Secure cloud-based research platform for whole slide biomarker analysis

Elizabeth Edwards, Dan Eversole, Peter Miller
Akoya Biosciences, Marlborough, MA

Background:

Spatial biomarker discovery and exploration tools for the assessment of multiplex immunofluorescence (mIF) tissue sections are often locally based, leading to bottlenecks in processing and difficulty in data-sharing.

Existing web browsing tools do not scale to the high-dimensional, high-complexity experiments used in biomarker discovery. Fast whole slide acquisition has only intensified these concerns.

Even as cloud-based collaborative workflows come online, organizations must continue their work in ongoing and extensive historical LAN-locked projects - migration to the cloud must accommodate these without disruption.

We demonstrate a solution that combines a LAN-based solution with a tightly integrated secure cloud hub, enabling users to operate in both environments as they transition to fully cloud-based workflows.

Methods:

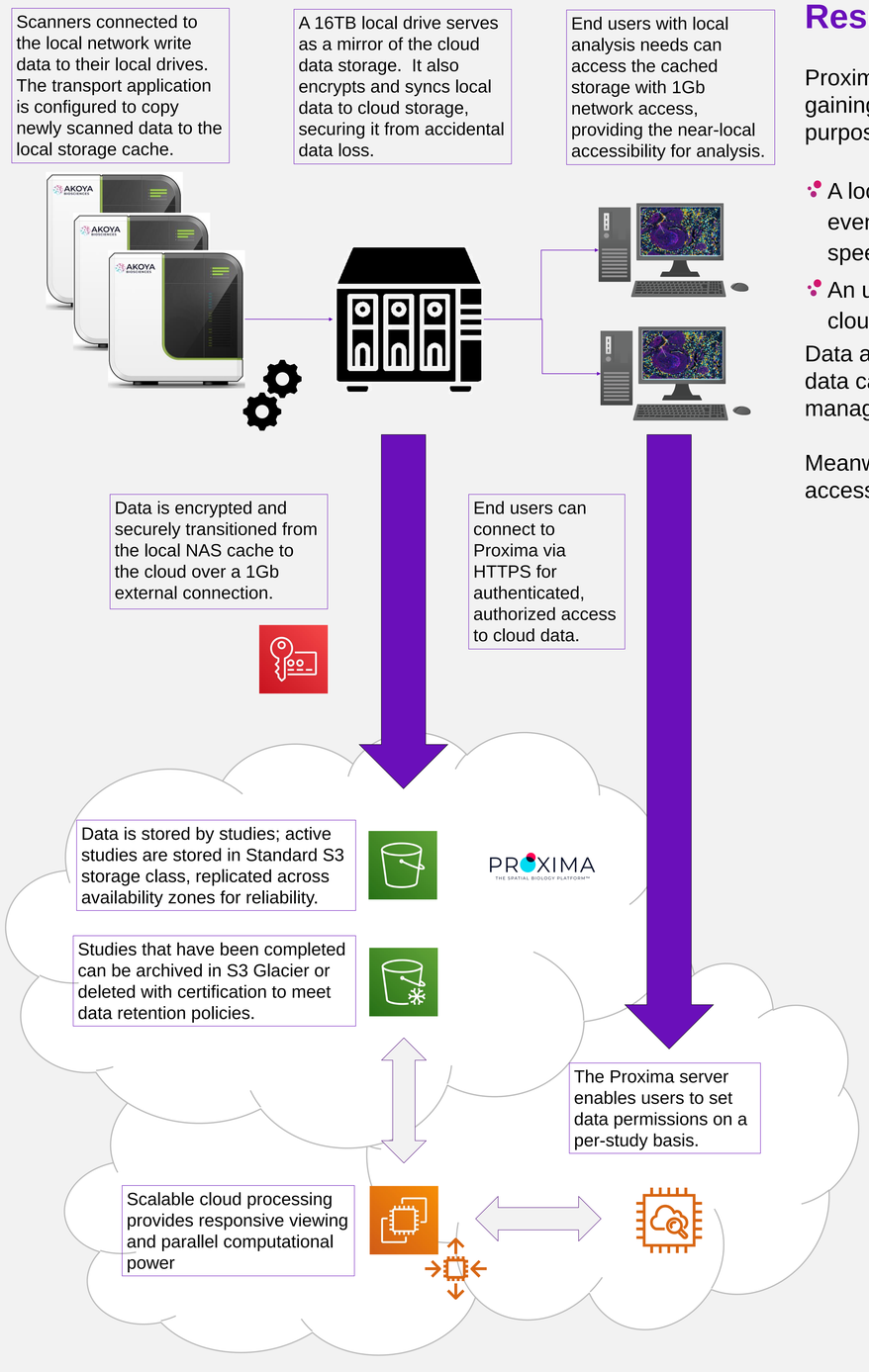
Digital slide analysis requires data to be proximate to the processor for speed and latency. Simple cloud storage of imagery for LAN analysis is prohibitive due to long download times.

