

Модуль индексирования LUNA

Примечания к выпуску

v.5.72.1

Содержание

LUNA Index Module v.5.64.0	3
LUNA Index Module v.5.62.3	4
LUNA Index Module v.5.59.0	4
LUNA Index Module v.5.57.0	5
LUNA Index Module v.5.53.0	6
LUNA Index Module v.5.51.6	6
LUNA Index Module v.5.51.4	7
LUNA Index Module v.5.51.0	7
LUNA Index Module v.5.47.4	8
LUNA Index Module v.5.38.1	8
LUNA Index Module v.5.36.5	10
LUNA Index Module v.5.34.0	10

LUNA Index Module v.5.64.0

Изменения LIM

- Добавлена возможность обновления существующего индекса вместо полного перестроения.

Обновление существующего индекса необходимо в целях экономии времени и ресурсов сервис Index Manager. Чтобы установить приоритет обновления вместо перестроения, нужно задать для настройки «rebuild_rules» > «default» значение `false`, что означает «не перестраивать, а обновлять». По умолчанию выполняется именно перестроение (значение `true`).

Обновление подразумевает под собой удаление старых и/или добавление новых биометрических шаблонов из/в собранный индекс. Из-за алгоритма индексирования удаление биометрического шаблона из собранного индекса приводит к деградации индекса. Чем больше биометрических шаблонов было удалено, тем хуже качество индекса, что приводит к снижению точности поиска по этому индексу. Настоятельно рекомендуется перестраивать индексы в случае удаления большого количества лиц из исходного набора данных.

Чтобы указать допустимое количество удалений биометрических шаблонов с момента создания индекса, при котором индекс будет перестраиваться с нуля, можно установить соответствующее значение для параметра «rebuild_rules» > «max_removal_for_rebuild». Например, если этот параметр установлен на 10, это означает, что можно удалить до 10 биометрических шаблонов, прежде чем индекс будет перестроен с нуля для обеспечения его эффективности. По умолчанию установлено значение «0», что означает «никогда не перестраивать с нуля». Рекомендуемый ориентир для перестройки индексов — удаление не более 10% от общего объёма биометрических шаблонов.

- В контейнерах сервисов LIM обновлена версия Python до 3.12.

Поддержка более старых версий Python прекращена.

Исправленные ошибки LIM

- Теперь при задании расписания выполнения процедуры планирования в [Cron-формате](#), день недели начинается с воскресенья.

Ранее день недели начинался с понедельника.

- Исправлена работа логирования сервиса Indexer, отвечающая за выполнение задач по индексации, обновлению индекса и другим подобным задачам.

Ранее при смене уровня логирования в Configurator, для вышеописанных задач все равно использовался уровень логирования «INFO».

LUNA Index Module v.5.62.3

Изменения LIM

- Дефолтная модель нейронной сети для извлечения биометрических шаблонов лиц была изменена с 59ой на 62ую.

При выполнении миграции настроек сервиса Configurator последней версии, настройка «DEFAULT_FACE_DESCRIPTOR_VERSION» **не будет автоматически обновлена**. При необходимости использования 62ой версии нейронной сети необходимо вручную изменить значение настройки «DEFAULT_FACE_DESCRIPTOR_VERSION», предварительно выполнив повторное извлечение существующих биометрических шаблонов с помощью 62 модели, используя задачу ["Additional extraction"](#).

Затем нужно пересоздать существующие индексы.

- Время ожидания ответа от сервиса Indexed Matcher уменьшено при загрузке нового индекса.

Исправленные ошибки LIM

- Исправлено возвращение кода ответа 500 при попытке запросов к ресурсам документации с косой чертой в конце (/1/docs/сpec/ или /1/docs/dev/).
- Исправлена проблема с зависанием блокировки при перезагрузке более новой версии индекса, которая могла возникать, когда сервис Indexed Matcher неожиданно останавливался/-перезапускался, например, из-за OOM.

