

Программный комплекс «Единая Медиаплатформа»

Функциональные характеристики

Москва, 2024

Аннотация

Данный документ содержит сведения о функциональных характеристиках экземпляра программного комплекса «Единая Медиаплатформа» (далее – Медиаплатформа).

В документе приведены следующие сведения:

- назначение и область применения Медиаплатформы;
- функции модулей Медиаплатформы.

Содержание

1 Общие сведения	7
2 Назначение программного комплекса	7
3 Описание программного комплекса.....	7
3.1 Состав программного комплекса.....	7
3.1.1 Модуль «Система загрузки аудиовизуального контента в сети Интернет»	8
3.1.1.1 Назначение модуля.....	8
3.1.1.2 Функциональные возможности.....	8
3.1.1.3 Способы и средства связи для информационного обмена между компонентами программного комплекса.....	9
3.1.2 Модуль «Система транскодирования аудиовизуального контента».....	9
3.1.2.1 Назначение модуля.....	9
3.1.2.2 Функциональные возможности.....	9
3.1.2.3 Способы и средства связи для информационного обмена между компонентами программного комплекса.....	10
3.1.3 Модуль «Внутренняя система управления задачами при обработке аудио визуального контента».....	11
3.1.3.1 Назначение модуля.....	11
3.1.3.2 Функциональные возможности.....	11
3.1.3.3 Способы и средства связи для информационного обмена между компонентами программного комплекса.....	11
3.1.4 Модуль «Микросервис диагностики файлов аудиовизуального контента»	12
3.1.4.1 Назначение модуля.....	12
3.1.4.2 Функциональные возможности.....	12
3.1.4.3 Способы и средства связи для информационного обмена между компонентами программного комплекса.....	13
3.1.5 «Модуль трансляции аудиовизуального контента».....	13
3.1.5.1 Назначение модуля.....	13
3.1.5.2 Функциональные возможности.....	13
3.1.5.3 Способы и средства связи для информационного обмена между компонентами программного комплекса.....	13
3.2 Режимы функционирования	14
3.2.1 Предполагаемая нагрузка.....	14

Термины и определения

Термин	Определение
Apache Kafka (Kafka)	Распределённый программный брокер сообщений с открытым исходным кодом, разрабатываемый в рамках фонда Apache на языках Java и Scala
Backend	Набор средств, с помощью которых происходит реализация логики программного комплекса
Frontend	Внешний по отношению к Медиаплатформе сервис, взаимодействующий с Медиаплатформой для получения, обработки, хранения и выдачи по запросу медиаконтента или других данных Медиаплатформы
Helm	Пакетный менеджер, облегчающий задачи настройки, установки и обновления приложений в Kubernetes
HLS	Коммуникационный протокол для потоковой передачи медиа на основе HTTP
Kubernetes	Открытое программное обеспечение для оркестровки контейнеризированных приложений: автоматизации их развёртывания, масштабирования и координации в условиях кластера
Mbps	Мегабит в секунду, показатель пропускной способности обработки
MIME	Спецификация для передачи по сети файлов различного типа: изображений, музыки, текстов, видео, архивов и др. Указание MIME-типа используется в HTML обычно при передаче данных форм и вставки на страницу различных объектов
NTP (Network Time Protocol)	Сетевой протокол для синхронизации внутренних часов компьютера с использованием сетей с переменной латентностью
RPM	Requests per minute — количество запросов на обработку видео в минуту. Например, «60 RPM» означает что в Медиаплатформу производится 60 запросов на транскодирование в минуту
RPS	Requests per second — количество запросов на обработку видео в секунду. Например, «1 RPM» означает что в Медиаплатформу производится 1 запрос на транскодирование в секунду, то есть, 1 RPM = 60 RPS
S3	Объектное хранилище. Его уникальность заключается в хранении огромного объема данных в исходном формате без иерархии и разбивки на отдельные каталоги. У хранилища S3 нет ограничений по масштабированию

