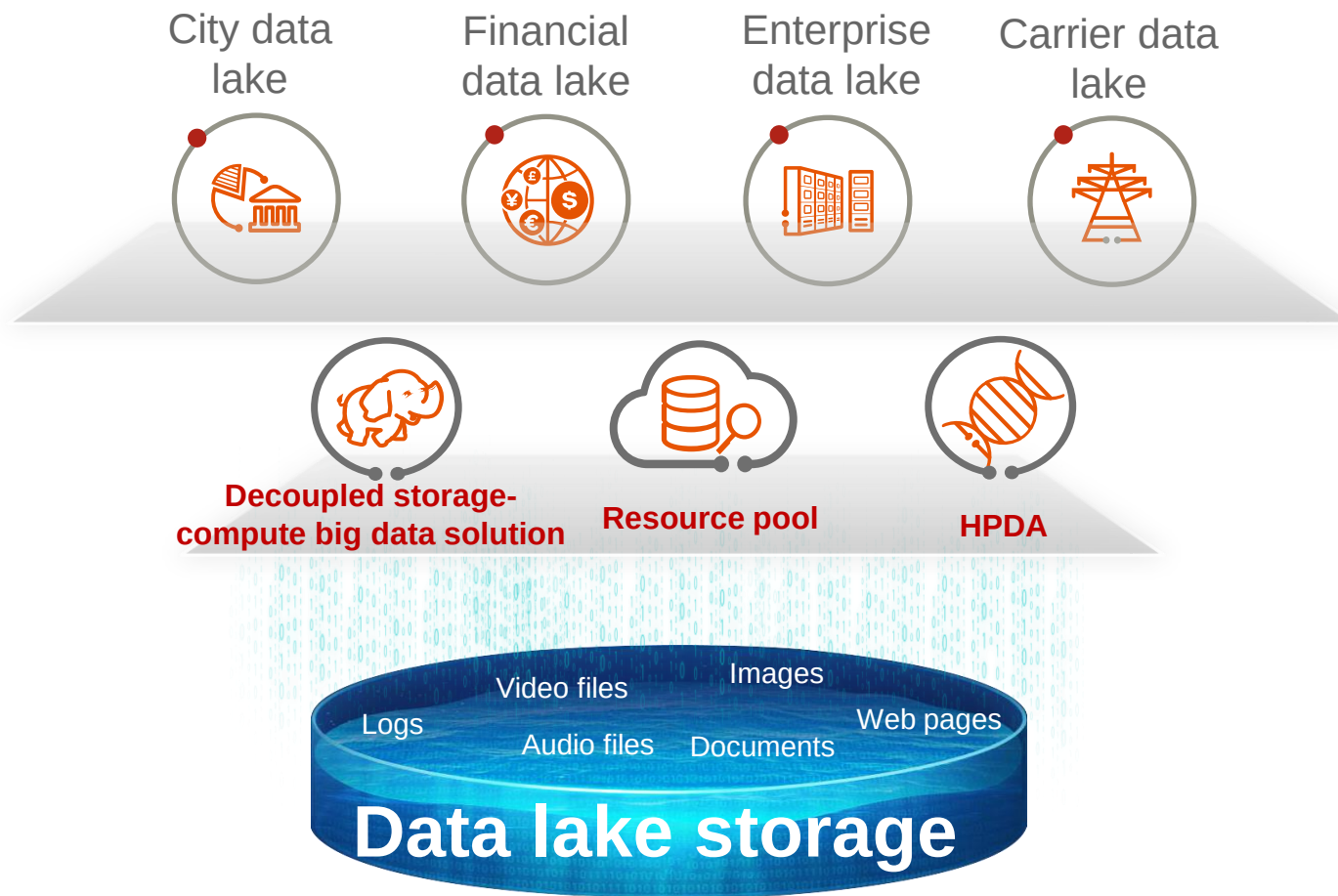
Hybrid workloads

Both high bandwidth and OPS required

Distributed storage as solutions.....



Data Lake Infrastructure to Address Data Diversity Challenges



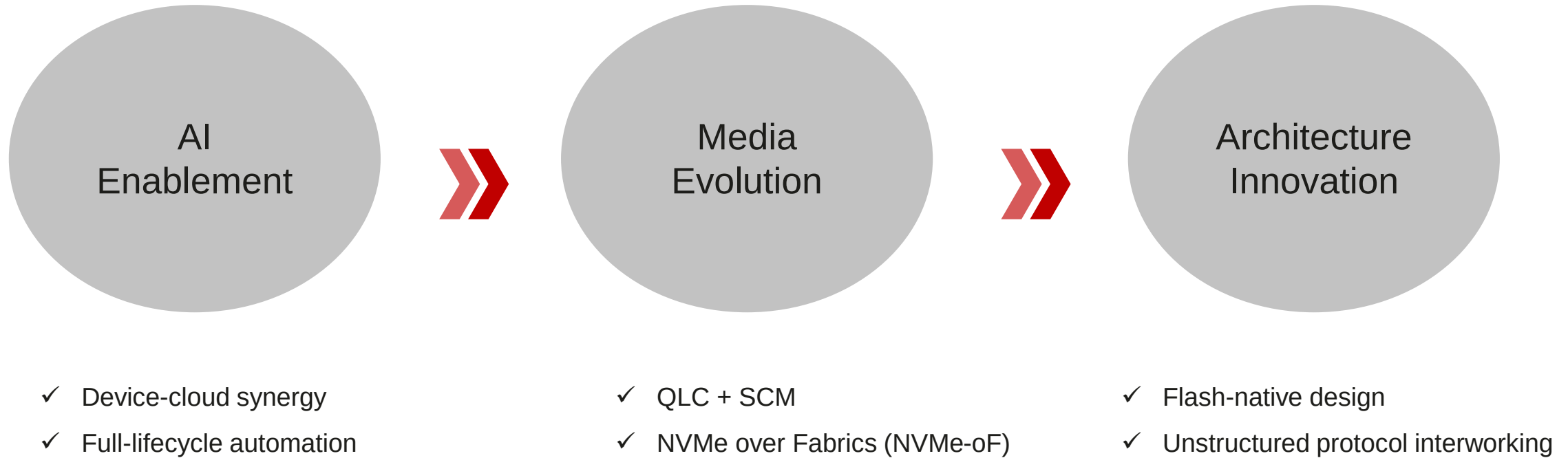
Vertical storage-compute decoupling
to break the ecosystem boundaries