LUNA Index Module v.5.59.0

Изменения LIM

- Добавлена возможность настройки расписания выполнения процедуры планирования в [Cron-формате](#), которая осуществляет проверку наборов списков, которые требуется проиндексировать.

Использование расписания дает точное знание о времени запуска индексирования по сравнению с периодическим запуском каждые несколько часов или минут. Это обеспечивает более предсказуемое и контролируемое выполнение индексации, а также позволяет избежать сдвигов во времени, которые могут возникать при периодическом запуске. Например, если сервис недоступен или возникли другие проблемы, периодический запуск может быть пропущен, что может нарушить планы обновления индекса. Однако, стоит учитывать, что периодический запуск имеет свои преимущества. Например, если требуется обновлять индекс каждый час без необходимости настройки сложного расписания.

Например, можно указать значение «0 0 * * 0», означающее, что индексирование будет выполняться в 00:00 каждое воскресенье.

По умолчанию параметр «`planning_schedule`» не задан.

Если задан параметр «`planning_schedule`», то параметр «`planning_period`» будет игнорироваться.

Если заданы оба параметра, то в приоритете будет параметр «`planning_schedule`».

- Добавлена возможность задать следующие переменные окружения при старте контейнеров LIM:
 - `LIM_MANAGER_MASTER_LOCK` — имя блокировки Redis для экземпляра-мастера сервиса Index Manager. Эта блокировка используется для обеспечения того, чтобы только один экземпляр Index Manager мог выполнять процедуры планирования и поиска.
 - `LIM_MATCHER_CONSUMER` — название группы потребителей Redis для запросов на сравнение.
 - `LIM_MATCHER_LOCK_PREFIX` — префикс для имени блокировки сервиса Indexed Matcher в Redis. Это позволяет избежать возможных конфликтов имен с другими пользователями Redis.
- Во все сервисы LIM добавлен новый ресурс `/metrics`, позволяющий собирать и сохранять метрики в формате Prometheus в виде данных временных рядов, которые можно использовать для отслеживания поведения сервиса.

По умолчанию сбор метрик отключен. Сбор метрик включается в параметре «`enabled`» в группе «`LUNA_SERVICE_METRICS`». Если сбор метрик отключен, то запрос к ресурсу `/metrics` вернет ошибку с кодом «12049» и сообщением «Forbidden, Resource is disabled».

Доступно два типа метрик:

- `request_count_total` — общее количество запросов
- `errors_count_total` — общее количество ошибок

Каждый из них имеет как минимум два лейбла для сортировки:

- `status_code` (или `error_code` для метрик ошибок)
- `path` — путь, состоящий из метода запроса и маршрута конечной точки

При необходимости можно добавить пользовательские типы меток, указав пару `label_name=label_value` в параметре «`extra_labels`» в группе «`LUNA_SERVICE_METRICS`».

См. подробную информацию в разделе [«Экспорт метрик в формате Prometheus»](#).

LUNA Index Module v.5.57.0

Изменения LIM

- В спецификацию сервиса Indexer добавлены версии API к маршрутам всех ресурсов кроме «`/version`» и «`/healthcheck`».

Исправленные ошибки LIM

- Исправлена ошибка, при которой сервис Python Matcher Proxy использовал 100% CPU при использовании LIM.
- Выполнение индексации при запуске теперь пропускается в случае, если с момента последней индексации не прошло время, указанное в параметре «`planning_period`» группы параметров «`LIM_MANAGER_INDEXING`».

Например, если «`planning_period`» = 600 секунд (10 минут) и через 5 минут после выполнения процедуры планирования LIM был перезапущен, то выполнение индексации при запуске не будет произведено.

- Добавлена офлайн версия сайта с онлайн-документацией в комплект поставки LUNA Index Module вместо предыдущих документов в формате HTML.