Термин	Определение
SaaS (программное обеспечение как услуга)	Модель обслуживания, при которой подписчикам предоставляется готовое прикладное программное обеспечение, полностью обслуживаемое провайдером. Поставщик в этой модели самостоятельно управляет приложением, предоставляя заказчикам доступ к функциям с клиентских устройств, как правило через мобильное приложение или веб-браузер
SAR (sample aspect ratio)	Отношение количества пикселей по горизонтали к количеству пикселей по вертикали (эти цифры указаны в разрешении видеофайла)
Thumbnail	Эскиз видео, уменьшенная до небольших размеров копия изображения для компактного представления пользователю
UNIX-time	Количество секунд, прошедших с полуночи (00:00:00 UTC) 1 января 1970 года
URI (Uniform Resource Identifier)	Унифицированный указатель ресурса — система унифицированных адресов электронных ресурсов, или единообразный определитель местонахождения ресурса (файла)
URL (Uniform Resource Locator)	Адрес, который выдан уникальному ресурсу в интернете
БД	База данных
Base64	Стандарт кодирования двоичных данных при помощи только 64 символов ASCII. Алфавит кодирования содержит латинские символы A-Z, a-z, цифры 0-9
Медиаплатформа	<p>Интегрированная система обработки медиаконтента. Система позволяет:</p> <ul style="list-style-type: none"> · проанализировать загруженный медиаконтент, оценить его качество; · обработать медиафайл, сохранить его в различном качестве и нескольких форматах; · создать серию эскизов для ознакомительного просмотра; · извлечь аудиодорожку из загруженного медиафайла
Контейнер	Формат файла, определяющей распределение аудио, видео, а в некоторых случаях и текстовой информации внутри него
Медиафайл	Компьютерный файл, содержащий аудио- и видеoinформацию
Под, (англ. pod)	Абстрактный объект Kubernetes, представляющий собой «обертку» для одного или группы контейнеров. Контейнеры в поде запускаются и работают вместе, имеют общие сетевые ресурсы и хранилище. Модули Медиаплатформы запускаются как поды Kubernetes

Термин	Определение
Транскодирование	Обработка медиафайлов, прямое цифровое преобразование одной кодировки в другую
Трикмод	Функция показа соответствующего Thumbnail-а во время движения мышки на полосе прокрутки видеофайла
Холст	Растровый холст, задающий размер и разрешение для цифровых изображений
Чанк	Часть медиафайла для удобства его загрузки и обработки. Обычно файл делится на чанки одинакового размера, кроме последнего чанка
Чарт	В Helm, совокупность yaml-файлов с шаблонами конфигураций, а также нескольких служебных файлов, используемая для описания ресурсов, необходимых для запуска приложения, инструмента или службы внутри кластера Kubernetes

1 Общие сведения

Наименование программного комплекса: «Единая Медиаплатформа».

Условное обозначение программного комплекса: Медиаплатформа.

2 Назначение программного комплекса

Программный комплекс «Единая Медиаплатформа» — интегрированная система обработки медиаконтента.

Медиаплатформа позволяет:

- проанализировать загруженный медиаконтент и оценить его качество;
- обработать медиаконтент, сохранить его в различном качестве и нескольких форматах;
- транслировать медиаконтент в различных качествах;
- создать серию эскизов для ознакомительного просмотра;
- извлечь аудиодорожку из загруженного медиафайла.

3 Описание программного комплекса

Медиаплатформа осуществляет следующие процессы:

- загрузку медиаконтента пользователя;
- анализ загруженного медиаконтента, оценку его качества;
- обработку медиафайла, сохранение его в различном качестве и нескольких форматах;
- создание серии эскизов для ознакомительного просмотра;
- извлечение аудиодорожки из загруженного медиафайла.

Медиаплатформа имеет программный интерфейс – API, позволяющий подключить к ней существующие и развивающиеся сервисы работы с медиаконтентом.

Компоненты системы можно развернуть как на внутренней технической инфраструктуре, так и получить по модели SaaS (программное обеспечение как услуга).

Язык исполнения: C++, GoLang.

Среда исполнения: Kubernetes.

3.1 Состав программного комплекса

Программный комплекс включает в себя:

- Backend, состоящий из компонентов:
 - «Система загрузки аудиовизуального контента в сети Интернет» (далее Uploader);
 - «Система транскодирования аудиовизуального контента» (далее Transcoder);
 - «Внутренняя система управления задачами при обработке аудио визуального контента» (далее TaskManager);
 - «Микросервис диагностики файлов аудиовизуального контента» (далее Prober);
 - «Модуль трансляции аудиовизуального контента» (далее – Модуль трансляции);
- серверную инфраструктуру;
- хранилище S3 и FileNep;
- внешние сервера HTTP и FTP;
- базы данных;

- интерфейсы передачи данных на Frontend.

