

Issue: Q3-2016

# Mediaflux Release News

## Overview:

In the past eight months, Mediaflux server software has been significantly enhanced and improved, with 61 releases of Mediaflux Server this year to date, including 226 enhancements, and 160 bug fixes.

The current server version is v4.4.039 (Oct. 5 2016).

The changes include everything from minor tweaks and refinements to significant capability enhancements across all areas of the system, such as access protocols, database performance, indexing, authentication, and much more.

Although each release includes notes that summarize the details, and provide a running history of the changes, we realize that it may be difficult to keep track of all details by monitoring the release notes alone.

Going forward, therefore, Arcitecta will release a newsletter at least once a quarter with a summary of key updates, enhancements and bug fixes across all areas of the Mediaflux ecosystem. In addition, subsequent issues will include updates on various GUI options, including Mediaflux Desktop, Mediaflux Explorer, etc.

In the sections below you will find a summary of selected highlights including enhancements and critical bug fixes. As always, please refer to full release notes and documentation for more details.

## Table of Contents

1) Installation Alerts: .....	2
2) Access Protocol Enhancements:.....	2
3) Plug-in API Topics .....	4
4) Email Service Execution and remote system monitoring .....	5
5) Metadata Encryption: .....	5
6) Indexing Enhancements: .....	5
7) Query Enhancements:.....	6
8) Server Folders – Enhancement .....	7
9) Assets/Content.....	7
10) Access Control / Authentication / Security: .....	10
11) Replication: .....	11

---

## Mediaflux Server -- Key Enhancements & Bug Fixes:

Note that the references below are selected, abbreviated highlights.

For listing of all enhancements, and bug fixes, or full documentation about the particular feature, please refer to the release notes and other documentation.

### 1) Installation Alerts:

- MF v4.4.039 (Oct 5, 2106): Change in default for symbolic links.  
The default for "symbolic-links" has been changed to "false" (previously defaulted "true") in the service "asset.import". Any scripts relying on that being "true" should be altered.

### 2) Access Protocol Enhancements:

There have been numerous enhancements to Mediaflux access protocols, including SMB, NFS, S3 and Glacier. For full details please refer to the release notes.

Below are some of the highlights:

#### a) POSIX Updates

- MF v4.4.039 (Oct 5, 2106): Point-in-time Mount Option
  - A network mount can now be mounted at a specific point in time with a new mount option "point-in-time", as shown in the following example:
    - `> posix.fs.asset.network.mount :name aa :imap aa :point-in-time "04-Oct-2016 21:16:48"`
    - The file system will then only show namespaces and asset (versions) that existed at that point in time.
  - The point in time for an existing mount point can be changed with a new service "posix.fs.asset.mount.point.in.time.set" as shown in the following example:
    - `> posix.fs.asset.mount.point.in.time.set :name aa :point-in-time "05-Oct-2016 22:03:16"`
  - In order for the client to see changes in the time-point as they are made the client should mount with no attribute caching enabled (mount option "noac").

#### b) SMB Interface Updates

- MF v4.4.010 (Mar 8, 2106): SMB Server
  - MF Server has been enhanced and is now capable of supporting Server Message Block (SMB) protocol clients. A new network type 'SMB' has been added, which can now be used to start a network capable of handling connections from SMB client.
- MF v4.4.024 (July 22, 2016): SMB Authentication
  - SMB interface now allows users to authenticate using Kerberos tickets.
  - SMB interface now allows packet signing in both NTLM and Kerberos authenticated sessions.

---

c) NFS Protocol:

- MF v4.4.030 (Aug 15, 2016):
  - NFS/SMB: The service "posix.fs.asset.network.mount" now allows control over proxy generation and content analysis per mount with new arguments "pgen", "gen-csum" and "analyze". If not specified, the defaults for NFS and/or the system apply.

d) S3 Protocol:

