Archer™ Analysis Unlimited—Cloud-based, scalable genomics analysis solution

Archer Analysis Unlimited provides a reliable genomics analysis platform that couples comprehensive regulatory compliance with stringent data protection capabilities.

Key points

- Easy-to-implement, cloud service hosted through Amazon Web Services (AWS) configured to maximize computing usage.
- Full support and maintenance by Archer means no hardware issues, along with a secure computing environment with automatic daily backups.
- Complete control over software updates allowing labs to use the latest version of Archer Analysis or continue to run their current version.

Archer Analysis Unlimited is a cloud-based computing solution

Archer Analysis is a secure, reliable software solution for analyzing genomic sequencing data from RNA, DNA, and circulating tumor DNA (ctDNA) assays that is available as:

- **Archer Analysis Unlimited**—cloud service hosted through Amazon Web Services (AWS)
- **Virtual machine**—installed on premise using a lab's hardware

Although multiple data security and privacy measures are built into both Archer Analysis Unlimited and the virtual machine, key differences exist between of the two options (Table 1).

Table 1. Key features of Archer Analysis Virtual Machine compared to Archer Analysis Unlimited.

Features	Virtual machine	Archer Analysis Unlimited
Customer support	Archer supports VM installation and troubleshooting	Complete remote bioinformatics and IT support
Version updates	Previous data does not port to new version	Available upon request, preserve previous data
Audit trail	Requires a local solution for capturing activity logs to ensure regulatory compliance	Activity logs are monitored and captured in the cloud
Data encryption	Encryption mechanisms are set up and maintained by lab	Supported and maintained by Archer at rest and in transit using industry standards
Hardware support	Machine installation, configuration, and upkeep are handled by lab	Access to scalable computing resources that are fully supported and secure
Disaster recovery	Managed by lab	Managed by Archer

Archer Analysis Unlimited is built for security and scalability

Hardware resource optimization: Labs with limited internal analysis support resources can access the best cloud computer resources through Archer Analysis Unlimited. This platform offers security, flexibility, and scalability for customers who need access to computing power on demand (Figure 1). Archer Analysis Unlimited leverages AWS to offer resource optimization:

- Auto-scaling of computer resources, allowing Archer Analysis Unlimited to provision the necessary amount of computer resources in minutes and scale up as the system's demands change.
- Cloud computing offers freedom from on-premises hardware limitations, which is particularly valuable for labs who work with terabyte- and even petabyte-scale data.

