VisionLabs LUNA Index Module

Release Notes

v.5.88.0

Contents

LUNA Index Module v.5.88.0

LIM changes

- Support for 65 neural network model for extracting face descriptors has been supported.

  The model is not included in the default distribution package. You need to request it and add it to the container separately.

- Now Redis in the LUNA Index Module delivery is password protected.

  The password is passed when starting the Redis container in the Docker Compose script. Also, in the documentation, the default password has been passed to the Redis container startup command.

  Since some of the LUNA Index Module settings require authentication to Redis, and Redis is now password-protected by default, when upgrading, you must **requiredly add a password** to the following settings:

  - LIM_MANAGER_DB
  - LIM_MATCHER_DB
  - LUNA_INDEXED_LIST_PLUGIN

  Otherwise, the services will not start

  For the initial installation, adding a password is not necessary if the `lim_settings.json` dump file from the distribution is loaded (it is done in the Docker Compose script).

LUNA Index Module v.5.81.2

LIM changes

- A request `remove indexes` has been added to the Video Manager service to remove indexes from the storage by a certain policy.

  The policy is specified in the query parameter and takes the following values:

  - `all` - delete all indexes,
  - `outdated` - delete only irrelevant indexes, i.e. for which there are newer indexes (by default).

LUNA Index Module v.5.76.0

LIM fixed errors

- Fixed excessive memory consumption of Indexed Matcher services that occurred when reloading the index.

LUNA Index Module v.5.75.0

LIM changes

- Improved logging in the LIM Matcher client.

  The logs now contain more information about connections, messages received, and errors occurring.

- Improved index search and index construction.

LUNA Index Module v.5.64.0

LIM changes

- The ability to update an existing index instead of a complete rebuild has been added.

  By default, a complete index rebuild is performed. If necessary, you can update the indexes instead of rebuilding them. To save time and resources, the Index Manager service can only take into account newly added or deleted descriptors. In this case, the Indexer service will not rebuild indexes from scratch. To prioritize upgrade over rebuild, you need to set the "rebuild_rules" > "default" setting to `false`, which means "do not rebuild, but update". By default, a rebuild is performed (value `true`).

  An update involves deleting old and/or adding new descriptors from/to the created index. Due to the indexing algorithm, removing a descriptor from the created index results in degradation of the index. The more descriptors that have been removed, the worse the quality of the index, resulting in lower search accuracy for that index. It is highly recommended that you rebuild your indexes if you remove a large number of individuals from the original dataset.

  To specify the number of times descriptors can be removed from index creation before the index is rebuilt from scratch, you can set the "rebuild_rules" > "max_removal_for_rebuild" parameter to an appropriate value. For example, if this setting is set to 10, it means that up to 10 descriptors can be removed before the index is rebuilt from scratch to ensure it is effective. The default is "0", which means "never rebuild from scratch". The recommended guideline for rebuilding indexes is to remove no more than 10% of the total volume of descriptors.

- The Python version has been updated to 3.12 in LIM service containers.

  Support for older versions of Python has been discontinued.

LIM fixes

- Now when setting the schedule of scheduling procedure execution in Cron-format, the day of the week starts from Sunday.

  Previously, the day of the week started on Monday.

- Indexer service logging responsible for indexing, index update and other similar tasks has been fixed.

  Previously, when changing the logging level in Configurator, the "INFO" logging level was still used for the above tasks.

LUNA Index Module v.5.62.3

LIM changes

- Default neural network model for face descriptor extraction was changed from 59th to 62nd.

  When migrating Configurator service settings of the latest version, the "DEFAULT_FACE_-DESCRIPTOR_VERSION" setting will not be automatically updated** **will not be automatically updated**. If you want to use the 62nd version of the neural network, you must manually change the value of the "DEFAULT_FACE_DESCRIPTOR_VERSION" setting by first re-extracting the existing descriptors with the 62nd model using the "Additional extraction" task.

  Then you need to recreate the existing indexes.

- Response latency from the Indexed Matcher service has been reduced when loading a new index.

LIM fixed errors

- Fixed response code 500 returning when attempting requests to documentation resources with a slash at the end (`/1/docs/spec/` or `/1/docs/dev/`).

- Fixed a lock hang issue when reloading a newer version of the index that could occur when the Indexed Matcher service was unexpectedly stopped/restarted, e.g. due to OOM.

