

EXACT SCIENCES

”Exact Sciences has implemented AWS Storage Gateway as the foundation for NGS Data Lake on AWS, **relying on the flexibility, scalability, ease of management, and native AWS integrations**. The solution scales indefinitely to provide robust processing for short term time-sensitive workloads and long term data lake capabilities, all while **lessening on-premises footprint in favor of relying on cloud native services.**”

Raj Jampa
VP, Infrastructure Operations

Problem

Producing large amounts of NGS data in labs around the world and need to send it to AWS for processing

Rapidly expanding global lab footprint and need to scale data transfer capabilities

Current on-premise solutions are complex to manage and expensive to procure and deploy

Solution

Deploy multiple Storage Gateway hardware appliances to transfer data from NGS lab machines to Amazon S3

Seamlessly access 100s of TBs of data in Amazon S3 for processing

Outcome

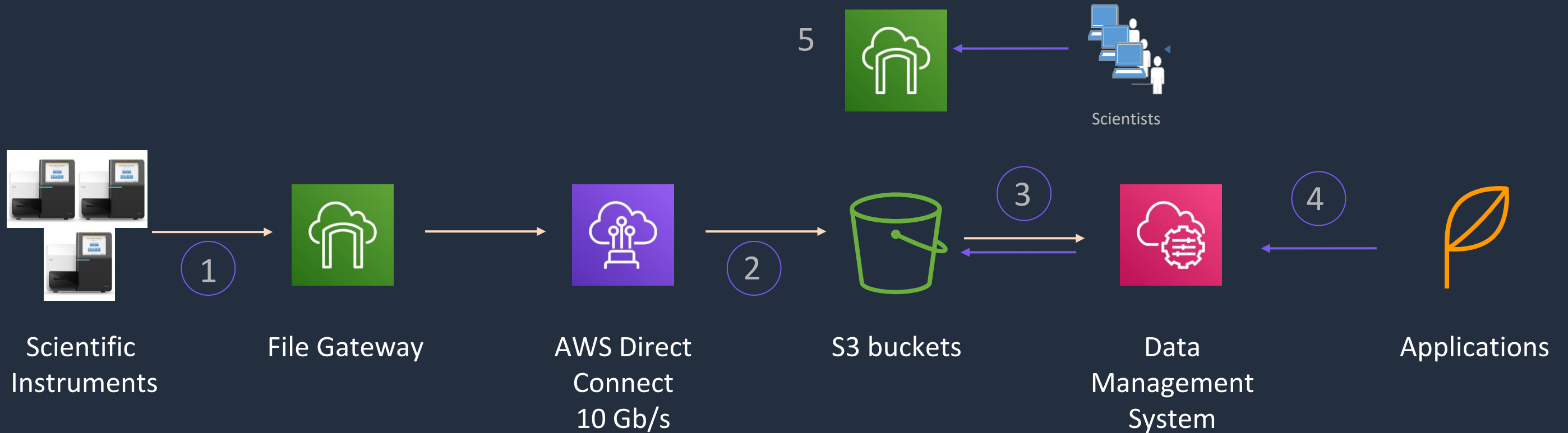
Scalable data ingestion solution that can be rapidly deployed

Reduce on-premises infrastructure management, cost, and automate data pipeline management

Accelerate data migration to cloud and decrease turnaround time of data processing

Seamless integration with AWS services

Exact Sciences: General Processing workflows



1. Instruments write raw data into File Gateway file share

2. File Gateway transfers files to S3 buckets

3. Data Management system scans S3 buckets regularly

4. Applications request data via Data Management system meta catalog

5. Files accessible by Scientists through another File Gateway