LUNA Index Module v.5.53.0

Изменения LIM

- Начиная с текущей версии LIM, необходимо использовать LUNA PLATFORM v.5.53.0 и выше. Более старые версии LUNA PLATFORM не поддерживаются текущей версией LIM.
- Добавлены новые настройки соединения с Redis Sentinel.

В секции «`sentinel`» настроек «`LIM_MANAGER_DB`» и «`LIM_MATCHER_DB`» добавлены параметры «`user`» и «`password`».

Для настройки «`REDIS_URL`» данные авторизации необходимо вводить следующим образом: `redis+sentinel://sentinel_username:sentinel_password@sentinel0:26379,sentinel1:26378/indexed_matcher?username=master_username&password=master_password`.

См. раздел «Использование Redis Sentinel» в руководстве администратора LIM для более подробной информации о Redis Sentinel.

- В сервисы Indexer, Indexed Matcher и Index Manager добавлен новый запрос «[get list of plugins](#)», позволяющий получить список импортированных плагинов и их статус.

Исправленные ошибки LIM

- Исправлена зависание сервиса Indexed Matcher при включенном обновлении индекса (параметр «`lim_matcher_refresh`» > «`enabled`»).

LUNA Index Module v.5.51.6

- Добавлена поддержка формата логирования [ECS](#).

Для использования нового формата необходимо задать значение «ecs» в настройке «format» секции «LIM_service_LOGGER».

При использовании значения «ecs» в логах будут возвращаться следующие поля:

- «http.response.status_code» — содержит код состояния ответа HTTP (200, 404, 500 и т.д.);
 - «http.response.execution_time» — содержит информацию о времени, затраченном на выполнение запроса и получение ответа;
 - «http.request.method» — содержит метод HTTP-запроса (GET, POST, PUT и т.д.);
 - «url.path» — содержит путь в URL-адресе запроса;
 - «error.code» — содержит код ошибки, если запрос завершается с ошибкой.
- Обновлены описания ошибок для запросов [из группы «tasks»](#) сервиса Index Manager.

Исправленные ошибки

- Исправлена ошибка, при которой во время миграции настроек LIM сбрасывались настройки Redis.

LUNA Index Module v.5.51.4

Исправленные ошибки

- Исправлена проблема с установкой блокировки для поиска/обновления индекса, вызывавшая ошибку «internal server error».
- Исправлена медленная работа сравнения по большим индексам при включенном обновлении индекса (настройка «LIM_MATCHER_REFRESH») и непрерывном добавлении лиц.

LUNA Index Module v.5.51.0

- Уменьшен размер Docker-образа lim-configs.
- В группы настроек «LIM_MANAGER_DB» и «LIM_MATCHER_DB» добавлена новая настройка «db_number», позволяющая указать номер БД Redis.

В настройке можно указать номер от 0 до 15. Каждый номер соответствует отдельной базе данных, что позволяет разделить данные.

- SDK обновлен до версии 5.16.0.

В LUNA Index Module добавлена поддержка 57-ой и 62-ой моделей нейронной сети для извлечения биометрических шаблонов лиц.

LUNA Index Module v.5.47.4

- Добавлена возможность использования Redis Sentinel.

В настройки сервисов Index Manager и Indexed Matcher добавлены новые группы параметров «LIM_MANAGER_DB.SENTINEL» и «LIM_MATCHER_DB.SENTINEL», содержащие следующие настройки:

- «master_name» — имя мастер-ноды Redis, которая будет отслеживаться и управляться Redis Sentinel (по умолчанию «index_manager» и «indexed_matcher»).
- «sentinels» — список адресов слушателей Redis Sentinel (по умолчанию не задан).

Также использование Redis Sentinel поддерживается в настройке «REDIS_URL» в группе настроек «LUNA_INDEXED_LIST_PLUGIN», которая отвечает за соединение плагина сравнения с Redis при вычитывании сложности запроса. Пример заполнения настройки «REDIS_URL»: `redis+sentinel://localhost:26379,localhost:26378/indexed_matcher`.

