



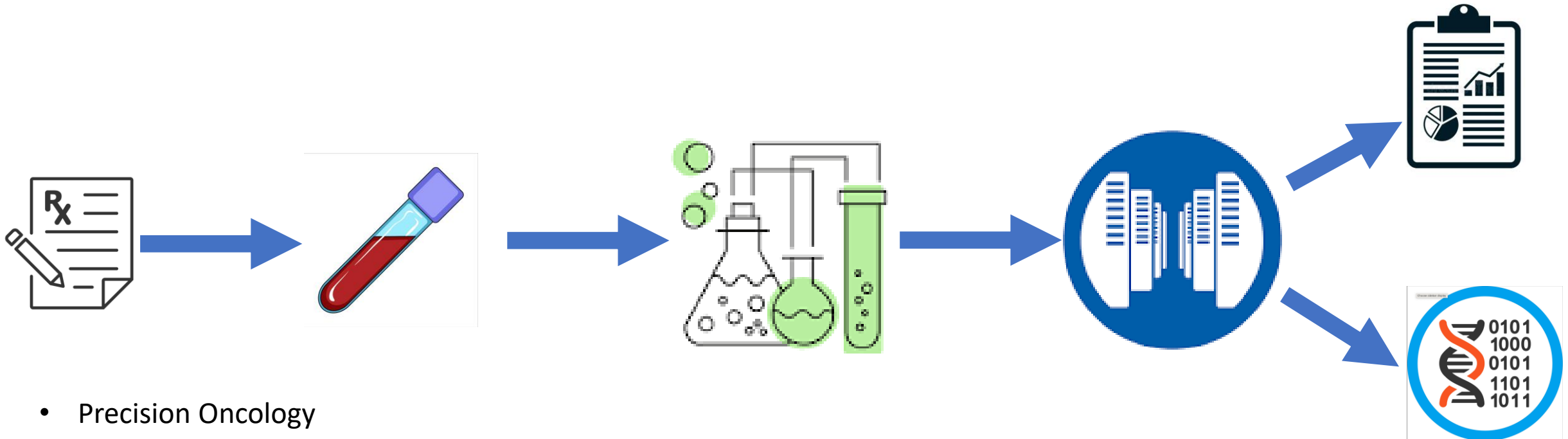
One Namespace to Bind Them

Adventures in a wide area single namespace with GPFS, S3 and crazy HPC people

12 November 2023

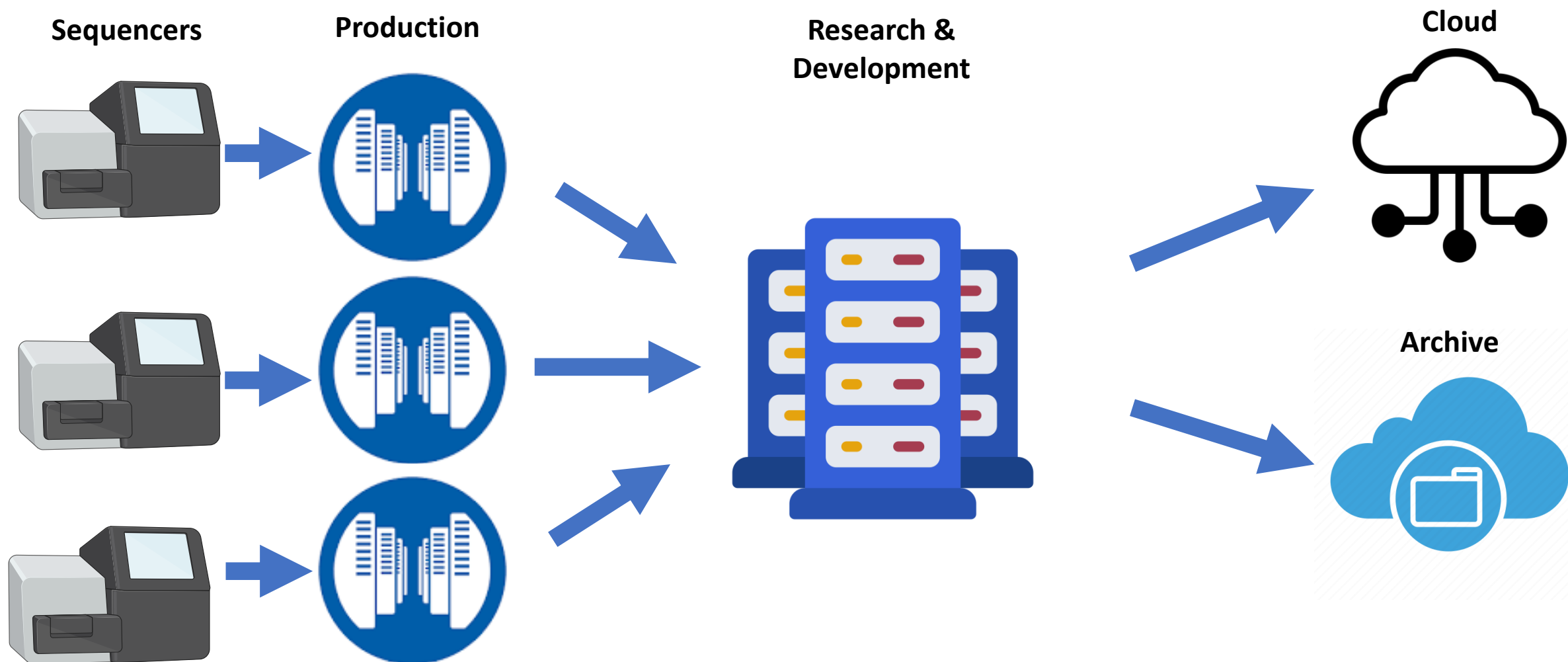
What is Guardant Health?

Conquering Cancer Through Data



- Precision Oncology
- Preventive Screening

Guardant Health HPC Dataflow



Guardant Health Systems



General Info:

- Over 1,000 CPU nodes