- MF v4.4.024 (Aug 29, 2016): S3 API - Enhancement:
  - We have added a new http processor type to the Mediaflux server that implements Amazon S3 API.
  - The new type is 's3' and a default url '/s3' is automatically created and setup during server start up. In order to use this new feature number of things have to be setup first and for more on that please refer to S3 API section in the administration guide. If you are using S3, please refer to the release notes for more details.
- MF v4.4.033 (Aug 9, 2016): S3 API - **Critical Bug Fix**:
  - Fixed a bug in S3 Data Store which resulted in files being left behind in the cache directory after it has already been uploaded to the S3 server. This would occur in two scenarios:
    - a) File was less than the upload part size setting for the store (default 5MB).
    - b) If the s3.max.cache.size argument was set to 0, in which case all files would be left behind in the cache.
- MF v4.4.029 (Aug 29, 2016): S3 API - **Critical Bug Fix**:
  - Fixes an error in MF 4.4.026, so that the S3 API interface now correctly accepts object metadata headers that are case sensitive.
  - If you are using S3, please refer to the release notes for more details.
- MF v4.4.022 (Jul 19, 2016): S3 Upload configurable -- Enhancement
  - S3 Data Store upload part size is now configurable and can be set to values that are higher than 5MiB.
  - S3 Data Store can be configured to automatically expire incomplete uploads after certain amount of days has passed since the upload was initiated but is still not completed.
  - It is now possible to use existing buckets to create the S3 Data Store.
- MF v4.4.028 (Aug 4, 2016): S3 Asset by Reference -- Enhancement
  - It is now possible to create an asset by reference to data in an (Amazon compliant) S3 source, by specifying the protocol "s3".
  - E.g.: `asset.create :url -by reference s3://arcitecta/myfolder/abrain.jpg`
- MF v4.4.034 (Sep 8, 2016): S3 Supports Clustering -- Enhancement
  - This means that the S3 API clients can connect to cluster nodes instead of the cluster controller, which in turn will increase the throughput of the whole system as the I/O generated by S3 API clients can now be distributed to cluster nodes.
    - After this is setup, capabilities should be added to the node so that it can take over the IO.

- 
- Please refer to the 'Mediaflux Concepts.pdf' document that comes with Mediaflux installation.
  - MF v4.4.037 (Sep 26, 2016) Pause-Resume S3 services – Enhancement
    - S3 API services can now be paused and resumed. All requests that were active at a time when service was paused will run to completion while any new request will be rejected until the service is resumed.

e) **Glacier Data Store Enhancement:**

- In addition to supporting S3 environments as a Mediaflux data store, we've also added support for Amazon Glacier.
  - The Glacier Data Store in Mediaflux enables policy-based data movement to Amazon Glacier storage, in addition to S3 and the other storage types Mediaflux supports.
  - To accommodate the latencies introduced by Glacier, the Glacier data store must have a local disk cache available to it, even if only a small one.
    - This is because we need to cache files locally before uploading them to Glacier in case people use Glacier Data Store as a backing store for a namespace. If we don't have that, any content analysis done on the asset would delay the process of creating an asset completely by 4 hours, which is how long it would take to pull down the asset from Glacier
    - Also, we need local cache when downloading assets from Glacier, because it is a lot more efficient to pull the asset down in one go than issue a Glacier download request for byte ranges which is what would happen without the cache.
  - For more information, please refer to the System Administration Guide.

### 3) Plug-in API Topics

- MF v4.4.021 – (Jul 8, 2016): Data Migration in Multi-tier environments - API Enhancement:
  - Developers: Plugin data stores can now implement data migration methods if that is supported by them. In order to support data migration, plugin data store implementation must indicate that the store is a multi-tiered, and it must override other PluginDataStore methods. Please see release notes for more details.
- MF v4.3.083 – (Mar 17, 2016): Plug-in Authentication Enhancement:
  - Plugins now have access to a range of new classes that can be used to authenticate users using tokens, or verify incoming message signatures under 'arc.mf.plugin.identity.\*'
- MF v4.3.083 (Mar 17, 2016) – Parallel IO Chunking – Enhancement
  - A parallel IO chunking framework has been added to the server.
  - It is able to process IO chunking operations queued for execution via a set of classes (arc.mf.plugin.io.chunking.\*) available through the Plugin API. See the release notes for more details.

---

## 4) Email Service Execution and remote system monitoring

- MF v.4.4.020 (4th, July 2016): Remote Monitoring - Enhancement
  - There are several new services introduced in this release that allow remote system monitoring (via e-mail). These are designed to support environments or situations where the only outgoing network connection is via e-mail, or communication to the system is via unreliable / intermittently connected networks.
  - Related to this are new services to allow one server to execute services in another via e-mail. These services allow the execution of services, including inputs and outputs (limited by mail box size limits).
    - For example, our new upcoming Health Monitoring System will use e-mail to periodically send health information to another server (via e-mail) for recording and monitoring the state of the sending system.