LUNA Index Module v.5.59.0

LIM changes

- The ability to configure the scheduling of the planning routine in Cron format has been added, which checks sets of lists that need to be indexed.

  Using a schedule provides accurate knowledge of the indexing execution time compared to periodic runs every few hours or minutes. This ensures a more predictable and controlled execution of indexing, as well as avoids time shifts that can occur with periodic runs. For example, if the service is unavailable or other issues arise, a periodic run may be skipped, which can disrupt the index update plans. However, it is important to consider that periodic runs also have their advantages. For instance, if the aim is to update the index every hour without the need for complex scheduling configurations.

For example, you can specify the value "0 0 * * 0", which means that indexing will be performed at 00:00 every Sunday.

By default, the "planning_schedule" parameter is not set.

If the "planning_schedule" parameter is set, the "planning_period" parameter will be ignored.

If both parameters are set, the "planning_schedule" parameter takes priority.

- The ability to set the following environment variables when starting LIM containers has been added:

  - `LIM_MANAGER_MASTER_LOCK` — Redis lock name for the Index Manager master instance. This lock is used to ensure that only one instance of the Index Manager can perform planning and searching routines.
  - `LIM_MATCHER_CONSUMER` — Redis consumer group name for matching requests.
  - `LIM_MATCHER_LOCK_PREFIX` — Prefix for the lock name of the Indexed Matcher service in Redis. This helps to avoid potential naming conflicts with other Redis users.

- A new resource `/metrics` has been added to all LIM services, enabling you to collect and save metrics in the Prometheus format as time series data that can be used to track service behavior.

  By default, metrics collection is disabled. Metrics collection is enabled in the "enabled" parameter in the "LUNA_SERVICE_METRICS" section. If metrics collection is disabled, then a request to the `/metrics` resource will return an error with the code "12049" and the message "Forbidden, Resource is disabled".

  There are two types of metrics available:

  - `request_count_total` — Total number of requests.
  - `errors_count_total` — Total number of errors.

  Each of them has at least two labels for sorting:

  - `status_code` (or `error_code` for error metrics)
  - `path` — Path consisting of the request method and the endpoint route.

  If necessary, you can add custom label types by specifying the `label_name=label_value` pair in the "extra_labels" parameter in the "LUNA_SERVICE_METRICS" section.

  See "Export metrics in Prometheus format" for details.

LUNA Index Module v.5.57.0

LIM changes

- API versions for routes of all resources except "/version" and "/healthcheck" have been added to the Indexer service specification.

LIM fixed errors

- Fixed the error where the Python Matcher Proxy service used 100% CPU when using LIM.

- Indexing at startup is now skipped if the time specified in the "planning_period" parameter of the "LIM_MANAGER_INDEXING" section has not passed since the last indexing.

  For example, if "planning_period" = 600 seconds (10 minutes) and 5 minutes after the execution of the planning procedure, LIM was restarted, then indexing will not be performed at startup.

- An offline version of the website with online documentation has been added to the LUNA Index Module distribution package, replacing the previous HTML documents.

LUNA Index Module v.5.53.0

Changes

- Starting from the current version of LIM, it is necessary to use LUNA PLATFORM v.5.53.0 and higher. Older versions of LUNA PLATFORM are not supported by the current version of LIM.

- New Redis Sentinel connection settings have been added.

  In the "sentinel" section of the "LIM_MANAGER_DB" and "LIM_MATCHER_DB" settings, the "user" and "password" parameters have been added.

  To configure "REDIS_URL", the authorization data must be entered as follows: `redis+sentinel://sentinel_username:sentinel_password@sentinel0:26379,sentinel1:26378/indexed_matcher?username=master_username&password=master_password`.

  See the section "Use Redis Sentinel" in the LIM administrator manual for more information about Redis Sentinel.

- A new request "get list of plugins" has been added to the Indexer, Indexed Matcher and Index Manager services, which enables you to get a list of imported plugins and their status.

Fixed errors

- Fixed the hanging of the Indexed Matcher service when the index refresh is enabled (parameter "lim_matcher_refresh" > "enabled").

LUNA Index Module v.5.51.6

Changes

- Support for ECS logging format has been added.

  To use the new format, you need to set the value "ecs" in the "format" setting of the "LIM_service_-LOGGER" section.

