
VoiceNavigator

Программные средства системы голосового
взаимодействия

STC-S450

Руководство по настройке

ЦВАУ.00291-01 95

Версия документа: 2-301019

СОДЕРЖАНИЕ

1. ВВЕДЕНИЕ	6
1. 1 Общие положения	6
1. 2 Требования к персоналу	6
1. 3 Соглашения и обозначения	6
1. 4 Товарные знаки	8
2. ОБЩИЕ СВЕДЕНИЯ	9
2. 1 Основные сведения о комплексе	9
2. 2 Служба технической поддержки	9
3. НАЗНАЧЕНИЕ И СОСТАВ ПРОГРАММНЫХ СРЕДСТВ	10
3. 1 Назначение программных средств	10
3. 2 Основные функциональные возможности	11
3. 3 Состав программных средств	12
3. 4 STC MRCP Server	12
3. 5 STC Resource Server	12
3. 6 Программный комплекс синтеза речи	13
3. 7 Программный комплекс распознавания речи	13
3. 8 Программное обеспечение синтеза голосов	13
3. 9 Программное обеспечение поддержки	13
3. 10 Список открытых компонентов, входящих в состав VoiceNavigator	14
4. ВАРИАНТЫ УСТАНОВКИ ПРОГРАММНЫХ СРЕДСТВ	16
5. ТРЕБОВАНИЯ К ВИДАМ ОБЕСПЕЧЕНИЯ	18
5. 1 Требования к вычислительным средствам	18
5. 2 Требования к сети передачи данных и безопасности	19
5. 3 Требования к подключению внешних устройств	23
5. 4 Типовые схемы и требования	23
5. 5 Защита программного обеспечения	31
5.5.1 Варианты ключей защиты	32

5.5.2 Установка ПО ключа защиты в Windows	33
5.5.3 Установка ПО ключа защиты в Linux	34
5.5.4 Сбор данных конфигурации сервера и активация программного ключа в Windows	35
5.5.5 Сбор данных конфигурации сервера в Linux для программного ключа	37
5.5.6 Активация программного лицензионного ключа в Windows и Linux через веб-интерфейс	38
5.5.7 Отключение лицензионного ключа	40
5.5.8 Проверка активации лицензий	40
6. ГРАММАТИКИ РАСПОЗНАВАНИЯ РЕЧИ	44
6. 1 Распознавание по закрытым грамматикам	44
6. 2 Распознавание по открытым грамматикам	46
6. 3 Пакеты для слитного распознавания	48
6. 4 Обучение системы распознавания слитной речи	49
6. 5 Установка и обновление графа слитного распознавания	50
6. 6 Установка и обновление графа слитного распознавания в Windows	50
6. 7 Установка и обновление графа слитного распознавания в Linux	52
6. 8 Удаление графа слитного распознавания в Windows	53
6. 9 Удаление графа слитного распознавания в Linux	54
7. СПОСОБЫ ИНТЕГРАЦИИ	55
7. 1 Интеграция по протоколу MRCP	55
7.1.1 Общая информация об интеграции через MRCP	55
7.1.2 Требования к форматам аудиоданных по MRCP	56
7.1.2.1 Формат выходных аудиоданных	56
7.1.2.2 Формат входных аудиоданных	56
7.1.2.3 Значения идентификатора формата аудиоданных	56
7.1.2.4 Параметры сигнала, передаваемого на распознавание	57
7.1.3 Взаимодействие STC ASR и голосовой платформы в режиме MRCPv1	57
7.1.4 Взаимодействие STC ASR и голосовой платформы в режиме MRCPv2	58
7.1.5 Передача грамматик для распознавания по MRCP	59
7.1.5.1 Передача закрытых грамматик по MRCP	59
7.1.5.2 Передача грамматик слитного распознавания речи по MRCP	59
7.1.5.3 Взаимодействие STC TTS и голосовой платформы в режиме MRCPv1	60

7.1.5.4 Взаимодействие STC TTS и голосовой платформы в режиме MRCPv2	60
7. 2 Интеграция по протоколу WEB API	61
7.2.1 Общая информация по интеграции через Web API	61
7.2.2 Требования к форматам аудиоданных через WEB API	62
7.2.2.1 Формат выходных аудиоданных	62
7.2.2.2 Формат входных аудиоданных	62
7.2.2.3 Значения идентификатора формата аудиоданных	62
7.2.2.4 Параметры сигнала, передаваемого на распознавание	62
7.2.3 Настройка взаимодействия через Web API в режиме работы с файлами	62
7.2.4 Настройка взаимодействия через Web API в режиме потока	63
7.2.5 Распознавание аудиофайлов через Web API	64
7.2.6 Синтез речи через Web API	66
7.2.7 Распознавание речи через Web API в потоковом режиме	67
7.2.8 Синтез речи через Web API в потоковом режиме	68
8. НАСТРОЙКА ПРОГРАММНЫХ КОМПЛЕКСОВ	69
8. 1 Настройка MRCP-сервера	69
8.1.1 Местоположение конфигурационных файлов	69
8.1.2 Конфигурационный файл MrcpServerConfig.xml	69
8.1.3 Конфигурационный файл MrcpServer.nlog	70
8.1.4 Управляемые параметры распознавания речи по MRCP	70
8.1.5 Управляемые параметры синтеза речи по MRCP	72
8. 2 Настройка Resource-сервера	73
8.2.1 Местоположение конфигурационных файлов	73
8.2.2 Конфигурационный файл resourceServerConfig.xml	73
8.2.3 Конфигурационный файл artifactStorageConfig.xml	75
8.2.4 Конфигурационный файл tts_config.xml	76
8. 3 Настройка статистики и мониторинга	77
8.3.1 Рекомендации по мониторингу	77
8.3.2 Настройка конфигурации SNMP и примеры запросов	78
8.3.3 Настройка конфигурации SNMP Trap	81
9. ЗАПУСК И ОСТАНОВКА СЛУЖБ	82
9. 1 Запуск и остановка служб в среде ОС Windows	82

9. 2 Запуск и остановка служб в среде ОС Linux	82
10. ПРОТОКОЛИРОВАНИЕ СОБЫТИЙ	84
10. 1 Общие положения	84
10. 2 Изменение уровней протоколирования событий	84
10. 3 Конфигурирование хранения log-файлов	84
10. 4 Структура пути к артефактам	84
10. 5 Формирование в артефактах NLSML-ответа	85
11. ТЕХНИЧЕСКАЯ ПОДДЕРЖКА	86
11. 1 Порядок обработки обращений в техническую поддержку вендора	86
11. 2 Передача данных для поддержки VoiceNavigator из Windows	86
11. 3 Передача данных для поддержки VoiceNavigator из Linux	88
11. 4 Требования к персоналу для обслуживания VoiceNavigator	89
11. 5 Рекомендации по обновлению VoiceNavigator	90
11. 6 Рекомендации по проведению сервисных и регламентных работ	90
12. ПРИЛОЖЕНИЕ А. ССЫЛКИ НА ОПИСАНИЯ ПРОТОКОЛОВ И СТАНДАРТОВ	95
13. ПРИЛОЖЕНИЕ Б. ПЕРЕЧЕНЬ ТЕРМИНОВ И СОКРАЩЕНИЙ	96

1. ВВЕДЕНИЕ

1.1 Общие положения

В руководстве приведены необходимые сведения о настройке программных средств системы речевого взаимодействия **VoiceNavigator**.

Данное руководство предназначено для системных администраторов – сотрудников организации, специалистов по обслуживанию системы, отвечающих за её работу в штатном режиме.

Настоящее руководство не заменяет учебную, справочную литературу и руководства от производителей операционной системы и общего программного обеспечения.

Сведения, приведённые в руководстве, относятся к версии 10.2 системы речевого взаимодействия **VoiceNavigator**.

1.2 Требования к персоналу

Персонал, производящий работу с программными средствами **VoiceNavigator**, должен иметь профессиональные навыки установки общего (ОПО) и специального (СПО) программного обеспечения.



Персонал, производящий установку и эксплуатацию программных средств **VoiceNavigator**, должен изучить следующий документ: «Программные средства речевого взаимодействия «VoiceNavigator» STC-S450. Руководство по установке. ЦБАУ.00291-01 97».

Персонал должен обладать следующими основными навыками и знаниями:

- администрирование операционной системы Linux или Windows (в зависимости от поставки);
- чтение log-файлов и настройки продукта **VoiceNavigator**.

Персонал должен обладать следующими опциональными навыками и знаниями:

- установка и настройка голосовых платформ и/или систем взаимодействия;
- обладание знаниями основных протоколов TCP/IP, ознакомление со стандартами MRCP, VoiceXML, SRGS, SSML, NLSML;
- диагностирование проблемы функционирования с возможностью определения зоны ответственности, снятия диагностической информации и устранения проблемы (в том числе на других системах, с которыми интегрирован **VoiceNavigator**) в рамках своей компетенции.

1.3 Соглашения и обозначения

В документе приняты следующие типографские соглашения:

Формат текста	Описание
Обычный	Основной текст документа
<i>Курсив</i>	Применяется для написания специального понятия при первом появлении в основном тексте и списке терминов

Формат текста	Описание
Полужирный	Применяется для написания наименований программных компонентов, а также наименований управляющих и информационных элементов интерфейса (заголовки, кнопки и т.п.)
<i>Полужирный курсив</i>	Применяется для написания имён файлов и путей доступа к ним
Моноширинный	Применяется для написания текстов конфигурационных файлов и исходных текстов программ
<u>Подчёркнутый синий</u>	Применяется для обозначения внешних ссылок и ссылок на разделы настоящего руководства
Элемент > Вложенный элемент	Выбор пункта меню. Например, текст Файл > Выход должен пониматься так: выбрать меню Файл , а затем — команду Выход из меню Файл
<метка>	Фрагмент текста, который должен быть заменён реальным значением. Например, в записи D:\<DirName>\file.txt текст <DirName> используется для обозначения названия каталога в файловой системе



Указания на действия, которые необходимо выполнить в обязательном порядке.



Примечания, рекомендации и прочие сведения, требующие особого внимания.



Ссылки на другие документы в основном тексте.

1.4 Товарные знаки

Названия компаний и названия продуктов, упомянутые в документе, являются собственностью их соответствующих владельцев.

Ни одна из частей этого издания не подлежит воспроизведению, передаче, хранению в поисковой системе или переводу на какой-либо язык в любой форме, любыми средствами без письменного разрешения общества с ограниченной ответственностью «ЦРТ».

2. ОБЩИЕ СВЕДЕНИЯ

2.1 Основные сведения о комплексе

Наименование: Программные средства системы голосового взаимодействия
VoiceNavigator

Условное наименование: STC-S450

Обозначение: ЦБАУ.00291-01

Изготовитель: Общество с ограниченной ответственностью «ЦРТ»

Адрес: 196084, г. Санкт-Петербург, ул. Красуцкого, д. 4, литера А

Телефон: +7 (812) 325-88-48

Факс: +7 (812) 327-92-97



Программные средства системы голосового взаимодействия **VoiceNavigator** представлены на сайте компании:

<https://www.speechpro.ru/product/sistemy-upravleniya-kachestvom-i-avtomatizatsii/voicenavigator>

2.2 Служба технической поддержки

Адрес службы сервисного обслуживания и технической поддержки в интернете:

Электронная почта: info@speechpro.com

Адрес в сети Интернет: <http://www.speechpro.ru/support>

При обращении в службу технической поддержки необходимо представить четкое описание возникшей проблемы.

Предварительно подготовьте следующую информацию:

- наименование и версия программного обеспечения;
- четкое описание возникшей проблемы;
- тип компьютера и сведения о его конфигурации;
- название используемой операционной системы и номер её версии;
- приложить log-файлы программного обеспечения, отображающие проблему.

3. НАЗНАЧЕНИЕ И СОСТАВ ПРОГРАММНЫХ СРЕДСТВ

3.1 Назначение программных средств

Система голосового взаимодействия (система IVR) является составной частью центра обслуживания вызовов (ЦОВ) и предназначена для маршрутизации звонков на основе информации, вводимой абонентом с помощью тонального набора или голоса.

Программные средства системы голосового взаимодействия **VoiceNavigator** являются составной частью системы IVR и предназначены для обеспечения взаимодействия абонентов с информационными системами ЦОВ (рис. 1). Общение абонента с системами ЦОВ происходит с использованием технологий синтеза и распознавания речи.