- Starting to use GPUs
- 70 petabytes of disk

Locations:

- Redwood City
- Palo Alto
- San Diego
- Barcelona
- London
- Shanghai
- Tokyo (once and future)
- & other sites to come

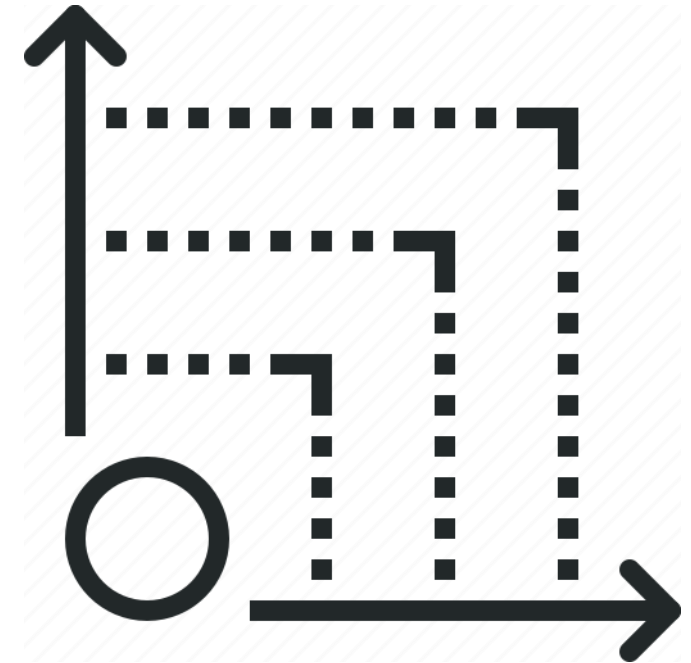
Guardant Health HPC Scaling Problems

Our problems

- Storage
- Data Center Space, Power & Cooling
- Computational Power
- Cost

The same old story for all of HPC, but...

