

SFD17 - Komprise
<https://www.komprise.com/>

*Kumar K. Goswami, CEO

Transforming Data Management without Disruption

[Komprise_Kumar.png]

Drowning in unstructured data. Identifying cold data and syphoning it off before it goes into the data workflow and data protection
Non-disruptive data management

Started in 2014
3 co-founders
3rd startup together
60 people, 40 in US and UK, 20 in Bangalore
Kaviza previously

15 - 30 minutes to get up and running
Hybrid SaaS - console in the cloud, observers running in VMs on-premises
“Green doughnut”
- data by age
- file analytics (size, types, top users, etc)
- Then set policies and see ROI based on the policy (customer enters their own costs)

Use cases
- capacity planning
- live transparent archive
- replication and data recovery
- data migration
- file access to object data

*Mike Peercy, Founder and CTO
Technology Deep Dive
- TMT - Transparent Move Technology
- Data access anywhere
- Scaling

Whiteboard
- Director
- Observer - does the watching of the storage (the sources)

Could use NFS, Object stores, Tape-backed storage, cloud target (S3-compliant, Azure) as a target
Move things you don't want to cheaper storage

Leaves a “breadcrumb” on the source storage - careful not to call it a stub -
Komprise Dynamic Link

“Distributed, software-only, hybrid SaaS” - doesn't have to be SaaS, Director can be on-premises

Call home for on-premises deployments

SaaS offering currently AWS

There is a recall limit policy - to potential keep an eye on recall costs

Periodic interval to move data to secondary storage? One week interval - this is a configurable parameter

Customers - 100TB and up, most of their customers in the PB+ range

The Observer is completely stateless. The only thing it stores are the credentials for the source devices

KCFS - Komprise Cloud File System - <https://patents.google.com/patent/US20170075907A1/en>

The premise is that the target storage is robust. Some sort of integrity check on the target storage would be good - trust but verify

System checks to make sure there's room before doing a bulk recall

Files not cached by the Observer on the way out

*Data access anywhere

Not in the hot data path

User wants old file, redirected to KCFS, data and metadata bought back from the target to KCFS in case they want it again

Everything runs as NFS or SMB

Can find files directly in targets, but they need permissions to access those in the cloud, for example

Customer supplies their own target storage (cloud, etc)

“Cloud Observer” over the target storage - read-only

*Scaling

Observer setup through the Director and configured through the Director

Observers still operate even if Director is disconnected

One of the Observers hosts the access address

Each Observer is responsible for a certain number of shares

Now over 70 customers

Customers trying to address GDPR? Added some GDPR features.

“Confined data” - which gives the ability to move data which falls under GDPR compliance out of the namespace for access - staged way of handling something like “the right to be forgotten”

Priced on source capacity that is being managed - don't pay for features such as replication

Subscription or perpetual license

You can do a move or a copy - doesn't need to just do a move

Deep Analytics (currently in beta) - distributed DB with one record for each file

Multi-tenancy would be separate Director/Observer grids - multi-site is on the roadmap

Departmental level can be done with Komprise Policy Groups

50/50 SaaS/On-premises model usage

Aim for quarterly releases at the moment, big releases twice a year on average