Общая схема работы программного комплекса представлена на рисунке 1.

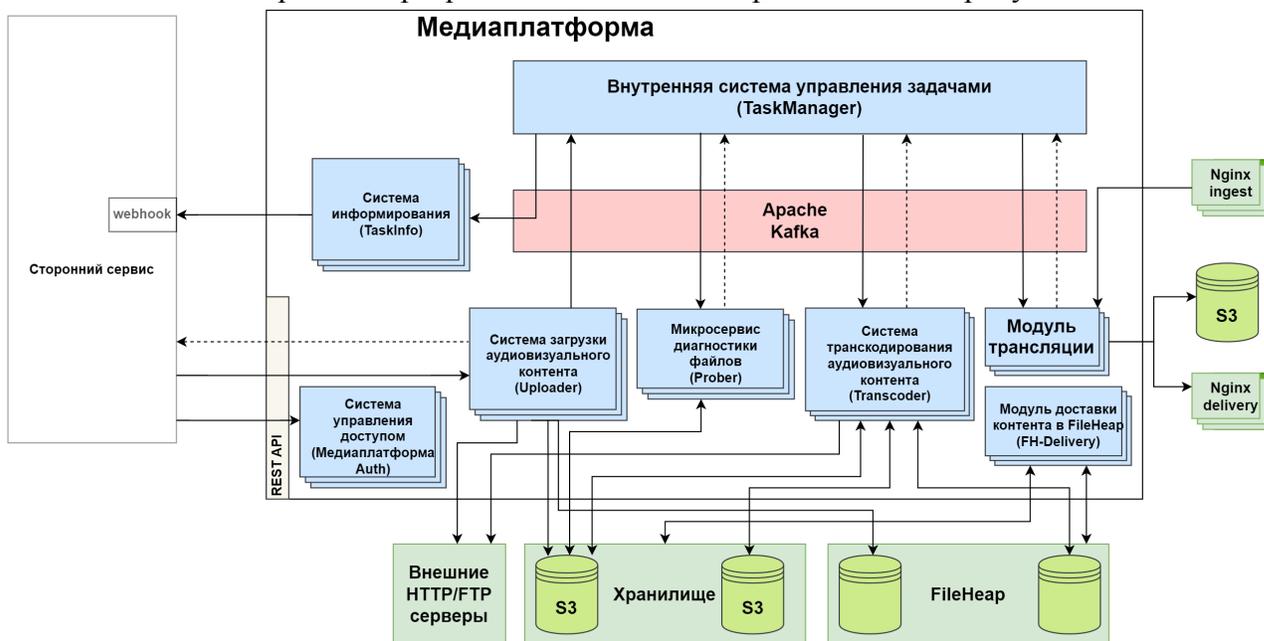


Рисунок 1. Общая схема работы Медиаплатформы.

Для взаимодействия компонентов Медиаплатформы используется брокер сообщений Apache Kafka.

Компонент-отправитель помещает сообщение в нужный топик, после чего получатель выполняет задачу в соответствии с параметрами, указанными в этом сообщении. После выполнения задачи получатель помечает сообщение как прочитанное и переходит к следующему сообщению в топике.

Компоненты Медиаплатформы являются Stateless микросервисами. Все обрабатываемые данные сохраняются в указанных на схеме хранилищах и базах данных.

Обмен данными происходит по протоколу HTTP(S) или через web-socket с использованием внутренней сети.

3.1.1 Модуль «Система загрузки аудиовизуального контента в сети Интернет»

3.1.1.1 Назначение модуля

Модуль предназначен для скачивания и приёма загружаемых пользовательских медиафайлов и сохранения их во внутреннем хранилище.

