Server Deployment
Best Practices
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Before You Deploy</td>
<td>4</td>
</tr>
<tr>
<td>Installation and Configuration</td>
<td>7</td>
</tr>
<tr>
<td>Gallery Setup and Administration</td>
<td>12</td>
</tr>
<tr>
<td>Maintenance/Ongoing Operations</td>
<td>14</td>
</tr>
<tr>
<td>Summary</td>
<td>16</td>
</tr>
<tr>
<td>About Alteryx</td>
<td>17</td>
</tr>
</tbody>
</table>
Alteryx Server provides a scalable server-based analytics solution that lets you publish, share, schedule, and scale analytic applications all in a secure and governed environment. Alteryx Server includes a collaborative, secure central repository system, called the Gallery, to provide analytic outcomes to consumers through customizable apps, reports, and dashboards. Additionally, Server allows organization to scale analytics for higher performance, faster delivery, and resilient automation while sharing the insights with decision makers. Finally, Server takes the load off IT while integrating with existing controls and services to enforce security and management of analytic resources ensuring only users with governed permissions are able to view and collaborate.

The goal of this whitepaper is to help IT leaders and Alteryx administrators understand how to deploy Alteryx Server in a scalable, secure, and efficient manner through best practices that have proven successful for many organizations. These best practices have been broken down into recommendations for the planning phase, the installation, configuration, and initial setup of the environment, and even recommendations for Gallery administration and routine maintenance. Before reading this white paper, it is recommended to review the Getting Started with Server guide.
Before You Deploy

Sizing

It is always recommended to work with your Alteryx representative to perform a sizing exercise to ensure the Alteryx Server environment is properly configured to support the number of users and jobs that will be executing in that environment. A properly sized environment will perform well for end users and jobs will execute quickly and efficiently. Additionally, Alteryx Server is licensed by the number of cores, so a sizing exercise is a key component that should not be overlooked.

Starter Environment Recommendations

Alteryx Server consists of four main components, the Controller, Worker, Gallery, and Database. These four components are discussed further in the Server documentation. In a simple, single machine deployment, all four components can be configured on the same machine. This is the default and simplest to deploy. For added processing capacity, a stand-alone Worker node can support more workflow executions for higher demand environments. Finally, separating out the Gallery to a separate node provides dedicated machines for these core components and allows each to be scaled independently in the future.

Starter Architectures

Minimum
- Default and simplest to deploy

Better
- Added processing capacity
- Easier to scale in future

Best
- Added processing capacity
- Easier to scale in future
- Dedicated nodes for individual components
Choose Appropriate Instance Types

With a properly sized environment decided on in conjunction with your Alteryx representative, it is important to configure servers with the correct number of cores. Alteryx Server is licensed by the number of physical cores. A server with 4 physical cores and 8 virtual cores is perfectly suitable for an Alteryx Server license of 4 cores. This is especially important in cloud environments which often list instance types by the number of “vCPUs” which are typically twice the number of physical cores. To illustrate the importance of this recommendation, consider deploying Alteryx Server on a cloud instance with 4 vCPUs but only 2 physical cores. In that scenario, the licensee would not be realizing the full compute power of the 4 core Alteryx Server license. This topic is covered in more depth in the article How Alteryx defines cores for licensing our products.

Firewall Considerations

When deploying Alteryx Server, it is important to verify the proper network ports are open and available to/from the server to ensure successful operations. Testing a single-node machine may not require any open ports when testing from the same machine. Ports may need to be opened to ensure proper operation when deploying a multi-node or user-facing environment. An up-to-date list of ports required can be found in the Alteryx Server System Requirements help documentation.
Development/Sandbox Environment

Alteryx recommends the use of a development environment for testing workflows, analytic apps, data connections, and product upgrades before implementing in the Production environment. This provides administrators a mechanism for testing new versions of Alteryx Server, and Designer users a sandbox for developing and testing new applications as part of the software development life cycle (SDLC) procedures. An Alteryx Server “Sandbox” environment should mimic the production environment, though it does not have to be identical in size. For further details on how to migrate assets from a Sandbox environment to a Production environment, see the Gallery API Migration documentation.