Unified data infrastructure to break
the resource boundaries

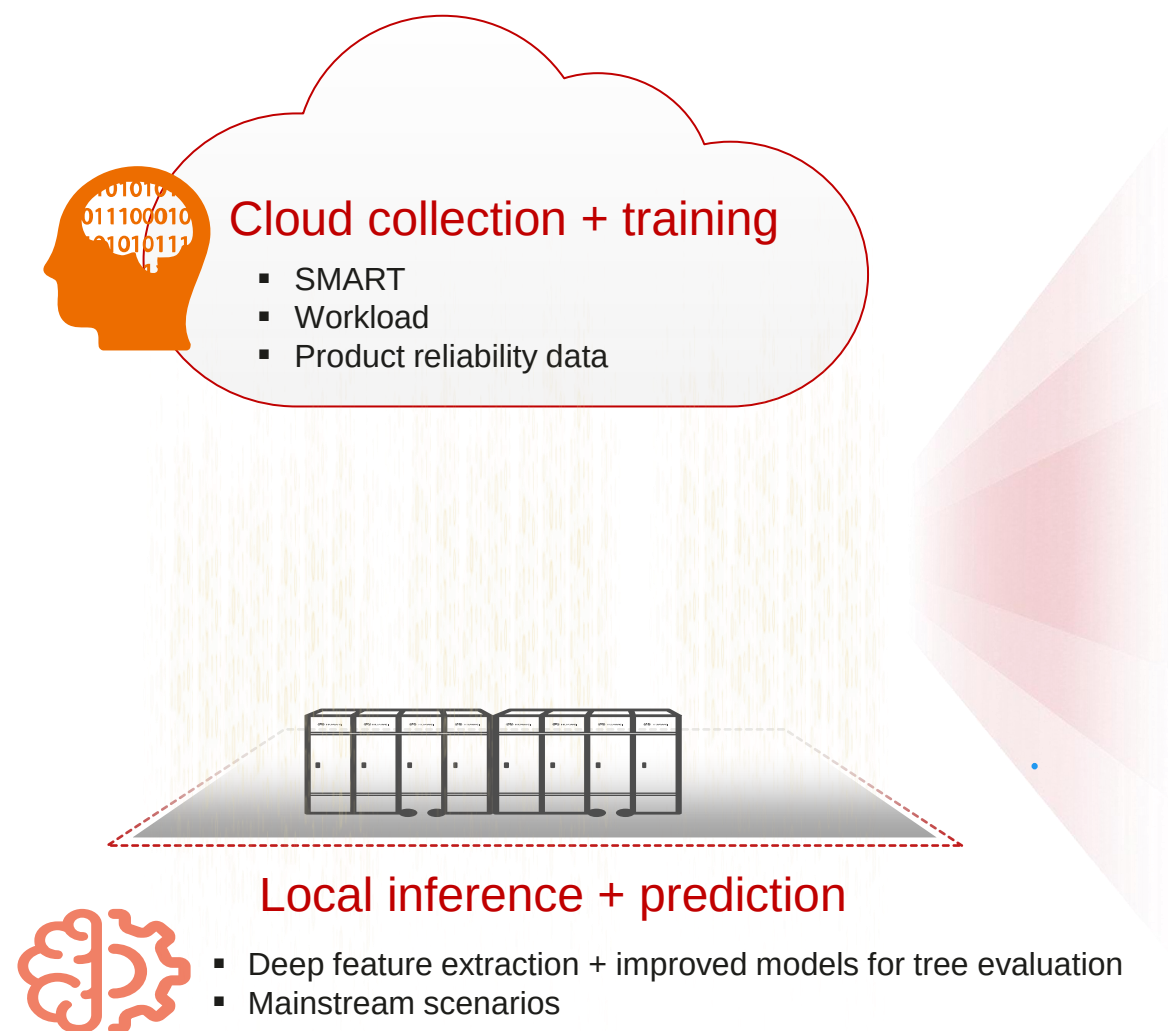
Horizontal data sharing to break the
protocol boundaries

Distributed Storage Trends...

Distributed Storage Growth This Decade



Edge-Cloud Synergy Continuously Expands Intelligent Storage



Faulty disk detection in advance

- Immediate solutions for over 90% of problems

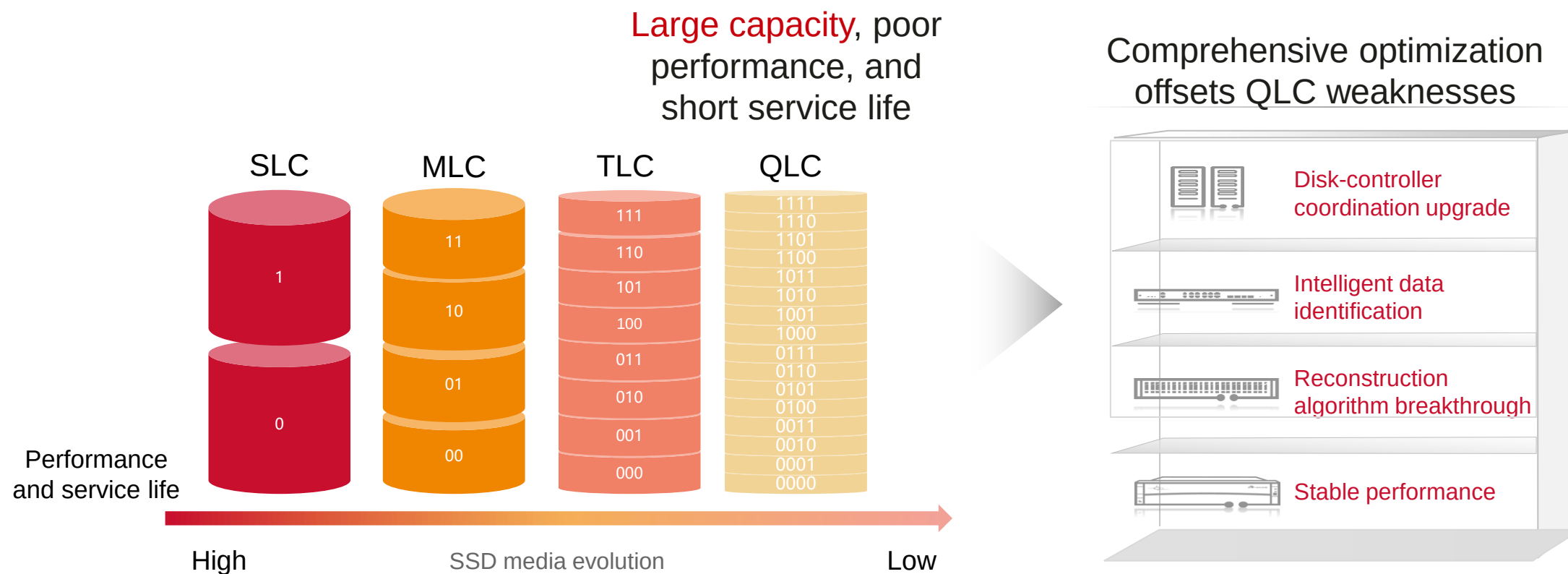
Performance optimization and prediction

- Adaptive to various performance models
- Monthly/daily/hourly tidal analysis to detect performance bottlenecks

