



Scaled Single Namespaces

preconceptions, functionality, limitations, current state, and the future plans

Agenda

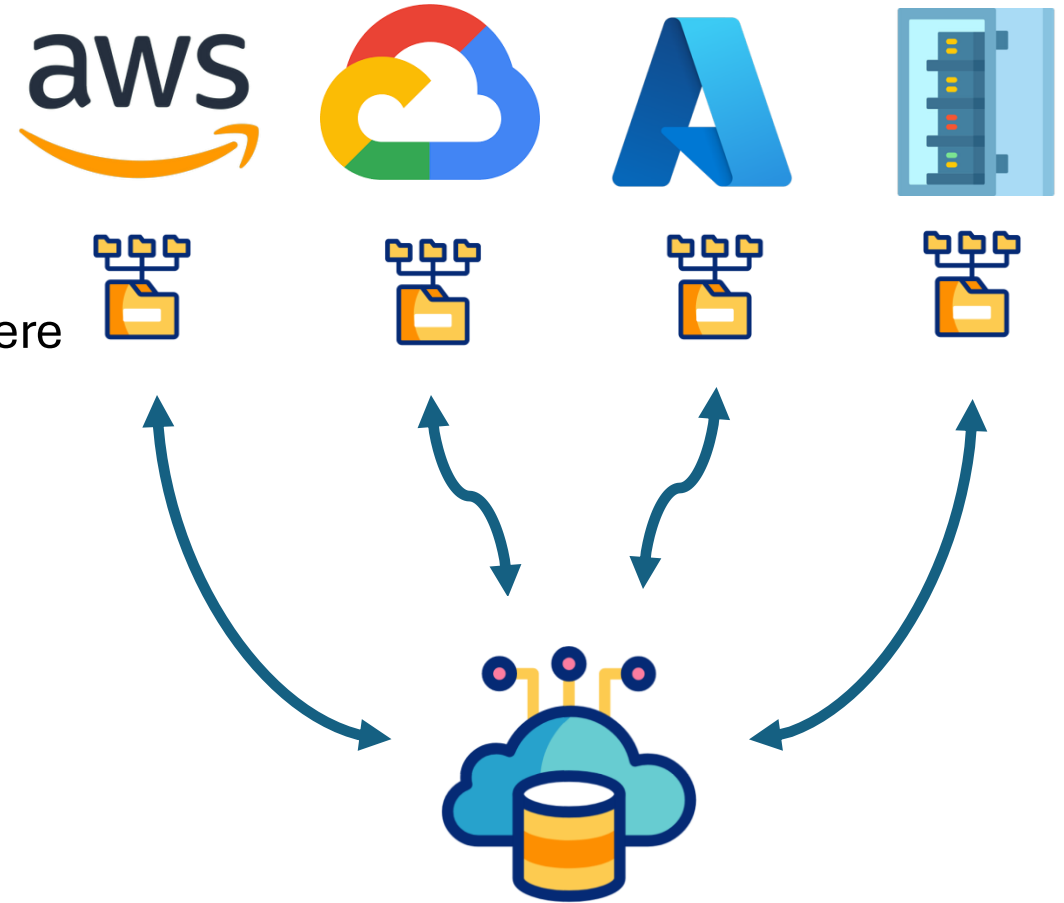
1. What is Single Namespace (SNS)
2. How does it work
3. Quirks, Gotchas, Dependencies, and Bugs
4. User Experience
5. User Interaction Scenarios
6. Status of the SNS work
7. Deployment Dependencies
8. SNS Working Group
9. Future Directions
10. Wrap up
11. Questions

Concept of a Single Namespace

All your data should be available everywhere in the same way
Once your data is everywhere, anything can be run anywhere

Single Namespace Concept

- What is a single namespace?
 - All data available everywhere
 - All locales look the same
- Data available everywhere, compute can run anywhere
- Compute locality becomes a business decision
- Interoperability between vendors
- Extreme Scalability
- Excellent Disaster Recovery



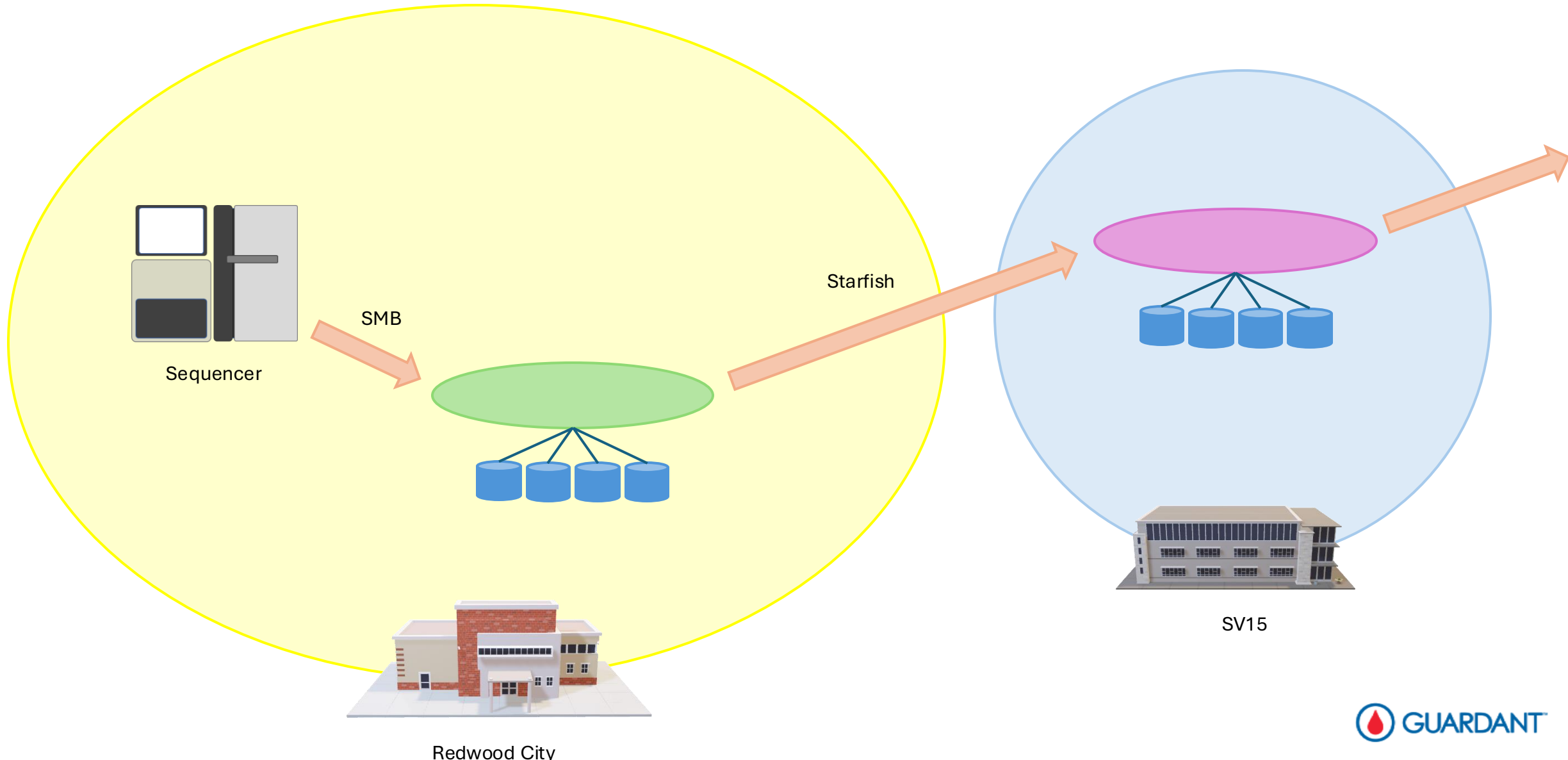
What does an SNS buy Guardant?

