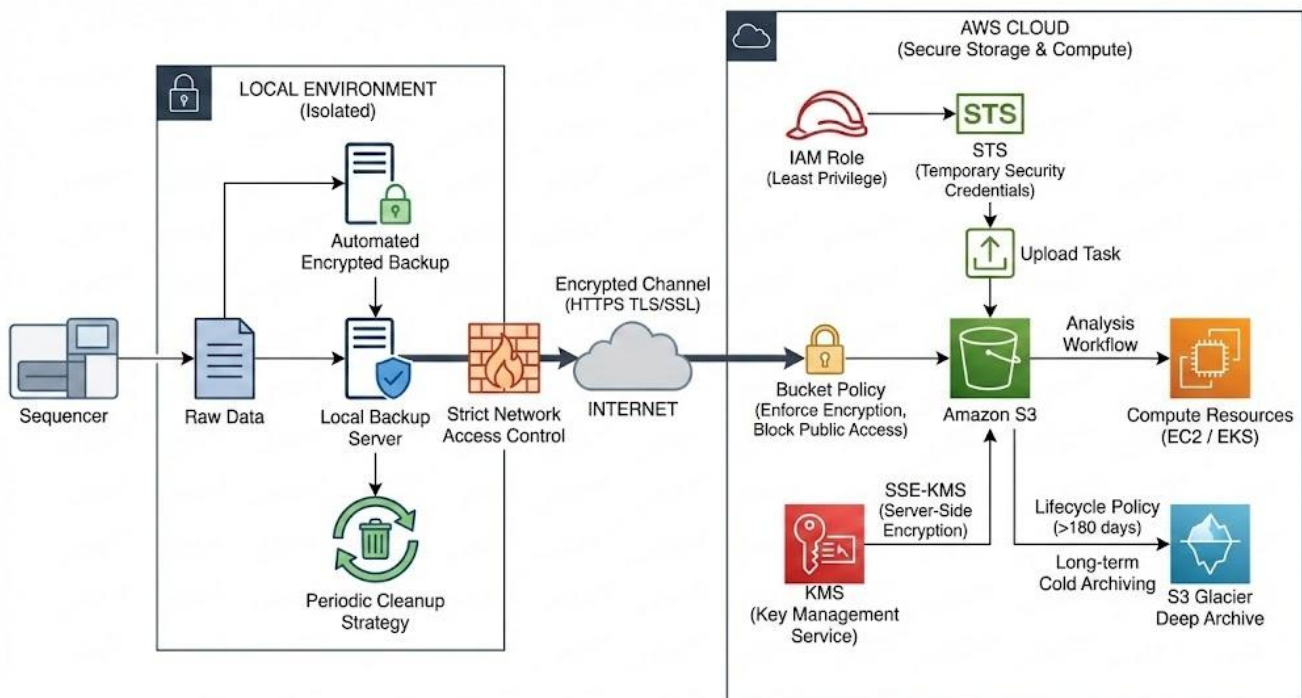


Ensuring Data Security and Compliance: A Comprehensive Framework for SNP Data Management in the Cloud

Updated 2026.01.29

IT System and Architecture Overview

Our process begins with the generation of raw data from sequencing instruments. Once the data is produced, it is automatically encrypted and backed up in a local secure environment that is completely isolated from the public network. The local backup server strictly adheres to network access control protocols, allowing only authorized internal access. During this stage, we implement regular local data cleanup strategies to free up storage space while retaining only the necessary active data for defined retention periods.



The data is then securely uploaded to the cloud via an encrypted channel. Throughout the transmission process, we use HTTPS (TLS/SSL) protocols to prevent data theft or tampering during internet transfer. Access permissions are enforced through **AWS Identity and Access Management (IAM)**, implementing the principle of least privilege. We

configure specific IAM roles for upload tasks, utilizing temporary security credentials to eliminate the risk of long-term key exposure.

Upon reaching the cloud, all core business data is stored in Amazon S3 object storage. For data storage, we enable **Server-Side Encryption (SSE-KMS)**, automatically encrypting static data with keys managed by **AWS Key Management Service**. Additionally, through finely configured S3 bucket policies, we enforce the requirement that all data must be encrypted and strictly prohibit public access, thereby establishing a strong data security boundary in the cloud.

