LeoFS

A Market Proven Parallel File System for Data Intensive Storage

ABOUT LEOFS

What is LeoFS?

With more than 1 EB in customer deployment, LeoFS is a market proven parallel file system developed to efficiently handle different types of I/O intensive workloads.

VM-aware and hybrid cloud enable, starting with tens of TB and up to tens of PB, LeoFS has helped hundreds customers worldwide.

Current largest cluster installed has more than 300 storage nodes, 95PB and over 200GB/s I/O throughput.

Worry-free with 24/7 customer support and cluster management.

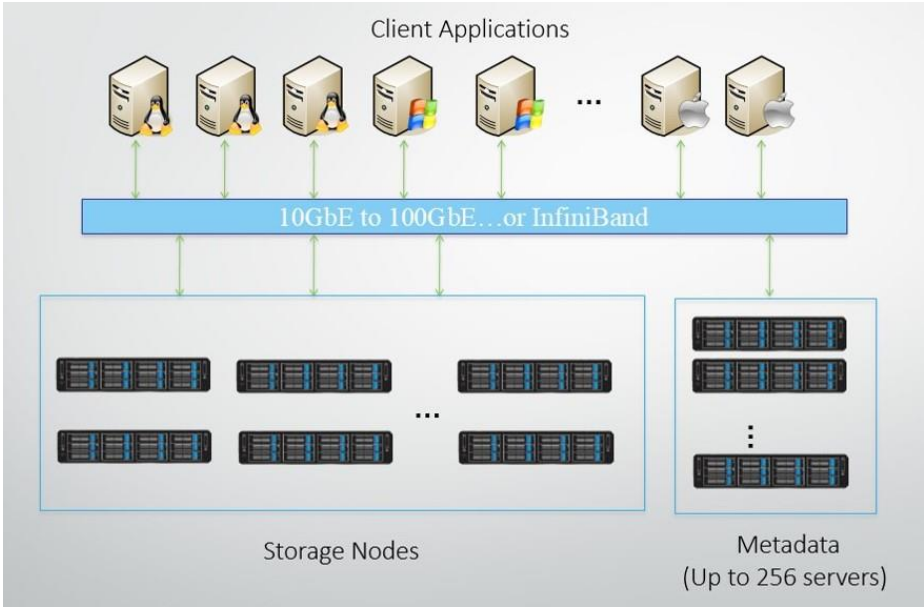
Why is LeoFS special?

With LeoFS, data files are transparently distributed over multiple nodes. By simply increase the number of servers and disks in the cluster, you can **seamlessly scale the file system's throughput and capacity to the needed level.**

LeoFS is right choice in building **a storage cluster with the best combined values** of high performance, scalable capacity, reliable data protection and affordable management cost.

All-in-One Storage for object, block, and file-based data.

System Architecture



PEACE OF MIND

With direct contact to the file system developers, LeoFS offers easy management with designated consultants. Our highly-trained and well experienced engineering team is available 24/7.

No matter how large is the cluster, our staff always take customer system monitoring a top priority.

Uncompromised Reliability

When hardware is ready, software installation and update can be done in one hour. **Adding servers for performance increase or capacity expansion requires no downtime nor the need to reboot.**

With graphical monitoring and administration, most cluster problems can be fixed remotely without operation interruption.

Options to choose Next Business Day Service Level Agreement, Remote or On-Site Support Warranty, Advanced Hardware Replacement.

Solution Guarantee

Whether your goal is to increase productivity or have a better ROI, **we guarantee usage satisfaction on all LeoFS clusters.**

U.S. Service Partner for High-Performance Computing

Data in Science Technologies

2160 Kingston Court, Suite B

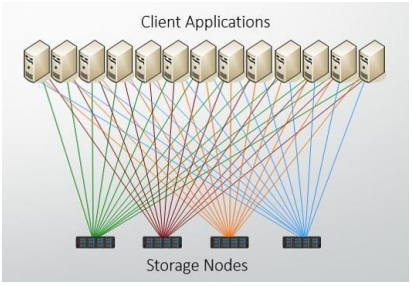
Marietta, GA 30068

Email Don Bailey at: dbailey@dstonline.com

DATA IN SCIENCE TECHNOLOGIES DST

BEST COMBINED VALUES

Without controllers nor gateways, the file system allows concurrent access between all application clients and storage servers in the cluster.