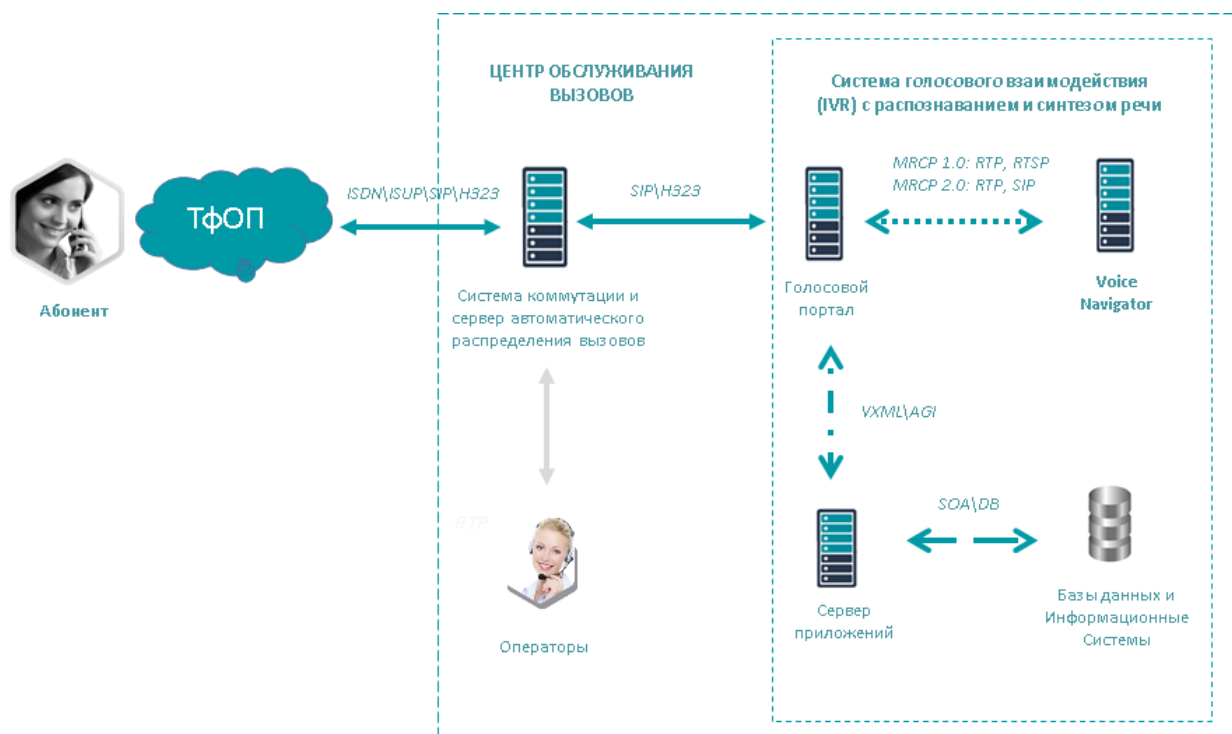


Рисунок 1 – Взаимодействие через MRCP с продуктом VoiceNavigator

Внедрение продукта позволяет создать функциональные и надёжные решения для выполнения различных задач, таких как:

1. Распознавание и синтез речи для центров обслуживания вызовов, систем голосового самообслуживания (IVR):
 - создание решений для маршрутизации обращений по голосовому высказыванию;
 - создание решений для самообслуживания на основе информации, вводимой с помощью голоса.
2. Создание решений по исходящему обслуживанию вызовов (информирование, оповещение, опросы и пр.) с распознаванием и синтезом речи.

Расширенным назначением ПО является система роботизированного голосового взаимодействия (для интеграции с системами iRPA), которая предназначена для помощи и контроля сотрудников на основе голосового общения.

Внедрение продукта позволяет создать функциональные и надежные решения для выполнения различных задач, таких как:

1. распознавание и синтез речи для компаний:

- онлайн распознавания разговоров с сотрудниками;
 - офлайн распознавания аудиофайлов;
 - роботизация операций на основе голосового интерфейса.
2. распознавание и синтез речи для сайтов, мессенджеров и мобильных приложений;
 3. распознавание и синтез речи для терминалов самообслуживания;
 4. создание других локальных и облачных решений, где общение с человеком производится через интерфейс с распознаванием и синтезом речи.

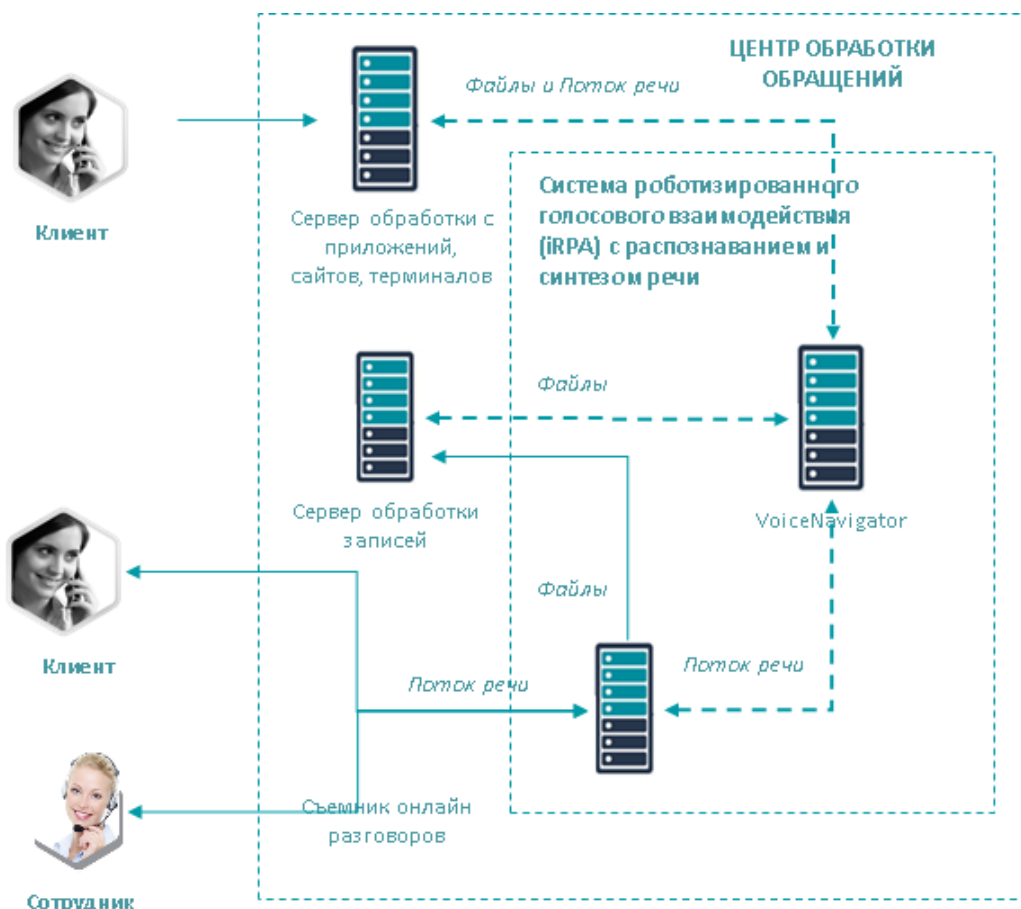


Рисунок 2 – Взаимодействие через Web API с продуктом VoiceNavigator

Принцип обработки запроса заключается в следующем:

- получение запроса через интерфейс интеграции;
- распознавание, синтез, понимание речи;
- предоставление ответа через интерфейс интеграции.



VoiceNavigator не несет в себе логики взаимодействия, а только обрабатывает запросы для распознавания и синтеза речи через интерфейсы взаимодействия (MRCP и Web API). Логiku взаимодействия с пользователями выполняют сторонние системы.

3.2 Основные функциональные возможности

Система выполняет следующие функции:

- системная поддержка систем виртуализации, ОС Windows и Linux;

- встроенные технологии распознавания речи (поддержка распознавания по грамматикам SRGS и слитное распознавание по большому словарю), синтез речи (поддержка SSML);
- поддержка голосовых платформ с поддержкой MRCP: Asterisk, Avaya, Cisco, Genesys, FortelT, Naumen и другие;
- поддержка протоколов интеграции MRCP v1 и v2, Web API (распознавание файлов, потоковое распознавание, синтез речи);
- поддержка языков распознавания речи: Английский, Казахский, Русский;
- поддержка языков синтеза речи: Английский, Казахский, Русский (7 голосов);
- расширенная система протоколирования;
- поддержка систем мониторинга SNMP v1 и v2, SNMP Trap.

3.3 Состав программных средств

Состав программных средств **VoiceNavigator** включает следующие программные комплексы: **STC MRCP Server**, **STC Resource Server**, программный комплекс синтеза речи **STC TTS** и программный комплекс распознавания речи **STC ASR**, которые представлены на рисунке 3.

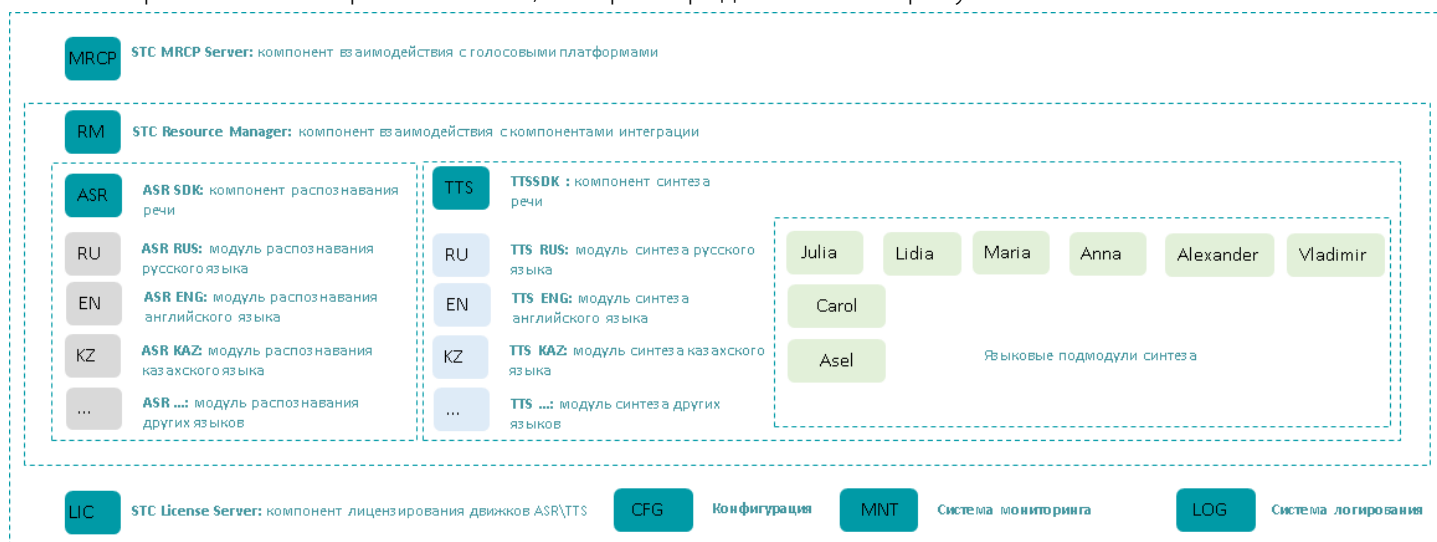


Рисунок 3 – Состав программных средств

3.4 STC MRCP Server

STC MRCP Server (MRCP-сервер) – программный комплекс, предназначенный для управления взаимодействием между голосовой платформой и программными комплексами синтеза и распознавания речи на основе протокола MRCP.

Программный комплекс **STC MRCP Server** состоит из работающей в виде службы программы **MrcpServerService**.

MRCP-сервер поддерживает протокол MRCP версий 1 и 2.

MRCP-сервер поддерживает форматы: VoiceXML, SRGS, SSML, NLSML.

MRCP-сервер совместим с библиотекой Open Source UniMRCP.

3.5 STC Resource Server

STC Resource Server – программный комплекс предназначен для управления взаимодействием между MRCP Server и программными комплексами распознавания и синтеза речи.

Программный комплекс **STC Resource Server** состоит из работающей в виде службы программы **STC ResourceServerService**.

3.6 Программный комплекс синтеза речи

STC TTS – программный комплекс синтеза русской, английской и казахской речи, предназначенный для использования в составе многоканальных систем голосового взаимодействия.

3.7 Программный комплекс распознавания речи

STC ASR – программный комплекс распознавания русской, английской и казахской речи, предназначенный для использования в составе многоканальных систем голосового взаимодействия.

3.8 Программное обеспечение синтеза голосов

Программное обеспечение синтеза голосов включает следующие голосовые базы дикторов:

Голоса для синтеза речи по правилам	Голоса на основе нейронной сети
Alexander8000 Anna8000 Vladimir8000 Lydia8000 Asel8000 Maria8000 Julia8000 Carol8000	Julia_8000n Anna_8000n Vladimir_8000n Julia_f8000n Anna_f8000n Vladimir_f8000n Голоса, содержащие в своём наименовании f требуют меньших ресурсов процессора, но более вариативны по качеству.

3.9 Программное обеспечение поддержки

Программное обеспечение поддержки включает: сервер лицензирования, систему конфигурации, систему мониторинга, систему логирования.