Capacity trend prediction

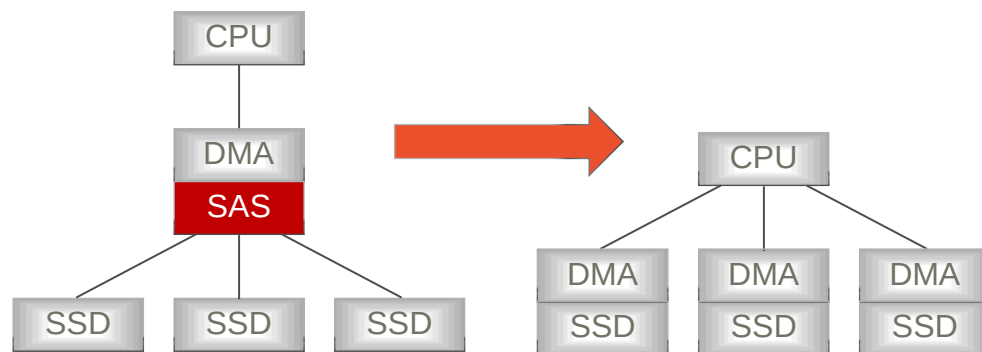
- Capacity usage analysis and prediction for flexible expansion
- Optimal solutions for capacity expansion to reduce TCO

Media: QLC Emerges as Mainstream Media for Diverse Storage Types



Protocol: NVMe Outperforms and NVMe-oF Pioneers the Network Protocols Innovation

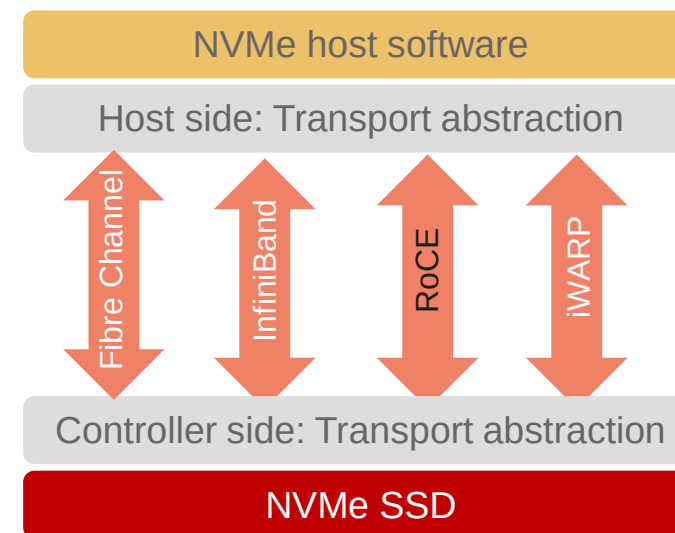
NVMe architecture eliminates transmission bottlenecks



SAS disks connect to CPU for data transfer only through the SAS controller

Each disk directly communicates with the CPU

NVMe-oF leads network protocol innovation

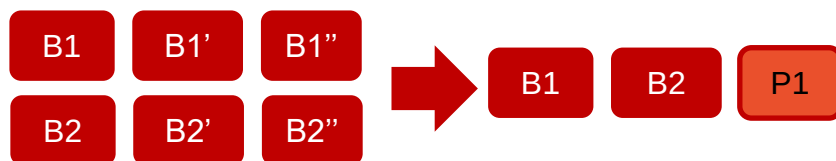


E2E NVMe architecture uses the same access method of local NVMe devices for remote NVMe access

Flash-Native Design Triggers Algorithm Innovation

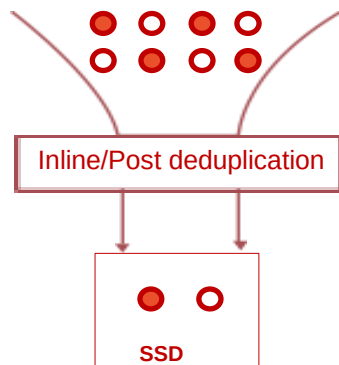
Cost optimization

Prioritize EC as copy mechanism is thrown away



Improves storage utilization and reduces write amplification

Multiple dedupe and compression mechanisms become mainstream configurations



- Adaptive variable-length deduplication
- Similarity-based deduplication algorithm
- Dedicated compression algorithm

Performance optimization

Algorithm offloading with dedicated chips

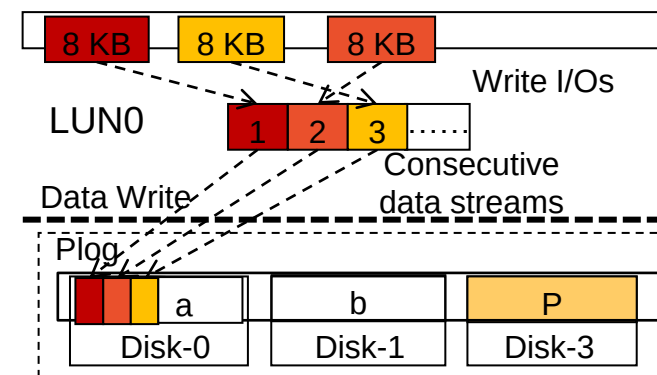


EC/encryption/compression offloading



TOE

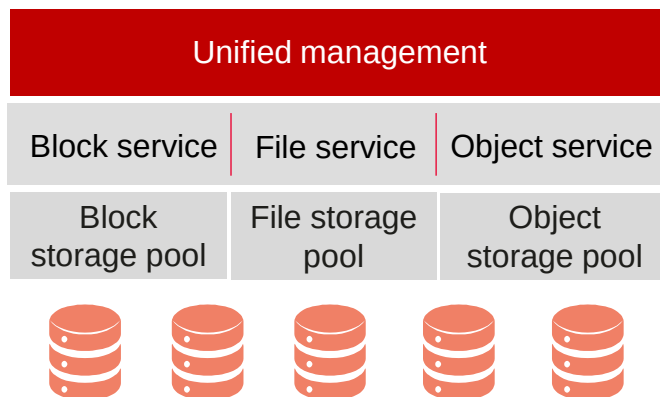
Intelligent aggregation scheme for EC with append