- **User experience is the identical across platforms**
- **All data is available to users**
- **Data movement and location is simplified**
- **Reduction in on prem storage**
- **Far more scalable for long term with new data**

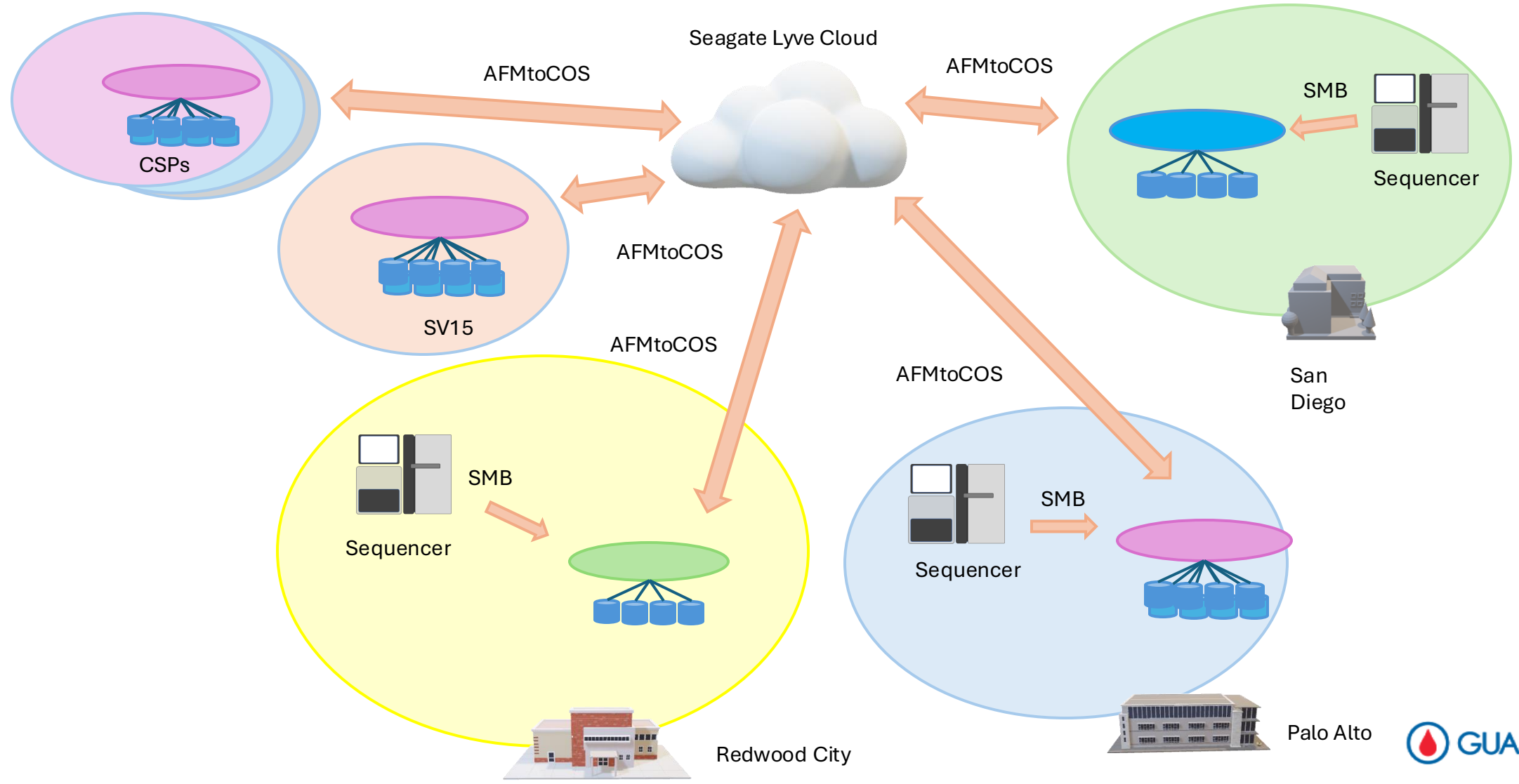
Single Namespace Functionality

- Data is written to a file system
- That data is sent to a backend
- Metadata is then propagated to the other file systems linked to the same backend
- Data is pulled down when accessed (hydrated)
- Data is removed and the metadata link goes back to the S3 object (dehydrated)
- Data changed data is all that is downloaded into the S3 and updated metadata is refreshed across file systems