## 5) Metadata Encryption:

- MF v4.4.031 (Aug 24, 2016):
  - It is now possible to encrypt metadata elements and attributes at rest, by setting a new document definition flag "encrypt" to true in the metadata definition. For example:
  - For Example: the following command creates two sub-elements "not" which is not encrypted, and "enc" which is encrypted.:
    - `asset.doc.type.create :type encrypted :definition < :element -name not -type string :element -name enc -type string -encrypt true >`
    - Please see the full release notes for additional details.

## 6) Indexing Enhancements:

### a) Binning Indexes: (Critical)

- MF v4.4.017 (June 2016) - The existing B-Tree indexes have been replaced for all date/time indexes with a new high-performance "binning" index.
  - Date indexes ("ctime", "mtime", "dtime" and any date indexes used by metadata document definitions) will remain as "btree" until they are recreated (using service "asset.index.recreate" or toggling the index off/on with a document definition), in which case they will be recreated as "binning".
  - All new date indexes are created as "binning". The type of the index can be seen when using "asset.index.describe". There is a new "algorithm" qualifier which shows the implementation type for the index.
- MF v4.4.027 (Aug 3, 2016) - Asset indexes using the binning index can be rebuilt for better performance:
  - The fan-out for binning indexes for dates and times has been changed to a factor of 10, which is significantly faster to remove entries (and query ranges) than the previous fan-out of 100. These changes will not apply until the "binning" indexes are rebuilt.

### b) Geospatial Indexing enhancements:

- MF v4.3.073 (Feb 21, 2016) – Spatial coverage query times have been accelerated with modifications to the spatial index. Queries of the following type are now significantly faster.;

- 
- e.g. > asset.query :where "geoshape inside inclusive rectangle [(-43.22,-121.242),(38.424,53.23)]".
  - In order to take advantage of this the geospatial index will need to be rebuilt:
  - > asset.index.recreate :type geoshape
- c) Labels Index enhancement
- MF v4.4.022 (Jul 19, 2016) – There is a new type of asset index "labels" that optimises the retrieval of assets by label query;
    - e.g. "asset.query :where label=PUBLISHED" The label index also optimises access to assets in a view that is constrained by a label.
- d) Index Controls enhancement
- MF v4.4.003 (Apr 6, 2016) – The service "asset.namespace.contained.asset.index.create" now has an option to control the number of levels to descend.
    - This allows indexing to a defined number of levels without forcing cascading for all sub-namespaces. For instance:
    - 'asset.namespace.contained.asset.index.create :namespace a :levels 5'
- e) New Audio Analyzer enhancement
- MF v4.4.035 (Sep 19, 2016) – There is new an inbuilt audio analyzer for supported formats: (WAV, AU, AIFF, AIFF-C, SND). Note: support for MP3 is planned for a future release.

## 7) Query Enhancements:

- a) MF v4.4.017 (June 2016) – Range Queries enhancement
- High performance range queries have now been added for dates.
  - > asset.query :where "mtime in range [01-Jan-2016,31-Jan-2016]"
- b) Mediaflux v4.3.084 (March 2016) – Set queries enhancement
- High performance set queries have now been added for asset names in
  - > asset.query :where "name in { 'Cat', 'Dog', 'Fish' }"
- c) Mediaflux v4.4.024 (March 2016) – New FAQL functions:
- There are new FAQL functions:
    - "datetoage", calculates the age (the time in some units - that is, the delta between now and the specified date). "round", rounds a decimal number to a number with N decimal places.
    - "rounddown", rounds a decimal number down to a number with N decimal places.
- d) Mediaflux v4.4.037 (26 Sep 2016) – New simplified text queries: Enhancement
- AQL: The text queries have been enhanced to simplify the expression of specific text to search.
    - Text types can be combined into a comma separated list. For example, to search for meta text or name text:
      - > asset.query :where mtext,ntext contains 'cat'

- 
- This replaces the former query construct:
    - > asset.query :where "mtext contains 'cat' or ntext contains 'cat'"
  - All of the forms forms of text can be listed.
  - Additionally it is possible to search all text excluding some specified text.
    - For example, to search all text except content text:  
> asset.query :where "text not ctext contains 'cat'"
    - The **not** clause can also have comma separated values:  
> asset.query :where "text not ctext,atext contains 'cat'"
  - These compact search forms result in faster text searching.