- Intelligent aggregation scheme for EC
- Incremental EC

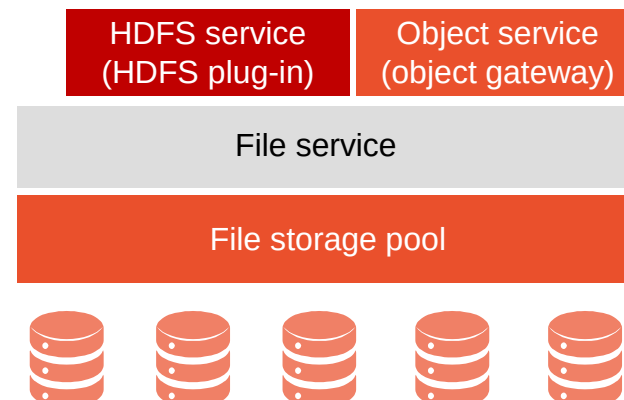
Seamless Data Convergence Makes Unstructured Data Interworking the Ideal Choice

Phase 1: Unified multi-service management



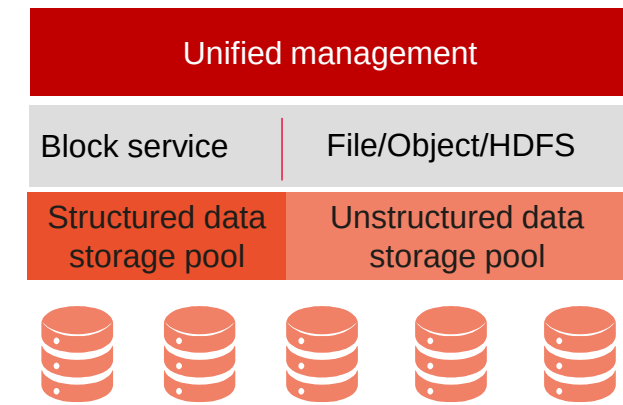
- One storage system provides multiple services for unified management and simplified O&M

Phase 2: Gateway protocol interworking



- One storage system provides multiple services, and some protocols interconnect for higher efficiency
- Gateway interworking limits performance with incomplete service capabilities

Phase 3: Uncompromised data interworking



- One storage system provides multiple structured and unstructured services for unified management
- Unified data allows for interconnection without affecting service semantics and performance

Few words about Huawei Distributed Storage.....



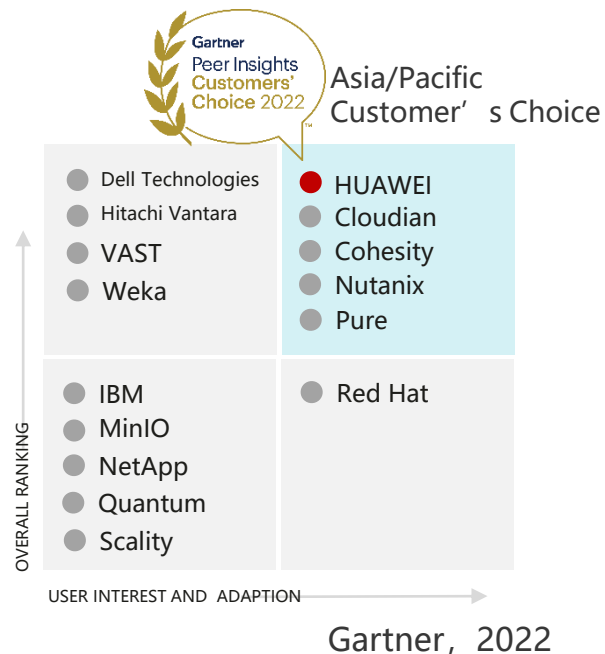
Huawei Continuously Invests in the Mass Storage Field and Wins Recognition from the Industry

Interop: Best of Show Award

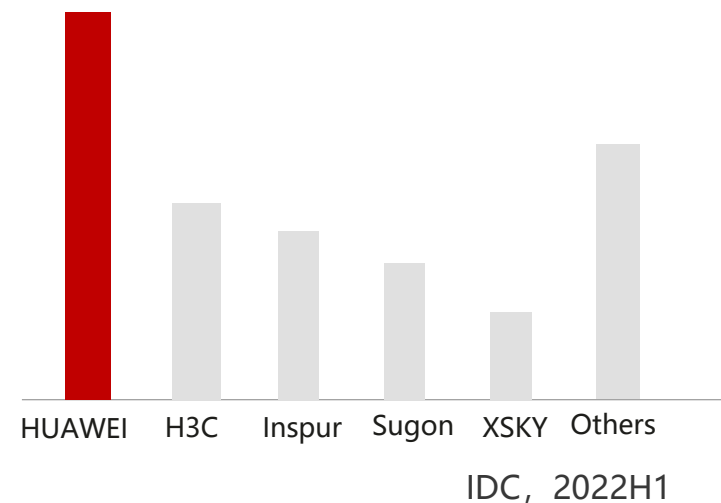


INTEROP A major event in global business technology that recognizes the most competitive products