Once the data has successfully persisted in S3, it triggers subsequent analytical workflows using elastic computing resources such as Amazon EC2 or Amazon EKS. For long-term data retention, we adhere to an established lifecycle strategy: all data will be retained in the cloud for the long term. Data exceeding 180 days after the completion of business cycles is automatically migrated from the S3 standard storage layer to the Amazon S3 Glacier Deep Archive service for deep cold archiving. This approach fulfills compliance and audit requirements while achieving minimal long-term storage costs.

1. Data Privacy

At the infrastructure level, we use Amazon Virtual Private Cloud (VPC) for logical network isolation, creating strict boundaries for production, testing, and development environments. This prevents unauthorized access and network-layer attacks.

For access control, we apply the principle of least privilege through AWS IAM, granting customized roles with fine-grained permissions for operations, development, and business systems. This limits access to only necessary data and resources, reducing the risk of data breaches.

All actions are logged via AWS CloudTrail and other services, enabling thorough security audits and incident tracing.

To protect data privacy, we employ technical measures and management processes to ensure sensitive data is accessible only to authorized personnel and systems. We do not share or transmit data to unauthorized third parties and collect only what is necessary for business operations, minimizing complexity and exposure to privacy breaches, thus adhering to compliance requirements like GDPR.

2. Data Security

We have implemented a comprehensive encryption system for data security, covering both static storage and dynamic transfer. For static data, we use Amazon S3's default server-side encryption (SSE-S3) with AES-256, ensuring data remains encrypted at rest. During transmission, all communications are secured with HTTPS/TLS protocols, preventing

data theft and ensuring end-to-end encryption.

To address vulnerabilities, we enhance security through multiple layers of network controls. We utilize AWS security groups and Network Access Control Lists (NACLs) to manage inbound and outbound traffic, allowing only authorized IPs and ports to access critical services. We also regularly update security patches and employ automated tools like AWS Trusted Advisor and Amazon Inspector to scan for vulnerabilities.

All configurations adhere to AWS security best practices, including IAM policies based on least privilege, and we log all API calls via AWS CloudTrail for audits. This layered protection minimizes risks of data breaches and unauthorized access, ensuring compliance and resilience in our workloads.

3. Backup and Disaster Recovery

For core business data storage, we've established a dual high-availability architecture for both object and relational data, ensuring durability and business continuity. Amazon S3 provides industry-leading reliability with up to 99.999999999% (11 nines) durability and a 99.99% SLA for availability, supported by a Multi-AZ architecture that redirects traffic in case of a single availability zone failure. We also enable cross-region replication to ensure safe data recovery during severe regional failures.

For relational databases, Amazon RDS supports online operations with built-in high availability. Its automatic failover can switch to a standby instance in a different availability zone within 60 seconds, ensuring zero data loss and minimizing operational downtime. We also enable automatic backups and point-in-time recovery, regularly saving snapshots and transaction logs. This approach meets compliance requirements and allows precise restoration to any historical point within minutes, protecting against data loss from accidental actions or logical corruption.

4. Compliance

Our data security and privacy framework strictly complies with international regulations like GDPR and CCPA and is based on a multi-layered management approach. It integrates compliance into business processes, using data classification and labeling to tier personal information and protect sensitive data. We practice data minimization, collecting only essential personal data and establishing clear processes for handling data subjects' rights to access, correct, or delete their information. We also conduct proactive privacy impact assessments before adopting new technologies to identify and manage privacy risks.

Conclusion

Utilizing AWS cloud services, our system employs a defense-in-depth approach to ensure data security throughout its lifecycle, from generation to archiving. It starts with automated encrypted backups of sequencing instrument data in an isolated local network, followed by regular cleanup to optimize storage. Data is securely transmitted to the AWS cloud via HTTPS/TLS, adhering to the principle of least privilege through IAM roles or temporary credentials, which mitigates long-term key exposure risk. Core data is stored in Amazon S3 with server-side encryption (SSE-KMS) and strict bucket policies for static data protection, with all operations logged for audits. Data analyses can safely trigger based on EC2 or EKS resources. Finally, data exceeding retention periods is automatically migrated to Amazon S3 Glacier for cost-effective archiving, ensuring compliance. This comprehensive approach creates a robust security framework for your data.

Information Security & Compliance Team

MolBreeding Biotechnology LLC