Система конфигурации **VoiceNavigator** (CFG) содержит файлы настроек для работы ПО, которые суммарно занимают несколько мегабайт и мало изменяются. В файлы конфигурации также входят грамматики слитного распознавания, которые могут занимать несколько Гигабайт и могут часто изменяться при тюнинге. При наличии у заказчика системы резервного копирования и защиты файлов от изменений производится настройка таким образом, чтобы при изменении конфигурации производилось резервирование, либо администратор оповещался об изменении конфигурации. При наличии резервной копии файлы конфигурации являются не критичным элементом при неисправности сервера. Для выполнения требований безопасности доступ к файлам конфигурации ограничивается на уровне операционной системы.

Для дополнительного анализа работы **VoiceNavigator** и улучшения качества распознавания и синтеза используется система логирования (LOG). **VoiceNavigator** постоянно добавляет отладочную информацию по работе ПО и артефакты, полученные при обработке запросов, в log-файлы. В зависимости от режима отладки и нагрузки эти данные могут занимать сотни мегабайт. Управление объёмом производится путём настройки ротирования в конфигурационных файлах. Объем log-файлов может быть

также ограничен имеющимся пространством жёсткого диска, при неправильной настройке ротирования файлы логирования могут занять все место, что может привести к проблемам функционирования ПО. Наиболее объёмными логами являются артефакты, полученные при обработке запросов. Обычно в режиме отладки объем артефактов на 1 обращение (порядка 3-4 шагов распознавания) составляет около 1,2 Мб (аудио) и 1,2 Мб (служебные данные для анализа). Для понимания общего объёма хранилища надо умножить среднюю длительность разговора на объём одной минуты и на среднее количество обращений в сутки. Включение логирования артефактов обычно обеспечивается на этапе опытной эксплуатации для отладки, рекомендуется включать логирование с ротированием на глубину до 15 суток, что позволит при выявлении проблемы с распознаванием быстрее расследовать ситуацию. При промышленном использовании в системах с высокой нагрузкой логирование артефактов рекомендуется отключать. При неисправности сервера log-файлы являются некритичными данными. Для выполнения требований безопасности доступ к log-файлам ограничивается на уровне операционной системы. При распознавании чувствительной информации имеется возможность не сохранять данные в артефактах посредством передачи от голосового приложения специальной команды.

Сервер лицензирования (LIC) выполняет функции контроля лицензий компонентов. Контроль реализован на базе продукта **Sentinel**.

В составе также имеется программный модуль, который предоставляет возможности по подключению к системам мониторинга по протоколу SNMP (MNT).

3.10 Список открытых компонентов, входящих в состав VoiceNavigator

Программное обеспечение содержит следующие открытые компоненты:

Компонент	Ссылка
nlog	https://nlog-project.org/
BSD License	
boost	http://www.boost.org/
Boost Software License - Version 1.0	http://www.boost.org/LICENSE_1_0.txt
Google Protobuffers	https://developers.google.com/protocol-buffers/
Creative Commons Attribution 3.0 License	https://developers.google.com/site-policies
jemalloc	https://github.com/jemalloc/jemalloc
Facebook, Inc	https://github.com/jemalloc/jemalloc/blob/dev/COPYING
boost::threadpool	http://threadpool.sourceforge.net/
Free	http://threadpool.sourceforge.net/license.html
boost::httpclient	http://cpp-netlib.org/

Free	https://github.com/cpp-netlib/cpp-netlib/blob/master/LICENSE_1_0.txt
jrtpplib	http://research.edm.uhasselt.be/jori/page/CS/Jrtpplib.html
Free	http://research.edm.uhasselt.be/jori/page/CS/Jrtpplib.html
breakpad	https://chromium.googlesource.com/breakpad/breakpad/
Copyright (c) 2006, Google Inc	https://chromium.googlesource.com/breakpad/breakpad/+/_master/LICENSE
concurrentqueue	http://dotnetpattern.com/csharp-concurrentqueue
Simplified BSD License	
net-snmp	http://www.net-snmp.org/
Мультилицензионная, CMU/UCD (аналог BSD License)	www.net-snmp.org/about/license.html

4. ВАРИАНТЫ УСТАНОВКИ ПРОГРАММНЫХ СРЕДСТВ

Перед установкой ПО необходимо подготовить оборудование и программное обеспечение. Выбор конфигурации зависит от количества каналов, используемых голосовой платформой.

При разворачивании программных средств **VoiceNavigator** рекомендуется использовать варианты, представленные ниже.

1. Вариант, при котором программные комплексы **STC MRCP Server**, **STC Resource Server**, **STC TTS**, и **STC ASR** устанавливаются на одном компьютере (рис. 4).

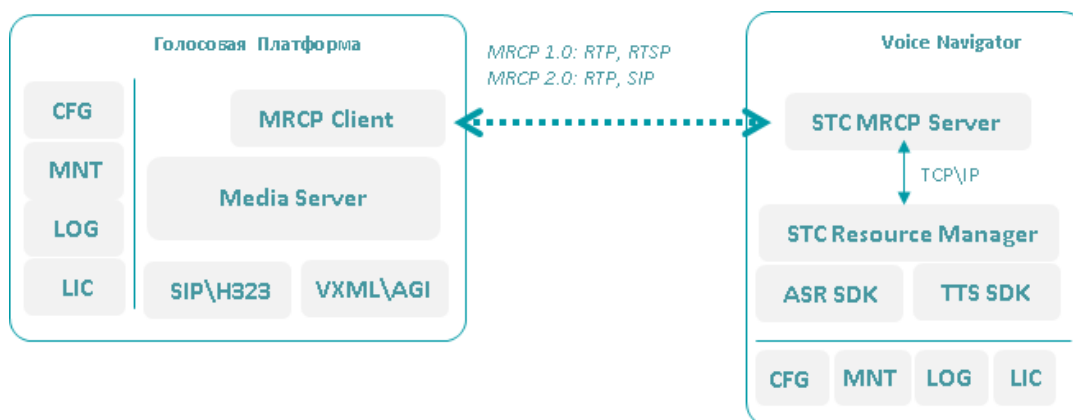


Рисунок 4 – Централизованное использование ресурсов VoiceNavigator

2. Вариант с распределённым размещением программных комплексов (рис. 5). На одном из компьютеров устанавливаются программные комплексы **STC MRCP Server**, **STC Resource Server** и **STC ASR**, а на другом – программные комплексы **STC MRCP Server** и **STC TTS**.

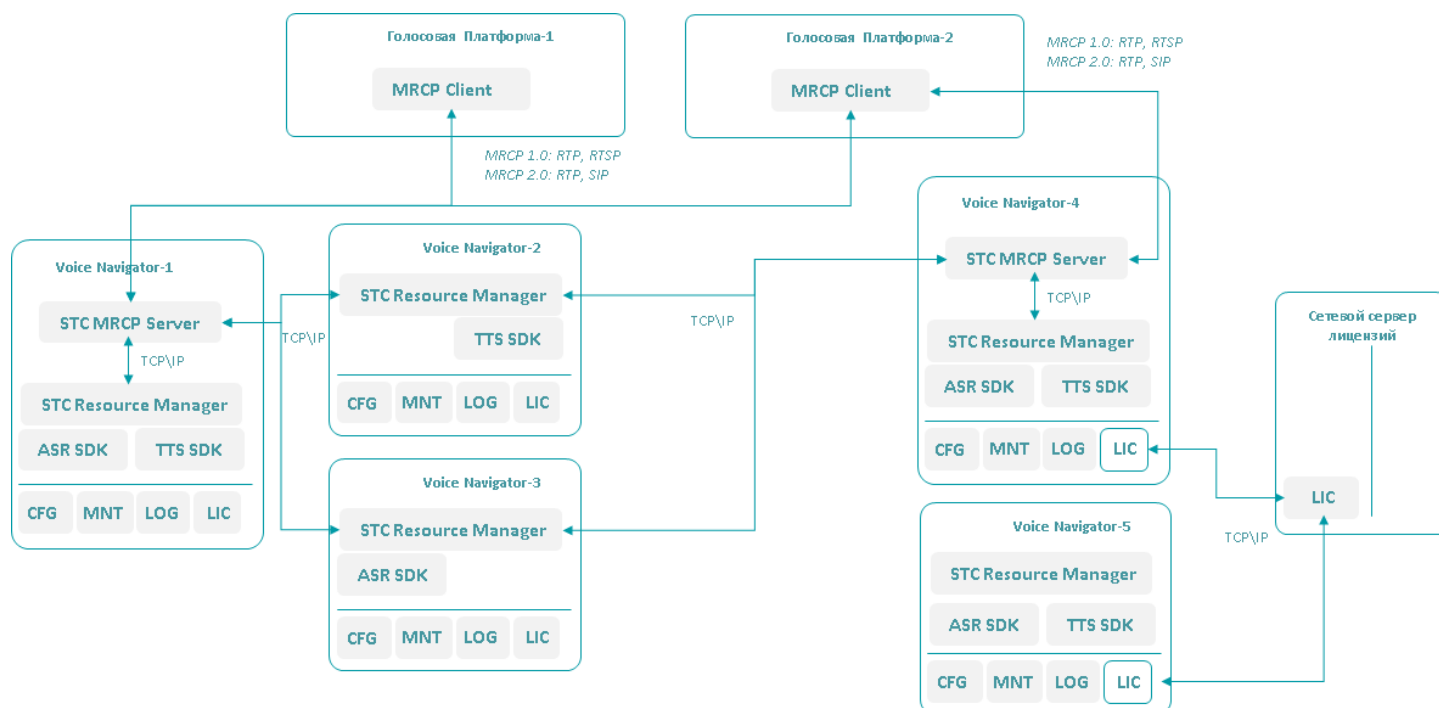


Рисунок 5 – Распределенное использование ресурсов VoiceNavigator

При количестве используемых каналов более 50 рекомендуется использовать вариант с раздельным размещением программных комплексов.

Резервирование ПО возможно с использованием одного из подходящих для эксплуатации вариантов:

- Разнесение компонентов с распознаванием и синтезом на разные сервера с дублированием элементов (например, по схеме N+1). При этом балансировка между компонентами производится средствами Голосовой Платформы.
- Резервирование данных на серверах средствами корпоративных специализированных систем (например, полного образа или только изменяемых данных);
- Резервирование средствами системы виртуализации путём копирования виртуальных серверов, использование отказоустойчивого СХД для работы и хранения копий серверов и данных, использовании технологии FaultTolerance, переноса машин средствами V-Motion с созданием снапшотов и пр.
- Резервирование путем разнесения ПО и его компонентов на несколько серверов и площадок (с использованием схем холодного или горячего резервирования).

При резервировании ПО чувствительным элементом является сервер лицензирования, при недоступности которого функциональность ПО нарушается. Для возможности гибкого переноса, архивирования виртуальных машин используется лицензия вида FQDN (привязка по доменному имени) или отдельного сервера с аппаратным сетевым ключом если вендор ПО отказывается от передачи FQDN. Также возможен вариант с холодным или горячим резервированием лицензий (описано далее).

Рекомендации по резервированию **VoiceNavigator**:

- наличие инструкций, дистрибутивов и резервных образов виртуальных машин (согласно принятым внутрикорпоративным документам);
- наличие дублирующего сервера;
- периодическое (не чаще чем раз в месяц) резервное копирование конфигураций;
- сохранение резервных грамматик при обновлении;
- опционально (на усмотрение заказчика) может рассматриваться резервное копирование записей или использование NAS для долговременного хранения;
- резервное копирование системных log-файлов.

5. ТРЕБОВАНИЯ К ВИДАМ ОБЕСПЕЧЕНИЯ

5.1 Требования к вычислительным средствам

Ниже приведены типовые требования к свободным вычислительным ресурсам, предъявляемые программными комплексами **STC TTS** и **STC ASR**. Требования зависят от количества каналов синтеза и распознавания речи, количества запущенных движков, объема грамматик слитного распознавания и голосов, рисунка диалога (частоте и длительности обращения к ресурсам).

Закрытые грамматики предназначены для распознавания речи по ограниченному набору возможных высказываний (обычно не более 1000 фраз). Используются в ситуациях, когда необходимо быстро сформировать грамматики и с высокой вероятностью распознать короткие высказывания человека из четкого набора слов/фраз.

Открытые грамматики (или распознавание по большому словарю) предназначены для распознавания речи по большому объему возможных словоформ (более 1 млн). Используются в ситуации, когда надо распознать длинные высказывания человека, выраженные в свободной форме, не имеющего четкого набора слов/фраз.