Gartner: Customers' Choice



No. 1 market share in China for 5 consecutive years

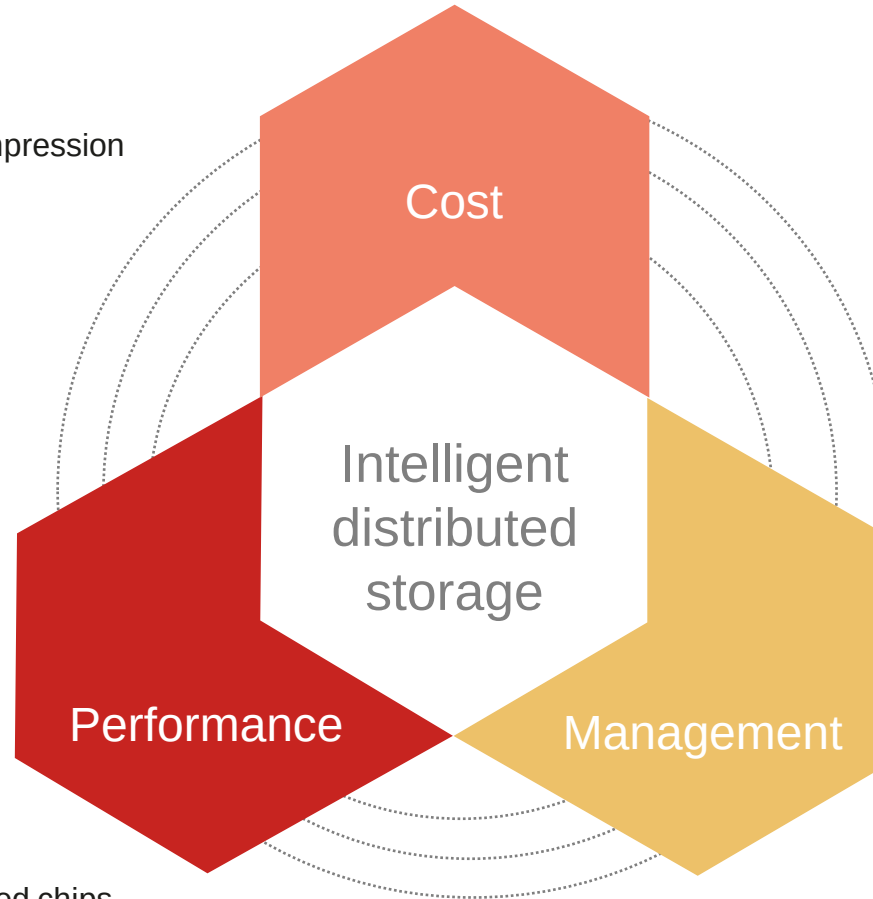


...getting back to Heraclitus.....



The Next Generation of Intelligent Distributed Storage Will Lead the Industry for the Decade

- High-density hardware
- EC + multi-dimensional deduplication and compression
- Unstructured data interworking



- All-flash hardware
- NVMe over Fabrics
- Algorithm is offloaded from CPU to dedicated chips

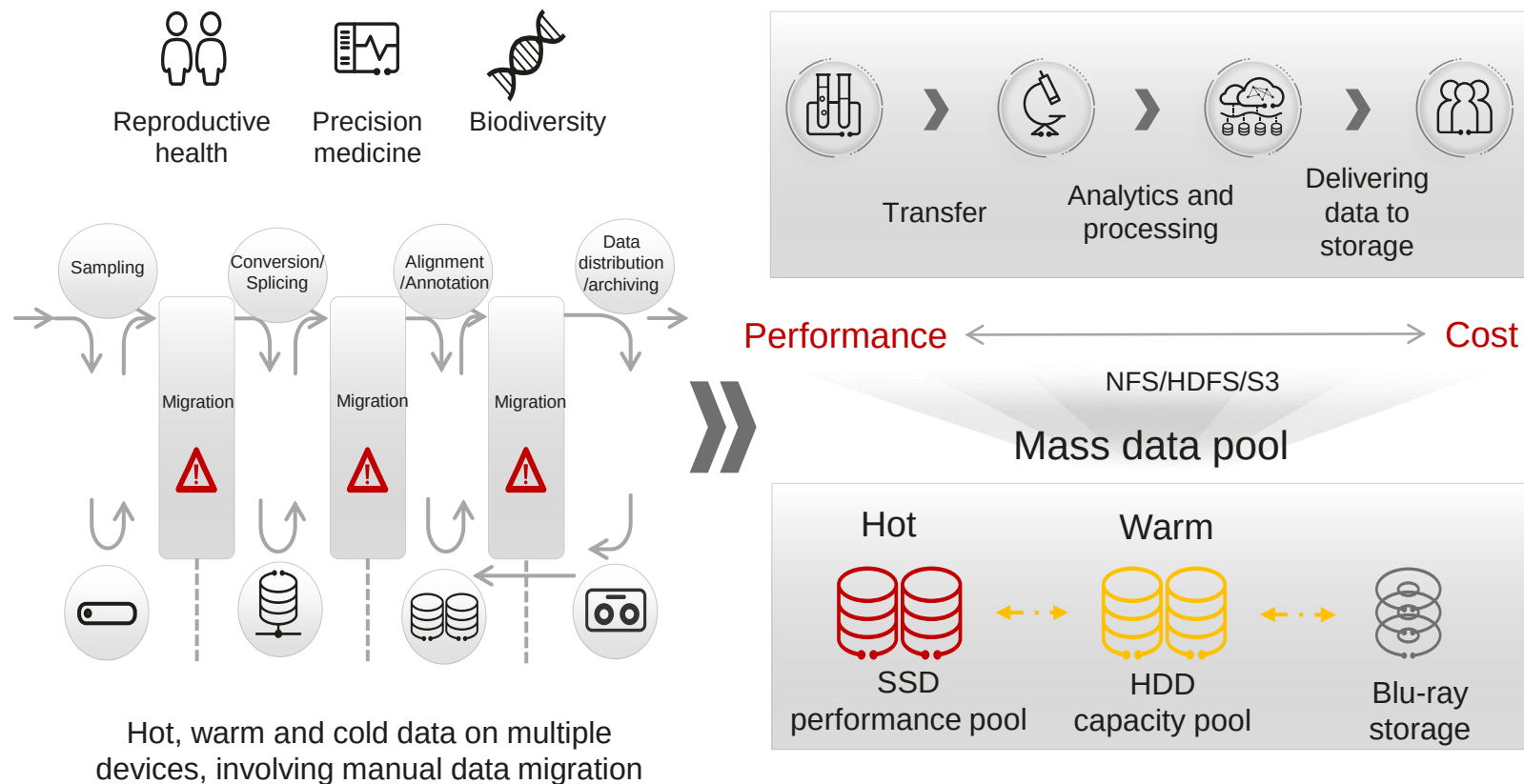
- Edge-cloud synergy
- Automated full-lifecycle management
- Unified multi-service management

Thank you.



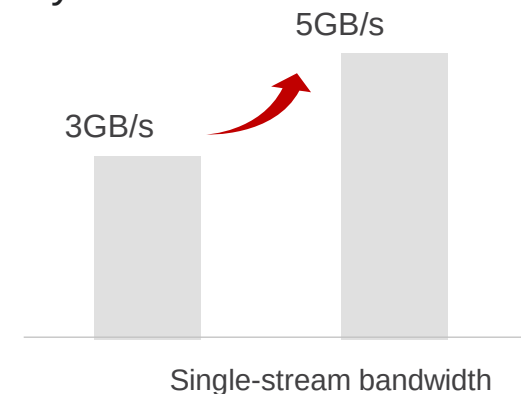
Some real-life user cases...

