Cubbit

Cubbit Hive

The Private Distributed Cloud

Jan 25th 2021

Gianluca Granero

Need: give easy tools to people vs. retain control of data

What do you need as a research institution when it comes to storage?

Protect very valuable data

You need to comply with security and regulatory issues that come with the nature of your institution itself, on what you can or cannot do with data.
You have to do whatever is needed to keep results of millions of funds invested.

Have an easy everyday life

Your workers must focus 100% on their job, and you would like to give them handy tools that every consumer enjoys.
You want people on value, and machines on “boring stuff” such as moving files and backup.

Control costs

Controlling and predicting costs is a challenge in every organization, and research has even more complex strings to pull to make it happen.
No major upfront upgrades should ever happen, if not based on existing funds and projects.

Minimize overhead

The head of a research group is a researcher her/himself, why should we give him even more headache then he already has, on how to move budget from one project to another?
The IT organization must be fluid and adapt to it, not the other way around.

Strict policy about tools vs. Handy tools

Those seem totally irreconcilable if you manage IT...

What you do

Apply strict policy.
You write down exactly what can be used, that will insure 100% safety, and ban everything else.

What workers do

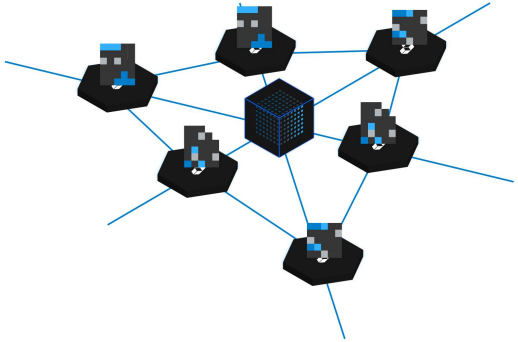
Subscribe to a freemium service for cloud storage, on their personal device.
You have no control at all.

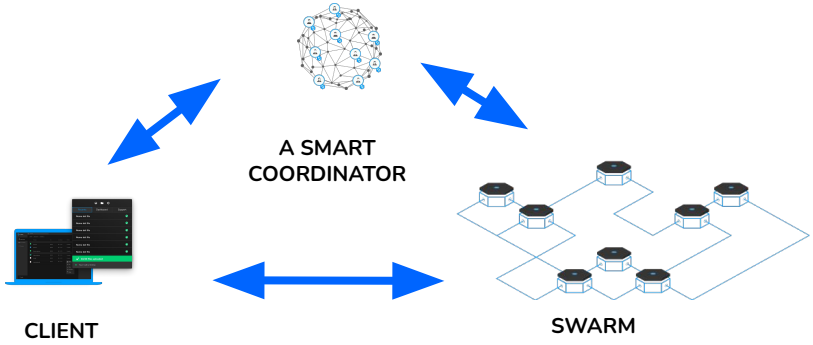
What is Cubbit?

Cubbit is a software platform that enables the full power of edge computing offering a **distributed cloud** solution.
The idea is simple: **a small central coordinator and a myriad of network cells spread across the network**.

On top of this infrastructure **many different products can be deployed**, leveraging the benefits of edge storage and computing.

First end-user products released are **cloud storage solutions**, with 2 different approaches: **sync and share and object storage**.





Cubbit's technology is made of:

Coordinator: a super-node that manages metadata, monitor the network and optimize performances.

Swarm: the peer to peer network based on devices that offer some space

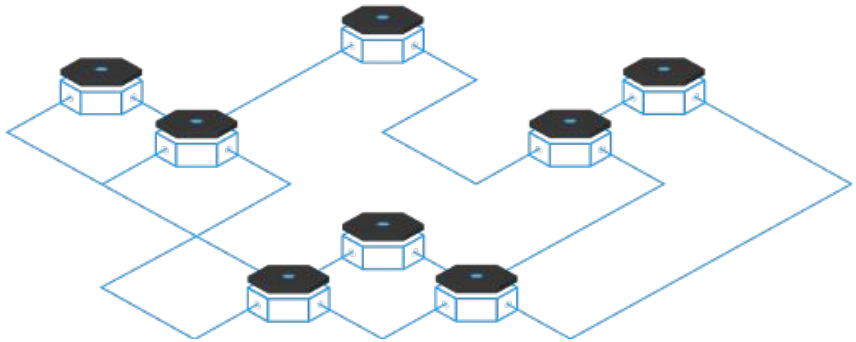
Client: software that the user can use on computers or other devices.