Минимальные требования к ресурсам на 1 сервер (при одном языке и голосе с минимальным объемом грамматик до 0,5 Гб):

Количество каналов распознавания речи	1-45	46-90
Количество ядер процессора	4	8
Объем оперативной памяти, Гб	4	8
Объем жесткого диска для установки грамматик, Гб	40	40

Минимальные требования к ресурсам на 1 сервер для онлайн распознавания с высокой нагрузкой и объемом грамматик до 4Гб:

Количество каналов распознавания речи	1-45	46-90
Количество ядер процессора	16	Установите 2й сервер
Объем оперативной памяти, Гб	24	
Объем жесткого диска для установки грамматик, Гб	60	60

Минимальные требования к ресурсам на 1 сервер для синтеза речи (1 язык, 2 голоса):

Количество каналов синтеза речи	1-50	51-100
Количество ядер процессора (технология Unit Selection)	4	8

Объём оперативной памяти, ГБ (технология Unit Selection)	4	8
Объём жесткого диска для установки голосов, ГБ (технология Unit Selection)	40	40
Количество ядер процессора (технология DNN)	12	24
Объём оперативной памяти, ГБ (технология DNN)	8	16
Объём жесткого диска для установки голосов, ГБ (технология DNN)	100	100

Каждый из компьютеров, на которых устанавливаются программные средства **VoiceNavigator**, должен включать в свой состав:

- процессор не хуже Xeon E5-26XX 2,4 Ghz;
- жёсткий диск с объёмом свободной памяти от 40 ГБ (зависит от планируемой нагрузки, устанавливаемых компонентов и необходимой длительности хранения логов);
- контроллер локальной вычислительной сети Ethernet, обеспечивающий скорость передачи данных не менее 100 Мбит/с;
- при использовании синтеза DNN рекомендуется включать набор инструкций AVX2 и AVX-512 (например, имеет новое поколение процессоров Intel Xeon серии Gold), не забывая про необходимость включения и про наличие поддержки технологией виртуальной машины.

Возможна установка на системы виртуализации. Рекомендуемая система виртуализации VNWare 6.7 и выше.



При использовании слитного распознавания количество одновременно используемых каналов распознавания определяется количеством ядер процессора.

5.2 Требования к сети передачи данных и безопасности

Программные средства **VoiceNavigator** допускают распределенную установку, при которой каждый компонент программных средств устанавливается на отдельный компьютер.

Взаимодействие компонентов рекомендуется обеспечивать через локальную сеть, в одном сегменте, с пропускной способностью не менее 100 Мбит/с по протоколу TCP/IP.

При использовании распознавания и синтеза речи необходимо учитывать состояние каналов связи (использование кодеков, задержки, эхо, реверберацию, потерю пакетов и пр.), внешние шумы и удалённости микрофона от говорящего, что может значительно (до 20-30%) снижать качество работы.

В случае кроссканальности (переконвертации кодеков G711-G723-G729-G711) качество может падать на 5-7%.

Требования к пропускной способности сетевого интерфейса компьютера включают в себя полосу пропускания, необходимую для управления комплексом **VoiceNavigator** и слежения за его состоянием. Расчёт нагрузки делается стандартным способом как при проектировании сети для гарантированного пропуска VOIP-трафика.



Для расчёта полосы пропускания (на 1 канал) в части VOIP может использоваться формула БвС = $РС * ППК / 8$, где БвС – количество байт в сэмпле, РС – размер сэмпла в секундах, ППК – пропускания кодека в битах/сек.

При использовании кодека G.711 в одном пакете, к примеру, 20 миллисекунд разговора необходимо выставить значение параметра БвС = $0,02 * 64000 / 8 = 160$.

Если использовать протокол Ethernet с шифрованием VPN, то к 160 байтам добавятся 18 байт служебной информации, плюс заголовки IP, UDP и RTP (20+8+12 байт), плюс для IPSec 50 байт. Суммарно имеем пакет размером 268 байт.

Полоса пропускания в этом случае составит 107,2 Кбита в секунду.

При отсутствии шифрования полоса пропускания составит порядка 84,6 Кбит в секунду.

Сетевые протоколы, которые обеспечивают корректную работу с **VoiceNavigator**, представлены на рисунке 6 (данные для работы с ОС не включены):

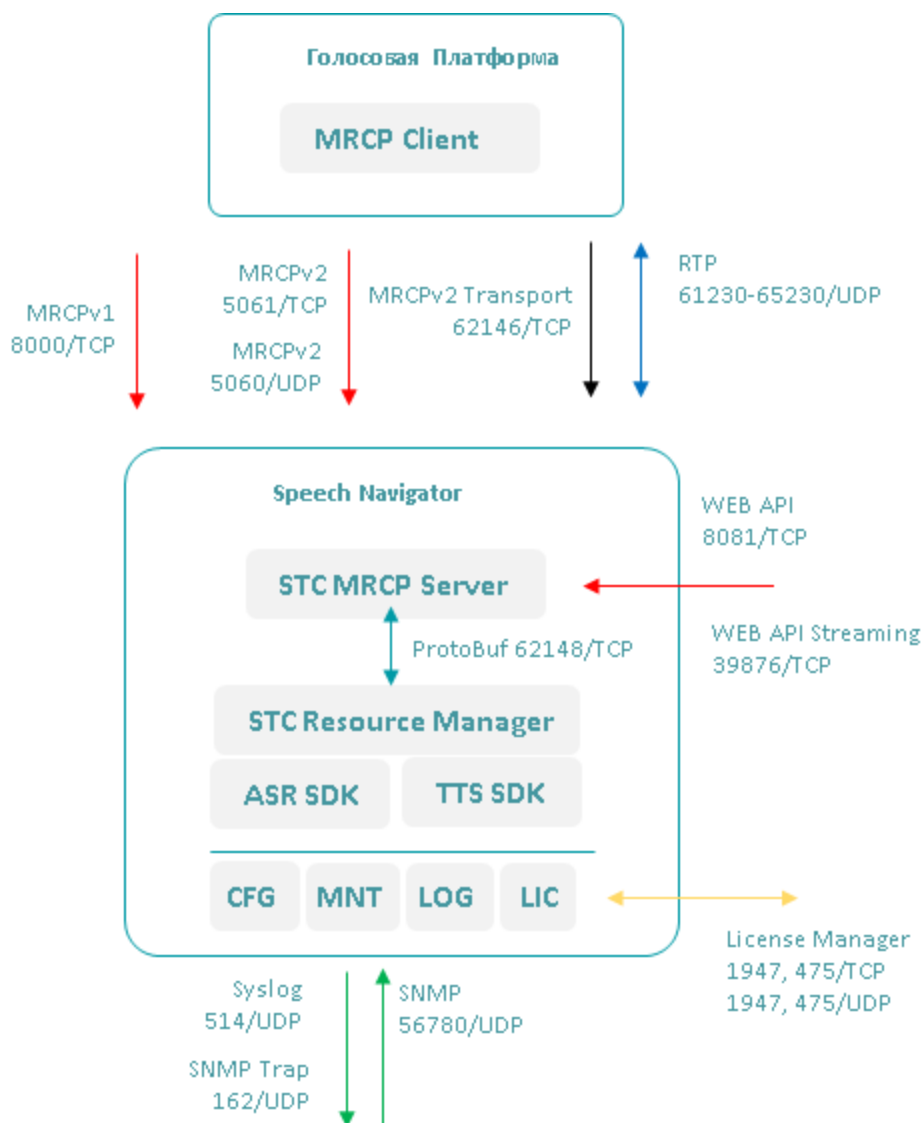


Рисунок 6 – Сетевые протоколы для корректной работы VoiceNavigator

В целях безопасности и прозрачности поведения ПО **VoiceNavigator** не изменяет системные настройки брандмауэра (файрвола).

Для обеспечения работоспособности ПО выполните настройку фаервола в соответствии с сетевыми маршрутами или отключите его (может быть использована как временная мера для первичной проверки корректности установки и функционирования).

Справочные консольные команды для Windows для работы с фаерволом представлены ниже.

Отключение фаервола через командную строку:

```
netsh advfirewall set allprofiles state off
```

Включение фаервола через командную строку:

```
netsh advfirewall set allprofiles state on
```

Открытие порта:

```
netsh advfirewall firewall add rule name="Open MRCPv1 Server Port 8000" dir=in action=allow  
protocol=TCP localport=8000
```

Открытие фаервола для конкретной программы:

```
netsh advfirewall firewall add rule name="Allow STC Resource Server Service" dir=in action=allow  
program="%ProgramFiles%\Speech Technology Center\Resource Server\Bin\ResourceServer.exe"
```

Экспорт/импорт установок фаервола:

```
netsh advfirewall export "C:\temp\FWConfSTC.wfw"  
netsh advfirewall import "C:\temp\FWConfSTC.wfw"
```

Сброс параметров фаервола к установкам по умолчанию:

```
netsh advfirewall reset
```

Справочные консольные команды для Linux для работы с фаерволом и системами безопасности представлены ниже.

Проверка статуса компонент:

```
sestatus  
systemctl status firewalld  
systemctl status iptables
```

Отключение ядра безопасности:

```
sudo setenforce 0  
sed -i "s/SELINUX=enforcing/SELINUX=disabled/" /etc/selinux/config
```

Включение ядра безопасности:

```
sudo setenforce 1  
sed -i "s/SELINUX=disabled/SELINUX=enforcing/" /etc/selinux/config
```

Остановка сервисов фаервола (в зависимости от используемого):

```
systemctl stop firewalld  
systemctl stop iptables
```

Запуск сервисов фаервола (в зависимости от используемого):

```
systemctl start firewalld  
systemctl start iptables
```

Запрет автостарта фаерволов (в зависимости от используемого):

```
systemctl disable firewalld  
systemctl disable iptables
```

Активация автостарта фаерволов (в зависимости от используемого):

```
systemctl enable firewalld  
systemctl enable iptables
```

Открытие порта, добавление правила (в зависимости от используемого):

```
firewall-cmd --permanent --zone=public --add-port=8000/tcp  
firewall-cmd --zone=public --add-port=61230-65230/udp  
iptables -I INPUT -p tcp --dport 8000 -m state --state NEW -j ACCEPT  
iptables -A INPUT -m state --state NEW -m udp -p udp --match multiport --dports 61230:65230 -j ACCEPT  
service iptables save
```

Перезагрузка правил без потери текущих соединений (в зависимости от используемого):

```
firewall-cmd --reload
```

Открытие порта в **selinux** (через semanage с установленным polycoreutils-python):

```
semanage port -a -t mrcpl_port_t -p tcp 8000
```

Перезагрузка для проверки внесённых изменений:

```
sudo shutdown -r now
```



С подробными инструкциями по работе с брандмауэрами можно ознакомиться в инструкции администратора операционной системы.



ПО **VoiceNavigator** корректно функционирует с антивирусными программами.

Однако при значительной нагрузке антивирусные программы могут негативно сказываться на производительности, пытаясь проверять все пакеты, которые пришли для распознавания или синтеза. Для детектирования данной ситуации рекомендуется временно приостанавливать функционирование антивирусных программ, а при подтверждении проблемы необходимо провести дополнительную настройку антивируса, в том числе запрашивая поддержку производителя антивируса либо резервировать дополнительные ресурсы (память, процессоры) для выполнения проверки.

Общие рекомендации для корректной работы голосовых сервисов в сети:

- программные средства **VoiceNavigator** допускают распределенную установку в локальной сети, при которой каждый компонент программных средств устанавливается на отдельный компьютер;
- взаимодействие компонентов должно осуществляться через локальную сеть с пропускной способностью не менее 100 Мбит/с по протоколу TCP/IP;
- все компьютеры системы рекомендуется устанавливать в одном сегменте локальной сети;
- требования к пропускной способности сетевого интерфейса компьютера включают в себя полосу пропускания, необходимую для управления комплексом **VoiceNavigator** и слежения за его состоянием;
- в сегменте сети между серверами должны быть открыты порты (организованы сетевые маршруты) для протоколов RTP, RTSP, HTTP;
- задержка доставки пакетов в сегменте сети между серверами должна составлять не более 20 мс;
- для распознавания речи в режиме диктовки голосовой трафик может захватываться со съёмника;
- для гарантирования бесперебойной работы съёмник рекомендуется подключать к двум 1Gbps портам;
- использование 711 кодека при использовании протокола MRCP;
- использование wav-файлов PCM (8 и 16bit в зависимости от используемого профиля).

Для корректной работы голосовых сервисов из облака рекомендуется обеспечивать следующие показатели каналов (в том числе VPN) передачи данных:

- потери пакетов при максимальной пропускной способности не более 2х на 1000 пакетов;
- пропускная способность каналов для 2Мб 1500 rps, 10Мб 4000 rps (при размере пакета 64 байта);
- пропускная способность канала не менее 90% от расчётной;
- задержка сигнала не более 35 мс (при 50% загрузке канала);
- джиттер в канале не более 20 мс;
- использование 711 кодеков при использовании протокола MRCP;

- использование wav-файлов PCM (8 и 16bit в зависимости от используемого профиля).

Для корректной работы голосовых сервисов по телефонным цифровым каналам (ISDN, SS7, ISUP) рекомендуются следующие показатели:

- отсутствие слипов в телефонном канале связи;
- отсутствия фазового дрожания цифрового сигнала;
- отсутствие эха и микширования в канале;
- отсутствие перегрузки и реверберации;
- соотношение сигнал/шум не менее 15 дБ;
- отсутствие перегрузки;
- отсутствие фоновых шума (музыка, "речевой коктейль");
- отсутствие искажений канала связи (искажения частотной характеристики; периодические наводки; импульсные помехи);
- энергия сигнала не менее 15 % от максимума.

