caringo®

The Cloud and Object Storage Platform

CIUK 2018

Company

2005 — Pioneer in software-defined storage
Founders early tech basis of EMC Centera

2010 — Caringo surpasses
100 customers

2016 — 100s of Billions of Files Stored

Recognized Leader

Unrivaled Integration

Enterprise Grade

Why Object?

Scalability + Accessibility + Durability + Intelligence

Why Object in HPC?

Scale <u>on your terms</u>, transform archives into assets.

Internal & External File Sharing

Built-in Tenant Management

Rivals Parallel FS Throughput

Scale to 100s of PBs & Billions of Files

NFS, S3 & SMB Support

Custom Metadata & Search

Performance Benchmark at STFC

- Parallel file system like performance for archiving and access to warm or cold data

S3 Aggregate Throughput

	Performance Requirements	Caringo Results	% Above Requirement
S3 Read	21.5 GB/s	35.0 GB/s	63%
S3 Write	6.5 GB/s	12.5 GB/s	92%

NFS Single Instance Throughput

	Performance Requirements	Caringo Results	% Above Requirement
NFS Read	150 MB/s	349 MB/s	132%
NFS Write	110 MB/s	392 MB/s	256%

S3 & NFS Access
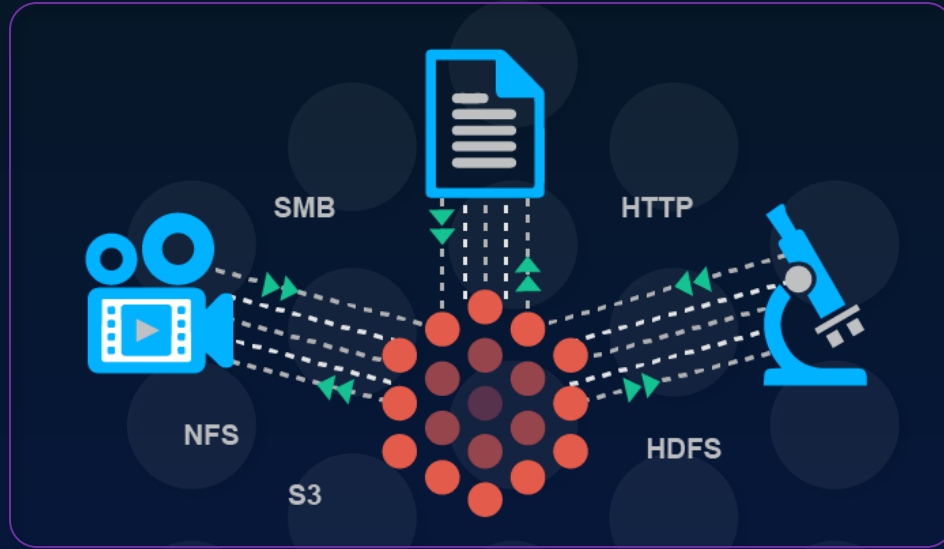
- Universal access to the same data via S3, NFS or HTTP

- Archive data using S3 APIs, read data back via NFS for simple, transparent, access from legacy tools

SwarmNFS – Centralize, Distribute, and Manage Data

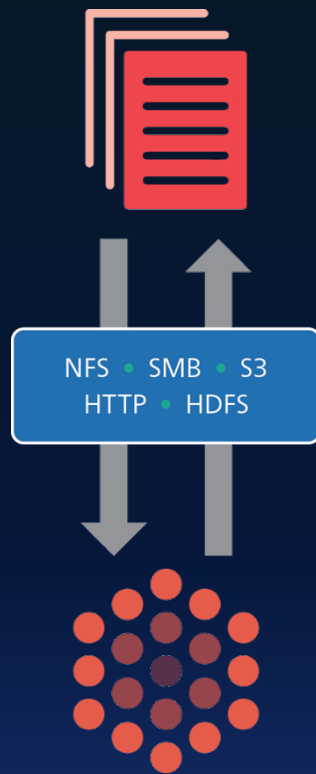## Completely new approach

- New class of protocol conversion
  - *Not a file system gateway or connector*
- Leverages Swarm's Elasticsearch integration

Think "Object NFS"

- Every object can be written and accessed via NFS, S3 or HTTP without restriction
- Unified address space – eliminates protocol and object silos