## 8) Server Folders – Enhancement

- a) MF v4.4.001 (Mar 29 2016) –
  - With this release a framework has been added for exposing access controlled server accessible file systems for clients to import from and export to.
    - This is commonly used to expose network accessible file systems and is fully supported by all file import and asset export services.
    - Mediaflux Desktop has been enhanced to take advantage of Server Folders, even in environments where the Desktop Integration Agent (DTI) is not available.
    - More information is available in the Reference Books Administration Guide under Server-side Data Access.

## 9) Assets/Content

- a) MF v4.4.039 (Oct 5, 2016) – Asset.Import – Enhancement
  - File system traversal is now up to twice as fast as previously
  - The default for "symbolic-links" is now "false" (was previously "true")
  - There is now an option to import "POSIX" metadata (equivalent to creating the files via NFS). This option imports UID, GID etc. (populating the "arc.posix.attributes" metadata document).
  - If importing file system metadata, then can now import symlinks (as symlinks)
  - Supports importing into a collection (as a collection hierarchy)
- b) MF v4.4.019 (Jun 30, 2016) – Content Copy – Enhancement
  - The process by which a content copy is selected has been enhanced with the ability to provides content copy selection policies.
    - The 'asset.content.copy.policy.\*' services provide the ability to create and manage policies which can be applied to namespaces.
- c) MF v4.4.015 (May 31, 2016) – Citeable Identifiers – Bug Fix
  - Fixed a problem with citeable identifiers whereby changing the CID associated with an asset using the service "asset.cid.set" would not properly remove the asset from the current parent collection properly meaning queries of the form "cid in .." would return incorrect results - asset would appear in both the old and new parent (cid)..

- 
- d) MF v4.4.022 (Jul 19, 2016) – Content Move – Bug Fix
    - Fixed a problem with the service "asset.content.move.to.namespace.store" where the "content.store" index was not properly updated and queries to find asset in specific content stores would fail.
  - e) MF v4.4.011 (Apr 29, 2016) – Content Move – Enhancement
    - The service "asset.content.copy.move" has been enhanced to recover the moved data to the original location if the inner database transaction fails to complete for some reason.
  - f) MF v4.4.001 (Mar 29, 2016) – Asset Export – Enhancement
    - The service 'asset.export.filesystem' has been substantially enhanced to support the generation of XML and CSV Manifest files, exporting Licences and the inclusion of namespace layout templates.
  - g) File Compilation Profiles – Enhancement
    - MF v4.4.014 (May 24, 2016) File Compilation Profiles (FCP) have been enhanced with the ability to import files 'by reference'. The new 'FILE AS REFERENCE' statement allows for ingest of server-accessible file content where the content remains in the source directory.
    - MF v4.4.012 (May 12, 2016) File Compilation Profiles have been enhanced so that any matching 'file name' separation pattern is matched.
      - 'file name' separation patterns now allow for the generation of multiple fragments per matching syntax.
      - File Compilation Profiles documentation has been enhanced to include a section on Document Merge directives and controls with examples explaining how to control the specifics of merging multiple instance of the same fragment derived from different metadata operations.
  - h) Namespace Behaviors – Enhancement
    - MF v4.4.018 (Jun 24, 2016) It is now possible to specify a behaviour for a namespace.
      - Behaviours are convention based strings that are interpreted by applications; For example, an application may use the behaviour to select a metadata presentation layout.
      - Behaviours can apply to only the namespace they are defined for, for all sub-namespaces, or both.
      - Behaviours can be added to the namespace when it is created using new arguments with asset.namespace.create or manipulated on existing namespace using the "asset.namespace.behaviour.\*" services.
      - Behaviours for a namespace, including inherited behaviours, are discovered using the asset.namespace.describe service.
  - i) Asset Metadata Layouts – Enhancement
    - MF v4.4.034 (Sep 08, 2016) There is a new concept for displaying asset metadata, an "asset metadata layout", to control:
      - what metadata is (or is not) displayed, and
      - in what order the metadata is displayed
    - The service "asset.namespace.create" has been enhanced to include a "meta-layout" argument, and there are new services:
-



