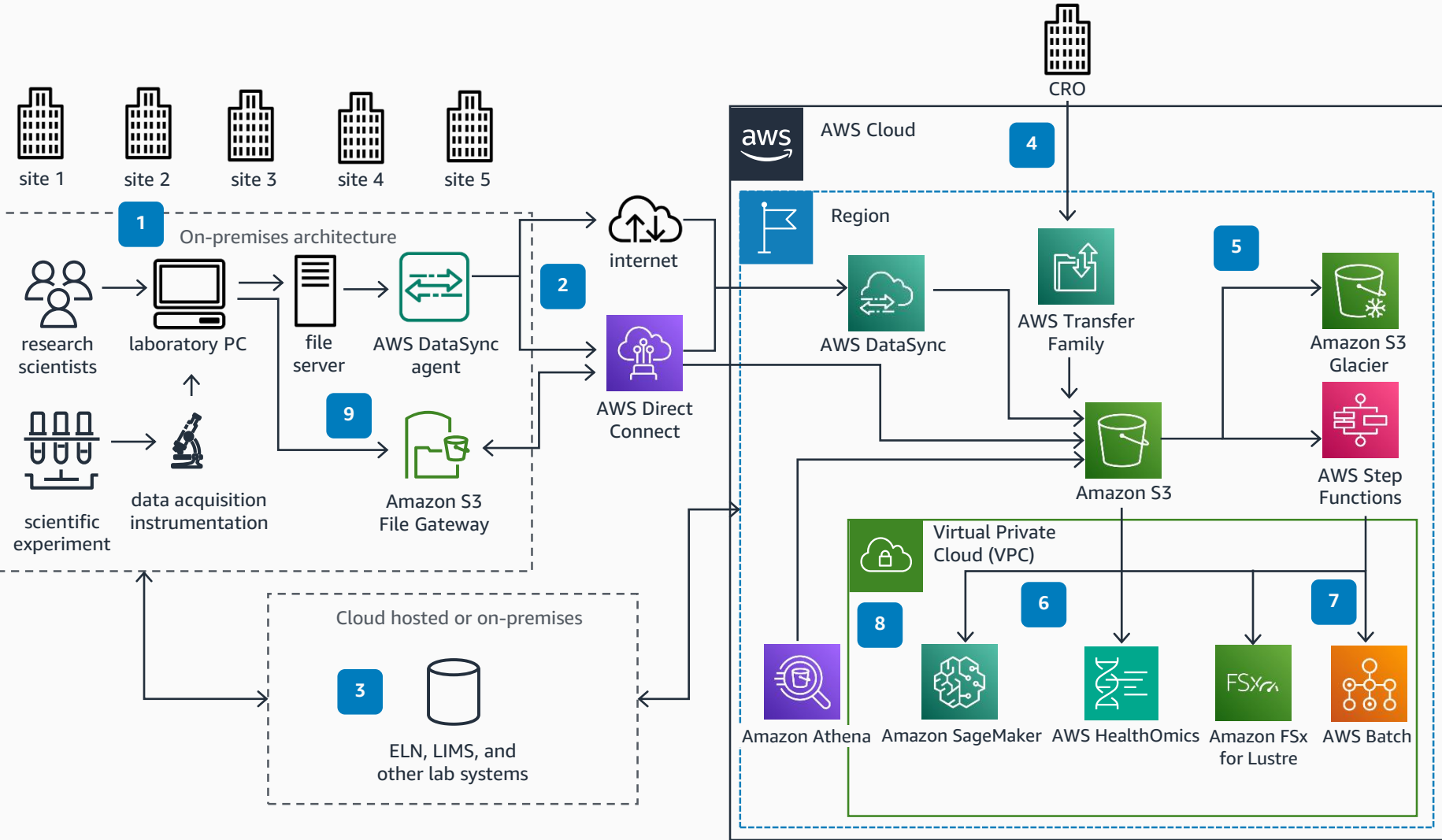


Guidance for Digital Connected Labs on AWS

This architecture diagram helps you learn how to connect file-based life sciences instruments and laboratory systems to the cloud and provide scalable data access and computing using Amazon Web Services (AWS).



- 1 A lab technician runs an experiment or test, and results are written to a folder on an on-premises file server. An **AWS DataSync** task is set up to sync the data from local storage to a bucket in **Amazon Simple Storage Service (Amazon S3)**.
- 2 Data is transferred to the **AWS Cloud** either through the internet, or through a low-latency direct connection that avoids the internet, such as **AWS Direct Connect**.
- 3 Electronic lab notebooks (ELN) and lab information management systems (LIMS) share experiment and test metadata bidirectionally with the AWS Cloud through events and APIs. Learn more about this integration in [Guidance for a Laboratory Data Mesh on AWS](#).
- 4 Partnering entities, like a contract research organization (CRO), can upload study results to **Amazon S3** by using **AWS Transfer Family** for FTP, SFTP, or FTPS.
- 5 You can optimize storage costs by writing instruments data to an **S3** bucket configured for infrequent access. Identify your **S3** storage access patterns to optimally configure your **S3** bucket lifecycle policy and transfer data to **Amazon S3 Glacier**.
- 6 Using **Amazon FSx for Lustre**, data is made accessible to high performance computing (HPC) on the Cloud for genomics, imaging, and other intensive workloads to provide a low millisecond-latency shared file system.
- 7 Bioinformatics pipelines are orchestrated with **AWS Step Functions**, **AWS HealthOmics**, and **AWS Batch** for flexible CPU and GPU computing.
- 8 Machine learning is conducted with an artificial intelligence and machine learning (AI/ML) toolkit that uses **Amazon SageMaker** for feature engineering, data labeling, model training, deployment and ML operations. **Amazon Athena** is used for flexible SQL queries.
- 9 For researchers using on-premises applications for data analysis and reporting, they view and access data in **Amazon S3** by using Network File System (NFS) or Server Message Block (SMB) through **Amazon S3 File Gateway**.