The Swarm

Peer-to-peer: nodes communicate with each other through p2p data-channels boosting up network performances.

Distributed Redundancy: based on Reed Solomon error correcting codes. Ensures high availability while maintaining a low storage overhead.

Recovery: the network is provided with a smart self-healing algorithm which recovers data automatically if needed

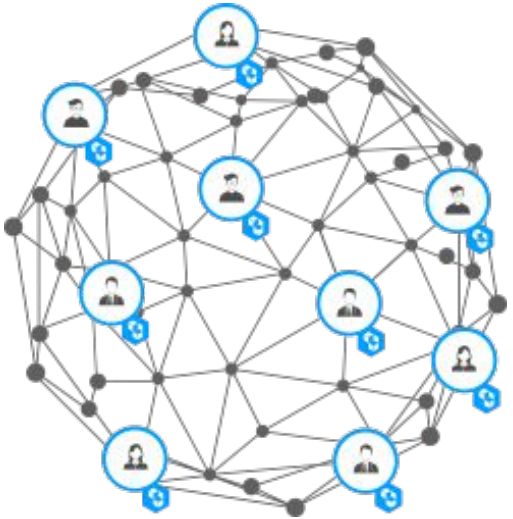


A central coordinator

A smart super(small)-node: a special node of the network that handles metadata and optimizes the overall performances

Optimizations: it employes machine learning algorithms to minimize latency while better distributing payloads across the swarm.

Monitoring: it monitors the network to resolve congestions and trigger recoveries.

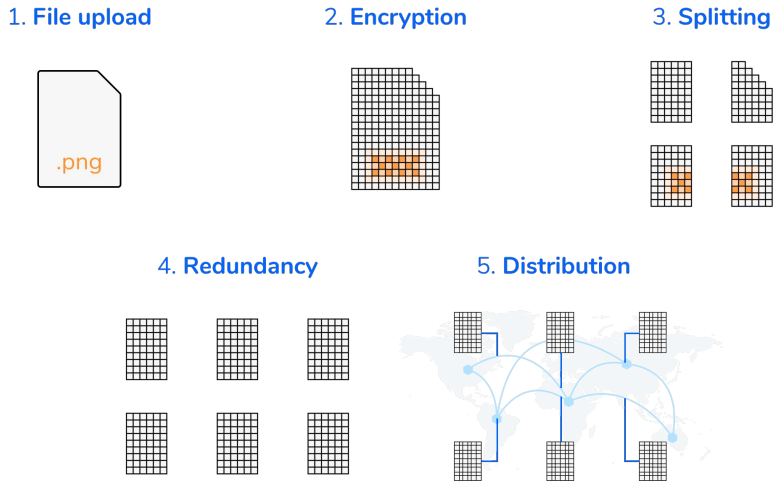


The path of a file

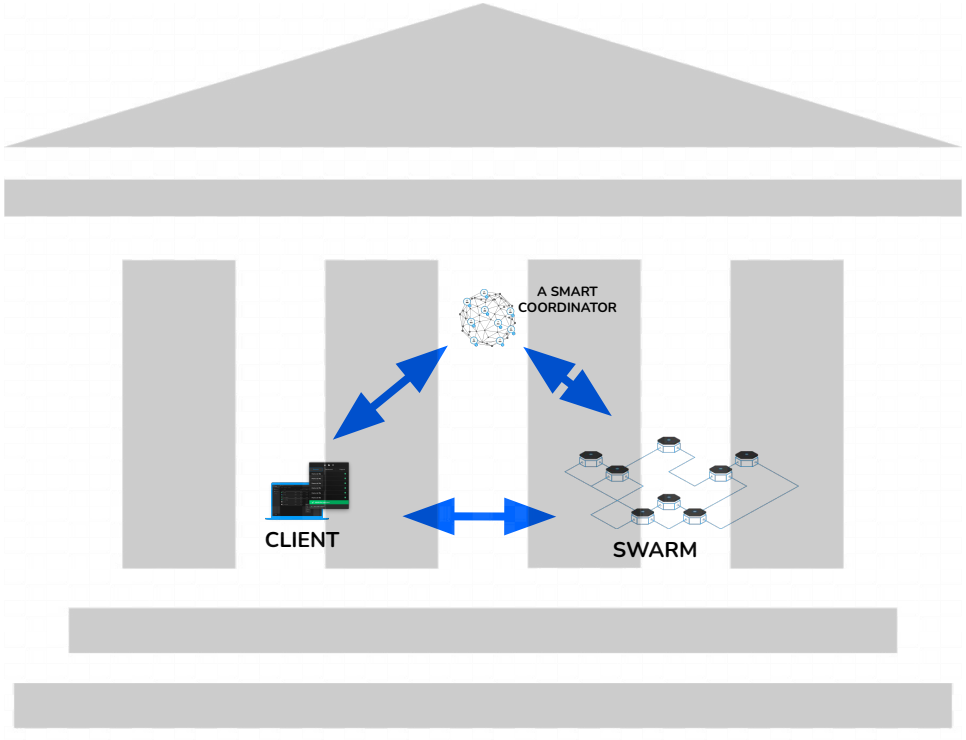
Enhanced security: each file is encrypted with a randomly generated key. This key is never stored on any super-node accessible from Cubbit.