Historical Data Flows



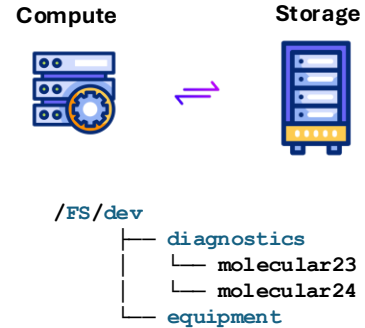
How Data Moves with SNS



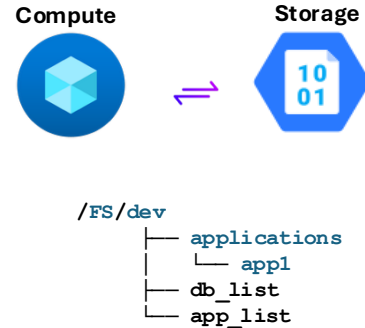
Isolated FS vs Single Namespace

Isolated File Systems

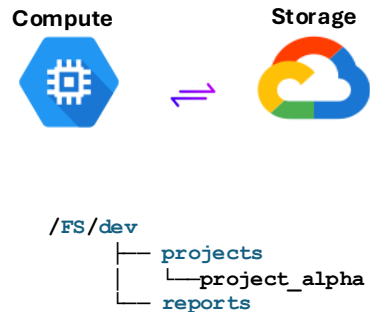
On Prem: San Fran



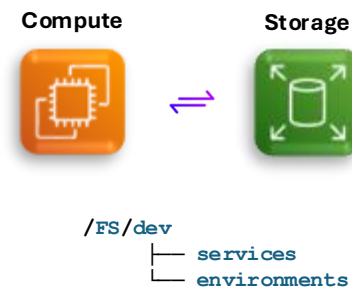
New York: Azure



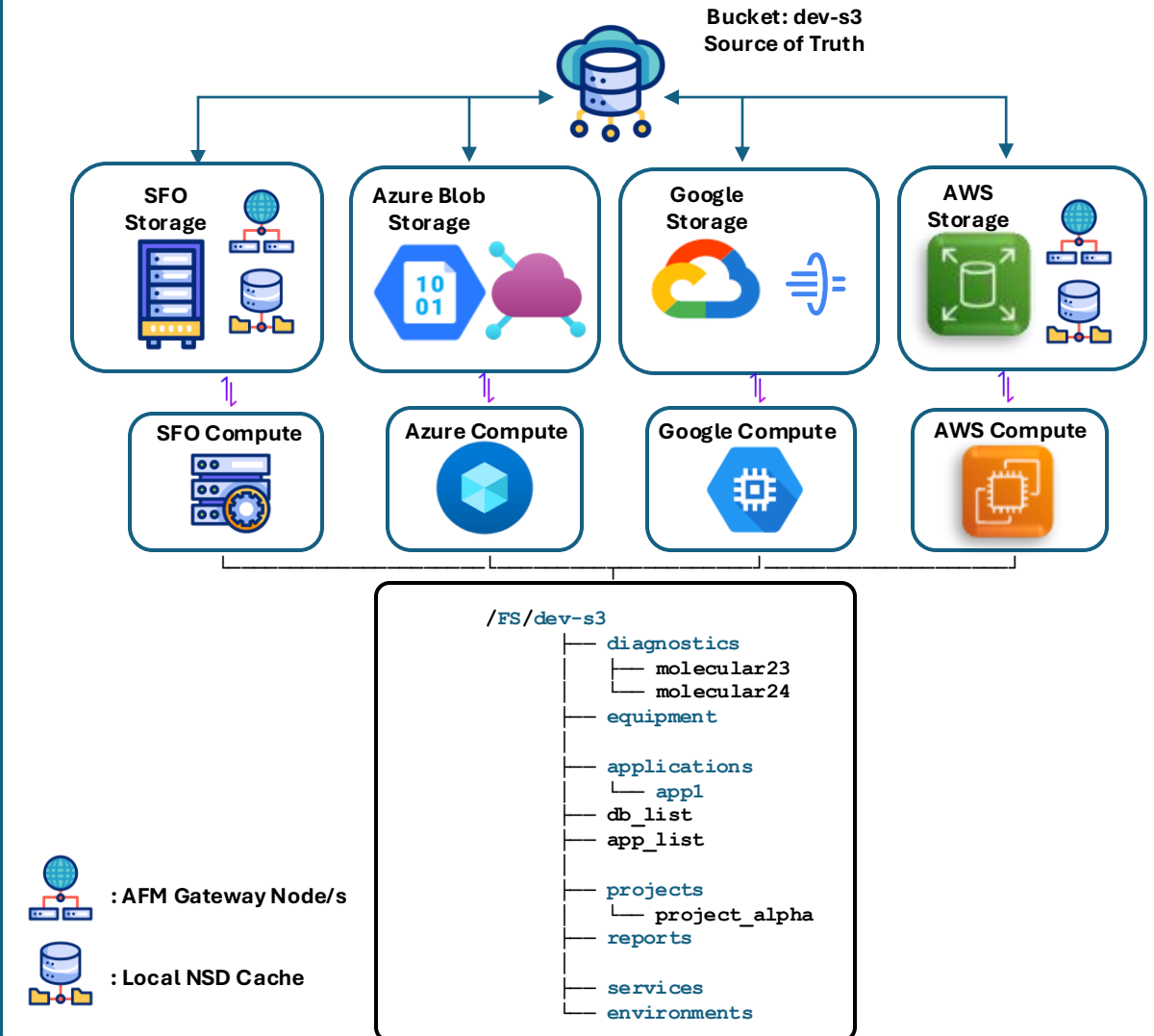
Austin: Google



Oregon: AWS



Single Namespace



- Without SNS, data from other clusters must be accessed by manually copying files or using tools like rsync between different compute clusters.
- With SNS, the same data is accessible from all clusters.

Guardant's SNS Implementation

- GPFS Front end
- Individual Filesets (*) tied to singular buckets
 - /ghds/bob-sns-fileset -> Seagate:/bob-sns-bucket
 - Entire FS cannot be tied to a single bucket
 - Must be done explicitly
- Data backend (today) in Seagate Lyve
 - S3 compatible
 - Two buckets raw and flowcentral
- Data moved by AFM (part of GPFS)
- Connected to dev, palo alto, test clusters and an instance in AWS

Functional Recap

- **File Systems become cache**
- **Backends become source of truth**
- **Policies dictate when to dehydrate and how much to rehydrate**
- **A data mover follows policies to move data**
- **Dehydrated data is only a stub in the file system pointing to backend**
- **Produces a user experience where all data is present**



Quirks, Gotchas, and Dependencies

No such thing as a free lunch

DEPENDENCIES!

- **SNS is wildly dependent on network bandwidth and latency**
 - Push 1 PB through a 10 gb/s pipe and it will take a while
 - Metadata will appear, but...
 - Latency can be a major issue if the network is congested
 - 100 gbs link from Seagate being deployed to dev cluster
 - IT has a plan for network revision under the SNS umbrella (dynamic routing, burstable BW, etc)
- **Settings! Can make or break a GPFS integration**
 - Force tight coupling and stability and performance will suffer

GOTCHAS!

- **GPFS cannot differentiate between different files for what to move (yet)**
 - Does have its own version of copycomplete.txt to denote when data is transferring in a directory
- **GPFS cannot preferentially move (or delay) certain files**
 - Change request submitted
- **There is no locking in SNS today: last writer wins!**
 - Versioning can be a solution, not yet supported, but planned
- **Object versioning is not yet supported in GPFS.**
 - It seriously screws up a file system for now (feature coming). Currently, we get last writer wins (I'll explain)
- **Dedicated AFM nodes are required: underspec'ing is a danger**
- **Modifying objects behind the back of SNS can cause corruption**

QUIRKS!

- **Loosely Coupled GPFS/SNS can recover from network loss**

However!

- They take time to reconcile
- This can cause conflicts
- This takes longer than merely uploading

Recap

- **There are still bugs** (have you met software?)
- **There are dependencies** (have you met infrastructure?)
- **There are gotchas!** (have you met coders?)
- **There are quirks** (have you met sysadmins?)
- **And there are those who repeat the obvious**

State of Now in Guardant's SNS!

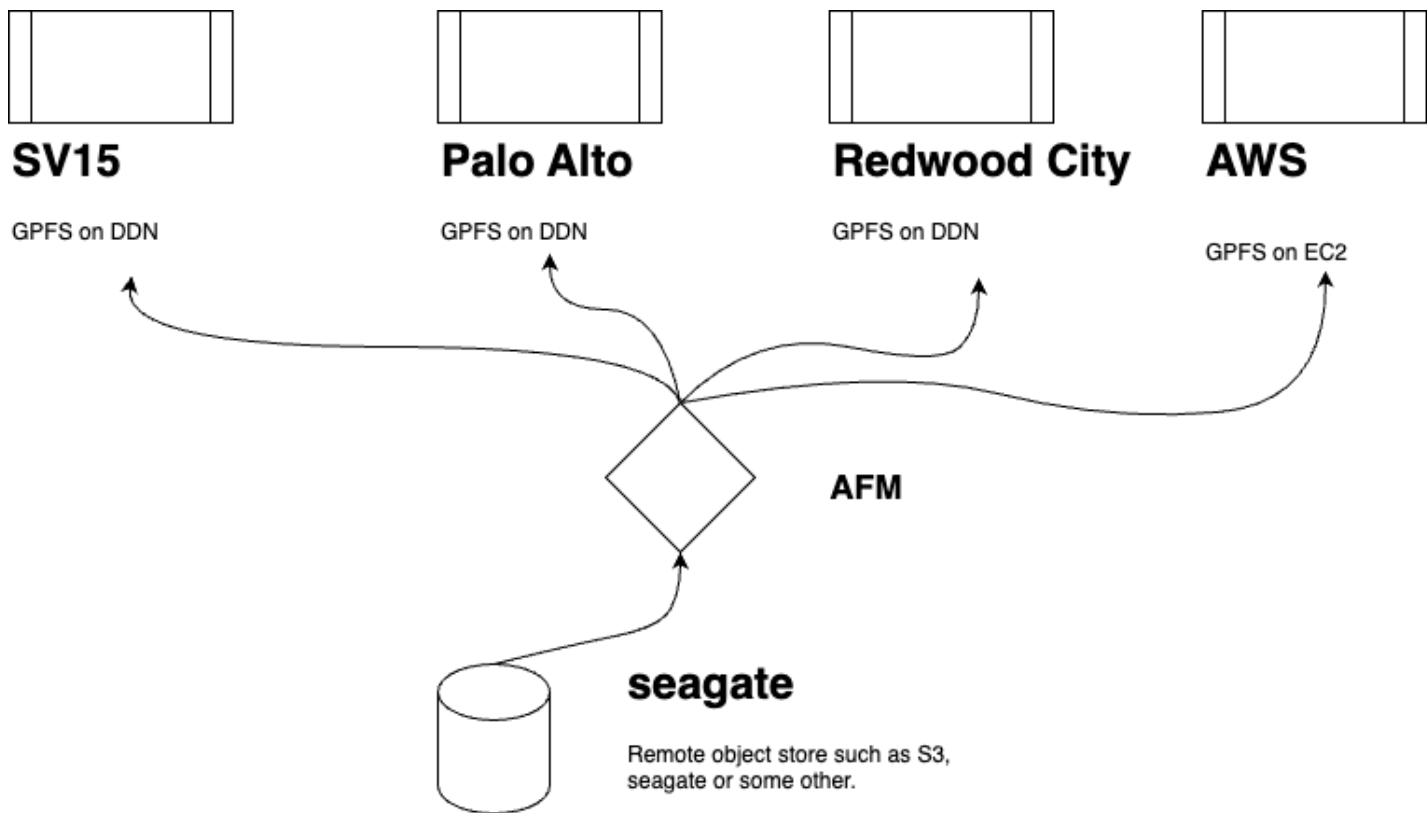
- **Streaming vs copy on close**
- **In conjunction with techdev:**
 - Ran flowcell from a sequencer in palo alto which wrote to SNS file system
 - Ran BIP in dev cluster in Santa Clara
- **Across account access tested in AWS**
- **Set up in bioinformatics AWS account**
 - Acting a lil weird, being debugged
 - Metadata works well, data...not so much

User Experience & Use Cases

A wide-angle photograph of a beach at sunset. The sky is a gradient of orange, yellow, and blue, with some clouds. The sun is low on the horizon, casting a warm glow. The ocean has gentle waves breaking onto a sandy beach. In the distance, there are dark rock formations and a few small figures of people on the beach. The overall mood is peaceful and serene.

a file system

But what does it allow?