Designer & Server Version Compatibility

Alteryx Server provides the best user experience for Designer users when the Designer version and Server versions match. Alteryx Server attempts to maintain backwards compatibility with previous versions of Designer, but not forward compatibility. Therefore, it is highly recommended to always keep the Designer version and Server version in sync.
Access to Data Connectors, Macros, and Licensed Data Packages

For Data Connectors, it is recommended to have the same version of the Connector and ODBC Driver across the Alteryx platform. A Designer user should have the same version of the connector and driver as the Alteryx Server. This will ensure that the functionality is consistent across the Alteryx ecosystem.

Alteryx Macros used in production should be stored on a network share and referenced using fully-qualified UNC paths. NTFS/Share permissions can be used to control access to production macros. Storing macros on a network share ensures all users and workflows that reference the macro are using the latest version of the macro. For customers with Data Packages, it is recommended to upgrade at the same time as the Server.

Installation & Configuration

Enable Gallery SSL

A default installation of Alteryx Server uses HTTP (unencrypted) to simplify the installation and configuration. We highly recommend enabling Gallery SSL (HTTPS) to encrypt the communication between the client and the Gallery. This enables the integrity and confidentiality of user sessions, as well as affirms the Server’s identity. The steps to enable Gallery SSL are documented in the online Help.
Set the Gallery Default Run Mode

The Gallery Run Mode setting can provide enhanced security controls over workflow execution of potentially malicious actions. By default, the Gallery Run Mode is set to Unrestricted, which means there are no restrictions, and that any workflow can execute. Configuring this setting to Semi-safe or Safe will further protect the Server environment by preventing workflows from reading or writing data to a location that is not within the workflow staging directory or executing workflows containing restricted tools. This is a global setting that applies to the entire Server environment, but it can be overridden by administrators at the individual workflow level. The Gallery Run Mode setting is described in more detail on the Gallery settings page.

Configure An Appropriate Number of Simultaneous Workflows

Each Alteryx Server Worker can process multiple jobs at a time. The maximum number of jobs that you want to allow to run simultaneously on a given machine is controlled by the Worker setting “Workflows allowed to run Simultaneously”. The recommended starting point for this setting is \( \frac{\text{# of Physical CPU Cores}}{2} \) for the Original Engine. For a Worker configured to run only AMP Engine jobs, the recommended setting would be \( \text{Simultaneous Workflows} = 1 \), as the AMP Engine enables multi-threaded processing. The Worker settings page has more information on this setting.
Configure An Appropriate Memory Limit

The Alteryx Server Worker executes workflows using the Alteryx Engine. The Engine will utilize available memory up to the Engine “Memory Limit” setting. In prior releases this setting was named “Default Sort/Join Memory Usage”. This value should be set appropriately based on the number of Simultaneous Workflows configured. A recommended starting point would be \( \frac{1}{2} \text{Total RAM} / \# \text{of Simultaneous Workflows} \). With this formula, each Workflow allowed to run simultaneously would get its fair share of memory, and the system would not be overly utilized. The value could be increased as needed as long as the memory utilization does not near full capacity as that could cause unintended poor performance. For more information, see the Engine settings page and the AMP memory usage documentation.

Persistence Options

The MongoDB database retains information about all executed jobs, as well as any uploaded files used in executing those jobs. The Persistence Options settings allow an Administrator to specify whether that data should be deleted, and if so, for how many days it should remain before being deleted. These settings may help reduce the amount of hard-drive space used. Alteryx recommends enabling the Persistence Options with a value of 180 days or less, depending on the organization’s requirements. See the Controller settings page for more information.
Job Assignment/Job Tags

The Worker is responsible for the execution of all workflows in Alteryx Server, whether run on-demand or scheduled to run on a recurring basis. With a default deployment of Alteryx Server, all components are enabled on a single server. As workloads increase and Workers are added to the Alteryx Server environment for increased concurrency, it is recommended to configure Job Tag(s) on each Worker.