Genome Sequencing: Parallel Access Speeds up Genetic Analysis

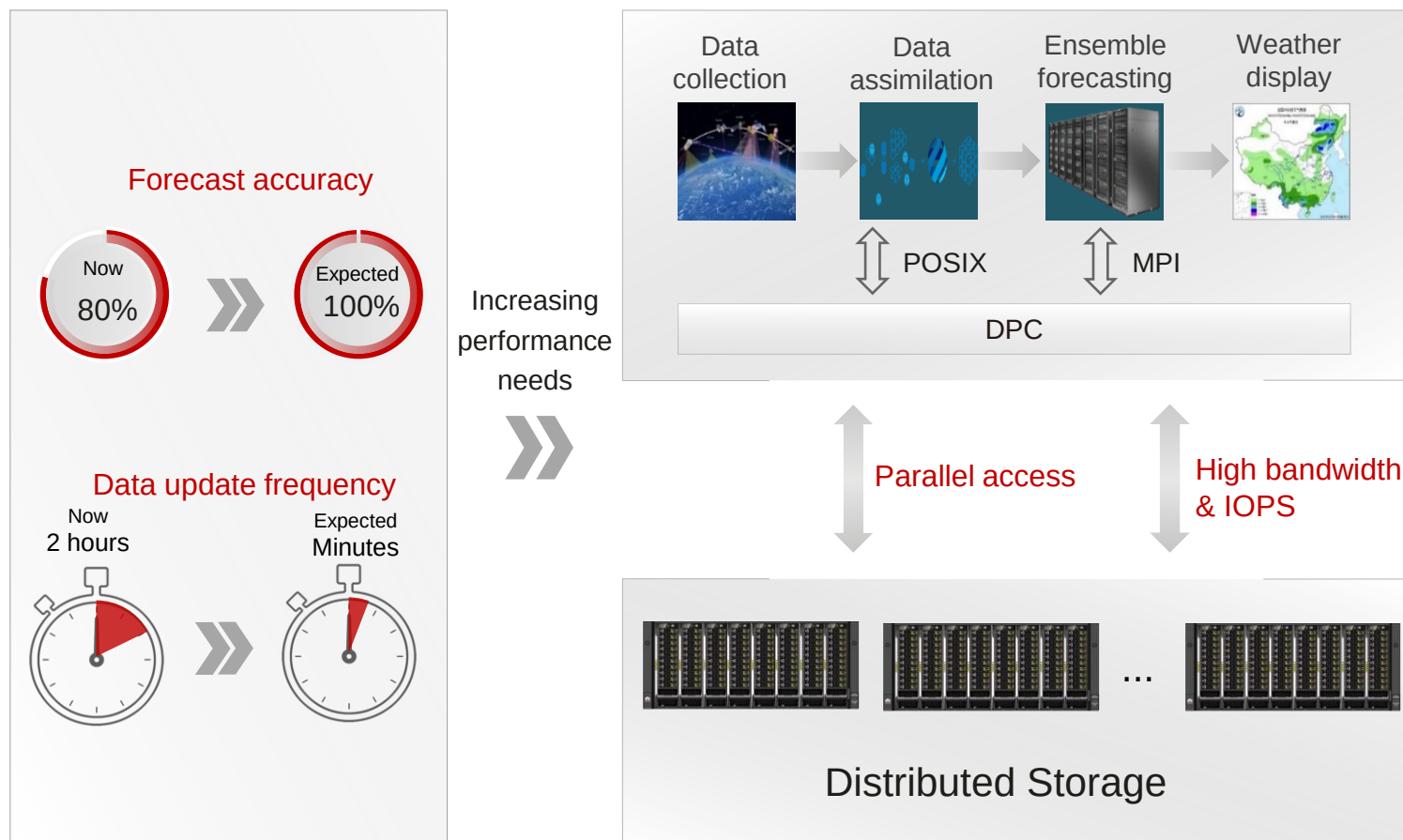


30%↑
analysis efficiency

NVMe Flash + DPC parallel client,
single stream bandwidth increased
by **66%**

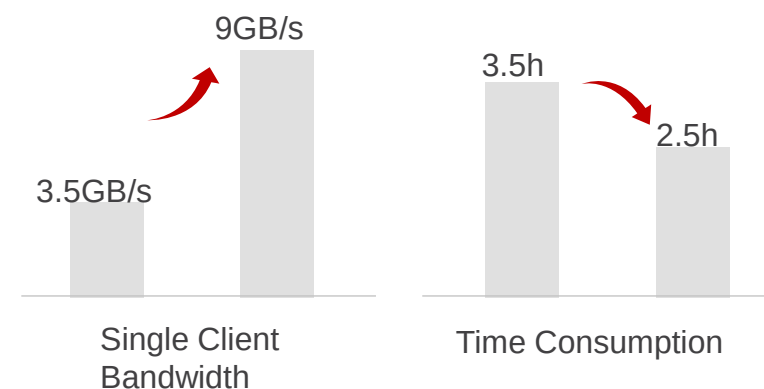


Weather Forecasting: Parallel Analysis of Massive Fata Improves the Timeliness of Forecasting