Direct ingest and access

- Distributed mount points
- Multi-tenancy and security

NFS • SMB • S3
HTTP • HDFS

Export Domains, Buckets and Collections as NFS Mounts

Content Management

Provision, manage, delegate, search, organize, protect

Key Points

- Delegation and self-service with RBAC
- Collections can become NFS shares
- Data protection policy management
  - Replication/Erasure Coding
  - Versioning
- Quotas for tenant, domain, and bucket
  - Storage: logical or raw
  - Bandwidth resets at first of the month (settable)
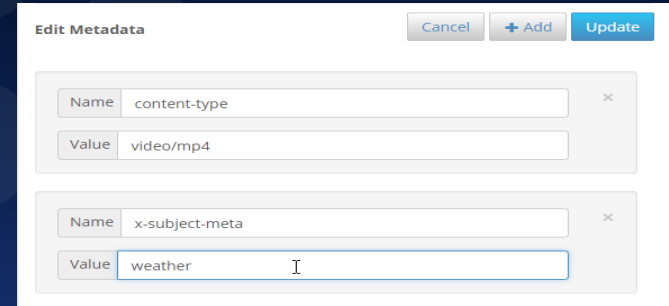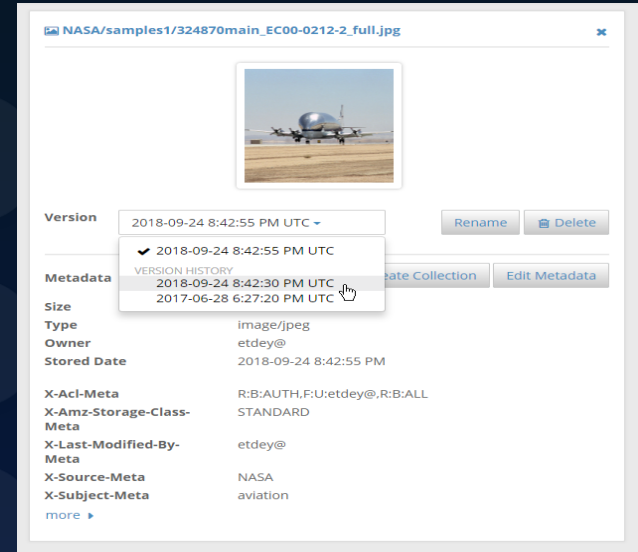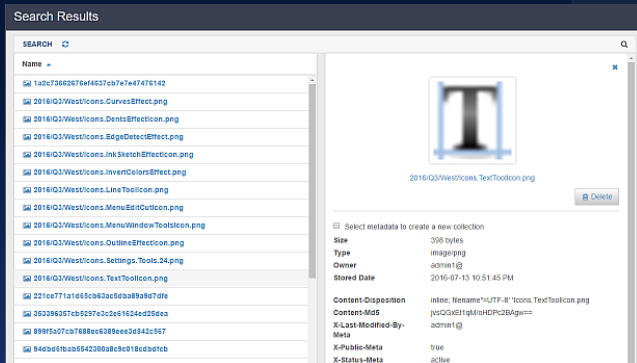
Collections – The New Object Search Paradigm

- Catalog and organise content via standard or custom metadata strings and create "collections" based on smart views

Simple Internal & External Sharing

- Share objects using a global URL accessible internally or via the web (HTTP or HTTPS)

- Set object and bucket permissions using a simple access policy wizard

Multi-Tenant Management

Secure

- Connects to LDAP, Active Directory, Linux PAM
- Token-based authentication
- Coarse to fine grained user access controls

Flexible system administration

- RESTful API & via web portal
- Comprehensive naming scheme with no bucket limits
- Use your domain

Interoperable

- HTTP, S3, HDFS, NFS, SMB
  - Full support for S3 versioning

Architecture Overview

Caringo Swarm™

Applications | Automated Tiering from Filers

HTTP	amazon web services™	S3	openstack™	hadoop	NFSv4	(Windows)	NetApp®

MULTI-TENANCY / ORCHESTRATION / SECURITY / METERING

SCALE-OUT STORAGE CLUSTER
Automated Content Management
Elastic Content Protection