- Увеличено значение по умолчанию для настроек «CONNECTION_POOL_SIZE» сервисов Index Manager и Indexed Matcher с 5 до 100.
- В контейнерах сервисов LIM обновлена версия Python до 3.11.

Поддержка более старых версий Python прекращена.

Исправленные ошибки

- Изменен приоритет ошибки вида «Couldn't find idx for » с «ERROR» на «WARNING».

Такая ошибка может возникнуть в случае, когда пользователь добавил новое лицо в список, а затем удалил его из списка в период между запусками обновлений индекса в памяти сервиса Indexed Matcher (по умолчанию 1 секунда).

См. подробную информацию о процессе обновления индекса в разделе «Обновление индекса в памяти» руководства администратора LIM.

LUNA Index Module v.5.38.1

- В настройки сервисов Indexer, Index Manager и Indexed Matcher добавлены группы настроек вида <service_name_HTTP_SETTINGS, содержащие настройки, отвечающие за обработку HTTP-подключений.

Доступны следующие настройки:

- request_timeout — продолжительность времени между моментом, когда новое открытое TCP-соединение передается на сервер. Значение (в секундах) — целое число, по умолчанию 60.

- `response_timeout` — продолжительность времени между моментом, когда сервер передает HTTP-запрос приложению, и моментом, когда HTTP-ответ отправляется клиенту. Значение (в секундах) — целое число, по умолчанию 600.
- `request_max_size` — насколько большим может быть запрос. Значение (в байтах) — целое число, по умолчанию 1 Гб для всех сервисов.
- `keep_alive_timeout` — тайм-аут поддержания активности HTTP. Значение (в секундах) — целое число, по умолчанию 15.

См. подробную информацию по следующей ссылке: <https://sanic.dev/en/guide/deployment/configuration.html#builtin-values>

- Добавлен тег `socket_address` в точки мониторинга серии построения индекса. Тег содержит значение в формате `<host>:<port>`.

Это позволяет фиксировать адрес сервера для конкретного экземпляра Indexed Matcher для мониторинга.

- Добавлено поле `duration` в точки мониторинга серии построения индекса с этапами `drop_index` (удаление индекса из памяти Indexed Matcher) и `load_index` (загрузка индекса в память Indexed Matcher). Данное поле означает длительность соответствующего этапа.
- В сервис Indexed Matcher добавлен новый ресурс `«/indexes»`, позволяющий получить набор индексов, загруженных в память сервиса. Для каждого индекса выводится его метаданные («label», «index_id», «account_id», версия БШ и т.д).
- Во все сервисы LIM добавлена настройка `FORMAT`, позволяющая указать формат логирования — `default` (стандартный формат вывода логов) и `json` (вывод логов в формате json).
- Теперь если процесс построения индекса прерывается из-за нехватки памяти, то в логах сервиса Indexer будет выдана ошибка `26109 Build process killed. It's probably due to running out of memory and the OS was triggering the OOM killer`.
- В сервис Indexed Matcher добавлена проверка существования списка при запуске, загрузке нового индекса в память и обновлении индекса в памяти. Индекс без существующего списка будет удален из памяти сервиса.
- Ускорены запросы на сравнение.

Ускорение было достигнуто за счет внутренних оптимизаций и изменения формата результата поля ответа. Раньше это был список следующих объектов `{"face_id": ..., "similarity": ...}`. Теперь результатом является список пар, где первый элемент пары — это `face_id`, второй элемент — `similarity`.

Примечание. Это изменение требует обязательного обновления сервиса Python Matcher Проху.

- В руководство администратора добавлены диаграммы последовательности основных процессов LUNA Index Module.

LUNA Index Module v.5.36.5

- Производительность сравнения по большому индексу была улучшена за счет некоторых внутренних оптимизаций.
- Во всех сервисах LIM отключена запись логов в файл по умолчанию (настройка `log_to_file` каждого сервиса).

LUNA Index Module v.5.34.0

- Обновлено руководство по установке LUNA Index Module.

