LucidLink filespaces

Security Model

LucidLink filespaces are built on a modern, general-purpose distributed file platform designed for the cloud. One of the fundamental tenets of its design has been a strong focus on security to provide the best-of-class solution for highly sensitive workloads.

LucidLink

Zero-knowledge encryption model

Various market surveys cite that one of the most significant challenges to public cloud storage adoption is concern over data security. The increase of remote users and distributed workloads can pose an even more significant threat to cybersecurity.

To address this challenge, we have taken a novel approach to data confidentiality in the cloud, one that minimizes the amount of trust required by all entities involved in storing, managing and transferring data. This means not making any assumptions about explicitly trusting the network infrastructure, the cloud storage providers or LucidLink itself.

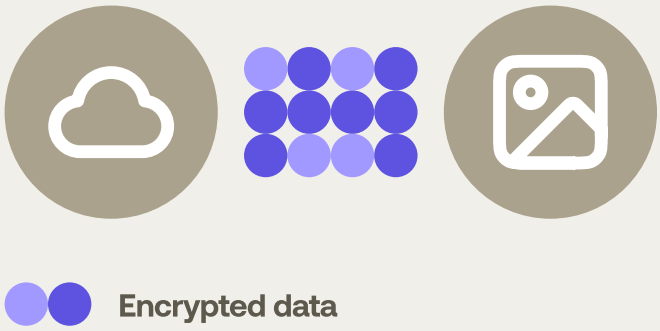
LucidLink's "zero-knowledge" encryption model embodies this approach, where the above providers know nothing about the data the customers store and transmit on their infrastructure. To achieve that, we use a strong end-to-end, full system encryption where all data is encrypted on the customer device and remains encrypted both in transit and at rest with only the customer in possession of the encryption keys. Most importantly, neither LucidLink nor the storage provider (e.g., Amazon S3) can "see" the data, allowing our customers to treat the entire hosting environment as a semi-trusted service. In other words, trusting the providers to reliably store and transmit their data but not with its contents. Note that all this is in contrast with server-side encryption typically employed by other cloud storage services, where data is encrypted at rest, but the service providers maintain the encryption keys and therefore have full access to the content itself. To emphasize this difference, it's worth pointing out that in LucidLink's encryption model there is no 'password reset' for a lost password, losing it leads to permanent data loss.

All file synchronization solutions replicate whole data sets on each client device. LucidLink's technology is fundamentally different as it streams data on-demand, thereby eliminating the need for storing unprotected copies of files on multiple devices.

LucidLink

Stream data on-demand

All the locally cached data and metadata on the client devices themselves are stored encrypted on the local disk. Simply disconnecting from the filespace prevents an attacker with physical access to the device from gaining access to the LucidLink filespace. This is the principle of no residuality.



LucidLink filespaces are based on a split plane architecture where the metadata and the data planes are managed separately. The metadata is synchronized through a central metadata service provided by LucidLink, while the data is streamed directly to and from the cloud or an on-premise object store.

This split plane design requires securing the metadata and the data independently. Every file and folder along with its metadata, like name, extended attributes and so on is fully encrypted. Each file has its own unique encryption key to provide isolation and minimize the attack surface. We support different ciphers and transport encryption modes to offer the best tradeoff between performance and security to address each customer's specific needs. By default, **we use the strongest AES-256 in GCM mode**, a form of authenticated encryption.

Using authenticated encryption has the added benefit that any malicious tampering or data integrity issues such as bit rot on the server side will be immediately detected upon access. This offers a high degree of guarantee and peace of mind that any accessed data is indeed genuine.

The metadata in a typical file system could be roughly divided into two tiers:

1. Space layout – the underlying data layout, i.e., information about where the various segments of the files are stored on the disk (or the cloud).
2. User metadata – the file system hierarchy, names of files and folders, access control lists, etc.