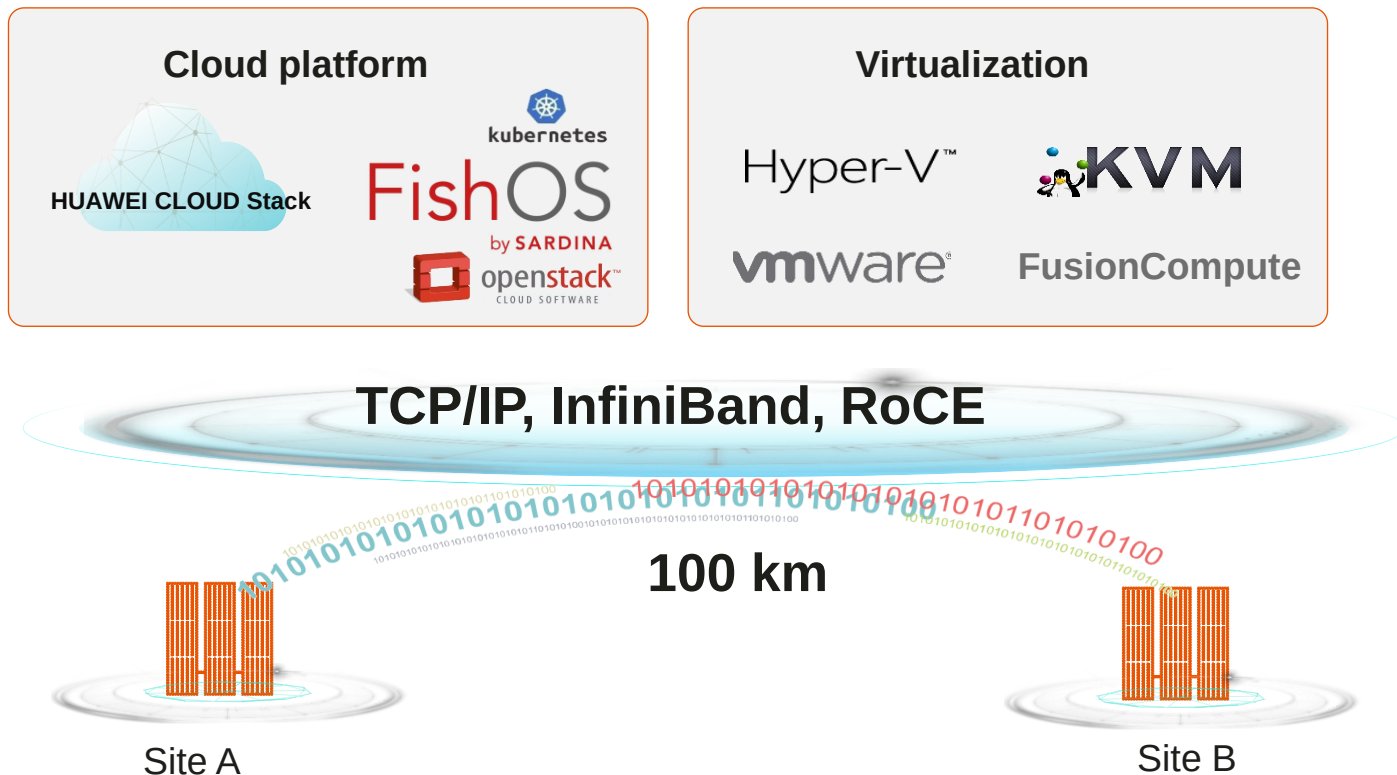


Forecasting period
30% ↓

Support for MPI-IO to allow multiple compute nodes to access the same file at 9 GB/s, with a 2.6-fold increase in single-client performance

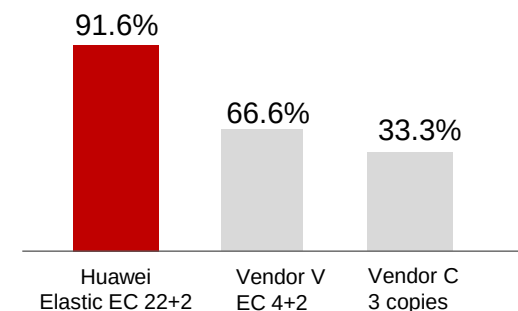


Cloud/Virtualization Resource Pool: Open Architecture with Business Continuity



Unique Elastic EC

No performance decrease compared with copies



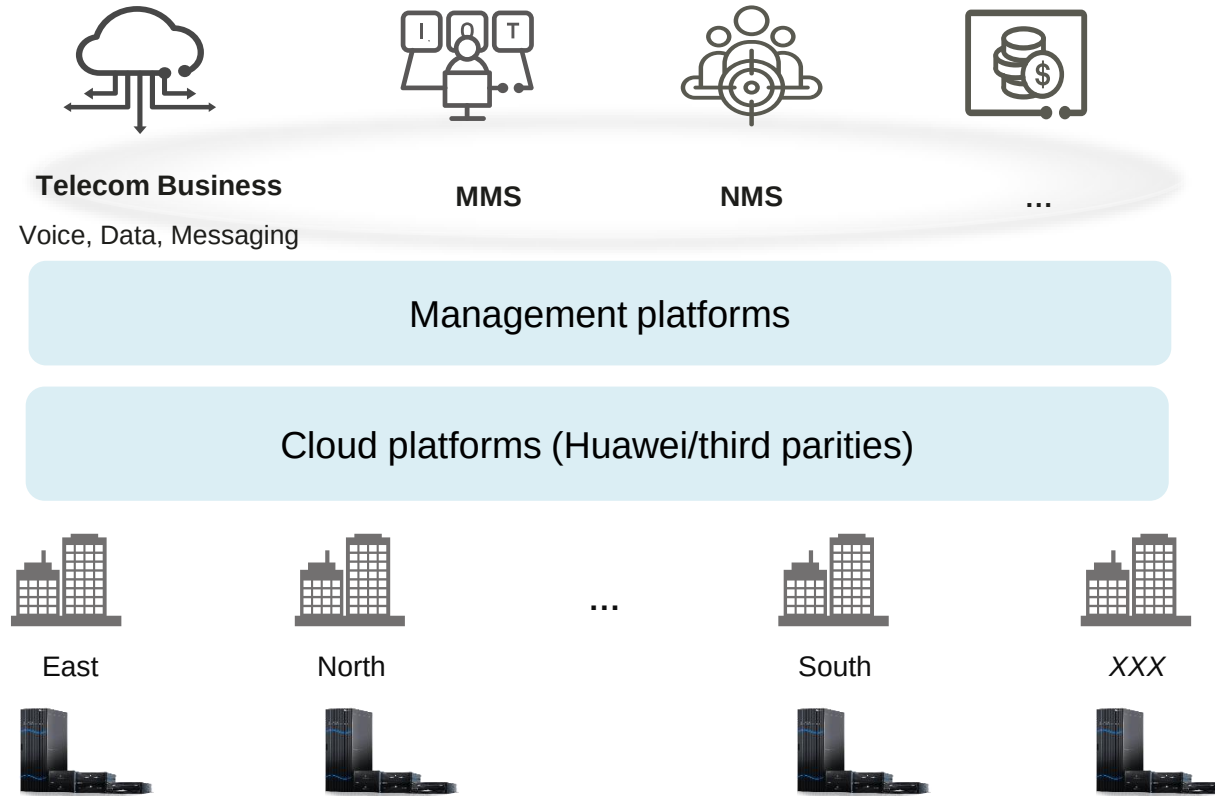
HyperMetro offering true active-active

RPO = 0, RTO ≈ 0

Cross-cluster load balancing vs. Ceph stretched cluster



Telco Operator Deploys 3500+ Scale-Out Block Storage Nodes in NFVi Scenarios



3500+ nodes commercially deployed at 8 sites

Disk capacity utilization >1x ↑

3-copy -> EC (N+2) (scaled deployment)

Seamless scalability

Automated tools + professional team, with deployment time reduced by 90%

Latency <1 ms

FlashLink acceleration, intelligent data distribution, global GC, and small I/O aggregation