Best I/O Throughput

Always saturate what the hardware can offer. Using commodity servers and HDDs, with dual 10GbE network, average throughput of a single SATA or SAS drive is 45~100MB/s. Most of our customers fulfill their throughput requirements without using expensive hardware such as SSD drives.

Reliable Data Protection

With N+M File-level Erasure Coding, LeoFS distributes data content on a file level based across different storage servers. When N+2 erasure code is applied, cluster can sustain operation with up to two simultaneous failures, whether it is an individual drive or a whole node.

While traditional hardware or software RAID needs to rebuild a whole physical drive, **LeoFS solution rebuilds only the files that are affected**, and it uses the entire cluster to rebuild. Hence, it delivers much faster data recovery, usually in a fraction of the time traditional RAID architectures require. **No downtime or reboot, recovery of one TB data usually takes less than 20 minutes.**

By directory base, the system offers optimum data protection plans for different files and **better capacity utilization, up to 90% with 16+1.**

Fully POSIX-Compliant

LeoFS is compatible with all software applications, x86 based servers and IP networks. No need to make application changes. **With POSIX interface, system can add clients and servers without downtime.**

The file system supports Linux kernels up to the latest version and Linux distributions including Debian/Ubuntu, SLES/OpenSuse, or RHEL/Fedora.

Computational Storage

Besides metadata management, LeoFS storage servers can also run client applications or computing tasks on the same physical nodes.

Such converged structure cost-effectively provide computational storage so **the cluster of servers become shared storage plus data processing units.**

Linux, Windows and macOS

Native clients, all kernel modules that do not require any patches.

- **Linux:** all versions from kernel 2.6 and up
- **Windows**: XP, Windows Server 2003 and up
- **macOS:** 10.5 and up

Highest Level Data Availability

The system's automated self-monitoring mechanism can single out failed hardware, either at the disk or server level.

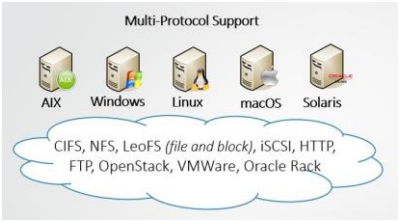
Once inactive hardware is detected, it will be isolated for read-only operation or taken out. The system will then start a self-healing process without any operation interruption.

Cluster or single server product, LeoFS is ideal for HPC, AI, and Big Data Analytics.

U.S. Service Partner for High-Performance Computing
Data in Science Technologies
2160 Kingston Court, Suite B
Marietta, GA 30068
Email Don Bailey at: dbailey@dstonline.com

ENTERPRICE FEATURES

Offering SAN and Object storage options, our system supports Hadoop, Oracle/SQL, VMware, KVM, and Xen. LeoFS works with LeoSAN, CIFS, NFS, HTTP, S3, Swift, iSCSI, HDFS, and Cinder.



Single Cluster Threshold

	Theoretical	Actual Deployment
Storage nodes	4,096	**333**
Metadata servers	256	**32**
System capacity	EB	**95PB**
Number of files	Unlimited	**50 Billion**