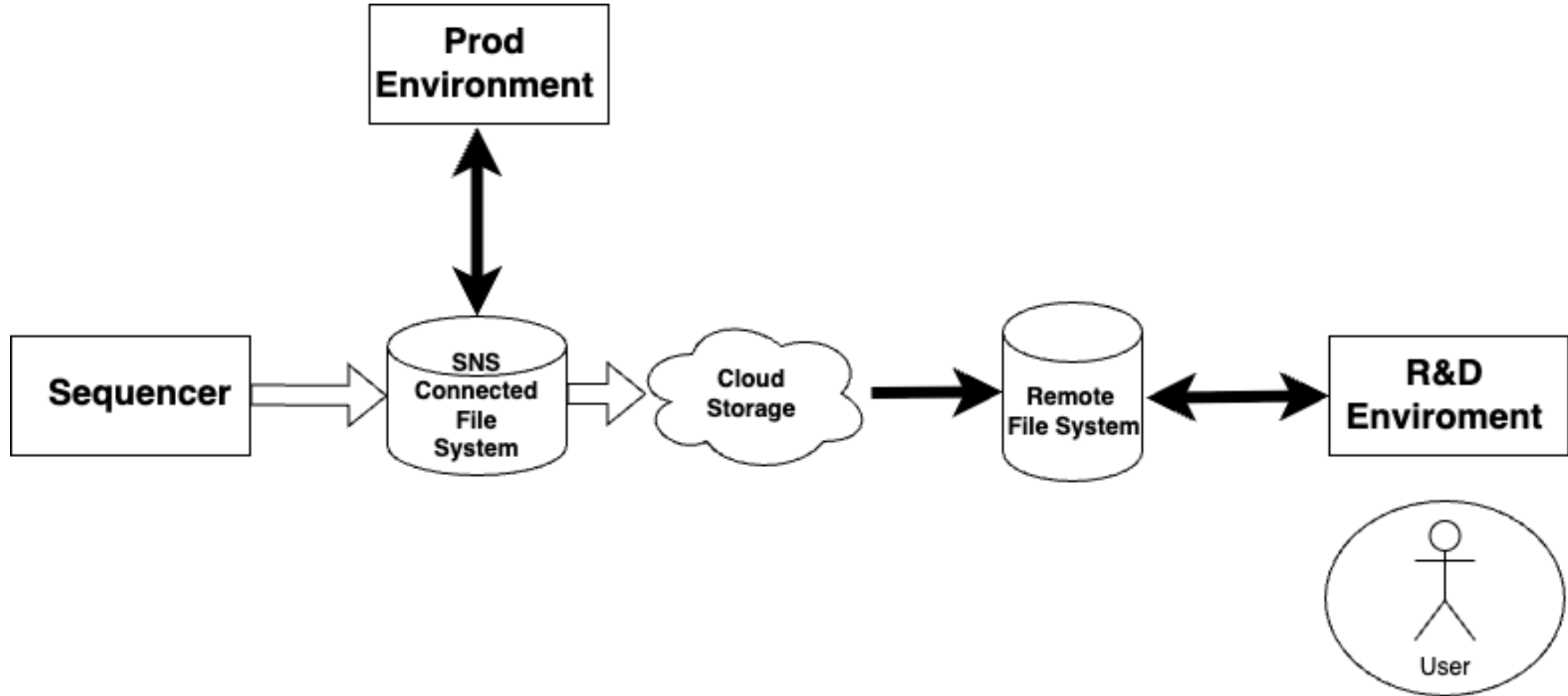
Example SNS

User Workflows

Workflow 1: standard at GH (a)

- Sequencers write to SNS enabled file system
- Computation run locally
- Data pushed to cloud storage
- Metadata becomes available in remote file systems
- Available in R&D, SQA, test & other environments

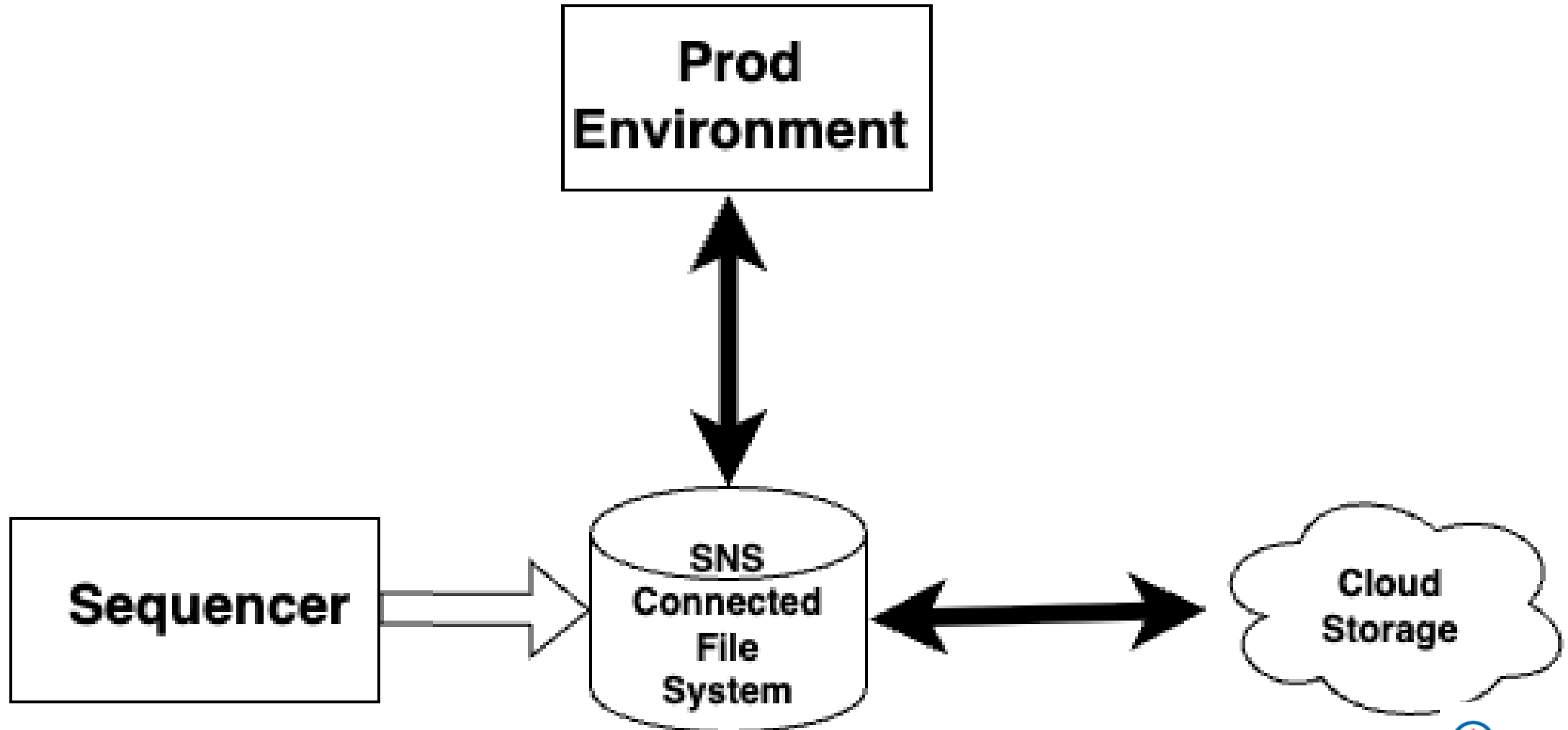
Workflow 1: standard at GH (b)



Workflow 2: clinical monitoring (a)

- Sequencers write to SNS enabled file system
- Computation run locally
- Long term analysis needed, require access to past samples
- Data pulled to local file system from cloud
- Analysis run, report issued
- All new data pushed to cloud, old data dehydrated
- Data available in R&D and other environments long term

