Defining a Log Collection Policy

For each log added to XpoLog, a Log Collection Policy must be used for defining how XpoLog server should collect the log information into its repository, and how long the logs should be archived. This can be a default Collection Policy or a previously defined Collection Policy. You can also define a new Collection Policy.

The Log Collection Policy criteria can be defined in the following tabs:

- **Members** – for selecting the logs that are collected into the XpoLog repository using this policy
- **Storage** – for defining where to store the log data, the maximum disk space that the policy can use for collecting data, how long to keep files in the storage to be available for searches before deleting them, and the email address of the administrator to notify when the maximum storage space is reached or if there is an error collecting data.
- **Collection Schedule** – for defining the frequency of bringing data into XpoLog
- **Archiving** – for defining the location of the archived data

Storage: XpoLog (Indexed) Data is stored in a Binary, non readable format and cannot be read or decrypted only by XpoLog. In case data is being tampered, XpoLog immediately alerts on the issue.
Archive: XpoLog archived Data is stored in compressed flat files. XpoLog runs a standard checksum (SHA-1/256/MD5) on the archive repository. In case data is being tampered, XpoLog immediately alerts on the issue.

To define a new Log Collection Policy:

1. In the XpoLog Manager main menu, select **Administration > Collection Policies**.
   The Collection Policies page opens.
2. Click the **New Collection Policy** button.
   The Add new collection policy page opens.
3. In **Name**, type the name of the Collection Policy.
4. In **Description**, type a short description of the Collection Policy.
5. Define the Collection Policy members. See *Defining the Collection Policy Members* section below.
6. Define the Collection Policy storage criteria. See *Defining the Collection Policy Storage Criteria* section below.
7. Define the archiving policy and security of the Collection Policy. See *Defining Archiving* below.
8. Click **Save**.
   The Collection Policy is saved and can be used for adding logs and log directories.

Defining the Collection Policy Members

In the Members tab, you can select the logs that are to use the Collection Policy.
In the Collection Policies page, select the Members tab:

1. In the page that appears, select the checkboxes of the logs that are to use this Collection Policy.

Defining the Collection Policy Storage Criteria

In the Storage tab, you can define where to store the collected data and other storage criteria.
In the Collection Policies page, select the Storage tab:

1. **Storage Repository -** browse to the location where to store the collected data, it is recommended to use a fast storage for this location. The default is the XpoLog internal data directory.
2. **Retention Policy -** in the Delete files older than, specify at what age files are to be removed from the repository.
3. **Email Notification -** specify a semicolon separated list of email addresses that an alert will be sent to on policy related failures.
4. **Data Encryption -** XpoLog stores its internal data in a propriety, non readable, model. It is possible to enhance it by activating an additional encryption on the repository (used algorithm AES CBC 128). By selecting an encryption algorithm, as of the next policy's execution data will be stored encrypted. Storing encrypted data causes an overhead when writing/reading data which may result in a performance decrease of data collection/index/search.

Defining the Collection Schedule

You can define the frequency of collecting data from the log: Daily, Weekly, Monthly, or Never. Depending on the frequency selected, parameters appear for specifying the collection schedule.
In the Collection Policies page, select the Collection Schedule tab:

1. **Set Frequency -** select the frequency of bringing data into the system: **Never**, **Daily**, **Weekly**, or **Monthly**.
   Set the parameters that appear, as relevant.
2. **Assigned Instance -** In case XpoLog is running in a clustered environment, with more than one processor node, this option allows to determine which of the processors will be responsible for the collection policy.

3. **Live Mode Collection Frequency -** activating Live Mode in the search console immediately executes collection from all relevant sources in order to fetch matching log records, in near real time, to the console.
   The frequency of collection while Live Mode is active is determined here. By default, the collection will run every 10 seconds as long as Live Mode is active in search.
   Pay attention: the frequency set here determines the collection frequency of logs which are part of the collection policy during Live Mode, and the load that may be seen on the sources while active. It is recommended that users will be guided to activate Live Mode on specific logs/folders/servers, and not on the entire environment, to avoid unnecessary load on multiple sources.

Defining Archiving

Data stored in an archive is for longterm storage of data, and unlike Storage data, is unavailable to the user for searching and viewing. However, archive data can be restored and added to XpoLog as a local log (note that it is a manual process).

Checksum algorithms for ensuring data integrity can be activated; supported types are SHA1 and MD5. The Checksum algorithm checks that there has been no data tampering. Execution of the checksum algorithm results in a signature, which is saved in a file location, so that the current signature can be compared with previous signatures. The Checksum result file location can be customized to any location that XpoLog can access (default is XpoLog internal repository.)

In the Collection Policies page, select the Archiving tab:

1. **Archive Repository:** select Local (**recommended**), Windows Network, Over SSH or AWS S3 Bucket. In **Archive Path**, browse to the location where to archive the collected data, it is not mandatory to use a fast storage for this location.
   Note: Local repository is recommended for large volumes of data, if you're not using Local then you'll be asked to select a connectivity account to the selected location.
2. **Archiving Policy**:
   a. Select the **Enable Archiving** checkbox to enable archiving collected logs; clear the checkbox to disable archiving.
   b. **Archive all data**, by default this option is selected. When initially enabled, the first archive execution will archive ALL data which is currently stored in XpoLog - note that this may take a very long time in case the repository is large. If not selected, then archive will take place from the first time of execution going forward.
   c. In **Delete archive files older than**, select the age that logs are automatically deleted from the archive.
3. **Archiving Security -** In Checksum algorithm, select a checksum algorithm for securing your archived logs: **None**, **MD5**, **SHA1 or SHA-256**.

Defining Data Forwarding

XpoLog instances support forwarding logs data over variety of protocols (Syslog UDP/TCP, HTTP/S, etc.) – the data can be received by either other XpoLog instances or any other supported device.

You can define predefined and generic data forwarders such as: XpoLog Forwarder, Syslog Forwarder, HTTP/S Forwarder, and Generic Forwarder. You can also define a couple of data forwarders that will run in parallel.

In the Collection Policies page, select the Data Forwarding tab:

1. In the page that appears, add new Data Forwarder. If this is the first time that XpoLog is configured to forward data then you will be asked to enter details that XpoLog can use to forward the requested data.
   Create the data forwarder and save it.