Теперь по умолчанию приведены команды для установки модуля в директории `«/var/lib/luna/»`, а не в директории `«/var/lib/luna/current/»`.

Хранилище индексов теперь создается в директории `«/var/lib/luna/»`, для упрощения процесса обновления.

- Добавлено [руководство по обновлению LUNA Index Module](#).
- LUNA Index Module теперь учитывает изменение версии биометрических шаблонов лиц. Это означает, что:
 - сервис Indexer выполняет построение индекса из биометрических шаблонов версии, указанной в настройке `«DEFAULT_FACE_DESCRIPTOR_VERSION»` сервиса Index Manager;
 - сервис Index Manager автоматически перестраивает индекс, если в нем не содержится информация о версиях биометрических шаблонов;
 - сервис Indexed Matcher загружает только те индексы, которые содержат биометрические шаблоны версии, указанной в настройке `«DEFAULT_FACE_DESCRIPTOR_VERSION»` сервиса Index Manager.

В связи с этим, в метаструктуру индекса (`«meta.json»`) добавлено обязательное поле `«descriptor_version»`. В запросах о получении информации об индексах [«get indexes»](#) и [«get most relevant indexes»](#) также возвращается параметр `«descriptor_version»`.

Важная информация по обновлению с предыдущей версий

После запуска сервиса Index Manager, он автоматически начнет перестраивание всех индексов, в которых отсутствует информация о биометрических шаблонах, т.е. всех индексов, созданных в предыдущей версии LIM. Перестройка индекса может занять длительное время, зависящее от количества лиц в списках. Для того, чтобы не выполнять длительный процесс перестроения индекса, можно добавить поле `«descriptor_version»` с соответствующей версией биометрического шаблона в файлы `«meta.json»` всех созданных ранее индексов перед запуском сервиса Index Manager.

Напоминание добавлено в раздел [«Important information»](#) нового руководства по обновлению LIM.

- Теперь по умолчанию сервис Indexed Matcher следит за изменениями в списках с лицами. В случае внесения новых изменений в список, сервис Indexed Matcher обновляет соответствующие индексы у себя в памяти, путем постепенного добавления небольшого количества биометрических шаблонов.

Использование данного функционала регулируется настройкой «enabled» секции «LIM_MATCHER_REFRESH» сервиса Configurator.

При обновлении кэшированного индекса сервис Indexed Matcher останавливает сравнение по этому индексу, но продолжает принимать новые запросы на сравнение для этого индекса. Благодаря тому, что в кэшированный индекс добавляется небольшое количество биометрических шаблонов, процесс сравнения выполняется с минимальным прерыванием. Однако следует учитывать то, что если в список слишком часто вставляются элементы (десятки и сотни добавлений), то это может вызвать существенную деградацию в скорости работы, вплоть до почти полной остановки процесса сравнения.

Если используется данный функционал, то нет необходимости и не рекомендуется выполнять частые перестроения индекса. Соответственно, рекомендуется увеличить период процедуры планирования (настройка «planning_period» секции «LIM_MANAGER_INDEXING» сервиса Configurator). Однако добавление новых лиц в кэшированный индекс происходит медленнее, чем перестроение индекса, поэтому нет смысла использовать данную функцию если в список было добавлено очень большое количество лиц. В таком случае проще заново перестроить индекс.

См. подробную информацию в разделе [«Cached index refreshing»](#) в руководстве администратора LIM.

- Оптимизирована загрузка биометрических шаблонов из сервиса Faces. Заголовок «Accept» изменен с application/x-flutbuf на application/x-msgpack.
- Добавлен «прогрев» (тестовый матчинг) первого созданного индекса перед началом его использования. Для последующих индексов «прогрев» не выполняется.

«Прогрев» выполняется после того как индекс загружен в память сервиса Indexed Matcher и сервис ожидает запроса на сравнение.

- Версия Python обновлена до 3.10. Поддержка остальных версий прекращена.