Zero knowledge: our technology ensures that only the final user can access his/her own files.

Transfer: the client splits each encrypted file in 36 encrypted chunks and spread them across the swarm



The Distributed Network
INSIDE your organization



Cubbit on-premises: easy UX

Authentication integration

Cubbit software solution will use the Single Sign On platform of the client which could be (Oracle SSO, Microsoft AD, Microsoft 365, Google Suite, generic LDAP and so on), supporting all major ones.
There will not be need for a new password to be managed, this will not add extra management to users and IT managers.

Simple Dashboard

With a simple tree structure management of groups can be given to IT managers and Business Unit managers, with full control, but without adding extra management effort.
Each manager will have the view, in his dashboard, of the users he controls.

Products and solutions currently deployed

PRODUCTS

Product	Function	Competitor	Status @Cubbit
#1 Sync & Share	File Sharing between customers and/or collaborators.	Dropbox	AVAILABLE ✔
#2 Object Storage	M2M application to store and access data.	Amazon S3	AVAILABLE (Beta) ✔

Future products based on Cubbit distributed technology

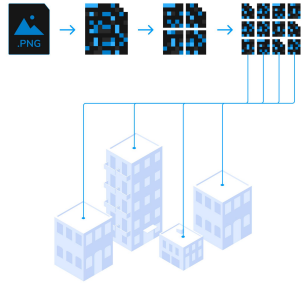
Cubbit final goal is to build a **complete suite of applications and provide all the major cloud and web services** with a distributed data-center solution. We have in our pipeline:

- Hosting
- Containerized applications
- Virtual Private Network
- Content Delivery Network

Cubbit On premises: Key features for critical and sensitive organizations

Data Sovereignty

Cubbit is an European (specifically Italian) entity, not subject to any other jurisdiction.



Data protection and security access

Cubbit is **fully GDPR compliant**. Data are encrypted on the client (and never travels unencrypted). Once encrypted, data are divided into chunks, redundant and sent to different nodes spread over the swarm.

Proprietary Software with some open-source

The stack of Cubbit is made of **proprietary software** written from scratch by Cubbit personnel, and integrated with few open-source tools (such as basic libraries, databases, network tools).
This is fully auditable.

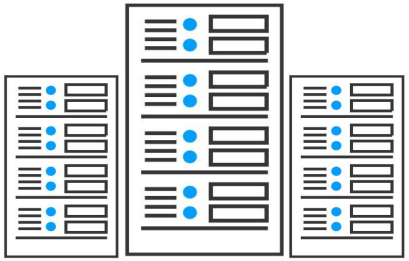
Permanent ledger of operations

You can enable a secure log of operations of files and establish retention of files themselves.
If so an user can eventually copy, but not permanently delete them you have the power to recover in a “time machine” style

Hardware/Operating system Neutrality

From day 0 Cubbit has been totally hardware independent, leveraging the best practices. There is **no specific hardware/operating system which is needed** ever to deploy Cubbit solution and stack.

Cubbit On premises: Smart (re/smart)use of hardware resources



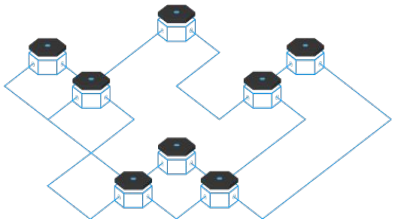
USE EXISTING NETWORK INFRASTRUCTURE

Since the centralized part is almost insignificant, this means we can **just plug nodes to the existing campus network.**

A node can be:

- A dedicated hardware (cubbit cell)
- An old pc/server (still good enough for storage)
- A portion of an existing workstation
  (just need to install the agent and reserve portion of HD)

Cubbit On premises: The power of the swarm



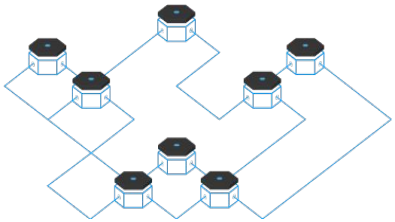
DATA CLOSE TO THE UTILIZER

The solution manages the **correct balance between geographical redundancy and low latency**.