When using the "ecs" value, the following fields will be used:

- "http.response.status_code" — contains the HTTP response status code (e.g., 200, 404, 500, etc.).
- "http.response.execution_time" — contains information about the time taken to execute the request and receive the response.
- "http.request.method" — contains the HTTP request method (GET, POST, PUT, etc.).
- "url.path" — contains the path in the request's URL.
- "error.code" — contains the error code if the request results in an error.

- Error descriptions for requests from the "tasks" section of the Index Manager service have been updated.

Fixed errors

- Fixed the error where Redis settings were reset during LIM settings migration.

LUNA Index Module v.5.51.4

Fixed errors

- Fixed the issue with locking for index search/update causing an "internal server error".
- Fixed slow matching work on large indexes when index updating is enabled ("LIM_MATCHER_-REFRESH" setting) and continuous addition of faces.

LUNA Index Module v.5.51.0

Changes

- The size of the Docker image of lim-configs has been reduced.
- A new setting "db_number" has been added to the sections "LIM_MANAGER_DB" and "LIM_-MATCHER_DB", which enables you to specify the number of the Redis database.

  In the setting, you can specify a number from 0 to 15. Each number corresponds to a separate database, which enables you to separate the data.

- SDK updated to version 5.16.0.

  Support for the new 57 and 62 neural network model for face descriptor extraction has been added to LUNA Index Module.

LUNA Index Module v.5.47.4

Changes

- The ability to use Redis Sentinel has been added.

  New sections "LIM_MANAGER_DB.SENTINEL" and "LIM_MATCHER_DB.SENTINEL" have been added to the settings of the Index Manager and Indexed Matcher services, containing the following settings:

  - "master_name" — Name of the Redis master node that will be monitored and managed by Redis Sentinel (by default "index_manager" and "indexed_matcher").
  - "sentinels" — List of addresses of Redis Sentinel listeners (not set by default).

  Also, the use of Redis Sentinel is supported in the "REDIS_URL" setting in the "LUNA_INDEXED_-LIST_PLUGIN" section, which is responsible for connecting the matching plugin with Redis when calculating matching cost of the request. Example of filling in the "REDIS_URL" setting in the "LUNA_INDEXED_LIST_PLUGIN" section: `redis+sentinel://localhost:26379,localhost:26378/indexed_matcher`.

- The default value for the "CONNECTION_POOL_SIZE" settings of the Index Manager and Indexed Matcher services has been increased from 5 to 100.

- The Python version has been updated to 3.11 in LIM service containers.

  Support for older versions of Python has been discontinued.

Fixed errors

- The priority of an error like "Couldn't find idx for " has been changed from "ERROR" to "WARNING".

  Such an error can occur when a user has added a new face to the list and then removed it from the list during the period between the launches of index updates in the memory of the Indexed Matcher service (by default, 1 second).

  For more information about the index update process, see the "Refreshing index in memory" section of the LIM administrator manual.

LUNA Index Module v.5.38.1

Changes

- Sections of the form `<service_name_HTTP_SETTINGS` have been added to the settings of the Indexer, Index Manager and Indexed Matcher services, containing settings responsible for processing HTTP connections.

  The following settings are available:

  - `request_timeout` — the duration of time between the instant when a new open TCP connection is passed to the server. Value (in seconds) is integer number, default 60.

- – `response_timeout` — the duration of time between the instant the server passes the HTTP request to the app, and the instant a HTTP response is sent to the client. Value (in seconds) is integer number, default 600.
- – `request_max_size` — how big a request may be (bytes). Value (in bytes) is integer number, default 1gb.
- – `keep_alive_timeout` — http keep alive timeout. Value (in seconds) is integer number, default 15.

See the following link for details: https://sanic.dev/en/guide/deployment/configuration.html#builtin-values

- New `socket_address` tag has been added to the index build series monitoring points. The tag contains a value in the format `<host>:<port>`.

  This enables you to fix the server address for a particular instance of Indexed Matcher for monitoring.

- New `duration` field has been added to index build series monitoring points with the stages `drop_index` (removing the index from the Indexed Matcher memory) and `load_index` (loading the index into the Indexed Matcher memory). This field indicates the duration of the corresponding stage.

- New resource "/indexes" has been added to the Indexed Matcher service, which enables you to get a set of indexes loaded into the service's memory. For each index, its meta information is output ("label", "index_id", "account_id", descriptor version, etc.).