To address these concerns, a unique hybrid solution was devised, combining LAN-based Network Attached Storage (NAS) mirror (cache) with bulk cloud storage as follows:

- A LAN-based transport app transitions data from the scanning hardware to the local mirror. This can occur at low-traffic time to minimize the local network impact.
- A gateway solution encrypts and stores the data in secure AWS S3 storage, preventing accidental local data loss.

Imagery is always close to processing (LAN or cloud-based); data and metadata are held on a HIPAA-compliant cloud.

A study-based browser enables viewing imagery and analysis results; administrators can invite collaborators across organizational and IT boundaries with data access defined by the owning organization. Containerized cloud image processing ensures consistent calculations regardless of local IT reconfiguration and updates.



Conclusion:

Biomarker discovery exploration in both LAN-based and cloud workflows is achieved using this architecture, enabling studies at scale.

PROXIMA
THE SPATIAL BIOLOGY PLATFORM™

Results: Continuing LAN-based workflows with infinite storage

Proxima enables organizations with current LAN-based workflows to continue working as planned while gaining secure, recoverable, infinite cloud-based storage for their studies. The 16TB NAS serves two purposes:

- A local mirror of the cloud data. Users are able to see all data stored in the cloud as if it were local, even when cloud storage exceeds the 16TB NAS limit. Active data remains cached to provide LAN-speed access.
- An uploader for new data. Data that arrives on the NAS is immediately encrypted and synced to the cloud.

Data arriving into cloud storage is catalogued by Proxima and stored in S3 buckets per organization; this data can be retrieved rapidly in case of accidental deletion on the 16TB local cache. Encryption is managed per organization for increased security.

Meanwhile, LAN-based workflows can continue using the NAS as a shared drive with local high-speed access.

Results: Transitioning to a cloud-based workflow

Cloud-based workflows provide multiple benefits, such as enabling collaboration across geographical and organizational boundaries, reproducible results in the face of changing IT environments, and scalable computing power without capital expense.

Still, the transition from a LAN-based workflow to a cloud-based workflow may need to occur gradually due to existing protocols and the need for validation. Proxima enables organizations to continue existing workflows and transition to the cloud only when ready; both cloud and LAN workflows can co-exist.

Even during LAN workflows, Proxima organizations can take advantage of shared access for image viewing and review. Once data is cloud-present, organizations can enable users for viewing.

- Organizations manage their cloud-based user pool; they can limit user access to within-organization only, or they can extend invitations to external collaborators.
- Data access is managed per-study; the study owner has the ability to invite users to view data. The ownership of the data is maintained by the organization.
- Users view whole slide imagery using their local browser with secure connections and authorization / authentication on a per-organization basis.

Additionally, owners can manage the life cycle of their studies; they can be readily available in S3 Standard storage, placed into S3 Glacier storage for data that is not actively used, but necessary to retain, or permanently deleted with certification according to data retention policies.

When ready, the organization can extend the viewing and collaboration experience to include image analysis.

- Auto-scaled, cloud-based processing can rapidly analyze whole slide data.
- Individual analysis protocols are containerized for robust, repeatable results, regardless of the end-user's operating system.
- Results are both viewable and explorable in a whole slide context for all users with privileges.

Akoya Biosciences, Inc., 100 Campus Drive, Marlborough, MA USA (855) 896-8401 www.akoyabio.com