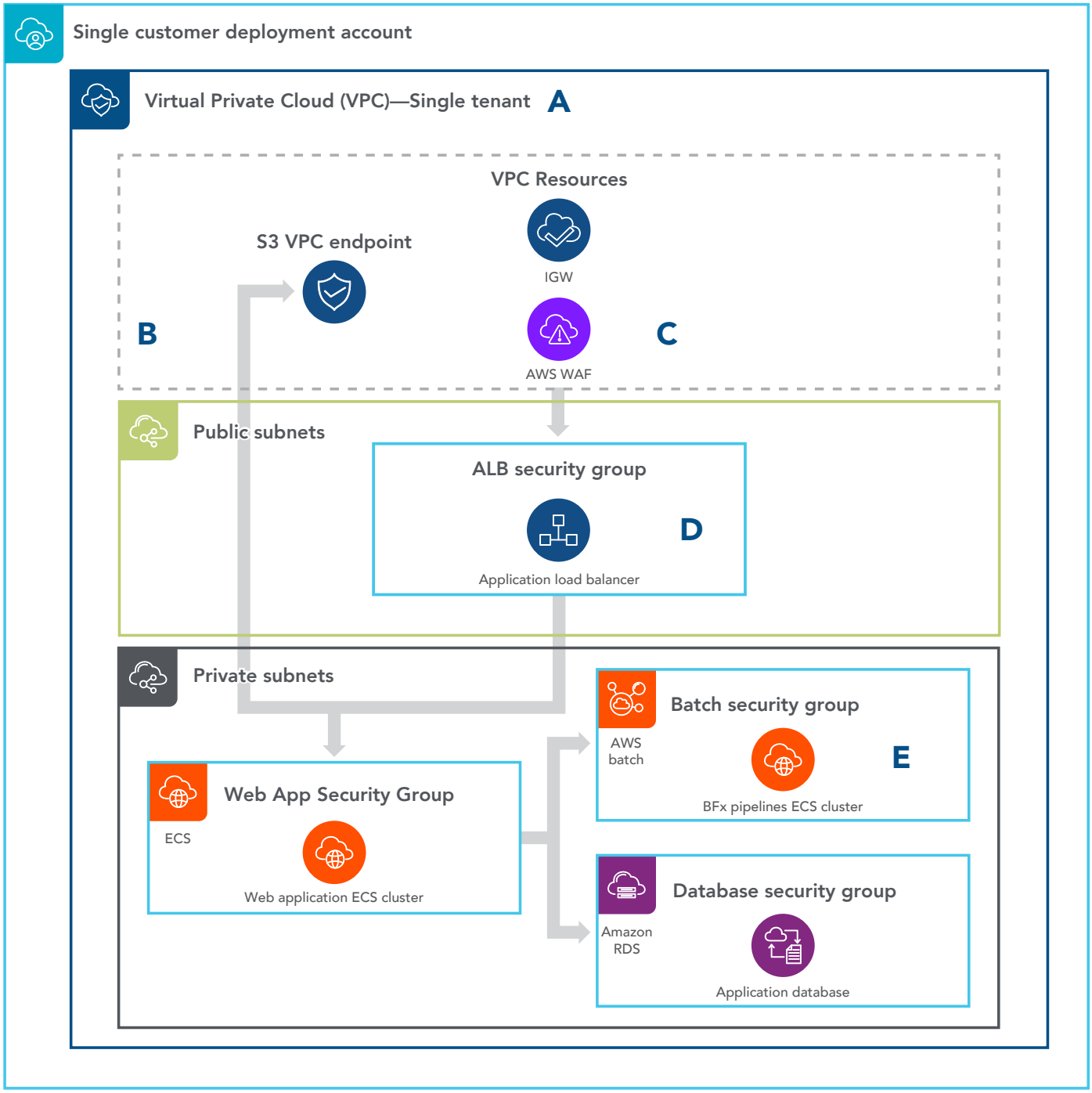


Figure 1. Overview of Archer Analysis Unlimited systems architecture. (**A**) Each lab uses a self-contained, isolated Virtual Private Cloud (VPC), where network traffic is segregated into appropriate subnets for public and private resources. (**B**) VPC endpoints are used to ensure traffic to public AWS endpoints remain within the private VPC network this ensures S3 file access is not performed over the public internet. (**C**) AWS Firewall is used to monitor and protect traffic to the Archer Analysis Unlimited Web application. Access to the web application can be limited to individuals using IP whitelists and geographic regions. (**D**) Application Load Balancer implements the latest secure TLS standards and encryption algorithms, and all requests are redirected to HTTPS/TLS. (**E**) Security Groups and NACLs are used throughout the deployment to further segregate and prevent unauthorized traffic to private resources.

Data isolation: Archer Analysis Unlimited supports high-performance, efficient processing of large genomic sequencing datasets. Data is processed in a private, secure environment on the cloud, offering a high degree of scalability for analysis and storage. The software runs on dedicated single-tenant computer instances, and pipelines are executed within the confines of the selected computer instances, which restricts access to data and resource consumption. Additionally, default controls are set to protect lab data from exposure. Labs can make the decision to share their data selectively and securely in a controlled manner.

Data encryption: The use of public cloud infrastructure can pose several security risks including the potential exposure of data to third parties, as well as theft or misuse of intellectual property. Archer Analysis Unlimited follows data protection best practices to ensure the safety and segregation of data when they are being transferred, processed, or moved to storage. These include:

- ***Encryption in transit:*** Archer Analysis Unlimited enforces the use of Transport Layer Security (TLS) (version 1.2) for communication with other instruments and infrastructure. This standard encrypts sensitive data as they are transferred over the internet in response to user requests. TLS 1.2 is used with the load balancer, which manages information between servers and endpoint devices, across all instances of Archer Analysis Unlimited.
- ***Encryption at rest:*** Data in Archer Analysis Unlimited are encrypted using the AES256 standard. This is the standard setting for Elastic Block Store (EBS) and S3 volume data and is enabled for all Archer Analysis Unlimited accounts by default.

Controlled software updates: Updates are not pushed automatically to a lab's instance. Labs control when to upgrade, and which version of the Archer Analysis is needed.

Accessible, fast deployment: Because Archer Analysis Unlimited relies on AWS infrastructure, it can be deployed and accessed quickly.

Automated data backup maintained by Archer

Given the sensitive nature of genomic data, ensuring the integrity of information is of the utmost importance. Archer Analysis Unlimited provides users with confidence in the accessibility of their data in the event of system failure. Implemented measures include a predetermined rolling automatic backup policy for Archer Analysis Unlimited that helps guard against data loss.