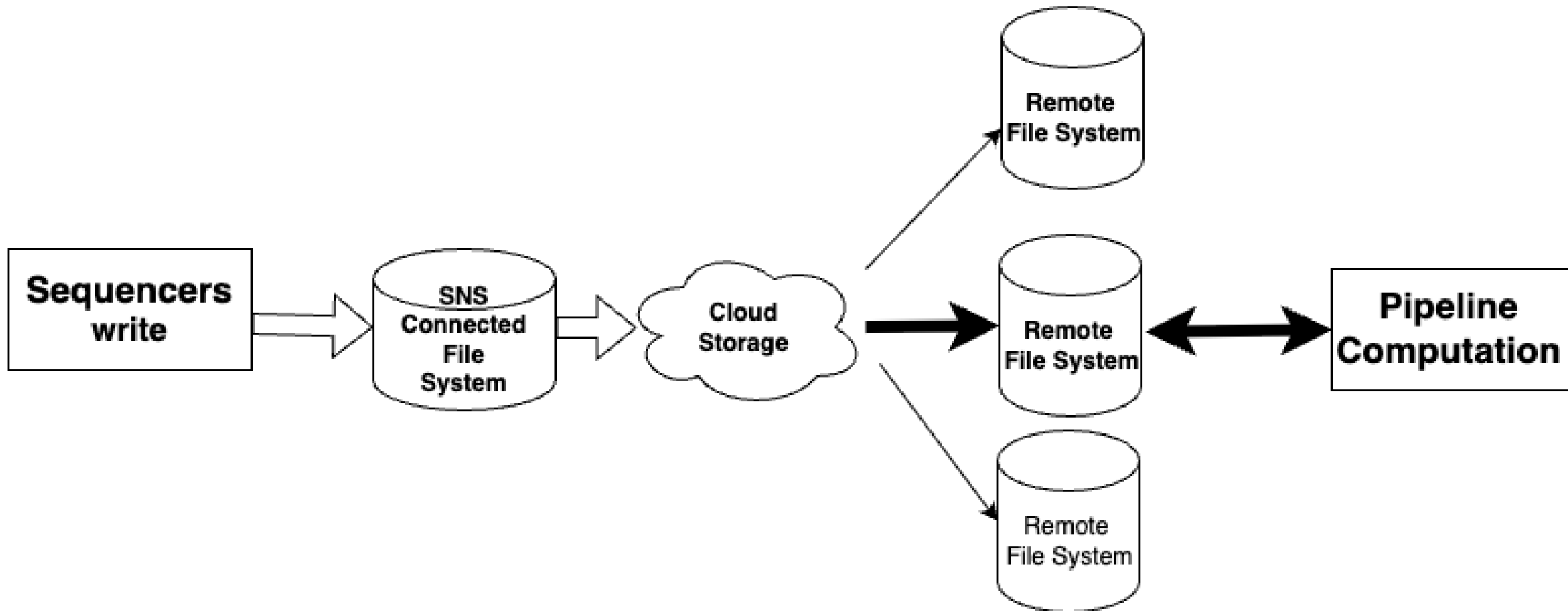
Workflow 2: clinical monitoring (b)



Workflow 3: No local computation (a)

- Sequencers write to SNS enabled file system
- No local computation
- Data pushed to cloud storage
- Metadata becomes available in remote file systems
- Computation run remotely
- Examples:
 - Guardant Health in a Box systems for overseas hospitals
 - Remote GH lab
 - Computational Maintenance

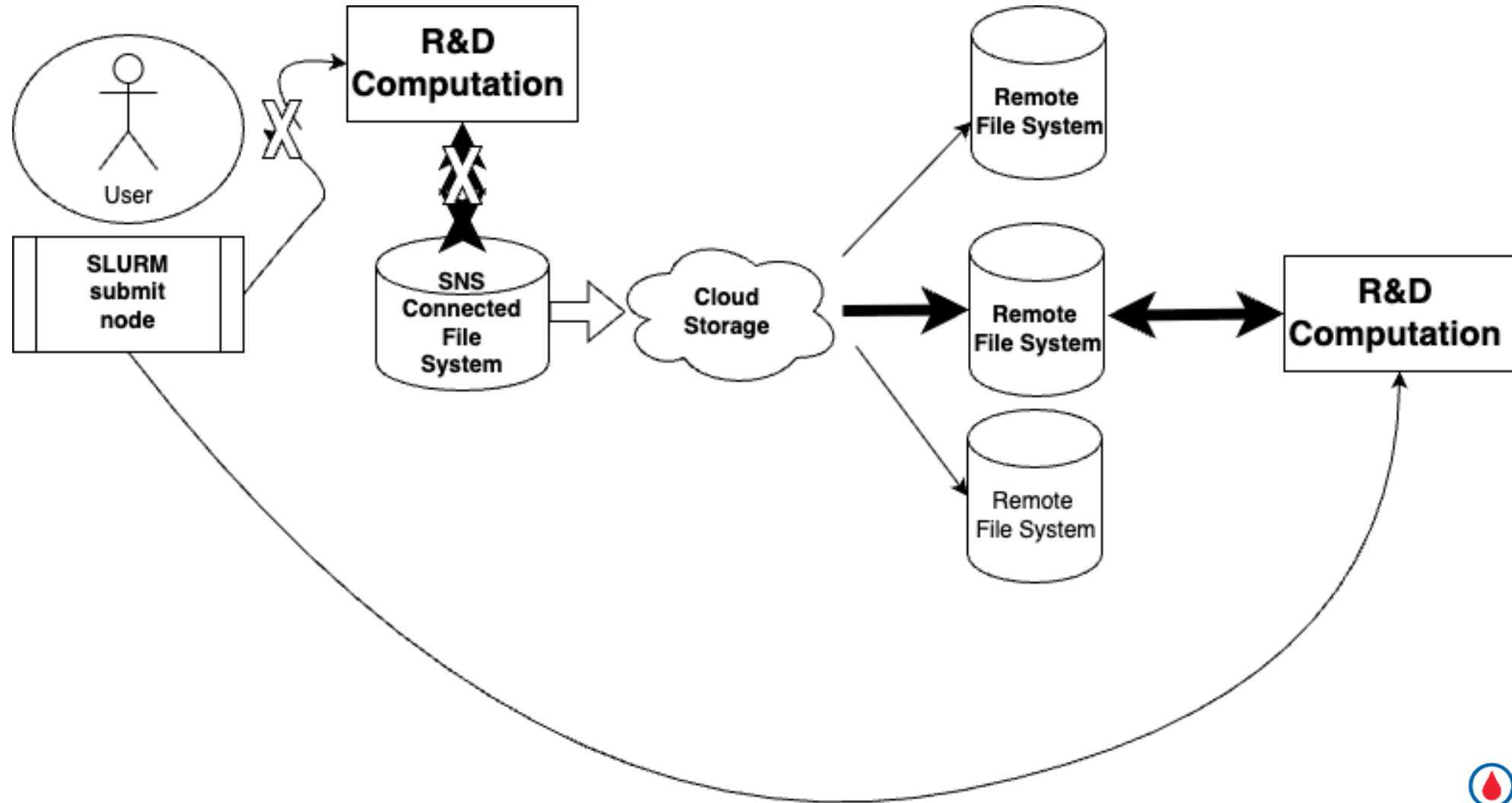
Workflow 3: No local computation (b)



Workflow 4: Insufficient Local Compute (a)

- User data in SNS attached file system
- No local computation available by deadline
- Jobs burst to remote cluster/cloud via slurm
- Data pulled to compute locale
- Computation run remotely