5.3 Требования к подключению внешних устройств

При установке программного комплекса **STC Resource Server** с синтезом речи **STC TTS** и распознаванием речи **STC ASR** на одном компьютере требуется один свободный USB-порт для установки HASP-ключа лицензирования. Один HASP-ключ содержит лицензии и для **STC ASR**, и для **STC TTS**.

При распределённой установке программных комплексов **STC Resource Server** синтеза речи **STC TTS** и распознавания речи **STC ASR** на разных компьютерах один свободный USB-порт для установки HASP-ключей лицензирования должен быть на каждом компьютере.

Есть вариант программной лицензий, который устанавливается локально (USB-порт не требуется).

5.4 Типовые схемы и требования

Все компоненты продукта **VoiceNavigator** могут функционировать на одном сервере по следующей схеме (рис. 7).

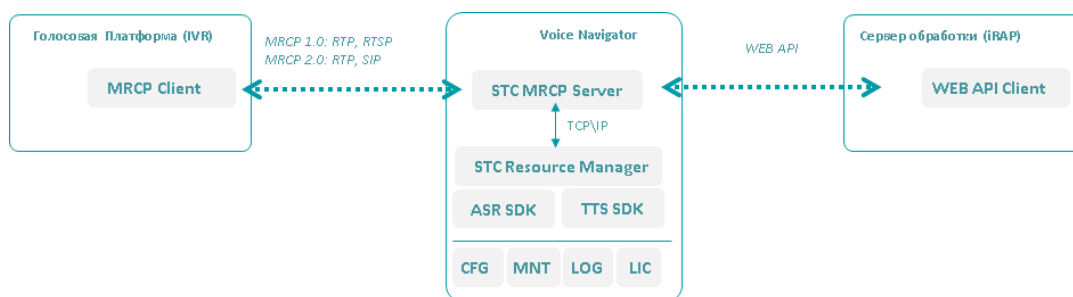


Рисунок 7 – Централизованное использование ресурсов VoiceNavigator

Рекомендованная конфигурация сервера представлена в таблице:

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
Рекомендуемый физический сервер с виртуальными машинами: сервер типа Xeon E5-26XX (2CPU 2,4 Ghz x 8 core); 64 Gb RAM; RAID 2x600 Gb HDD 15K + 2x1Tb HDD 7K; double power										
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания и система синтеза речи (все в одном) для MRCP	1	16	24	300	100	1000	70	До 15 одновременных распознавания на двух графах слитного распознавания + 15 одновременных синтезов
				16	24	300		1000		



Приведены примеры конфигурации. Точные параметры рассчитываются по запросу после анализа задачи и планируемой нагрузки.

В виду разной степени утилизации ресурсов сервера и принципов работы (обработка потоков речи или файлов) могут быть проблемы с прогнозированием ресурсов, поэтому рекомендуется разделять сервера под IVR и под iRPA, как представлено на рисунке 8.

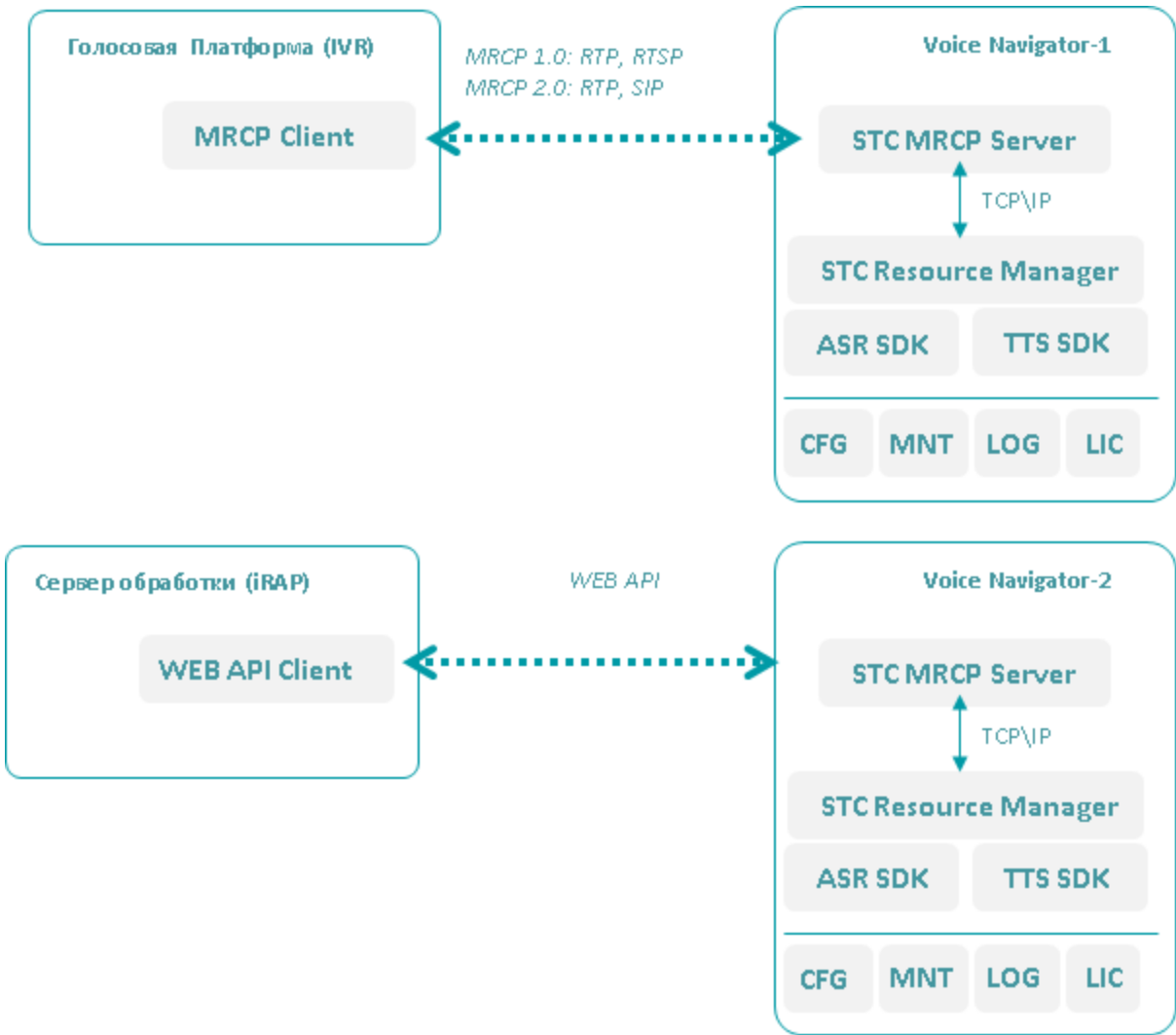


Рисунок 8– Распределенное использование ресурсов VoiceNavigator. Вариант 1

Рекомендованная конфигурация серверов представлена в таблице:

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
Рекомендуемый физический сервер с виртуальными машинами: сервер типа Xeon E5-26XX (2CPU 2,4 Ghz x 8 core); 64 Gb RAM; RAID 2x600 Gb HDD 15K + 2x1Tb HDD 7K; double power										
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания и система синтеза речи (все в одном) для MRCP	1	16	24	300	100	1000	70	До 15 одновременных распознавания на двух графах слитного распознавания + 15 одновременных синтезов

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания и система синтеза речи (все в одном) для Web API	1	16	24	300	100			До 15 одновременных распознавания на двух графах слитного распознавания + 15 одновременных синтезов
				32	48	600		1000		

Основная проблема использования – это трудности при планировании ресурсов для ASR и TTS, а также варианты обслуживания, когда при работах с ASR автоматически затрагивается TTS. Во избежание данных конфликтов рекомендуется выносить ASR и TTS на разные сервера (рис. 9).

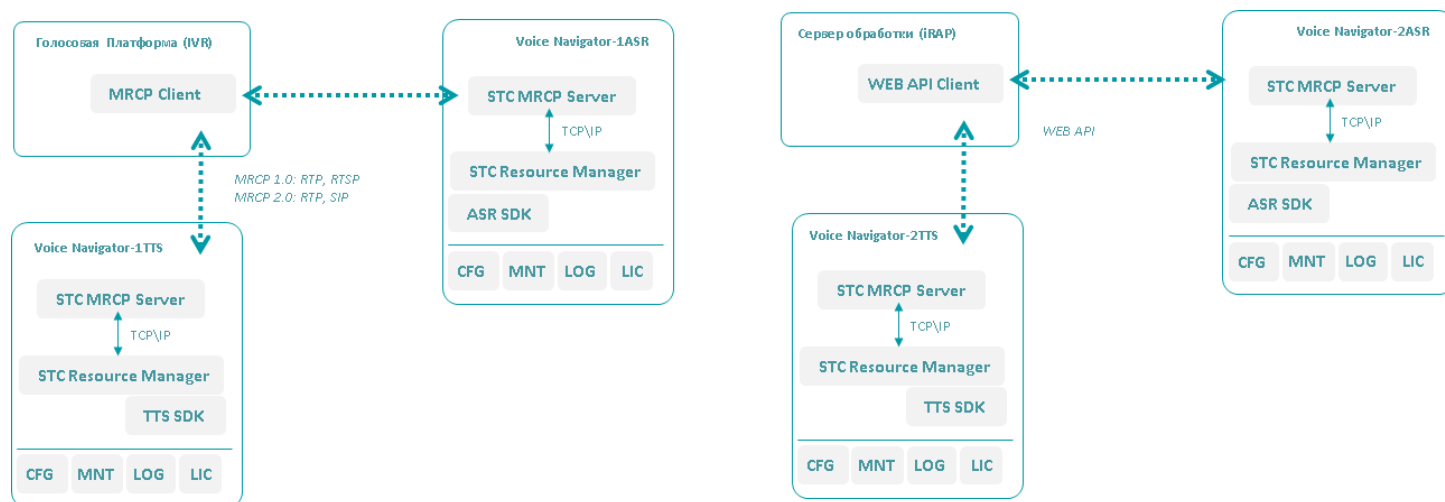


Рисунок 9– Распределенное использование ресурсов VoiceNavigator. Вариант 2

Рекомендованная конфигурация серверов представлена в таблице:

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для MRCP	1	16	16	200	100	1000	70	До 50 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для MRCP	1	8	16	200	100			До 50 одновременных синтезов

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения логов-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для Web API	1	16	24	200	100			До 40 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для Web API	1	8	16	200	100			До 50 одновременных синтезов
				48	72	800		1000		

При необходимости масштабирования или резервирования рекомендуется добавлять нужное количество серверов (рис. 10). Рекомендуется планировать количество серверов из расчета vCPU 8, 16 для удобства использования V-Motion и постепенного наращивания ресурсов.

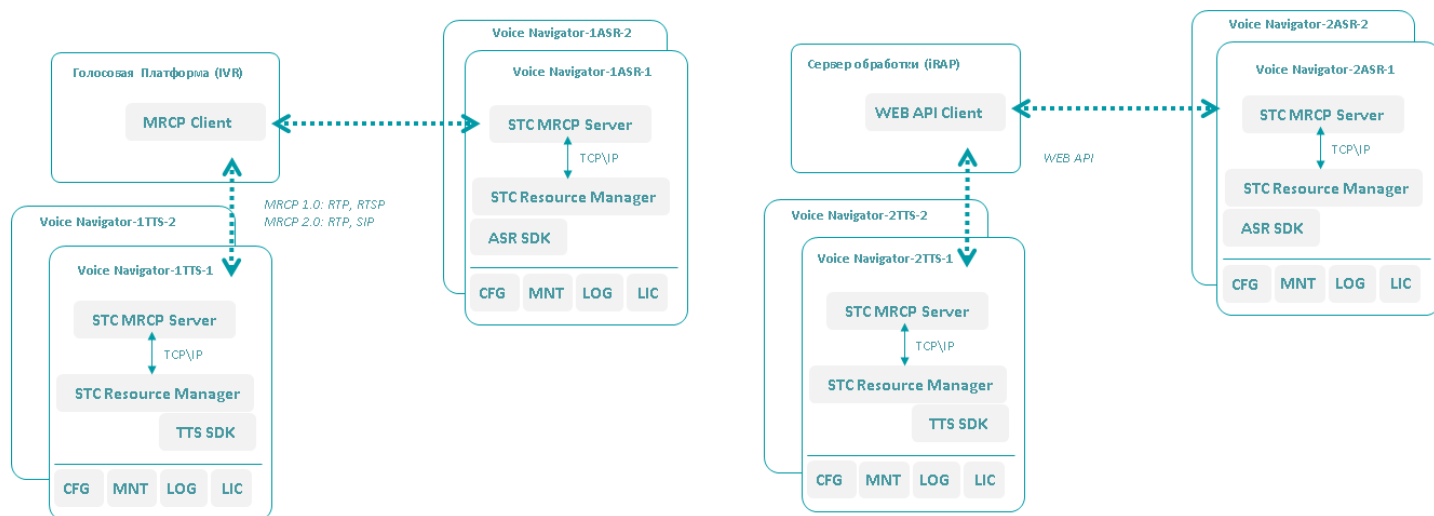


Рисунок 10 – Распределенное использование ресурсов VoiceNavigator. Вариант 3

Рекомендованная конфигурация серверов представлена в таблице:

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения логов-файлов)	vHDD2, iops	Примечание
Рекомендуемый физический сервер с виртуальными машинами: сервер типа Xeon E5-26XX (2CPU 2,4 Ghz x 8 core); 64 Gb RAM; RAID 2x600 Gb HDD 15K + 2x1Tb HDD 7K; double power										
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для MRCP	1	16	16	200	100	2000	70	До 50 одновременных распознавания на двух графах слитного распознавания

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для MRCP	1	8	10	200	100			До 25 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для MRCP	1	8	16	200	100			До 50 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для MRCP	1	4	8	200	100			До 25 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для Web API	1	16	24	200	100			До 40 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для Web API	1	10	16	200	100			До 20 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для Web API	1	8	16	200	100			До 50 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для Web API	1	4	8	200	100			До 25 одновременных синтезов
				74	114	1600		2000		

При большом масштабировании могут появиться трудности при планировании количества лицензий на ASR/TTS. Для решения этой проблемы предлагается выносить сервер лицензий на отдельный сервер (рис. 11).