3.1.1.2 Функциональные возможности

Модуль обеспечивает:

- подключение различных хранилищ данных;
- загрузку файлов пользователями из сторонней системы, использующей пользовательский интерфейс, во внутреннее хранилище S3, FileHear или локальное хранилище целиком или чанками;
- скачивание файлов пользователями из сторонней системы, использующей протокол FTP или HTTPS, во внутреннее хранилище S3, FileHear или локальное хранилище целиком или чанками;

- возобновление прерванной загрузки файла в случае обрыва соединения, ошибки;
- многопоточную загрузку файла;
- безопасное получение и сохранение учётных данных для доступа к сторонним системам, использующим протоколы FTP или HTTPS, в зашифрованном виде;
- безопасное получение и сохранение учётных данных различных внешних систем (организаций - пользователей Медиаплатформы) для доступа к внешним хранилищам FileNear, в зашифрованном виде;
- передачу информации о файле в другие компоненты Медиаплатформы для последующей обработки файла Медиаплатформой;
- масштабирование Модуля для управления нагрузкой на оборудование;
- связь с другими компонентами Медиаплатформы через брокера сообщений или другие интерфейсы;
- мониторинг внутреннего состояния Модуля;
- сбор аналитической информации;
- обработку ошибок;
- логирование;
- сбор и визуальное отображение метрик Модуля;
- отправку событий в шину данных.

3.1.1.3 Способы и средства связи для информационного обмена между компонентами программного комплекса

Модуль взаимодействует со следующими компонентами Медиаплатформы:

- внешние сервера HTTP и FTP - после обращения из сторонней системы к пользовательскому интерфейсу с заданием, исходный медиафайл скачивается Модулем с внешнего сервера;
- хранилище S3 или FileNear – после загрузки или скачивания медиафайла пользователем Модуль помещает полученный медиафайл в хранилище (раздельно в FileNear для каждого пользователя);
- «Внутренняя система управления задачами при обработке аудио визуального контента» – Модуль посредством брокера сообщений Kafka уведомляет этот компонент о состоянии загрузки;
- брокер сообщений Kafka.

3.1.2 Модуль «Система транскодирования аудиовизуального контента»

3.1.2.1 Назначение модуля

Модуль выполняет обработку (транскодирование) медиафайла после анализа, проведённого компонентом «Микросервис диагностики файлов аудиовизуального контента» на основе профилей обработки, настроенных в компоненте «Внутренняя система управления задачами при обработке аудио визуального контента».

3.1.2.2 Функциональные возможности

Модуль обеспечивает:

- загрузку файла из хранилища;
- преобразование файла из одного цифрового формата в другой (транскодирование) в соответствии заданными профилями транскодирования файлов по параметрам: тип медиаконтейнера, размер, битрейт, кодек и другие:
 - декодирование и кодирование файла в различные форматы;
 - демультимплексирование и мультимплексирование файла;

- исправление интерлейса при конвертации;
- конвертацию HDR видео;
- манипуляции с содержимым:
 - наложение специального изображения на видеокадр;
 - обрезка чёрных полос;
 - автоматическое выравнивание громкости;
 - нарезка/склейка видео/фрагментов видео;
 - «приклеивание» заставки к исходному контенту;
- поддержку файлов с несколькими аудиодорожками;
- поддержку субтитров в форматах teletext, DVB subtitles;
- формирование нескольких экземпляров файла в различных качествах;
- подготовка файла к просмотру;
- формирование VOD-плейлиста для проигрывания файла;
- формирование обложки видеофайла;
- формирование миниатюр для показа на полосе прокрутки во время перемотки видеофайла;
- передачу информации о файле в другие компоненты Медиаплатформы для последующей обработки файла Медиаплатформой;
- загрузку преобразованных файлов в хранилище;
- масштабирование Модуля для управления нагрузкой;
- связь с другими компонентами Медиаплатформы через брокера сообщений или другие интерфейсы;
- возможность настраивать Модуль под разные стратегии работы с медиаконтентом;
- мониторинг внутреннего состояния Модуля;
- сбор аналитической информации;
- обработку ошибок;
- логирование;
- сбор и визуальное отображение метрик Модуля;
- отправку событий в шину данных.

3.1.2.3 Способы и средства связи для информационного обмена между компонентами программного комплекса

Модуль взаимодействует со следующими компонентами системы:

- компонент «Внутренняя система управления задачами при обработке аудио визуального контента» – Модуль посредством брокера сообщений Kafka получает от данного компонента задачи на анализ загруженных пользователями медиафайлов и возвращает результаты такого анализа;
- хранилище S3 – исходный медиафайл для анализа Модуль берет из хранилища, в которое этот медиафайл поместил компонент «Система загрузки аудиовизуального контента в сети Интернет»;
- хранилище S3 или FileNhear - преобразованный (транскодированный) медиафайл Модуль загружает в хранилище;
- брокер сообщений Kafka.