Job Tags can be configured based upon many factors, such as geographic region, data source locality, department or team, or many other aspects unique to your business. When configuring Job Tags, we recommend the Job Tag be assigned to a minimum of two servers to eliminate any single point of failure.

An Alteryx Server environment should have a minimum of two Workers with “Run unassigned jobs” enabled. This ensures jobs that are not tagged do not sit in the queue indefinitely. Enabling this setting on a minimum of two Workers prevents this becoming a single point of failure.
SERVER DEPLOYMENT BEST PRACTICES

Cancel Long Running Jobs

Alteryx Server can process both on demand jobs, and jobs scheduled that execute on a recurring basis. On rare occasions, it is possible a scheduled job that executes quickly can take much longer due to data source or network latency, which prevents other jobs from running. Alteryx Server includes a setting at the Worker level to “Cancel jobs running longer than” a certain threshold. This setting applies to scheduled jobs only. You can use this setting to force jobs to cancel after a certain amount of time has passed. This will free up system resources that might otherwise be taken up by unintentionally long running jobs.

Use Quality of Service Settings

In an environment with multiple workers, the Quality of Service setting can be used to direct higher priority jobs to specific workers for workload management capabilities. Picture an Alteryx Server deployment that consists of several business-critical jobs and many less important jobs. Configuring a dedicated Worker to run higher priority jobs via the Quality of Service setting prevents the business critical jobs from waiting in the job queue to execute. For environments with a standalone Controller, consider configuring the machine to also act as a Worker with Quality of Service setting QoS=6. This allows the Controller to process workflow validation jobs which are typically very quick and utilize minimal resources. The benefit of this approach is Designer users publishing workflows to the Gallery will have a dedicated Worker for validation purposes without experiencing delays waiting for their workflow validation to complete. More details on the Quality of Service setting can be found in the Worker settings page.

In an environment with multiple workers, the Quality of Service setting can be used to direct higher priority jobs to specific workers for workload management capabilities.
Gallery Setup & Administration

Review Gallery Access Controls

The Gallery Default Role setting defines the default role given to Users and Groups that are not assigned an explicit user or group role. The Default Role is set to Viewer, which allows Users the ability to log in to the Gallery, view, and run Workflows on the homepage and in Districts. To enhance security of the Server environment, consider setting the Default Role to “No Access”, which prevents users and groups that have not been explicitly assigned a role from being able to login to the Gallery.

For additional security, consider disabling the Users Can Register setting. For Built-in or SAML authentication configurations, this setting allows new users to sign up via a form on the Gallery. When the signup form is disabled, new users must be explicitly added by an Administrator. More details on the Default Role and Users Can Register settings can be found on the Gallery configuration page.

Run As and Workflow Credentials

In many cases, workflows will be built that access file shares, and databases that support integrated Windows authentication. It is likely that the Windows account configured to run the Alteryx Server service will not have permissions to read and write to these locations. If there are workflows that require credentials to run and authenticate against a network file share or database, it is a good practice to set a Run As account at the Worker level and use Workflow Credentials. Workflow Credentials override the Run As account settings and can be set for each workflow execution. For more details, see the Run As and Workflow Credentials pages.
Managing How End Users Save and Share to Gallery

One of the primary benefits of Alteryx Server is the ability to publish Workflows from Designer to a Server environment for collaboration and sharing. Users with Artisan access can save Workflows, Apps, or Macros from their local Designer to their company’s private Gallery by adding the private Gallery into their designer “Save As” menu. It is recommended to package only the required workflow assets when uploading to the Alteryx Gallery.

End users should use Collections to share Workflows and Analytic Apps with other individuals or groups of people. Groups can include Active Directory groups or groups that are defined within the Alteryx Server. Note, users will need an appropriate role and permissions to access assets in a Collection.