LucidLink

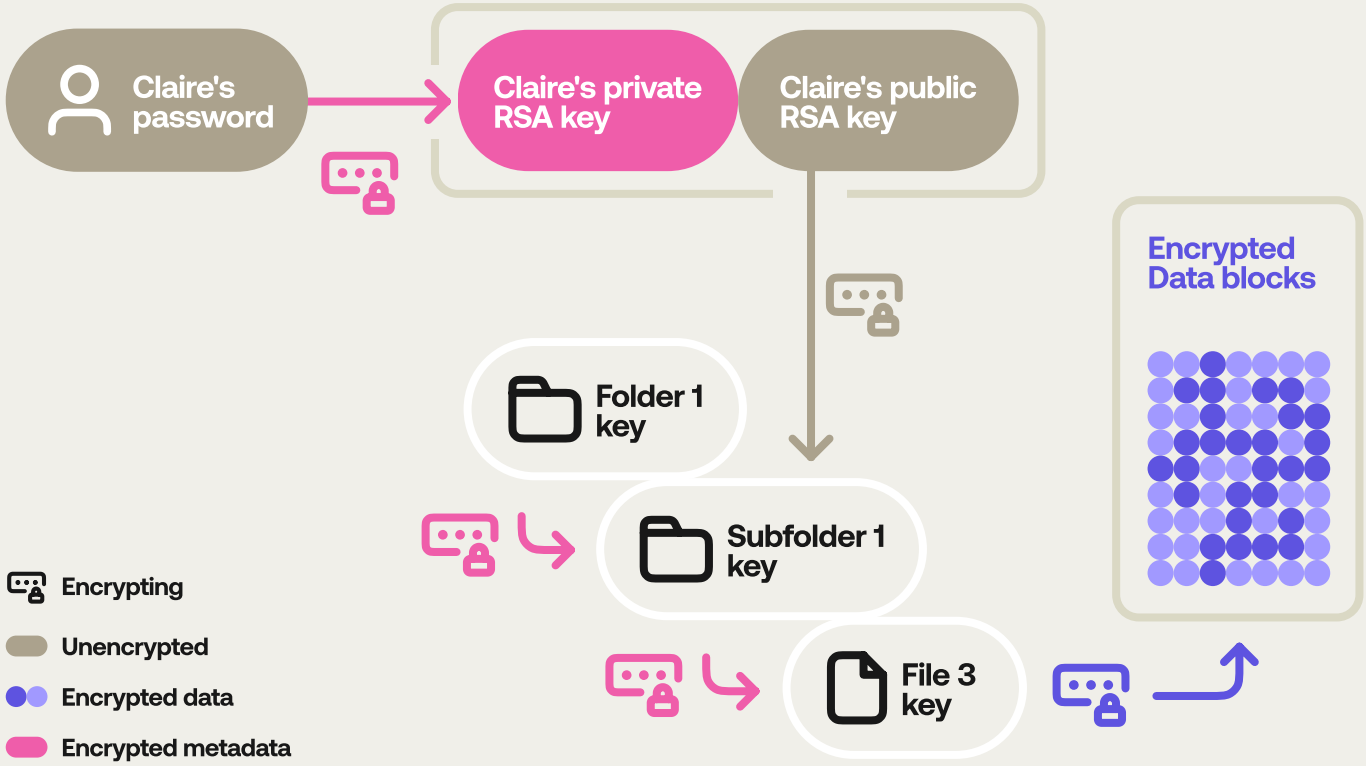
LucidLink encrypts all the user metadata. The space layout portion of the metadata remains unencrypted as it doesn't contain any user-generated and hence potentially confidential information. At the same time, it allows the LucidLink metadata service to perform core operations like snapshots, garbage collection, and others. Each file or folder's encryption key is used to encrypt its data and metadata. The encryption key is itself encrypted with the key of its parent folder. This forms a chain of encryption keys that goes all the way to the root folder key, which is then wrapped by a master key. We never send or store encryption keys in plaintext and never reuse encryption keys or initialization vectors (IV). Large files are split into blocks, and each block is encrypted separately. Whenever a write within a file block occurs, the block is re-encrypted in its entirety using the encryption key for the file and a new IV.

Separately, each user who has access to the filespace has a public-private RSA key pair. The public RSA key is stored in plaintext in the metadata. The private RSA key is encrypted with the user password, where the password is first key- stretched, using PBKDF2 to reduce vulnerability to brute force attacks. The filespace administrator gives access to each individual user to a certain set of folders (or files), called shares. For each share, every user keeps the encryption key for that folder (or file), which is itself encrypted with the public RSA key of that user.

This maintains isolation from other users and partitions access to the file space. This way each user can only decrypt the subtree from the share he is given access to but cannot see anything above the filesystem tree (see Fig 1). Additionally, the administrator has access to the master key to perform various management operations.

All network connections between the client devices, the metadata service, and the object store are separately secured via TLS, by default. This results in double encryption. First, the metadata and data are encrypted using the mechanism described above. Then the communication channel is itself encrypted again to provide the highest degree of security, e.g., perfect forward secrecy.

LucidLink



Note that most modern computing platforms perform AES encryption in hardware, allowing it to occur at line speed without sacrificing performance.

To meet the most stringent security requirements of our customers, all cryptographic operations are implemented with the industry standard FIPS 140-2 compliant OpenSSL. Various security settings and modes can be configured during initialization of the filespace. In specific cases such as when running in secure environments, transport and/or filespace encryption could be turned off entirely.

LucidLink believes that cloud file services will fundamentally change how businesses store and access information. To deliver on this vision we've put data security at the core of our solution to provide unparalleled protection for the businesses' most critical assets.

George Dochev
CTO

LucidLink