3.1.3 Модуль «Внутренняя система управления задачами при обработке аудио визуального контента»

3.1.3.1 Назначение модуля

Модуль является центром реализации бизнес-логики Медиаплатформы и предназначен для оркестрации работы компонентов программного комплекса.

3.1.3.2 Функциональные возможности

Модуль обеспечивает:

- регистрацию и авторизацию внешних информационных систем (организаций-пользователей Медиаплатформы);
- задание и настройку профилей транскодирования файлов по параметрам: тип медиаконтейнера, размер, битрейт, кодек, исправление интерлейсинга, конвертация HDR, обработку субтитров и другие;
- конфигурацию подключения хранилища FileNep (раздельно для каждого пользователя);
- задание и настройку групп профилей конвертации с возможностью: независимой настройки параметров, указания приоритетов использования профилей внутри группы;
- управление процессом предоставления контента сторонней системе, использующей пользовательский интерфейс, с доступностью каждого качества сразу после транскодирования;
- уведомление внешних информационных систем (организаций-пользователей Медиаплатформы) о промежуточных результатах транскодирования;
- управление манипуляциями с содержимым: наложением специального изображения на видеокадр, нарезкой/склеивкой видео/фрагментов видео, "приклеиванием" заставки, обрезкой чёрных полос, автоматическим выравниванием громкости;
- применение правил обработки аудиовизуального контента, которые сформированы с учетом бизнес-логики сервиса;
- повышение приоритета отдельной выполняемой задачи;
- контроль исполнения процедур загрузки, обработки, хранения и предоставления аудиовизуального контента;
- передачу и получение сообщений другим компонентам Медиаплатформы через брокера сообщений;
- мониторинг внутреннего состояния Модуля;
- сбор аналитической информации;
- обработку ошибок;
- логирование;
- сбор и визуальное отображение метрик Модуля;
- отправку событий в шину данных.

3.1.3.3 Способы и средства связи для информационного обмена между компонентами программного комплекса

Модуль взаимодействует со следующими компонентами системы:

- «Система загрузки аудиовизуального контента в сети Интернет» – от этого компонента Модуль получает информацию о загрузке исходного видео;

- «Микросервис диагностики файлов аудиовизуального контента» – после получения информации о загрузке видео Модуль ставит для этого компонента задачу анализа загруженного видео;
- «Система транскодирования аудиовизуального контента» – после анализа загруженного контента Модуль ставит этому компоненту задачу на обработку видео;
- «Модуль трансляции аудиовизуального контента» - Модуль ставит для этого компонента задачу конвертации потока данных;
- брокер сообщений Kafka.

3.1.4 Модуль «Микросервис диагностики файлов аудиовизуального контента»

3.1.4.1 Назначение модуля

Модуль обеспечивает выполнение анализа исходного медиафайла, сбор необходимой информации для формирования задания на дальнейшее транскодирование.

3.1.4.2 Функциональные возможности

Модуль обеспечивает:

- загрузку файла из хранилища;
- диагностику файла аудиовизуального контента для определения сценария последующего транскодирования;
- диагностику целостности файла;
- извлечение и проверку метаданных файла;
- диагностику возможных проблем с последующей обработкой файла; Файлы, не соответствующие критериям качества и безопасности, не должны передаваться другим компонентам Медиаплатформы для последующей обработки;
- формирование отчета о результатах диагностики с параметрами, включающими:
 - ссылку на файл в хранилище;
 - тип контейнера файла (MP4, TS, AVI и тому подобные);
 - количество и параметры аудио/видео дорожек внутри контейнера;
 - кодеки, которыми обработаны аудио/видео дорожки;
 - предпочтения по использованию аудио/видео дорожек;
 - наличие/отсутствие чёрных полос;
 - уровни громкости аудио дорожек;
- передачу отчета о результатах диагностики и информации о файле в другие компоненты Медиаплатформы;
- масштабирование Модуля для управления нагрузкой;
- связь с другими компонентам Медиаплатформы через брокера сообщений или другие интерфейсы;
- мониторинг внутреннего состояния Модуля;
- сбор аналитической информации;
- обработку ошибок;
- логирование;
- сбор и визуальное отображение метрик Модуля;
- отправку событий в шину данных.