- If a site experiences catastrophic system failure, Archer Analysis Unlimited customers can restore lost data from a predetermined rolling backup.
- Genomic data from Archer Analysis Unlimited is backed up in data centers chosen by the lab.
- Processing and storage of genomic data are guaranteed to be in a lab's region of choice.

In addition, customers can specify a region of choice when creating accounts, ensuring that genomic data is handled in accordance with the regulatory requirements in the appropriate area.

Organizational compliance support

Labs using Archer Analysis Unlimited have access to additional functionalities that can support their regulatory compliance needs.

HIPAA compliance: The Health Insurance Portability and Accountability Act (HIPAA) establishes standards for the security and privacy of protected health information. IDT passed its third-party HIPAA audit, and Archer Analysis Unlimited is compliant with all associated standards.

- Beginning with Version 7.0, IDT standard terms of service for AAU include a built-in business associate agreement (BAAs). Customers are required to comply with HIPAA and are responsible for ensuring that their research activities comply with the requisite administrative, physical, and technical requirements.
- IDT also has a Business Associated Agreements (BAAs) in place with AWS for Archer Analysis Unlimited, which further supports HIPAA compliance for customers.

GDPR compliance: General Data Protection Regulation (GDPR) establishes the standards for data protection in the processing of data of EU residents.

- Beginning with Version 7.0, IDT's standard terms of service include a data processing addendum that would allow for the lawful processing by IDT of customers' personal data. As data controllers, European customers are ultimately responsible for ensuring they have a GDPR-compliant practice in place and should use Archer Analysis Unlimited in a manner that ensures their compliance.

Auditing: Transparency, accountability, and reproducibility are important for successful genomics research. To that end, IDT has implemented various functionalities in Archer Analysis Unlimited to help our customers monitor and report on their activities.

- All actions occurring within the software are tracked and recorded. Administrator user activity is recorded in the same logs as normal user activity.
- Metrics and logs from all AWS instances used by Archer Analysis Unlimited are collected and stored using AWS CloudWatch.
- Critical file integrity is monitored.
- The system is monitored for malicious software.
- The system also captures and logs information from the queue manager software.

Data portability and exit management: Archer Analysis Unlimited allows labs access to their data whenever they need it. Moreover, data can be exported from an instance when needed. IDT complies with contractual requirements regarding destroying lab data when contracts end, while maintaining compliance with applicable regulations.

- Archer Analysis uses standard file formats for output and storage of genomic data. This prevents vendor lock-in and ensures interoperability with sequencing instruments and other platforms in the lab.
- Customers are the owners of their processed data and can transfer them to an Amazon S3 bucket for storage. They can also delete the data after decommissioning a site.

Complete control of logins and permissions

Archer Analysis Unlimited is designed to ensure that only authorized users have access to an instance. Labs can create and configure individual accounts that are password accessible. Password policies (i.e., length and complexity) can be modified depending on a lab's needs. Additional log-in policies can be implemented for added data security and controlled access.

Log-in policies: Archer Analysis Unlimited is a fully contained and secure environment for users to run genomic data analyses via their unique accounts. The system comes with a default set of complexity requirements for selecting strong passwords. Labs can change these criteria by modifying the configuration file. This can be accomplished by sending a service request to customer support.

As an added security measure, a maximum of five incorrect log-in attempts are allowed before the account is deactivated. To regain access, users have to respond to pre-selected security questions. Alternatively, the account can be re-activated by an administrator. The system also records invalid log-in attempts and monitors them for suspicious user activity.

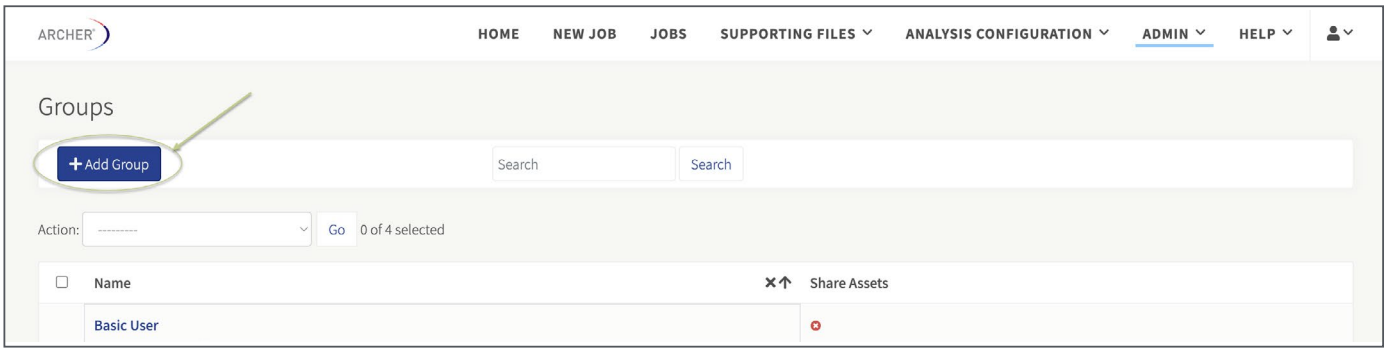
Two-factor authentication: For even greater security, labs can implement two-factor authentication.