Professional Support

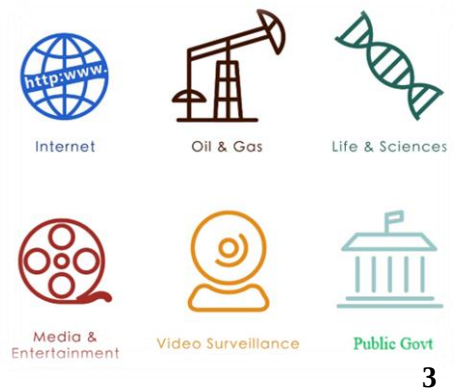
- **Cluster Monitoring:** Free support access via emails, phone and live chat. Our consultants can remotely access the system and run diagnostics to ensure cluster condition. On-site support is also available to keep customer business seamless.
- **Software Maintenance and Update:** Once installed, enjoy free software upgrades and access to a vast suite of enterprise features, such as clone snapshot and LeoSync.
- **High Quality Hardware:** All hardware including replacements must go through pre-configuration testing so installation and repair will done in a time sensitive manner.

Customer Industries

- **Oil and Gas**
- **Scientific Computing**
  - Genomics
  - Cryo-electron Microscopy
  - Satellite Imaginary and Observatory
  - Geographical Data and Mapping
  - Meteorology/Climate
- **Higher Education**
- **Media and Entertainment**
- **Telecom and Internet**
- **AI and Big Data**
- **Video Surveillance**

Selected Features

- Load balance switch, hardware evenly share system workload
- Runs on platforms such as x86, OpenPOWER, ARM, and Xeon Phi
- Re-export through Samba, NFS, FTP, HTTP, LeoSAN or iSCSI
- Support for group/user ACLs and quota
- Fully active network with automatic failure detection
- Supports Infiniband, GigE, multiple subnet and bonding
- Cold data sanity check, automatic repair, no downtime
- WORM directory, avoid modification of saved data



U.S. Service Partner for High-Performance Computing

Data in Science Technologies
2160 Kingston Court, Suite B
Marietta, GA 30068
Email Don Bailey at: dbailey@dstonline.com

DATA IN SCIENCE TECHNOLOGIES DST

STATISTICS

Metadata Cluster

Standalone or embedded with storage servers. **Designed to fulfill any performance requirements, extremely high IOPS or file open rate.**

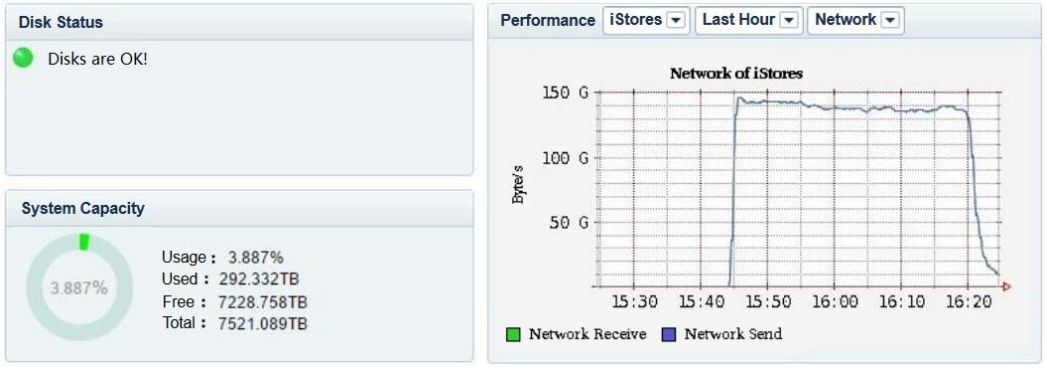
Scalable up to 128 pairs (or 256 nodes). In a testbed with 8 storage servers (each with 24*2TB 7200r SATA), 160 client processes and dual 10GbE network, a metadata pair can provide a sustained file creation rate of more than 20,000 creates per second.