we serve patients, not scientists, and costs matter



Guardant HPC Abstraction Project

Mostly not talking about the project

- Planning to discuss at BioIT (if accepted)

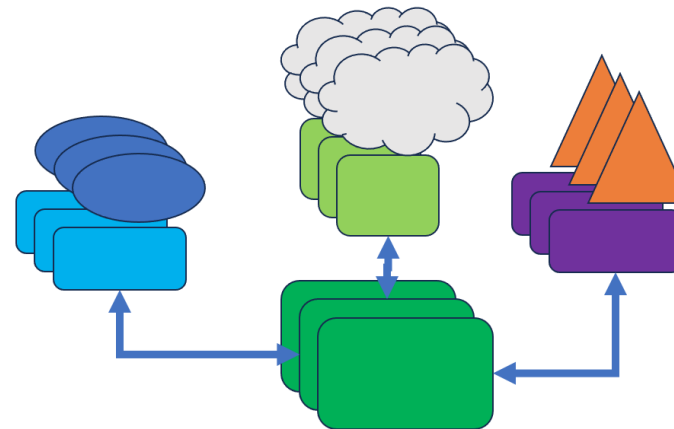
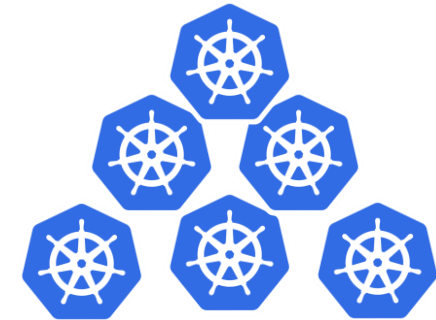
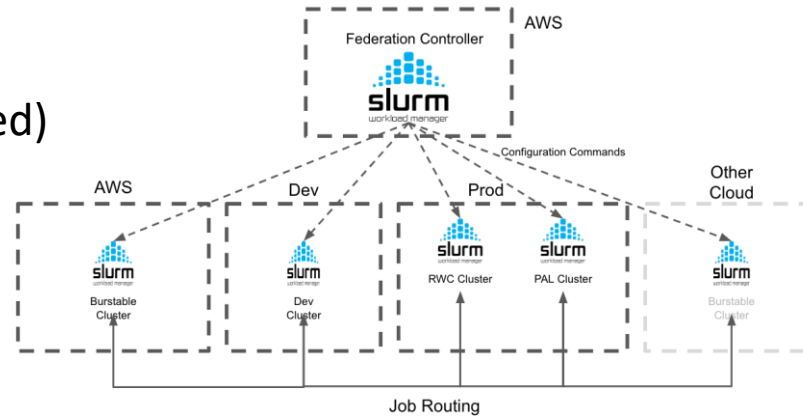
Why do we abstract?

- *To reduce costs (first and foremost)*
- *Improve & reduce development time*
- *Increase resiliency, flexibility*
- *Business decision where jobs run*