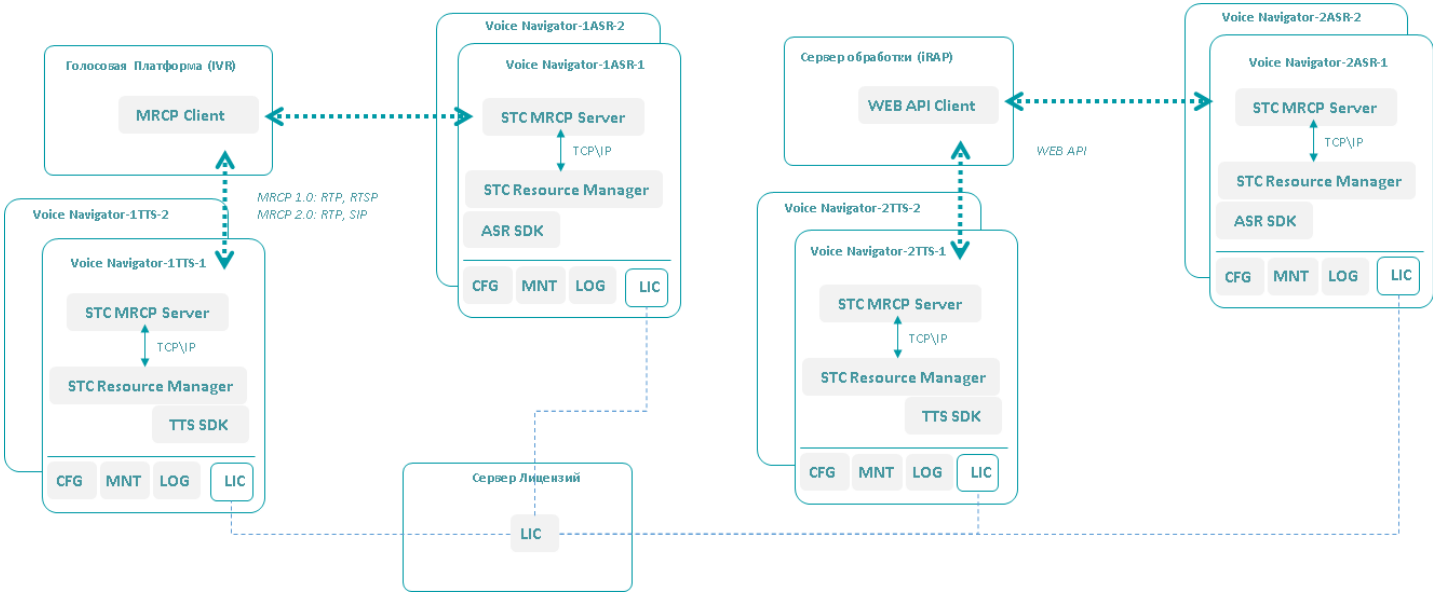


Рисунок 11– Распределенное использование ресурсов VoiceNavigator. Вариант 4

Рекомендованная конфигурация серверов представлена в таблице:

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для MRCP	1	16	16	200	100	2000	70	До 50 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для MRCP	1	8	10	200	100			До 25 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для MRCP	1	8	16	200	100			До 50 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для MRCP	1	4	8	200	100			До 25 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для Web API	1	16	24	200	100			До 40 одновременных распознавания на двух графах слитного распознавания

Системное ПО	Прикладное ПО	Назначение	Кол-во	vCPU	vRAM, Gb	vHDD1, Gb	vHDD1, iops	vHDD2, Gb (СХД для хранения лог-файлов)	vHDD2, iops	Примечание
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система распознавания речи для Web API	1	10	16	200	100			До 20 одновременных распознавания на двух графах слитного распознавания
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для Web API	1	8	16	200	100			До 50 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	VoiceNavigator 10.2	Система синтеза речи для Web API	1	4	8	200	100			До 25 одновременных синтезов
CentOS 7.2 x64 Windows 2012 x64	Sentinel	Сервер лицензий	1	2	2	40	100			
				76	116	1640		2000		

Для резервирования сервера лицензий используются следующие схемы:

1. Горячее резервирование. Сервер лицензий устанавливается рядом с основным. В конфигурациях клиентов лицензий прописываются основной и резервный сервера (рис. 12).

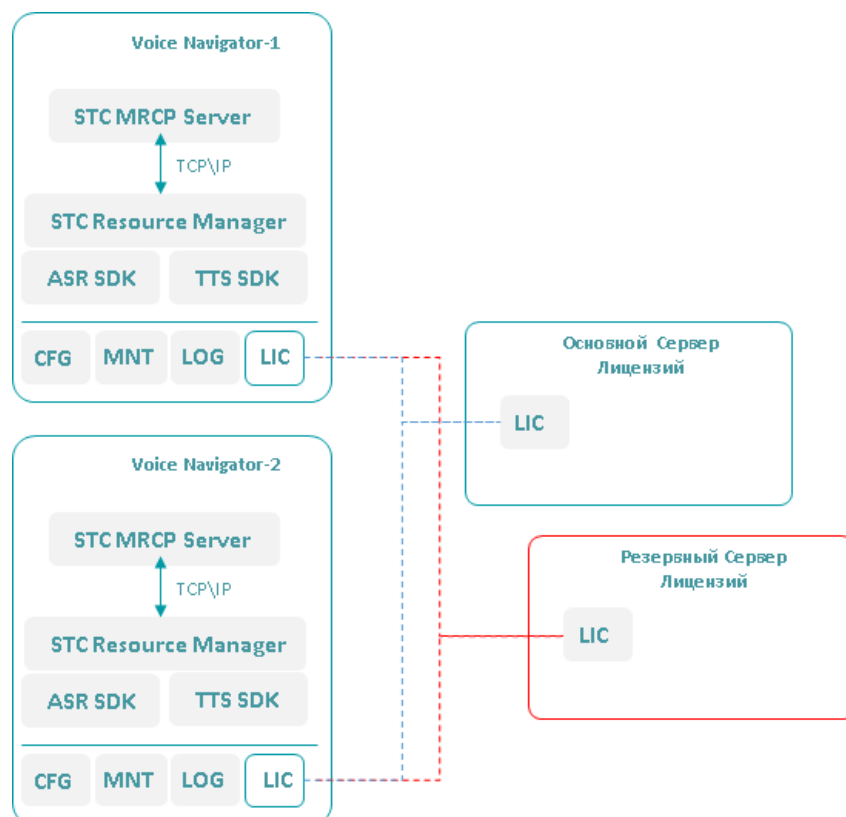


Рисунок 12 – Горячее резервирование. Вариант 1



Количество лицензий на основном сервере всегда должно соответствовать количеству лицензий на резервном сервере.

Вариантом горячего резервирования является наличие дублированной площадки, которая работает в режимах балансирования. Обращения одновременно направляются, но суммарное количество лицензий не может превышать количество лицензий на основном сервере (рис. 13).

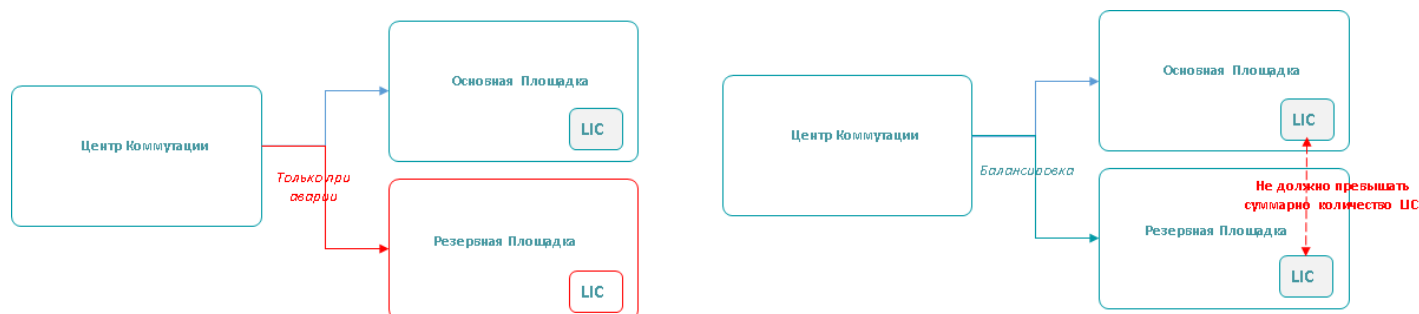


Рисунок 13 – Горячее резервирование. Вариант 2

- Холодное резервирование. Резервный сервер включается только после выключения основного (рис. 14).

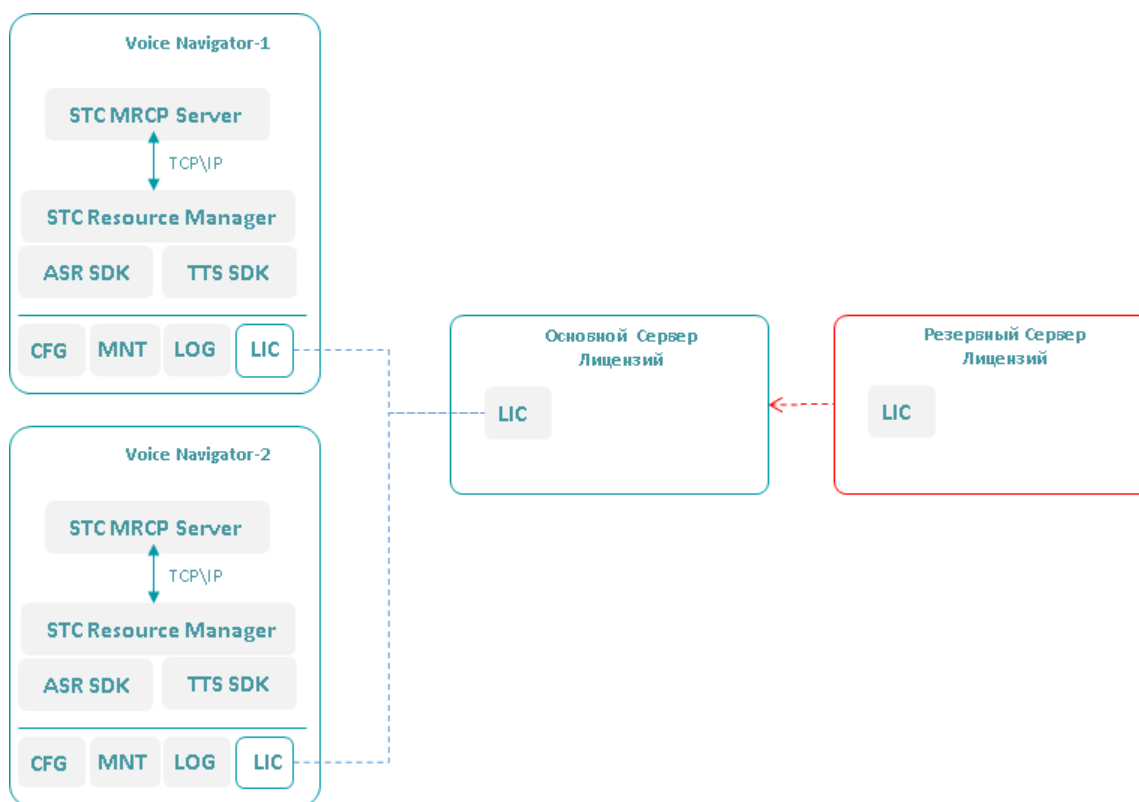


Рисунок 14 – Холодное резервирование

5.5 Защита программного обеспечения

Защита программного обеспечения системы голосового взаимодействия **VoiceNavigator** от несанкционированного копирования и использования осуществляется с помощью аппаратного или программного ключа защиты HASP.