Linear Throughput Increase

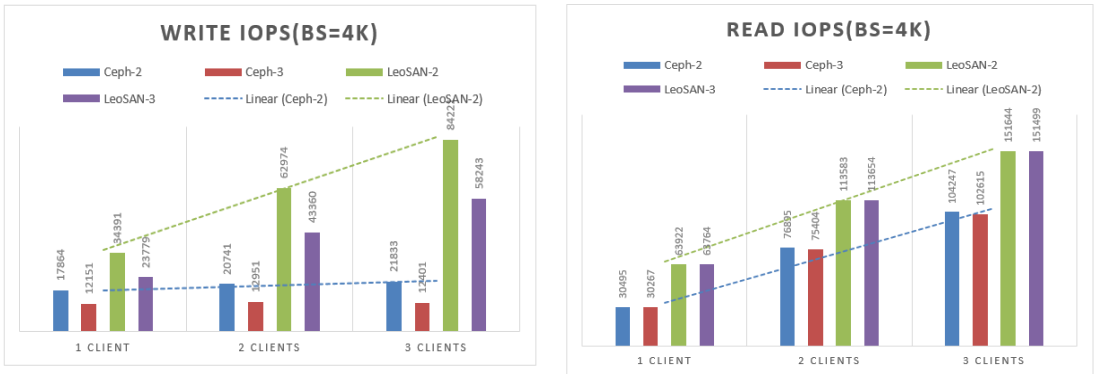
Unlike competition must use high-end hardware to boost performance, **with only commodity servers and in-expensive HDDs,** LeoFS provides predictable and sustainable throughput without deterioration overtime.

With highest level of scalability, system throughput will always increase when adding servers and drives.

Customer on-site 7.5PB, near 150GB/s throughput (102 4U 24-bay storage nodes, dual 10GbE network, 4TB SATA)



Block device SSD IOPS vs. Ceph (3 2U 12-bay storage nodes, dual 10GbE network, 240GB SSD, 3 client)



U.S. Service Partner for High-Performance Computing

Data in Science Technologies

2160 Kingston Court, Suite B

Marietta, GA 30068

Email Don Bailey at: dbailey@dstonline.com

REFERENCE

Comp Chart

Features	Isilon Nitro	IBM Spectrum Scale	Lustre	LeoFS
Snapshots	Yes	Yes-Complex	No	**Yes**
Independent capacity/performance scaling	No	No	No	**Yes**
Scale to thousands of nodes	No	Yes	Yes	**Yes**
QoS	Yes	No	No	**Yes**
N+M Data Protection	No	No	No	**Yes**
Encryption	Yes	Yes	No	**Coming**
S/W only, H/W independent	No	Yes	Yes	**Yes**
IB & GbE Support	No	Yes	Yes	**Yes**

Happy Customers

We take pride in the fact that most of our customers in PB usage started with only a few hundreds of TB.

- First Oil & Gas customer in 2009, I/O throughput 2x greater than StorNEXT FS
- Higher Education: University of Florida, Georgia Southern University
- Scientific Computing: Direct Electron (San Diego) on electron detection for biological molecules
- Video Surveillance: Dante Security (New York)

Software as a Service

Annual service fee

- No capacity limit, $5,000/storage node, any application servers and any metadata servers
- Current special: 1 PB free usage if replacing competitive file systems such as Lustre, GPFS, or BeeGFS

Option to choose cluster or single server products

- Cluster: Starts with 2 nodes, and up to thousands
  - CPU: Intel Xeon E5-2630 V4 * 2
  - Mother Board: Supermicro X10DRL-i
  - HBA: LSI SAS 9300-8I
  - System Disk: 480GB SSD * 2 + 240GB SSD * 2
  - Storage Disk: 2TB to 10TB HDD, SATA or SAS
  - RAM: 128GB
  - Network Port: 2 * 10 GbE or 40 GbE or Infiniband
- Single Server: from 12-bay to 36-bay drives
  - CPU: Intel Xeon E5-2620 V4 * 2
  - Mother Board: Supermicro X10DRL-i
  - HBA: LSI SAS 9300-8I
  - System Disk: 480GB SSD * 2 + 240GB SSD * 2
  - Storage Disk: 2TB to 10TB HDD, SATA or SAS
  - RAM: 64GB
  - Network Port: 2 * 1GbE or 10 GbE