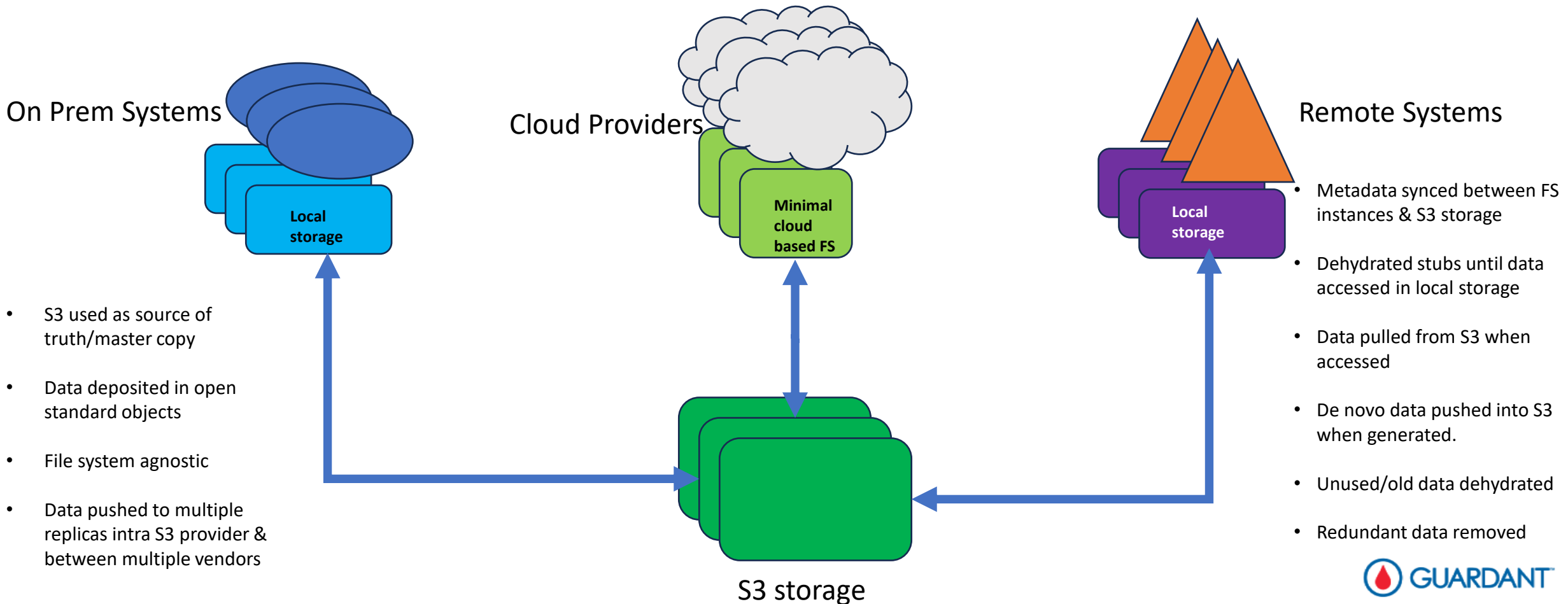
Three Subprojects:

- SLURM Federation
- Single Data Namespace
- One to be disclosed: k8s related

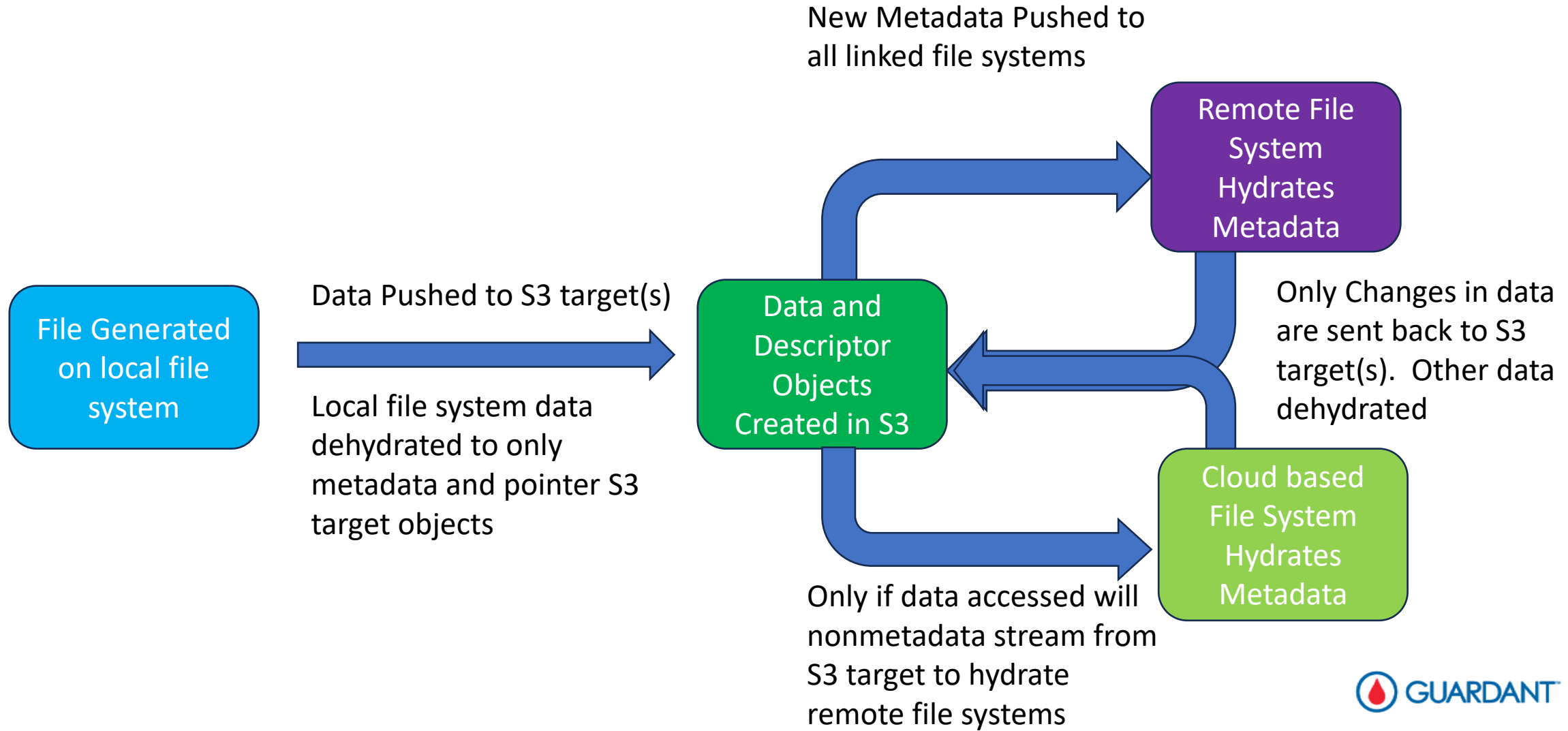
Today's topic is the single namespace



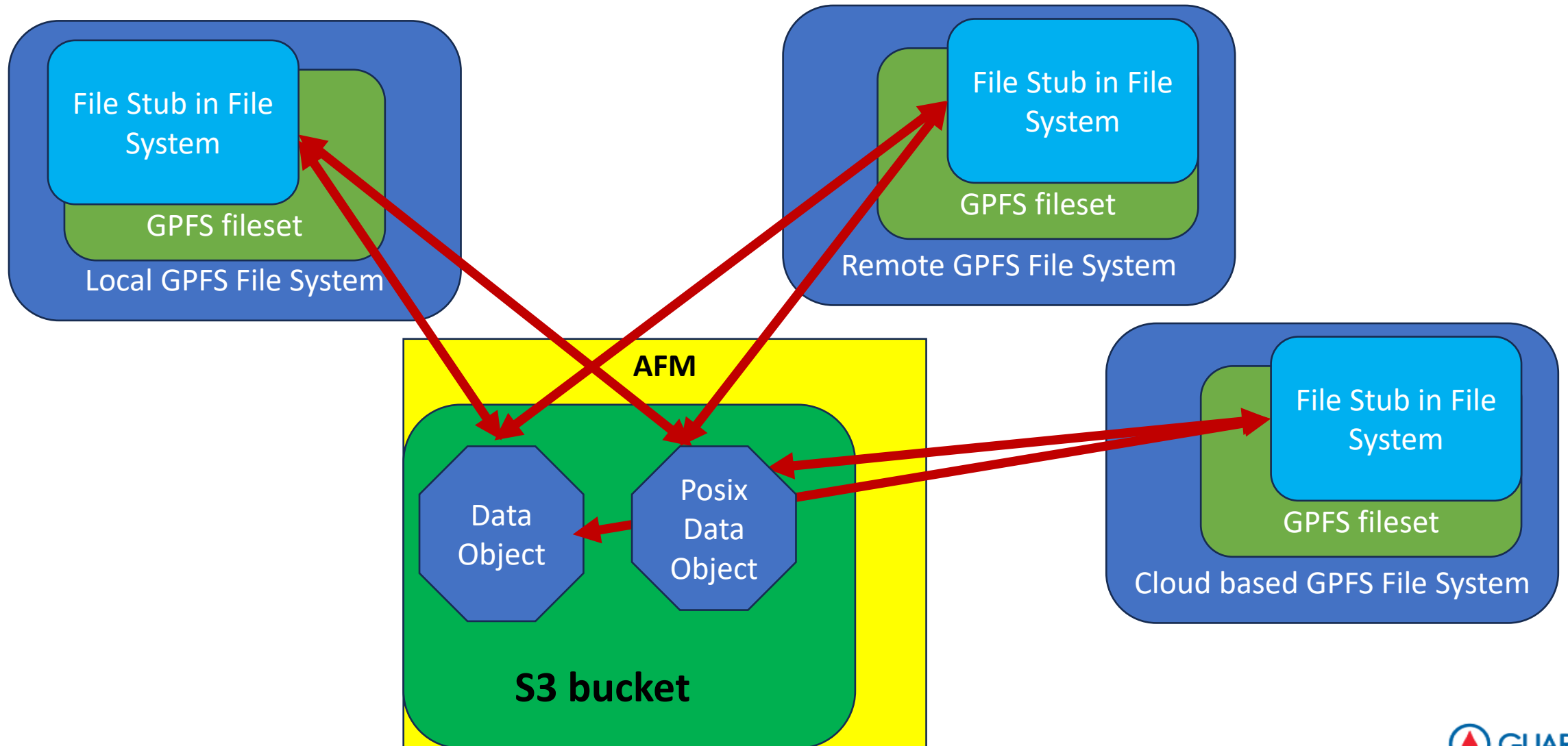
Conceptual Layout of Single Namespace



Event Driven View of Single Name Space



Drax How of Single Namespace



Tested and Functional, but...

The Good!

- It works! Running pipeline against alpha version
- Tested between Palo Alto, Santa Clara, Redwood City, and AWS
- Fantastic for disaster recovery
- Simplifies infrastructure

The Bad!

- Random disconnects
- AFM (any flavor) cannot handle billions of files to ingest for first time. Must be set up by another tool initially
- AFM needs a freakin multimode queue. HA is past the early Oughts.

The Ugly?

- Last writer wins (not so bad for us...but...)
- There is a *long, ugly* list of borken to fixen
- IBM has been slow (from Guardant POV) to embrace



Path	Size	Files	Dirs
/ghdevhome	284.7T	88.7M	11.1M
/ghds	12.6P	782.2M	57.7M
/ghsfa	22.9P	1.2G	29.8M