---

## 10) Access Control / Authentication / Security:

### a) Namespace Access Control Lists (ACL)

- This change only affects installations using Namespace Access Control Lists (ACL) to restrict access and permissions for Namespaces.
  - MF v4.3.083 (Mar 17, 2016) was released with a change to the way that Namespace Access Controls are interpreted.
  - MF v4.4.021 (Jul 4 2016) includes a bug fix for the namespace permission assignment outlined in the release notes and should be the version adopted if Namespace Access Controls are used in your environment
  - MF v4.4.021 (Jul 4 2016) - There are two new services for adding to/removing from existing namespace ACL permissions

### b) LDAP:

- MF v4.4.017 (Jun 17, 2016) – Enhancement: LDAP domains now have a new argument "**role-namespace**" that allows separation of mapped groups to roles within the server.
- MF v4.4.014 (May 24, 2016) – Bug Fix: Fixed NullPointerException in the authentication system when initializing a new database. This problem was introduced in 4.4.013.
- MF v4.3.080 (Mar 6, 2016) – Enhancement: The service "authentication.ldap.domain.create" and "authentication.ldap.domain.modify" have a new argument "keep-alive" which if true, will periodically access the user's LDAP context to keep the context alive. This is useful for LDAP servers that automatically close sessions after inactivity, and thus avoid socket close exceptions.

### c) Identity Maps:

- MF v4.4.014 (May 24, 2016) – Enhancement: Identity map has a new attribute 'auto-create-roles'.
  - If set to true any unknown group ID coming through the NFS interface will automatically create a new role. The new role will be given the name in the format 'gid\_<group\_ID>', e.g. if the GID is 1234 new role will be named 'gid\_1234'.
- MF v4.4.014 (May 24, 2016) – Enhancement: Service 'posix.fs.identity.map.import.from.unix.local' has a new argument "create-user-group-roles"
  - Some systems have a group that matches username, i.e. GID is identical to UID. If this argument is set to true, this service will create a role for each user group and automatically map that group to the newly created role. It will also assign that role to the matching user. GIDs are harvested from /etc/passwd rather than /etc/group. Defaults to false.

### d) Domain Aliases

- MF v4.4.013 (May 16, 2016) – Enhancement: Domains can now have aliases.
  - Domain alias is unique on the whole system and its name cannot clash with any other alias on all domains and it cannot clash with any domain names either, i.e. it is not possible to have an alias 'SomeAlias' if another domain has that alias or there is a domain with name 'SomeAlias'.

---

#### e) Secure Identity Tokens

- MF v4.3.070 (Feb 8, 2016) – Enhancement: Secure identity tokens have been enhanced to support secret key based message authentication.
  - The service "secure.identity.token.create" now has an additional boolean argument, "create-auth-key", which allows the creation of an additional secure random string that can be used to create message authentication codes (MACs). Because these tokens can be used as actor IDs there is a new service, "secure.identity.token.actor.get" to provide the actor ID from a token. For security reasons, if the token does not exist or does not have an additional key, the service returns nothing.

#### f) Administration Services

- MF 4.4.001 – Mar 29, 2016): Admin Services – Enhancement: A collection of administration services has been added to manage the existing 'server.url.restriction' server property.
  - 'server.url.restriction.\*' services provide a convenient way to add, remove and reset the list of urls which are allowed for import and export operations.
- MF 4.4.001 – Mar 29, 2016): Installation Security – Enhancement: As of this release the server property 'server.url.restriction', which restricts the locations where a service with a :url argument can read or write data from, has been provided a default value.
  - Previously the absence of a value in the 'server.url.restriction' server property meant that there were no restrictions governing where a file can be accessed. An empty 'server.url.restriction' server property will now default to: <mediaflux>/volatile/import-export. If your current scripting or workflows require access to a server accessible url configured through the :url argument the path MUST be included in the 'server.url.restriction' server property.

#### g) MF v4.4.014 (May 24, 2016) – Audit System Security enhancement

- The Auditing services "audit.count" and "audit.type.exists" have had their access level changed to more closely reflect the nature and intent of the services.

### 11) Replication:

- MF v4.4.035 (Sep 19, 2016): Asset Replication in Pull Mode – Bug Fix
  - Fixed a problem with asset replication using "pull" mode replication where replicated asset might have zero length data.
  - This did not effect "push" or "ref" mode replication.
  - This same issue could affect "asset.get" and "asset.content.get" when getting content for remote assets.