3.1.4.3 Способы и средства связи для информационного обмена между компонентами программного комплекса

Модуль взаимодействует со следующими компонентами программного комплекса:

- «Внутренняя система управления задачами при обработке аудио визуального контента» – Модуль посредством брокера сообщений Kafka получает от этого компонента задачи на анализ загруженных пользователями медиафайлов и возвращает результаты такого анализа;
- хранилище S3 – исходный медиафайл для анализа Модулем берется из хранилища, в которое этот медиафайл поместил компонент «Система загрузки аудиовизуального контента в сети Интернет»;
- брокер сообщений Kafka.

3.1.5 «Модуль трансляции аудиовизуального контента»

3.1.5.1 Назначение модуля

Модуль выполняет обработку (транскодирование) потока данных на основе профилей обработки, настроенных в компоненте «Внутренняя система управления задачами при обработке аудио визуального контента».

3.1.5.2 Функциональные возможности

Модуль обеспечивает:

- прием потока данных по протоколам HTTP/RTMP/SRT;
- независимое от других модулей управление;
- конфигурацию подключения хранилища S3;
- преобразование потока из одного цифрового формата в другой (транскодирование) в соответствии заданными профилями транскодирования потока по параметрам: тип медиаконтейнера, размер, битрейт, кодек;
- исправление интерлейса при конвертации;
- наложение специального изображения на видеокادر;
- формирование нескольких экземпляров потока в различных качествах по протоколу HTTP Live Streaming (HLS);
- формирование m3u8 плейлиста для потоков;
- передачу информации о потоке в другие компоненты Медиаплатформы для последующей обработки потока Медиаплатформой;
- загрузку файлов в хранилище;
- связь с другими компонентами Медиаплатформы через брокера сообщений или другие интерфейсы;
- возможность настраивать Модуль под разные стратегии работы с медиаконтентом;
- мониторинг внутреннего состояния Модуля;
- сбор аналитической информации;
- обработку ошибок;
- логирование;
- сбор и визуальное отображение метрик Модуля.

3.1.5.3 Способы и средства связи для информационного обмена между компонентами программного комплекса

Модуль взаимодействует со следующими компонентами системы:

- компонент «Внутренняя система управления задачами при обработке аудио визуального контента» – Модуль посредством брокера сообщений Kafka получает от данного компонента задачи на обработку потока данных и возвращает результаты обработки;
- хранилище S3 - преобразованный (транскодированный) медиафайл Модуль загружает в хранилище;
- брокер сообщений Kafka.

3.2 Режимы функционирования

Программный комплекс поддерживает следующие режимы функционирования:

- штатный – компоненты системы выполняют все свои основные функции;
- аварийный – один или несколько компонентов не выполняют своих функций.

В штатном режиме функционирования программный комплекс обеспечивает:

- работу пользователей в режиме 24 часа в день, 7 дней в неделю;
- выполнение общих функциональных требований.

В аварийном режиме система сохраняет состояние для последующего восстановления без потери данных.

Все компоненты должны иметь развернутые работоспособные копии, чтобы в случае отказа в короткий срок могли быть восстановлены.

Для всех технических компонентов развернутого комплекса необходимо обеспечить регулярный и постоянный контроль состояния и техническое обслуживание.

3.2.1 Предполагаемая нагрузка

- Формат загружаемых видео - MOV, MP4, MPG, AVI, FLV, 3GPP, WebM, MPEGPS, OGV, MKV, TS, HLS/Fragmented MP4.
- Формат потоков данных/контейнеров - HTTP/MPEG TS, HTTP Live Streaming/fmp4, RTMP, SRT.
- Максимальное разрешение загружаемых видео - UHD (до 4200×2160 пикселей).
- Видеокодеки - ProRes, HEVC (H.265), MPEG-1, MPEG-2, WMV, AVC (H.264), Theora, VP-8, VP-9, VC-1, AV1.
- Аудиокодеки - AAC LC, HE-AAC, AC3, E-AC3, MP3, Opus, Vorbis, WMA, PCM.
- Максимальный размер загружаемых видео - 24 Гб (настраиваемый параметр, может быть изменен после тестирования в рабочем окружении).