- Two-factor identification works with most code generators (e.g. Google Authenticator).
- When turned on for an instance, two-factor identification is globally enforced for all users within the instance.

Managing user access and groups: Archer Analysis Unlimited offers role-based access to the software and data. Individuals within a lab can be registered as either administrators or basic users. There are also options to define custom user groups and assign additional permissions and data access granularity. Software permissions can be assigned per user or per group (Figure 2).

A

B



C

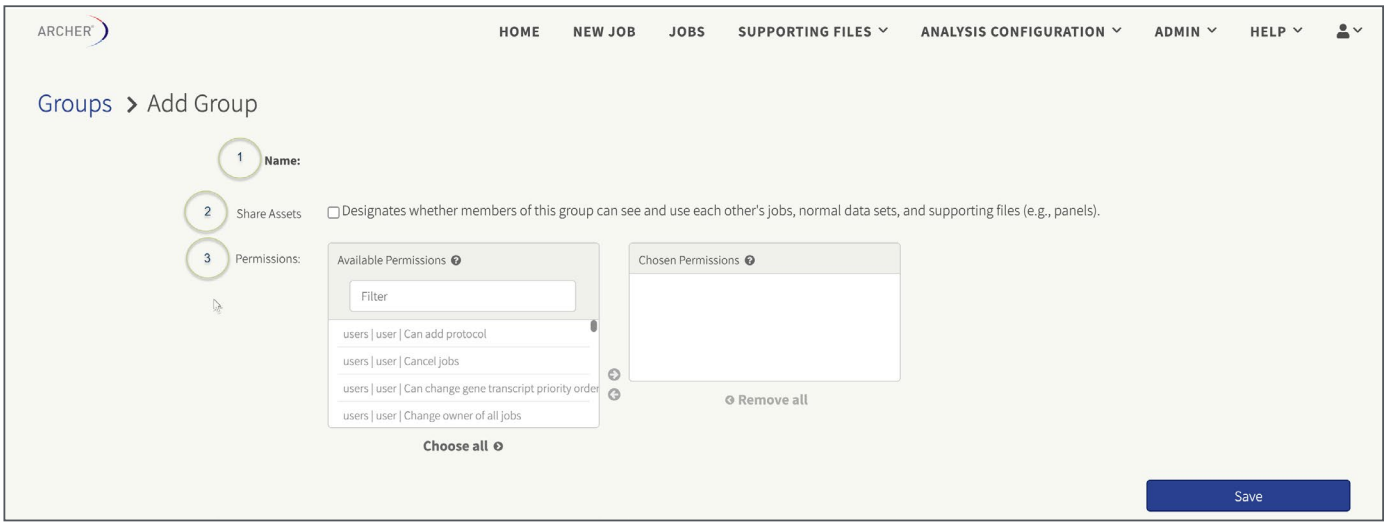


Figure 2. Archer Analysis Unlimited allows for easy control of user access and groups. (**A**) Group management can be quickly accessed in the main menu bar of the Archer Analysis Unlimited platform. (**B**) The 'Groups' page allows account administrators to determine group sharing permissions. (**C**) Sharing permissions and groups can be easily updated on the 'Add Groups' page.

Administrators have several privileges over basic users including:

- Individual user and group management
- Access to experimental features
- Visibility into all jobs in the system with the ability to assign jobs to different users
- Management of genomic data in GTF and VCF formats, as well as other standard file formats
- Only authorized users can access data; access can also be restricted to a limited range of IP addresses and specified countries upon request.

Conclusion

Archer Analysis Unlimited offers secure genomic data analysis with enhanced security options, data encryption, support for organizational compliance needs, and automated disaster protection. These benefits make Archer Analysis Unlimited an ideal option for labs that require an analysis software that is secure, scalable, and easy to implement.

Archer™ Analysis Unlimited—Cloud-based, scalable genomics analysis solution

For more information, go to: IDTDNA.COM/ARCHER

Doc ID: RUO23-2023_002 09/24

INTEGRATED DNA TECHNOLOGIES ARCHER™