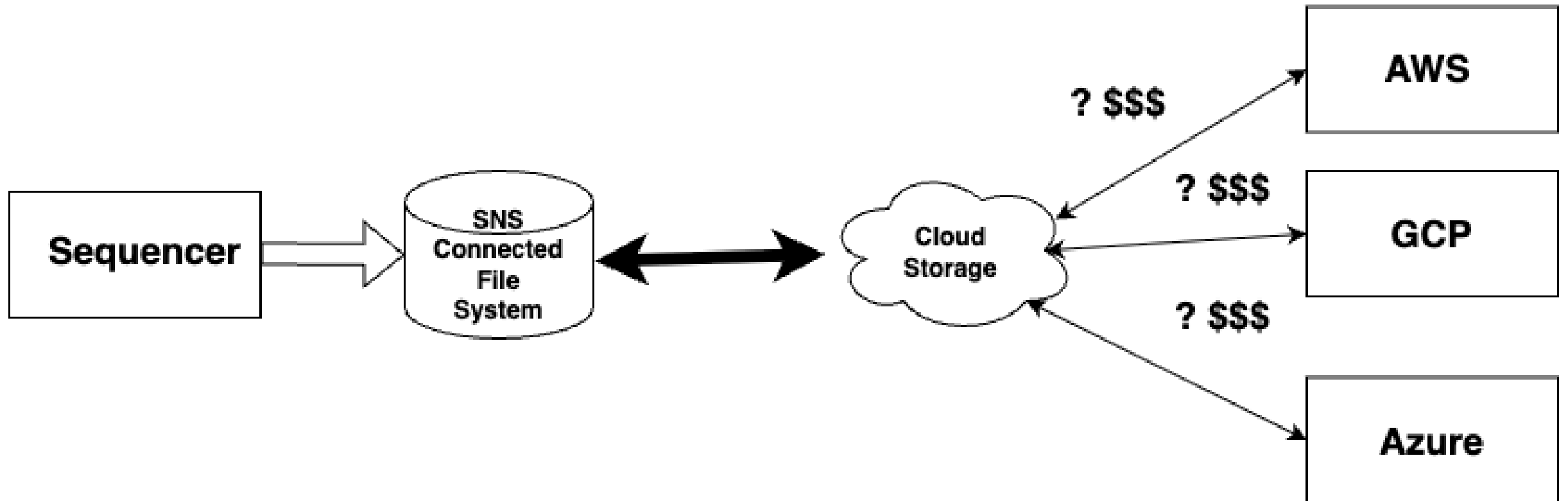
Workflow 4: Insufficient Local Compute (b)



Workflow 5: Workload Arbitrage (a)

- User data in SNS attached file system
- No local computation available by deadline
- Jobs burst to remote cluster/cloud via slurm
- Data pulled to compute locale
- Computation run remotely
- Distribution of workloads based on cost

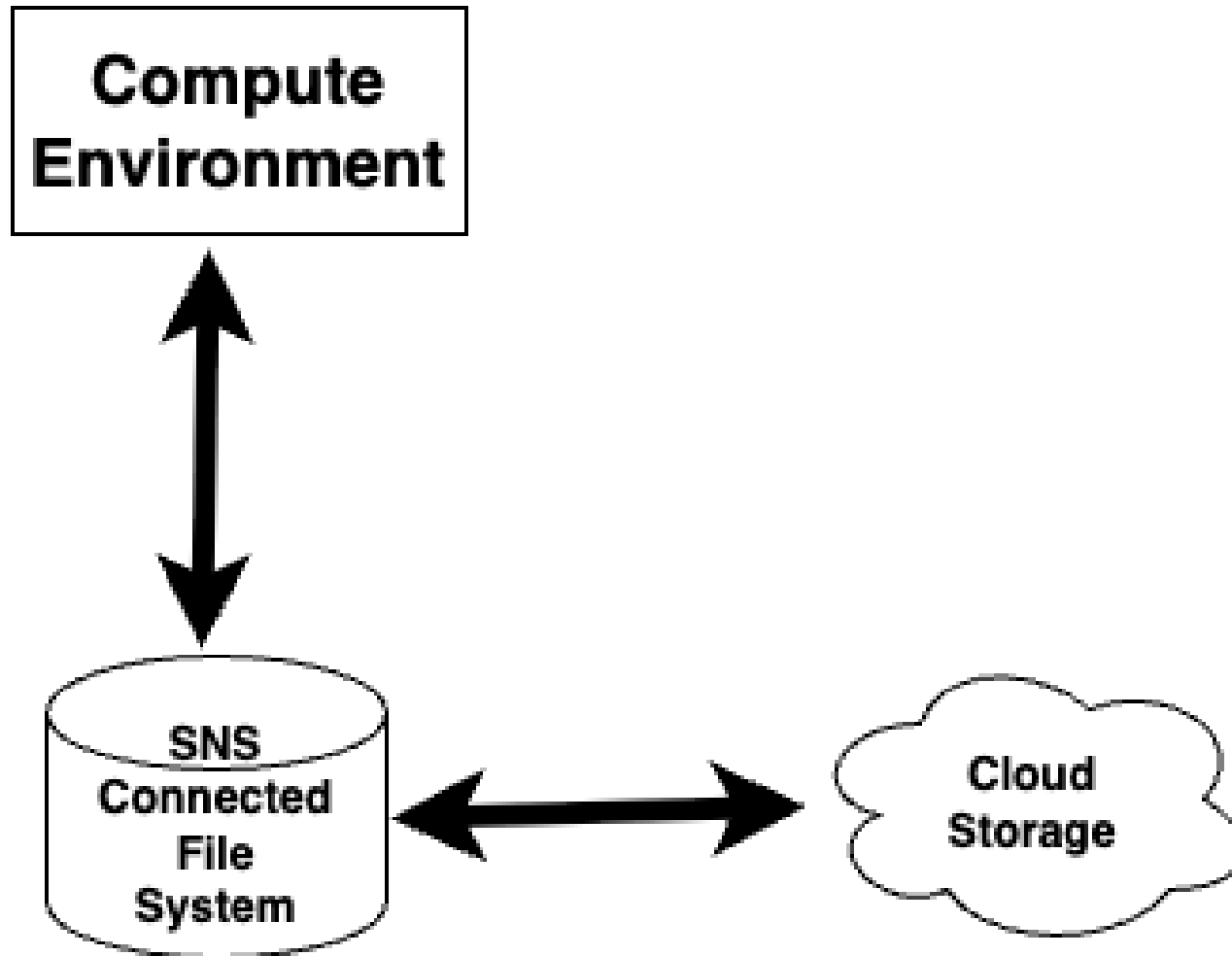
Workflow 5: Workload Arbitrage (b)



Workflow 6: Ultimate Data Mine(a)

- All data in cloud storage
- Massive study needed
- Data far larger than practical to hold in single file system
- Data pulled to local storage in rolling fashion
- Compute run locally
- Rolling hydration/dehydration

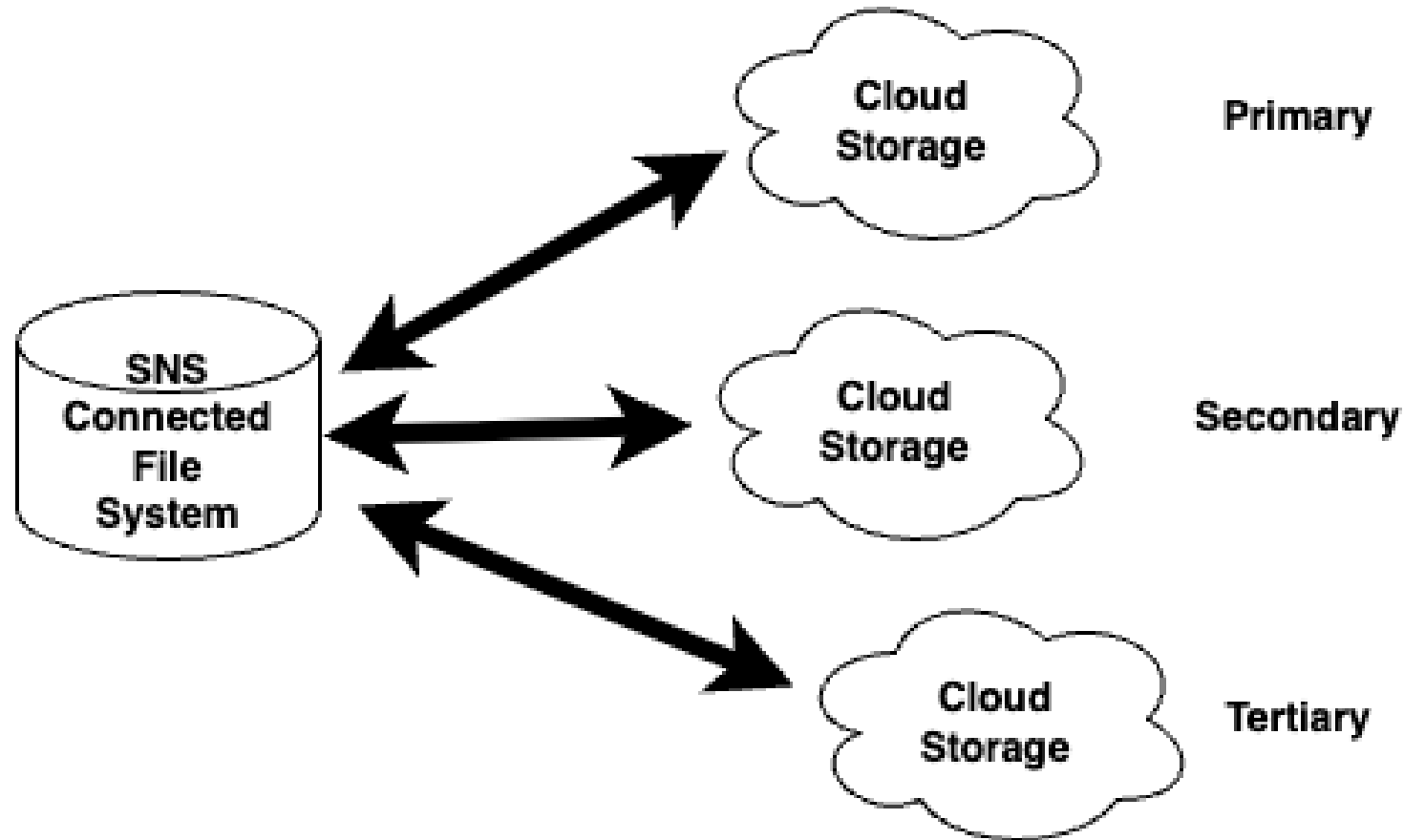
Workflow 6: Ultimate Data Mine (b)