На компьютеры, где будут устанавливаться ресурсы синтеза и распознавания, необходимо установить драйвер ключа защиты HASP.



Лицензионные ключи заказываются перед установкой ПО у менеджера по продажам.



Актуальная информация и драйвера защиты находится на сайте вендора <https://safenet-sentinel.ru/>.

5.5.1 Варианты ключей защиты

HASP-ключ определяет доступную заказчику функциональность и количество активных каналов синтеза и распознавания речи.



Если перед включением компьютера, на котором установлена система голосового взаимодействия **VoiceNavigator**, HASP-ключ не был установлен в порт USB или в процессе работы операционной системы был извлечён, работа системы **VoiceNavigator** блокируется.

Для возобновления работы системы необходимо убедиться в том, что ключ защиты установлен в порт USB компьютера, и запустить / перезапустить работу системной службы **STC Resource Server** для комплекса распознавания речи и/или для комплекса синтеза речи или произвести перезагрузку компьютера.

В зависимости от требований заказчика HASP-ключ может быть аппаратным (USB-брелок) или программным (предполагающим активацию продукта с помощью специальных процедур). Независимо от того, какой ключ был выбран (аппаратный или программный), ключи также бывают следующих типов (варианты ключей могут быть миксированы, т.е. один ключ может включать в себя несколько типов):

- аппаратный ключ предоставляется в виде USB-брелока. Использование возможно при условии доступности USB-порта для подключения ключа защиты. Привязка лицензий производится к конфигурации самого аппаратного ключа и не зависит от конфигурации сервера. Для корректной работы в системе виртуализации при использовании технологии V-Motion необходимо обеспечить переадресацию USB-порта средствами системы виртуализации. При неисправности сервера аппаратный ключ может быть физически переключен в другой сервер. Резервирование обеспечивается дублирующим ключем с идентичным набором лицензий, которые приобретаются со скидкой (зависит от варианта горячего или холодного резервирования). На работу ключа не влияют: перемещение виртуальных машин по технологии V-Motion, восстановление из резервной копии, смена доменного имени;
- программный ключ (с привязкой к конфигурации сервера) предоставляется в виде файла специальной конфигурации, активируемого на целевом сервере. Программный ключ привязывается к аппаратной конфигурации сервера. При изменении элемента конфигурации физического/виртуального сервера или перемещении виртуального сервера на другой хост, или восстановлении из резервной копии программный ключ деактивируется. Повторная активация программного ключа невозможна (только заказ нового). Смена доменного имени не влияет на работу ключа;
- программный ключ FQDN (с привязкой к имени сервера) предоставляется в виде файла специальной конфигурации, активируемого на целевом сервере. Программный ключ привязывается к доменному имени сервера. При изменении доменного имени сервера или появлении в сети сервера с идентичным именем программный ключ деактивируется. Повторная активация программного ключа невозможна (только заказ нового). На работу ключа не влияют:

перемещение по технологии V-Motion и восстановление из резервной копии. В виду ограничения в защите лицензий данный ключ распространяется ограничено;



Чтобы в Windows узнать доменное имя для лицензии FQDN, используйте оболочку **PowerShell** и через командную строку дайте команды:

```
$name = "$env:COMPUTERNAME.$env:USERDOMAIN"  
$name
```



Чтобы в Linux узнать доменное имя для лицензии FQDN, используйте команду **hostname -fqdn**. Параметр задаётся параметром **HOSTNAME** в файле **/etc/sysconfig/network** или предоставляется DHCP-сервером.

- локальный ключ устанавливается на тот же сервер, что и программное обеспечение;
- сетевой ключ устанавливается на любой сервер в составе комплекса. Данный ключ используется при построении распределённых систем, когда модули устанавливаются на нескольких серверах;



Для компьютера, выполняющего роль сервера лицензий, и виртуальных машин с установленными компонентами системы **VoiceNavigator** должны соблюдаться следующие условия:

- открыты 475, 1947 порт UDP и TCP;
 - разрешен транзит пакетов на широковещательные адреса.
- временный ключ, ограничивающий работу программного обеспечения временными параметрами. Время работы продукта может быть любым (1 год, 1 месяц и т. д.) и оговаривается индивидуально с заказчиком на основании договорных отношений;
 - демонстрационный (триальный) ключ. Это разновидность временного ключа защиты, но с дополнительными ограничениями: время работы продукта 90 дней, работа по 4 каналам, 2 голоса синтеза («Юлия» и «Владимир»). Для активации демонстрационной лицензий системы **VoiceNavigator** необходимо связаться с представителями компании ООО «ЦРТ»;
 - ключ с поддержкой виртуальных машин или без их поддержки;
 - пожелания по виду ключа предоставляются заказчиком.

5.5.2 Установка ПО ключа защиты в Windows

ПО ключа защиты устанавливается автоматически при установке **VoiceNavigator**.

Если у вас есть аппаратный ключ, то его необходимо заранее установить в свободный USB-порт и убедиться, что он подключен.

Для установки в ручном режиме (например, для установки на сервере лицензий) запустите исполняемый файл **HASPUserSetup.exe**, размещенный в подкаталоге **\Hasp\Driver**.

Установщик драйвера предложит подтвердить лицензионное соглашение и произведёт установку.

В некоторых случаях при предоставлении триального ключа может быть предоставлен специальный файл **haspdinst.exe**. Для его установки используйте команду **haspdinst.exe -i**, а для удаления – **haspdinst.exe -r**. Для вывода всех возможных команд запустите файл **haspdinst.exe**.

Запустите сервис **Sentinel LDK License Manager** (рис. 15):

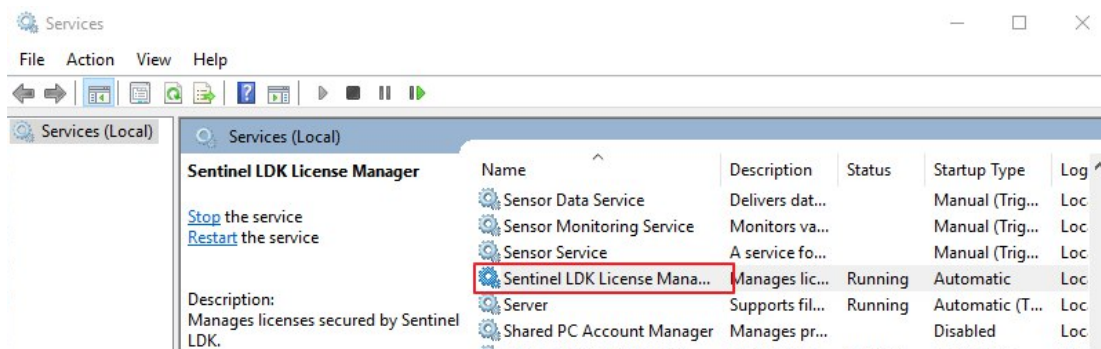


Рисунок 15 – Запуск сервиса Sentinel LDK License Manager

5.5.3 Установка ПО ключа защиты в Linux

ПО ключа защиты устанавливается автоматически при установке **VoiceNavigator**.

Если у вас есть аппаратный ключ, то его необходимо заранее установить в свободный USB-порт, и убедиться, что он подключён, проверив список устройств.

Для установки в ручном режиме (например, для установки на сервере лицензий) установите файл командой `rpm -Uvh aksusbd-7.63-2.i386.rpm`, размещённый в подкаталоге **\VoiceNavigator**.



Если есть проблемы с запуском **aksusbd** (после установки) и переустановка не помогает, выполните команды:

```
yum install -y initscripts
yum install -y compat-libstdc++-33.i686
rpm -Uvh aksusbd-7.63-2.i386.rpm
```

Для удалённого доступа к веб-интерфейсу администрирования ключами используйте настройки в файле `/etc/hasplm/hasplm.ini` (рис. 16):

```
[SERVER]
name = sn-test-perform-cetos7
idle_session_timeout_mins = 720
pagerefresh = 3
linesperpage = 20
accremote = 1
enablehaspc2v = 0
old_files_delete_days = 90

enabledetach = 0
reservedseats = 0
reservedpercent = 0
detachmaxdays = 14
commuter_delete_days = 7
disable_um = 0

requestlog = 0
loglocal = 0
logremote = 0
logadmin = 0
errorlog = 0
rotatelogs = 0
access_log_maxsize = 0 ;kB
error_log_maxsize = 0 ;kB
zip_logs_days = 0
delete_logs_days = 0
pidfile = 0
passacc = 0

accessfromremote = 1
accessstoremote = 1
bind_local_only = 0 ; 0=all adapters, 1=localhost only
```

Рисунок 16 – Настройки для удалённого доступа к веб-интерфейсу

Запустите или перезапустите сервис **aksusbd**:

```
service aksusbd start
service aksusbd restart
```

Проверьте статус работы сервиса **aksusbd**:

```
service aksusbd status
```

Проверьте, что сервис прослушивает порты:

```
netstat -natup | grep aksusbd
```

В некоторых случаях при предоставлении триального ключа может быть предоставлен специальный архив с файлами, представленными на рисунке 17.

```
bin
dinst
dunst
EMUrl.properties
haspvlib_71511.so
haspvlib_x86_64_71511.so
pkg
Unlocked_20190415_153213(2).v2c
```

Рисунок 17 – Архив файлов



При установке новых версий драйвера защиты название сервисов может изменяться с **aksusbd** на **hasplmd**.

Запустите файл **dinst** (рис. 18):

```
# ./dinst
Update .RPM on this system
Подготовка... ##### [100%]
Stopping aksusbd (via systemctl): [ OK ]
Обновление / установка...
  1:aksusbd-7.80-1 ##### [ 50%]
Created symlink from /etc/systemd/system/multi-user.target.wants/aksusbd.service to /etc/systemd/system/aksusbd.service.
Created symlink from /etc/systemd/system/multi-user.target.wants/hasplmd.service to /etc/systemd/system/hasplmd.service.
Очистка / удаление...
  2:aksusbd-7.63-2 ##### [100%]
Install VLIB
Install V2C
Install EMUrl
```

Рисунок 18 – Запуск файла dinst

Проверьте, чтобы порты были открыты (рис. 19):

```
# netstat -natup | grep hasplmd
tcp        0      0 0.0.0.0:1947          0.0.0.0:*            LISTEN      22133/hasplmd
tcp        0      0 192.168.20.221:1947  192.168.20.154:59641  ESTABLISHED 22133/hasplmd
tcp        0      0 127.0.0.1:1947       127.0.0.1:42100      ESTABLISHED 22133/hasplmd
udp        0      0 0.0.0.0:1947        0.0.0.0:*            22133/hasplmd
```

Рисунок 19 – Проверка портов

Для удаления ПО ключа защиты используйте файл **dunst**.

5.5.4 Сбор данных конфигурации сервера и активация программного ключа в Windows

При использовании программного ключа защиты в Windows необходимо провести активацию программного обеспечения.

Для этого запустите программу активации с помощью файла **RUS_COOEА.exe**, расположенного в подкаталоге **\\HASP\\Remote Update System** прилагаемого дистрибутивного диска.

В открывшемся окне (рис. 20) выберите вкладку **Collect Status Information** и нажмите кнопку **Collect Information**. Укажите путь к папке, где необходимо сохранить генерируемый файл. В поле для ввода имени файла укажите имя генерируемого файла и нажмите кнопку **Сохранить**.

Во вкладке **Collect Status Information** появится сообщение об успешном завершении операции, а в папке, указанной для сохранения, появится сгенерированный файл.