ELASTICSEARCH

SWARM PORTAL

MANAGEMENT
- Configuration
- Monitoring
- Reporting

CONTENT
- Content viewing
- Collections
- Metadata entry
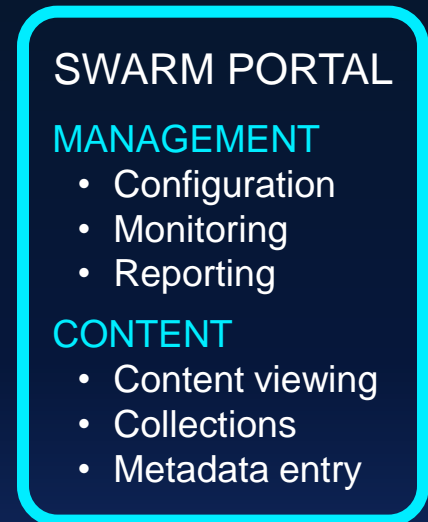
Standard x86 servers or VMs

caringo®

Storage Cluster Overview
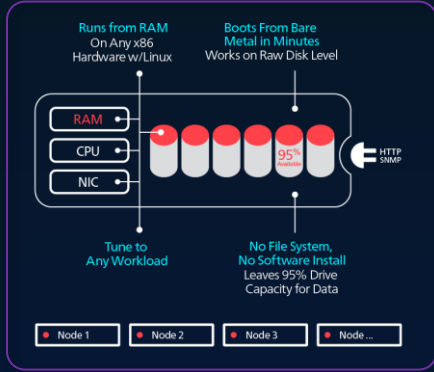
Complete Software Appliance

No file system, RAID, LUNs or DBMS required

Boots from Bare Metal; No Additional Setup

Runs from RAM; 95% capacity for files

Flexible Deployment

Adaptive – Bio-mimicry, self-managing, market behavior
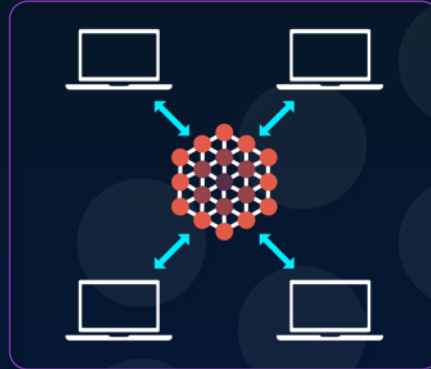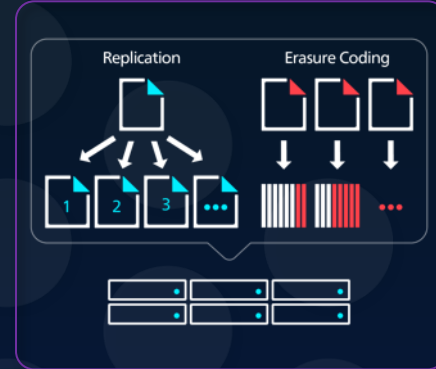
Symmetric storage nodes perform all functions
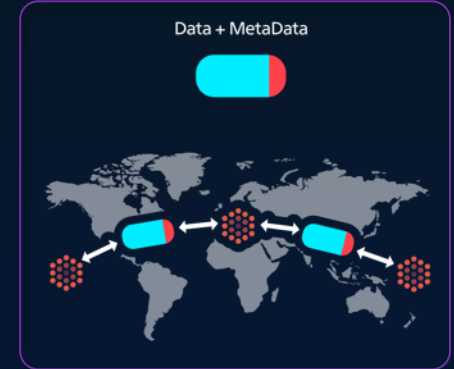
Darkive™ adaptive power conservation

Continuous Protection

EC and replicated files together (Patented)

Feeds and subclusters for replication and DR

Fast proactive recovery

AES-256 volume encryption

Instant Search and Analysis

Files contain all system and custom metadata

Unlocks data from application & location

Integrity seals, WORM, legal hold

STFC Case Study

Geospatial data, goal to store it all

JASMIN - "super data cluster", more bandwidth to disc than normal clusters
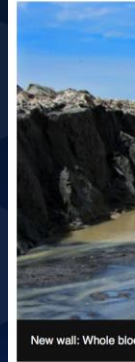
Benefits Achieved with Caringo

- Equivalent performance to parallel file systems for S3 throughput

- NFS and S3 access to the same objects

- Easy management of thousands of tenants and billions of objects

- Simple internal and external file access and file sharing

Kaikoura: 'Most complex quake ever studied'
By Jonathan Amos
BBC Science Correspondent
23 March 2017

Science & Environment

Sentinel satellites to monitor every volcano
By Jonathan Amos and Rebecca Morelle
BBC science reporters
19 April 2017

New wall: Whole block

The big earthquake most complex ever

GETTY IMAGES

Countries that have limited resources to monitor their volcanoes will benefit most

Science & Technology Facilities Council