99.9999% system reliability

E2E DIF consistency check

Comprehensive subhealth check and healing

Tolerates cabinet-level failures + Storage + App-level DR



Economical

Next-Gen High-Density Hardware Architectures for Mass Data

OceanStor Pacific 9950 High-density performance model



5 U, 80 NVMe SSDs, 160 GB/s

OceanStor Pacific 9550 High-density balanced model



5 U, 120 disks, 1.68 PB

Intelligent data tiering

Cross-sectional area:

65%↓

Component design

Half-palm NVMe SSDs

Heat dissipation efficiency:

30%↑

Heat dissipation design

Carbon fiber pad + VC heat spreader
Aviation-standard counter-rotating fans
Dual-layer air channel

Disk utilization:

91.6%

Algorithm design

Elastic EC
E2E DIF

Maintenance efficiency:

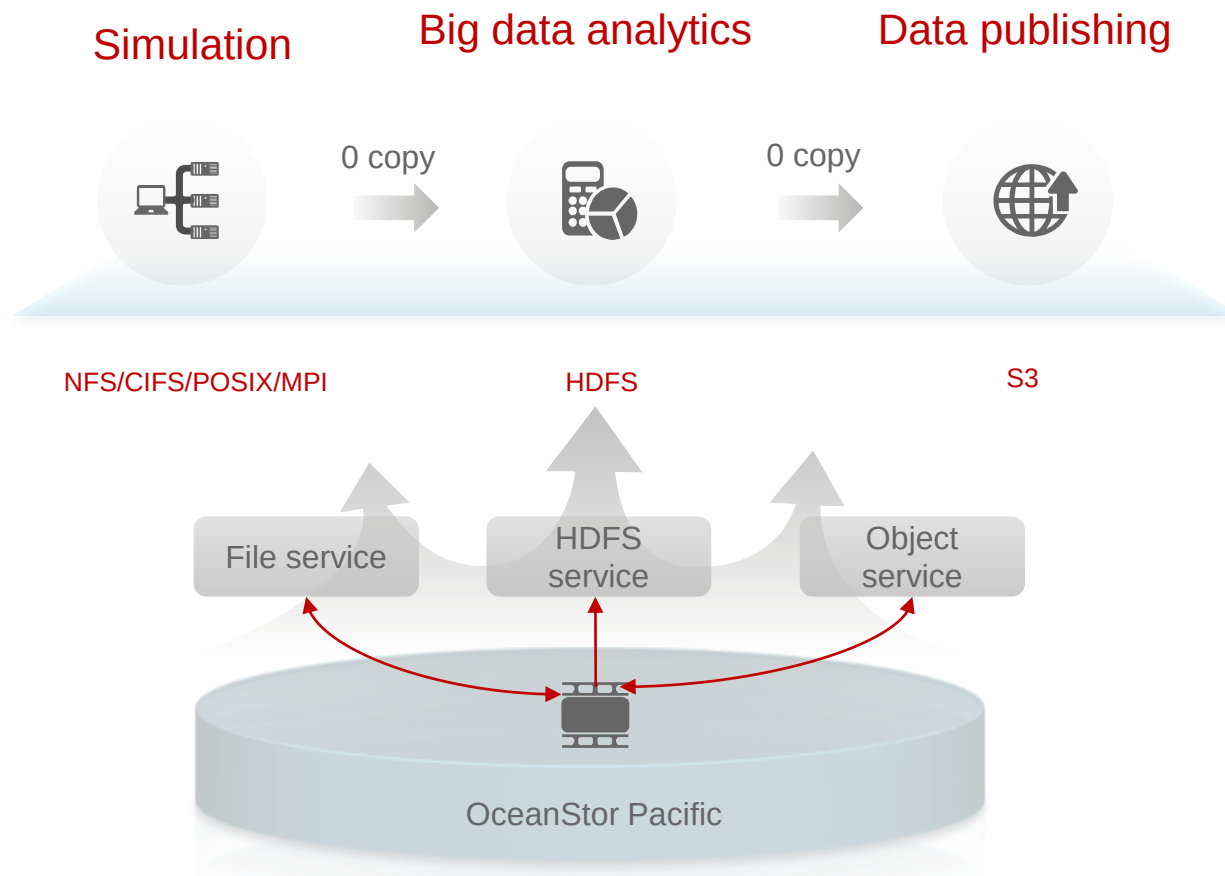
20%↑

System design

FRU
Bidirectional drawer slide
Tank chain



Efficient Multi-Protocol Interworking Accelerates Data Mobility and Service Innovation



Autonomous driving training

Mass data import, training, and analytics with zero migration



Precision medicine

Genetic data imported into data lakes, analyzed, and exported from data lakes with merely one data copy involved



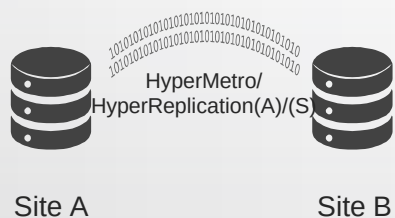
Electronic cheque imaging

Coexistence of old and new service systems and gradual object-based reconstruction

Everlasting

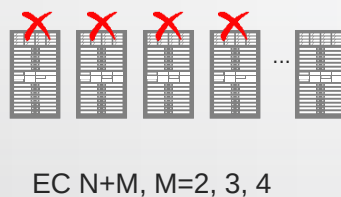
Multi-Level Reliability Mechanism for 99.9999% Availability

Solution level



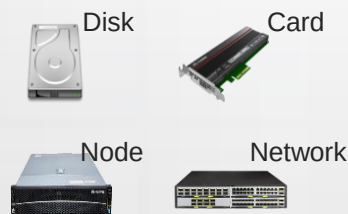
- Gateway-free **active-active** solution within 100 km
- **Remote replication**, second-level RPO
- **Cross-Sites DR**, up to 12 active sites

Cabinet level



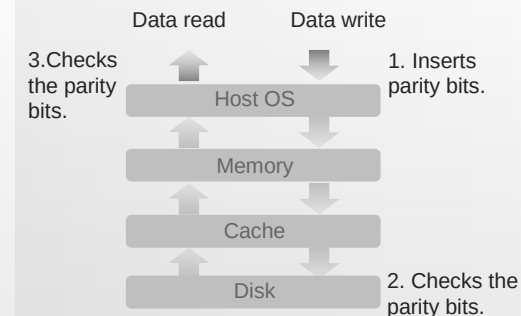
- **Zero data loss or service disruption** even if 4 nodes fail at the same time
- Fast node failure detection and **failover in 10s**

Device level



- Comprehensive **subhealth detection and preprocessing** for fault risk elimination in advance

I/O level



- **DIF-based E2E consistency** check for data integrity