The distributed nodes are by definition closer to where data are produced. This implies that **uploads are faster**, moving the data to close nodes and the swarm will spread accordingly afterwards.

The central coordinator can also move data copies closer to the user ahead of its utilization, according to heuristics specific to the domain, significantly reducing the latency.

Cubbit On premises: The power of the swarm



DATA CLOSE TO THE UTILIZER

The solution manages the **correct balance between geographical redundancy and low latency**.

The distributed nodes are by definition closer to where data are produced. This implies that **uploads are faster**, moving the data to close nodes and the swarm will spread accordingly afterwards.

The central coordinator can also move data copies closer to the user ahead of its utilization, according to heuristics specific to the domain, significantly reducing the latency.

Cubbit On premises: Business resilience

The Cubbit distributed infrastructure is intrinsically meant to help organizations to survive events. The level of redundancy is established upstream - by default 50% - and can be modified on going.

Disaster recovery

Given the correct redundancy, when a fully operating network of the organization is operative, the instantiation of a new coordinator will bring back a fully operating swarm.

The system administrator will hold the necessary keys, in a safe place, needed to initiate the process.

The coordinator is limited in needed hardware resources and can be run on any virtual or physical standard system.

Business continuity

Cubbit Coordinator has a low data intensity and requires few space to be replicated.

Given for granted an operative network inside the organization for the surviving locations after the disaster, there will be the need for hot replicated coordinator's nodes.

Tools used to perform the replica are included in the standard Cubbit distribution and based on well consolidated open-source tools.

Cubbit On premises: external vs internal protection

Organizzations usually do a good job protecting themselves from direct attack from outside.

Most threats are located in 2 main areas:

- **Internal attacks:** with Cubbit there's no point of phisical hacking because every file inside the company is fragmented and encrypted.

- **Social engineering**: now that smart working has a high impact security has to focus mostly on user's device. Recent history has proven that forbidding use of consumer tools (Dropbox/google drive) could result in use of it on personal devices, with a flow of data completely out of control.

Providing an on premise version of Cubbit greatly helps to keep data flow under under control.

Eco footprint: Swarm vs centralization in costs

Cubbit published a **green paper** that describes the **positive environmental impact of distributed solutions**.

With Cubbit solution only a very small coordinator is “centralized” and it will never outgrow an existing datacenter, every other piece of hardware has one unique and simple constraint: a network plug.

No extra power needed
There is no need of redundant power, network cables, switches, routers connectivity (and so on) per each node, if the node fails for any reason, the swarm will adapt.

No maintenance needed
There is no need for redundant equipment, that would add extra costs, maintenance and 24/7 requirements. It is totally unnecessary in a swarm, that will leverage a “field network” which already exists and does not add significant power or stress on top of it.

Devices re-used
Cubbit on premises can run on obsolete hardware for office use, saving also some disposal issues for organizations.

1PB ON CUBBIT

= -40,000KG CO2/Y

5-10 time less than centralized competitors using our cells
see our greenpaper
https://arxiv.org/abs/1803.06973*

Coping with research institutions budget issues

We proposed a model for the public Swarm which really was disruptive and worked.

Join with 2TB device -> Get 1TB lifetime cloud with no recurring costs

INDEMAND

Cubbit: $0 a Month, Secure Cloud Storage Hub

Turn any hard drive into a privacy-first cloud: sync, back-up & share files from anywhere, anytime.

Cubbit
1 Campaign | Los Angeles, United States

€1,015,780 EUR by 3,307 backers

€410,750 EUR by 1,691 backers on Mar 31, 2019 with another platform

♡ FOLLOW

Simple interaction idea

The central organization can put up the coordinator and the basic swarm

Each department/unit/group can contribute with existing or new hardware, based on its needs, budget and Capex it has.

Give 2TB -> Get 1TB

Join the swarm!

We are looking for partners for the firsts academic on-premise deployments

Please contact us:
gianluca.granero@cubbit.io

https://cubbit.io