- New `FORMAT` setting has been added to all LIM services, which enables you to specify the logging format — `default` (standard log output format) and `json` (json log output).

- Now if the index building process is interrupted due to lack of memory, the error 26109 `Build process killed. It's probably due to running out of memory and the OS was triggering the OOM killer` will be displayed in the logs of the Indexer service.

- Indexed Matcher service now includes a check for the existence of a list upon launching, loading a new index into memory, and updating the index in memory. An index without an existing list will be removed from the service's memory.

- Matching requests have been speeded up.

  The speedup was achieved through internal optimizations and changing the response field result format. It used to be a list of the following `{"face_id": ..., "similarity": ...}` objects. The result is now a list of pairs, where the first element of the pair is `face_id` and the second element is `similarity`.

  Note. This change requires a mandatory update of the Python Matcher Proxy service.

- Sequence diagrams of the main processes of the LUNA Index Module have been added to the administrator manual.

LUNA Index Module v.5.36.5

Changes

- Performance of matching on a huge index was improved with some internal optimizations.
- In all LIM services, logging to a file is disabled by default (setting `log_to_file` of each service).

LUNA Index Module v.5.34.0

Changes

- The LUNA Index Module installation manual has been updated.

  Now by default there are commands to install the module in "/var/lib/luna/" directory instead of in "/var/lib/luna/current/" directory.

  The index storage is now created in the "/var/lib/luna/" directory to simplify the upgrade process.

- The LUNA Index Module upgrade manual.
- LUNA Index Module now takes into account changes in the version of face descriptors. It means that:

  - the Indexer service builds an index from descriptors of the version specified in the "DEFAULT_-FACE_DESCRIPTOR_VERSION" setting of the Index Manager service;
  - the Index Manager service automatically rebuilds the index if it does not contain information about versions of descriptor;
  - the Indexed Matcher service loads only those indexes that contain descriptors of the version specified in the "DEFAULT_FACE_DESCRIPTOR_VERSION" setting of the Index Manager service.

  In this regard, the mandatory field "descriptor_version" has been added to the index metastructure ("meta.json"). The "get indexes" and "get most relevent indexes" requests also return the "descriptor_version" parameter.

  Important information for upgrading from previous version

  After starting the Index Manager service, it will automatically start rebuilding all indexes that do not contain information about descriptors, i.e. all indexes created in a previous version of LIM. Rebuilding the index can take a long time, depending on the number of faces on the lists. In order to avoid the lengthy process of rebuilding the index, you can add the "descriptor_version" field with the corresponding version of the descriptors to the "meta.json" files of all previously created indexes before starting the Index Manager service.

  Reminder added to "Important information" section of the new LIM upgrade manual.

- Now, by default, the Indexed Matcher service monitors changes in lists with faces. If new changes are made to the list, the Indexed Matcher service updates the corresponding indexes in its memory by gradually adding a small number of descriptors.

  The use of this functionality is controlled by the "enabled" setting of the "LIM_MATCHER_REFRESH" section of the Configurator service.

  When a cached index is updated, the Indexed Matcher service stops matching on that index, but continues to accept new match requests for that index. By adding a small number of descriptors to the cached index, the matching process is performed with minimal interruption. However, it should be taken into account that if elements are inserted into the list too often (dozens and hundreds of additions), then this can cause a significant degradation in the speed of work, up to an almost complete stop of the matching process.

  If this functionality is used, then it is not necessary and not recommended to perform frequent index rebuilds. Accordingly, it is recommended to increase the planning routine period ("planning_period" setting of the "LIM_MANAGER_INDEXING" section of the Configurator service). However, adding new faces to the cached index is slower than rebuilding the index, so it makes no sense to use this function if a very large number of faces have been added to the list. In this case, it is easier to rebuild the index again.

  See "Cached index refreshing" section in the LIM administrator manual.

- Downloading descriptors from the Faces service has been optimized. The "Accept" header has been changed from `application/x-flutbuf` to `application/x-msgpack`.

- "Warming up" (test matching) the first created index before starting to use it has been added. Subsequent indexes are not "warmed up".

  "Warm-up" is performed after the index is loaded into the memory of the Indexed Matcher service and the service is waiting for a matching request.

- The Python version has been updated to 3.10. Support for other versions has been discontinued.