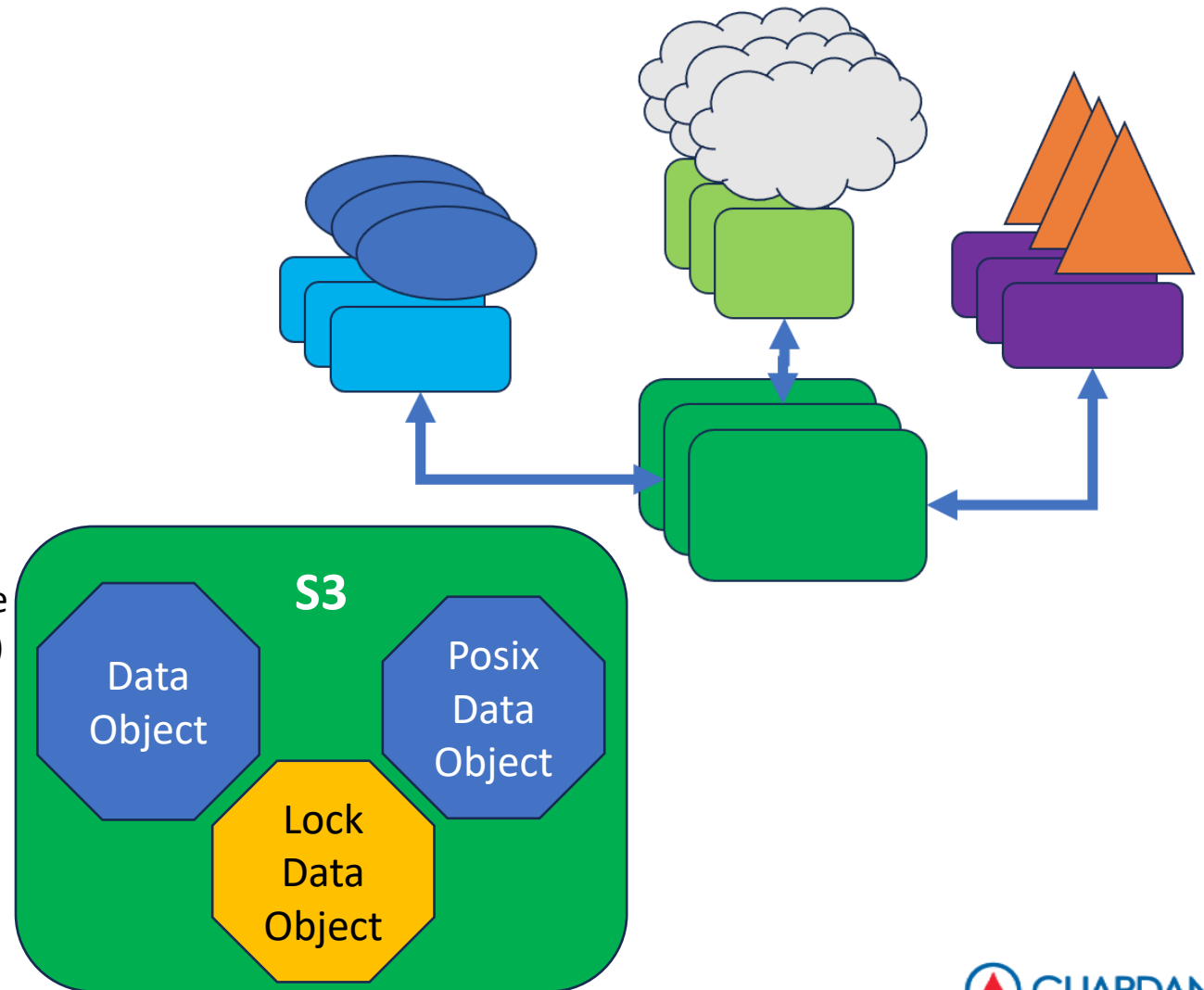
Next Steps for Guardant Health

The BETA!

- Stand up 2.5 PB beta for our dev & SRE teams
- Using Seagate's S3 targets (rstor/lyve cloud)
- Beta will last 1 year

And Then?

- Work with IBM on enhancements, such as...
 - A proposed lock object
 - Multiple parallel buckets by different providers for single file
 - Multinode, multi threaded, parallel AFM queue
 - Improved CSI drivers (parallel, multi node, etc.)
 - CSI driver integrated with single namespace
 - More...
- If IBM moves too slow, but Guardant is all in...



Wrap-up, Conclusions and Parting Comment

- **Single namespace does appear to work but is not production ready *yet*.**
- **Holds enormous promise for data anywhere. Once the data is anywhere...**
- **Final thought: objects in S3 are not IBM specific format**
 - **Another other file system reads the format and presents the data from S3 into their file system:
Cambridge Computing's Virtual FS**
 - **If one company has done it...**