Workflow 7: Data Replication (a)

- All data in cloud storage
- Multiple copies needed
- Data actively written to and replicated
- Single 'stub' points to multiple cloud data copies

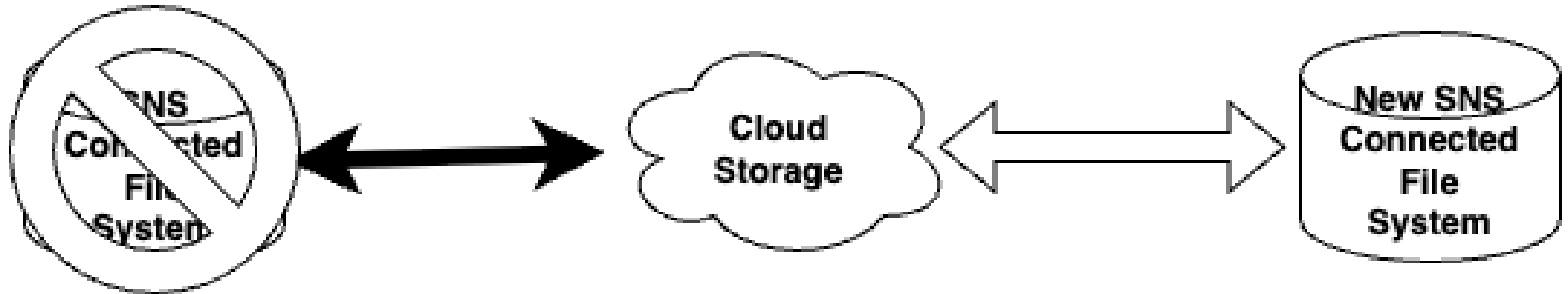
Workflow 7: Data Replication (b)



Workflow 8: Disaster Recovery (a)

- All data in cloud storage
- Local File System Destroyed
- New file system created and attached
- Metadata absorbed and file system ready for use
- No data in cloud lost

Workflow 8: Disaster Recovery (b)



Workflow 9: Multisource Science (a)

- All data in cloud storage
- Data from different organizations
- Each separate buckets
- Added to large working file system temporarily
- Access granted by groups/roles but appears as single FS

Workflow 9: Multisource Science (b)