Scheduling

Another benefit of Alteryx Server is the ability to schedule jobs to run on a recurring basis, so results are readily available for business users when they need them. The ability to schedule is protected by two settings: a universal setting to enable scheduling, and individual permissions for each user. Schedules can be shared with other users and groups in Collections. Anyone you share the schedule with can edit (permissions allowing) and see historical execution results. To remove contention of resources, a Gallery admin might consider leveraging the Server Usage Report or other monitoring tools to view Scheduled hot times and spread out the jobs scheduled to run.
Tagging and Prioritization

Alteryx Server can be scaled up for throughput or can be scaled out by adding workers for concurrency. Whenever possible, it is recommended to scale out as users can take advantage of Job Tags and prioritization (Quality of Service: 0=Low, 1=Medium, 2=High, 3=Critical). The ability to assign Job Tags and Prioritize Jobs is controlled by individual permissions for each user.

Maintenance/Ongoing Operations

Logs

Use the Audit Logs for monitoring (who did what, where, when, how, why). Any updates to AppInfo (Workflows), Collections, Credentials, Subscriptions and Users will trigger the creation of an AuditEvent record. You can return these records via a public Admin API endpoint. For more details, see the Auditlog API endpoint.

Service, Gallery and Engine logs should be used for troubleshooting needs. Service logs contain communication between components and startup and shutdown of processes on the Controller and Worker machines. Gallery logs contain Gallery processes and functions, including errors found when running an Analytic App that uses interface tools, and schedule migration records. Engine logs contain the output of messages created when a workflow runs, as well as the timestamps when tools run in a workflow. More information can be found on the Server logs page.
Backup Practices

Any production environment should have a reliable backup strategy. Virtual Machine snapshots are important as they are a quick failsafe roll back point in case anything goes wrong. This is especially helpful before reconfiguring Alteryx Server or performing a product upgrade. To perform a snapshot on an Alteryx Server machine, make sure the Alteryx Service is stopped.

Backing up MongoDB consistently on a scheduled basis is key to ensure you minimize data loss and downtime.

It is recommended to back up nightly during non-peak hours to minimize potential impact to users and amount of data loss in the case of a crash or another failure. If backing up nightly is not feasible, scheduling weekly backups is also an option. The important thing is to set a regular schedule. The Alteryx Server MongoDB backup procedure is documented on the MongoDB backups page.

Monitoring

All production environments should be consistently monitored to ensure they continue to perform at their peak level. The Diagnostics page of the Gallery provides visibility into user and asset information on the Gallery, as well as machine utilization in the Server environment. Use the Diagnostics page to monitor the number of users, assets, collections, studios, credentials, schedules, and jobs on the Server. You can also view the number of worker machines that are connected and identify constraints in CPU and Memory resources impacting performance. The CPU and Memory information on the Diagnostics page is not persisted and reflects data over the last hour. For a more robust monitoring solution, consider a third-party monitoring tool that can be used to monitor the infrastructure and the Alteryx Service application. For more detailed information, see the Diagnostics page. In addition to the Diagnostics page which reports live statistics from the environment, it is also helpful to periodically run the Alteryx Server Usage Report which can provide a summary of user and job activity on the Server over a period of time. The Server Usage Report can output data to a Tableau dashboard or summary files in Excel and PDF format.
Summary

This white paper shows IT leaders and Alteryx administrators how to plan out and implement an Alteryx Server deployment that will provide end users a robust and secure environment for publishing, sharing, and running Alteryx workflows and analytic applications. These “Best Practices” are not an exhaustive list, nor required in every deployment, but they do present some of the most common steps taken in successful deployments from customer engagements across industries.

Additional resources that can help provide an overview of Alteryx Server, how it is configured, managed, and utilized can be found in the Server documentation and Knowledge Base Community.

If you have questions about any of the recommendations listed in this article, please reach out to your Alteryx representative for further discussion.
About Alteryx

As a global leader in analytic process automation (APA), Alteryx unifies analytics, data science and business process automation in one, end-to-end platform to accelerate digital transformation.

Organizations of all sizes, all over the world, rely on the Alteryx Analytic Process Automation Platform to deliver high-impact business outcomes and the rapid upskilling of their modern workforce.

Visit alteryx.com