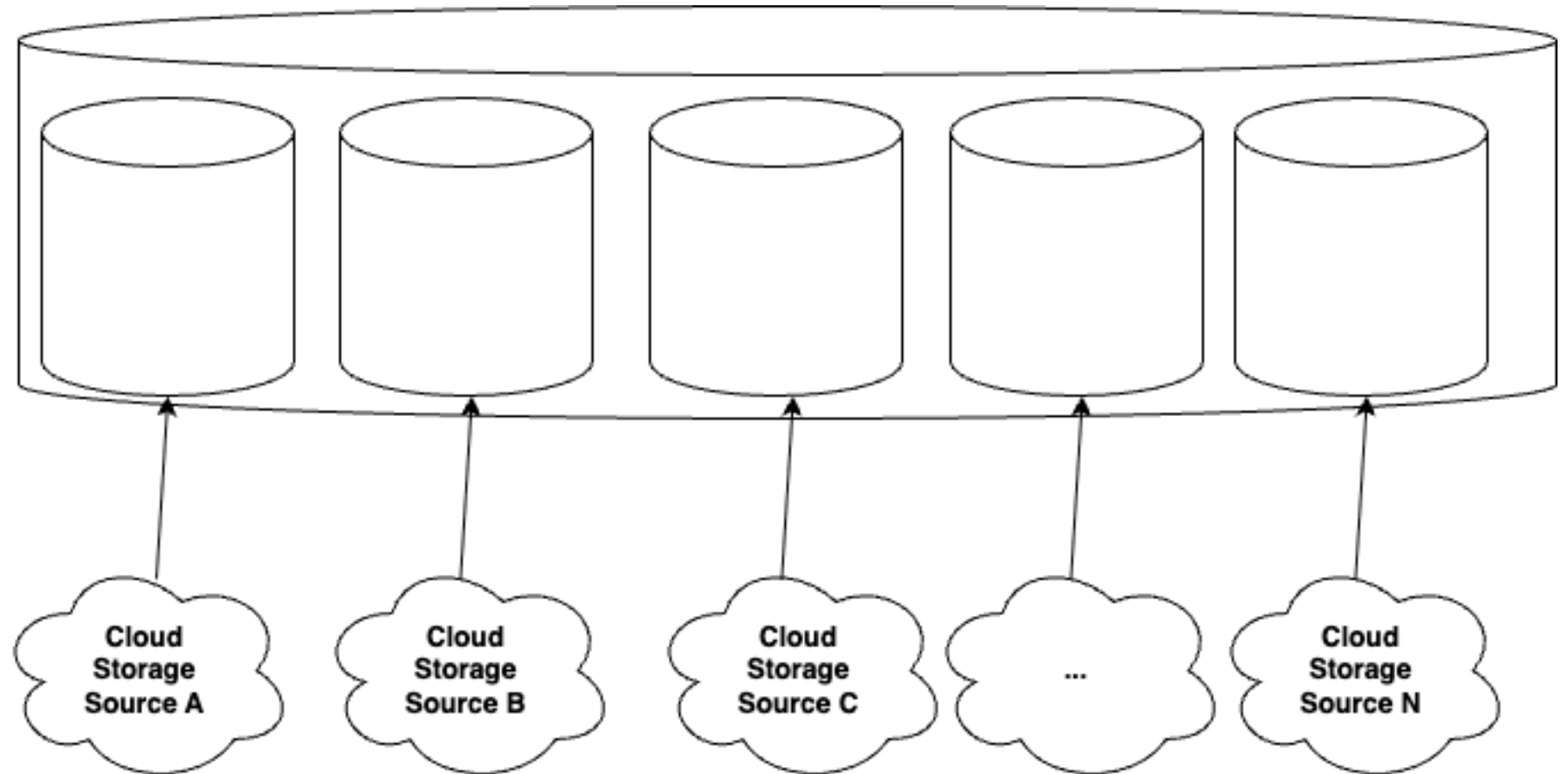
Monolithic FS

Attached to SNS

multiple file sections

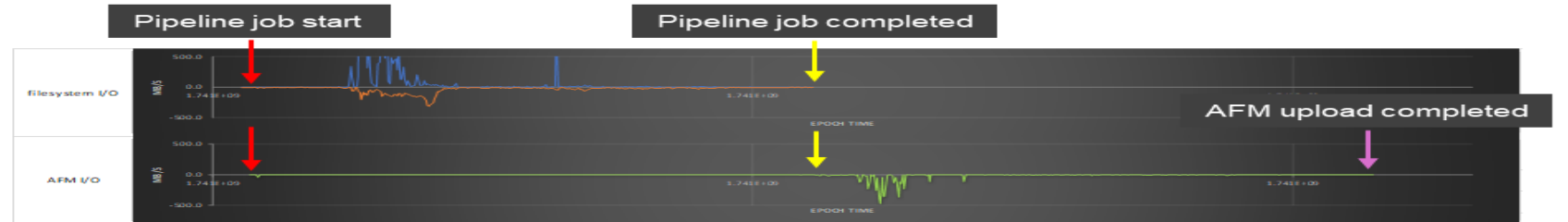
RBAC

Multiple Sources

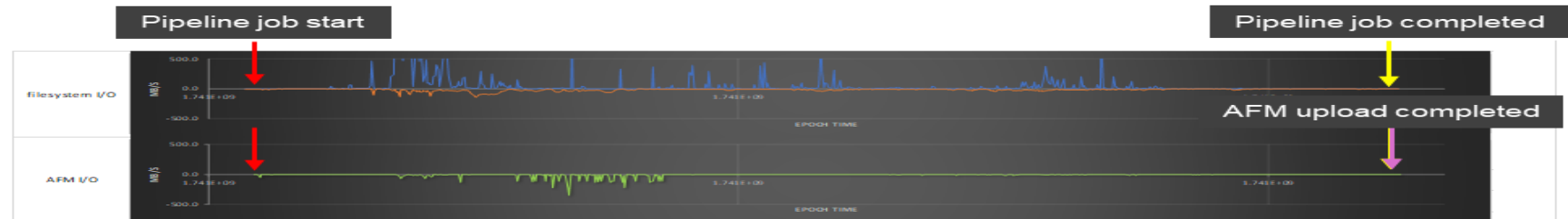


The streaming vs on close tests

```
afmObjectSyncOpenFile=no  
afmNumFlushThreads=32  
afmNumReadThreads=64  
afmNumWriteThreads=64
```



```
afmObjectSyncOpenFile=yes  
afmNumFlushThreads=32  
afmNumReadThreads=64  
afmNumWriteThreads=64
```



```
afmObjectSyncOpenFile=yes  
afmNumFlushThreads=32  
afmNumReadThreads=64  
afmNumWriteThreads=64
```





Prod Deployment Dependencies

We are here ➡

- RHEL must be completed
- Seagate data move completion
- AWS testing
- Network upgrades
- GPFS upgrade (and testing)
- Seagate data reorg

Single Namespace Working Group

Single Namespace Working Group

- **Technical group to meet and set a standard for compatibility between vendors**
- **Now on 10th meeting: hosted at Guardant Health on August 27th and 28th**
- **Presentations and documents:**
 - <https://github.com/SingleNameSpace/sns/wiki/Presentations-and-Papers>

- **Next step: writing the specification**

- Divided into 5 subcategories
 - Frontend
 - Backend (*)
 - Interaction
 - API
 - Security (*)

We are working on a logo.
promise.

Meeting schedule:

8/27/25: Guardant/PA
9/24/25: netapp
10/23/25: Genentech
11/18/25: SC25/St Louis

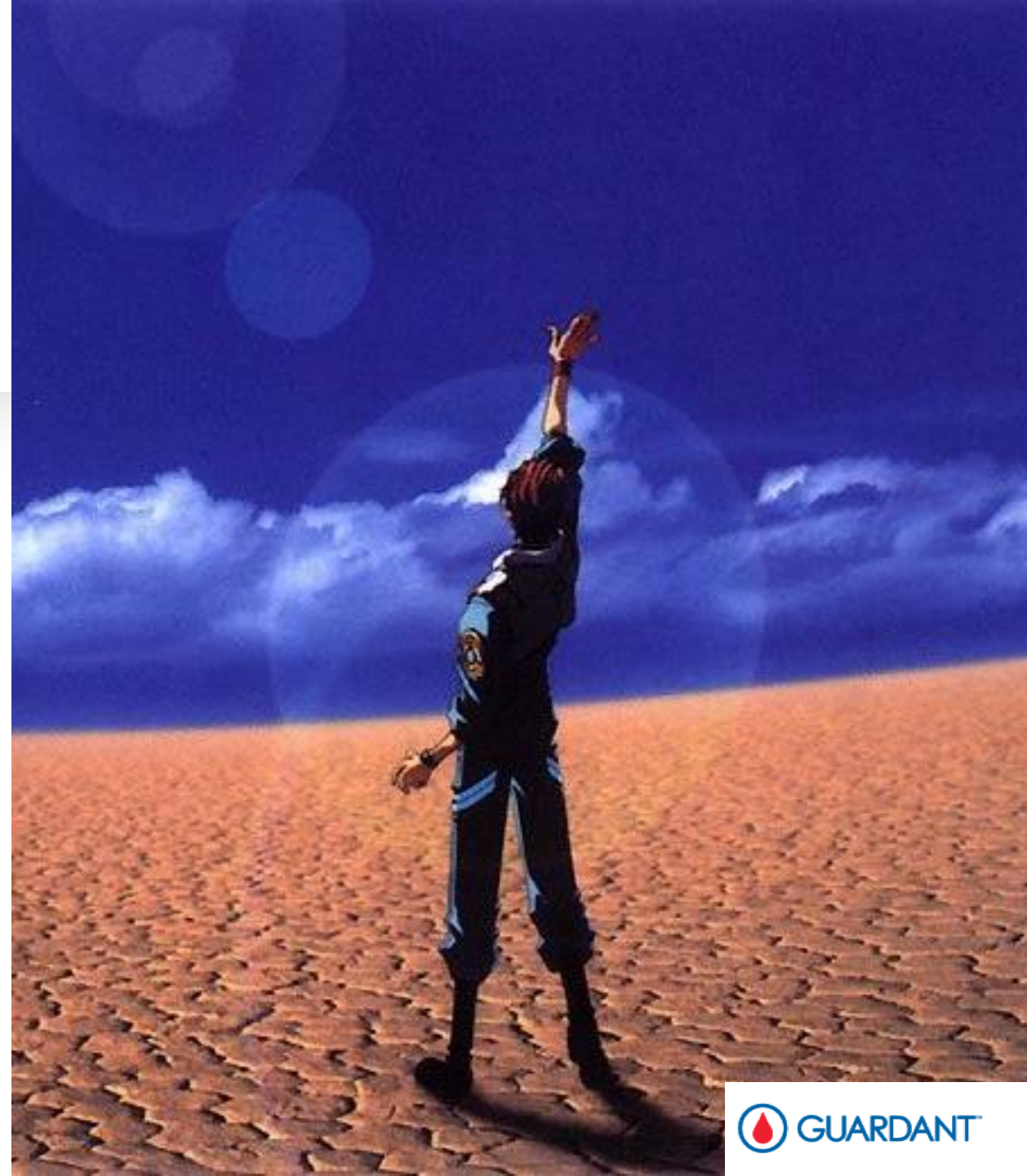
- **Estimated Draft Finished in August 2025**
- **Proposed Standards Body: OASIS**
- **Launching User Group at SC25: Single Namespace User Group (SNUG)**

Single Namespace Working Group Today



Future Directions

- **Multiple backends**
 - GPFS 5.2.x allows for multiple backend objects targets for a single file stub
 - Improves resiliency
 - Allows for cheaper options
 - Needs to be tested!
- **Other Clouds**
 - POC with GCP planned 2H25 (post RHEL)
 - Azure also planned FY26
 - Testing with specialist clouds
 - Intent to allow maximum flexibility



Wrap up

- SNS buys Guardant quite a bit
- Today is using GPFS + Seagate
- SNS has its quirks
- Work continues
- There are dependencies for prod deployment (mostly people time)
- SNS Working Group moving forward
- *We have a path forward*

A photograph of a modern, multi-story building with a glass facade and a flat roof. A vibrant rainbow is visible in the overcast sky above the building. In the foreground, there is a parking lot with several cars, including a silver sedan and a dark SUV. A tall, slender light pole stands near the building. The overall scene is dimly lit, suggesting an overcast day.

Questions?