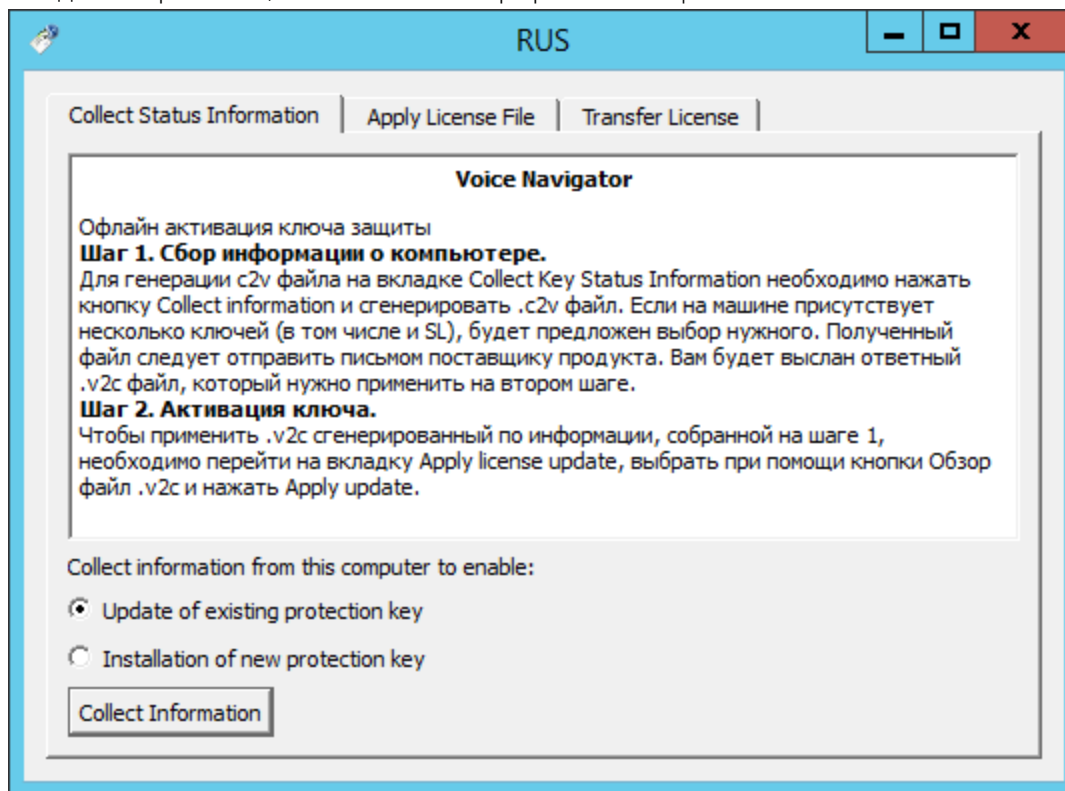


Рисунок 20 – Сбор информации в файл c2v

Отправьте сгенерированный файл по электронной почте на адрес менеджера по продажам. В ответ Вам будет отправлено письмо с файлом активации вида **<имя_файла>.v2c**. Сохраните полученный файл.

После получения файла активации повторно запустите программу активации. В открывшемся окне (рис. 20) выберите вкладку **Apply License Update**. При помощи кнопки **Обзор** укажите расположение файла активации. Во вкладке **Apply License Update** нажмите кнопку **Apply Update**. В случае успешной активации во вкладке **Apply License Update** появится сообщение об успешном завершении операции.

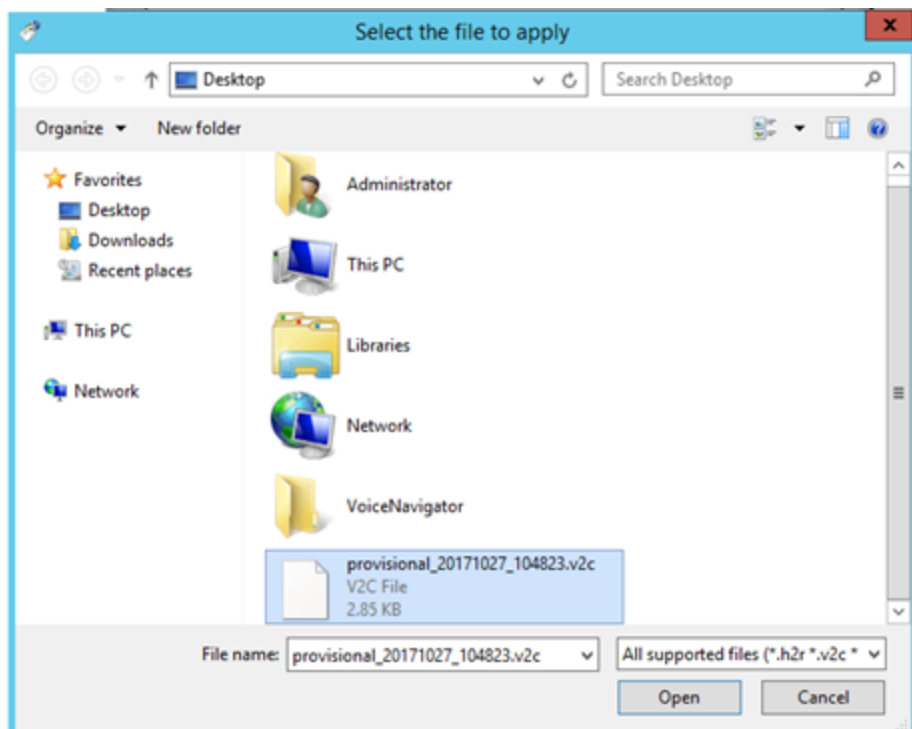


Рисунок 21 – Активация лицензии файлом v2c

5.5.5 Сбор данных конфигурации сервера в Linux для программного ключа

При использовании программного ключа защиты в Linux необходимо провести активацию программного обеспечения.

Для этого запустите программу **HaspTool**, расположенную в подкаталоге **VoiceNavigator**. При необходимости разрешите запуск командой `chmod +x`. Далее нажмите клавишу **f** (save fingerprint info) и введите желаемое имя файла (рис. 22):

```
# ./HaspTool

=====
----- STC Hasp Update Tool -----
=====

Please, press i-hasp information, f-save fingerprint info, c-save c2v file, u-update v2c file, e-exit.
f

Getting FingerPrint Info:Operation completed successfully

<?xml version="1.0" encoding="UTF-8" ?>
<hasp_info>
  <host_fingerprint type="SL-AdminNode" crc="3805267294">MXhJSSvEmDbM2BCyMYo2YCn/KKjGtGHBQ1yQ+8FnAJQ5BECgDQsK8W5Ahk2zAIAM7wgCWFCIMa7QqKs1JwB4G17BhkKMaQ
  AuDrMohYmJMGx2tXlyqKxwAhBtG5WLasRAg6BK480UGwNzYTCRKYUYAahEmDnSAJanAcAoRMSYRkUepKkSFHDNSj3</host_fingerprint>
</hasp_info>

Please enter c2v file name:
server_1.c2v

FingerPrint information stored into file: server_1.c2v

Please, press i-hasp information, f-save fingerprint info, c-save c2v file, u-update v2c file, e-exit.
e

Press ENTER to close application
```

Рисунок 22 – Активация программного обеспечения

По завершении файл с собранными данными появится в том же каталоге.

Отправьте сгенерированный файл по электронной почте на адрес менеджера по продажам. В ответ будет прислано письмо с файлами активации вида:

- файл **<имя_файла>.v2c** – активированный ключ. Сохраните файл;
- файл вида **haspvlib_*.so** – библиотека вендора, с помощью которой ключ активировался.

Положите файл **haspvlib_*.so** (библиотека вендора) в каталог **/var/hasplm/**.

5.5.6 Активация программного лицензионного ключа в Windows и Linux через веб-интерфейс

Для активации программного лицензионного ключа через файл ***.v2c** можно использовать веб-интерфейс менеджера лицензий.

Перед использованием убедитесь, что интерфейс доступен, а хост с менеджером лицензий находится в доверенной зоне браузера. При необходимости произведите дополнительные настройки, например для локального хоста <http://127.0.0.1:1947> (рис. 23).

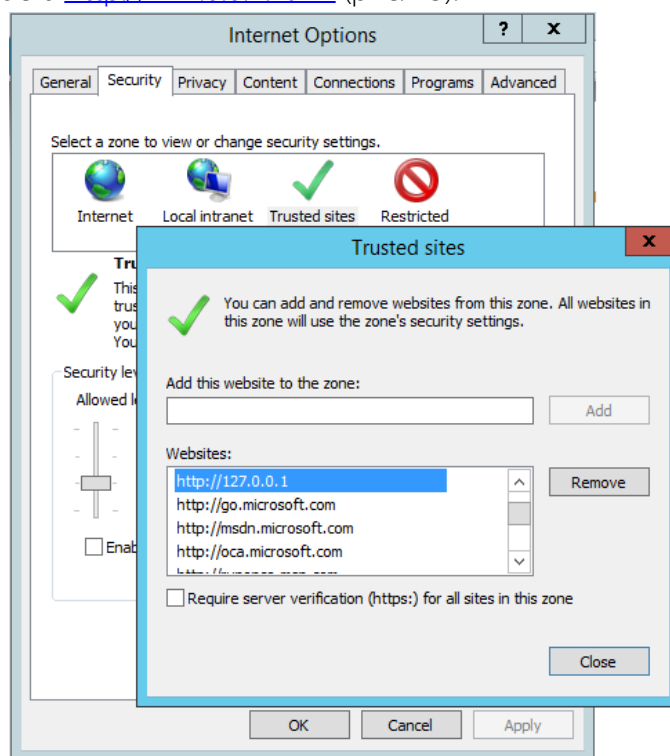


Рисунок 23 – Установка дополнительных настроек

Зайдите в веб-интерфейс менеджера лицензий http://имя_хоста:1947 (рис. 24).

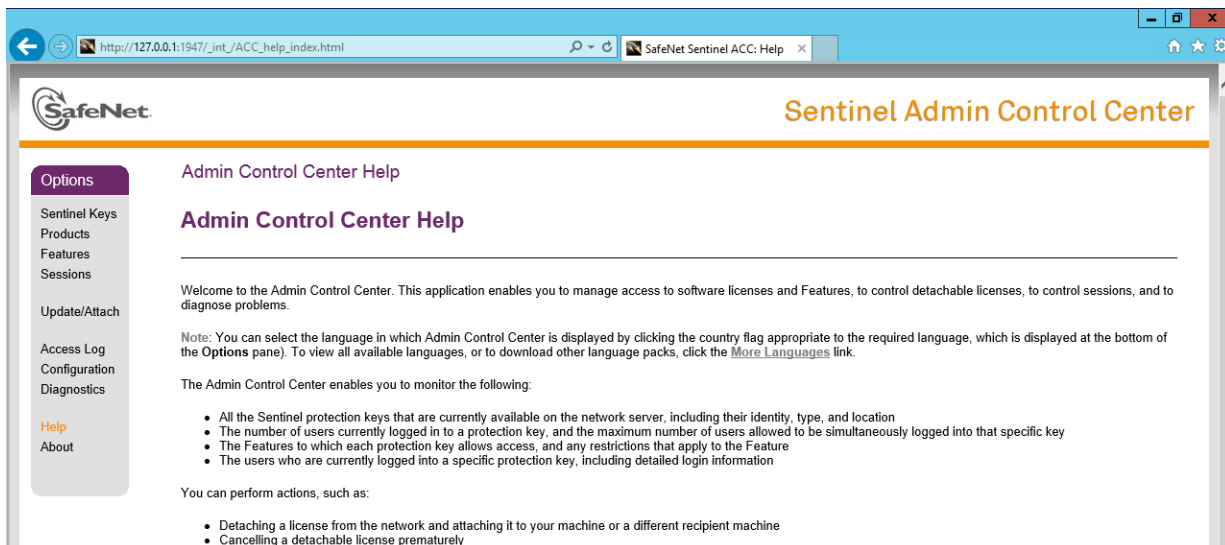


Рисунок 24 – Веб-интерфейс менеджера лицензий

Зайдите в меню **Update/Attach**, загрузите файл (рис. 25):

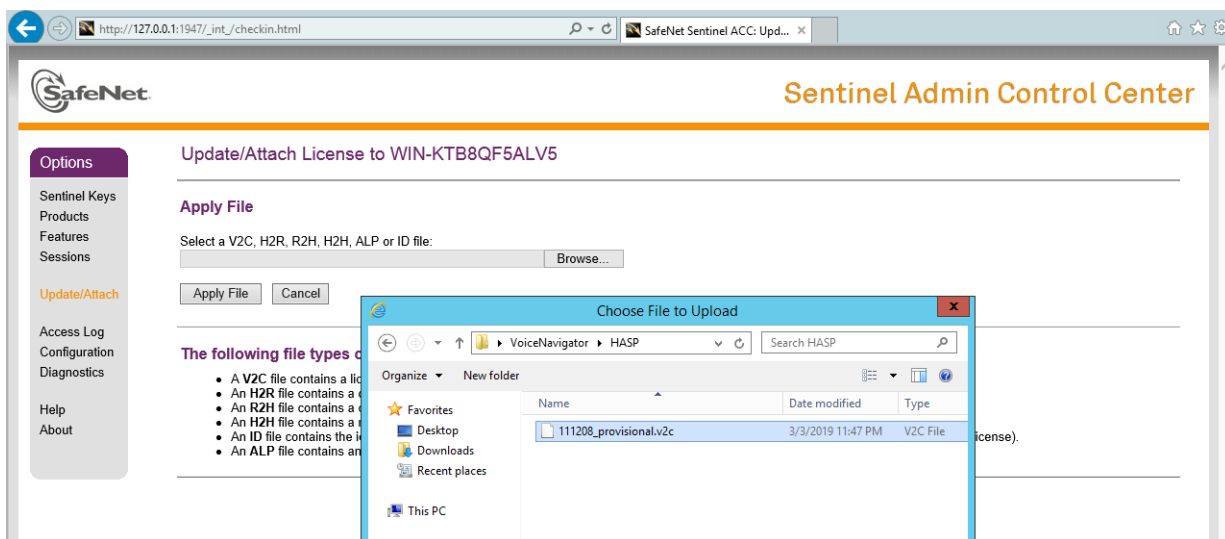


Рисунок 25 – Веб-интерфейс менеджера лицензий

Проверьте, что ключ установлен в меню **SentinelKeys**:

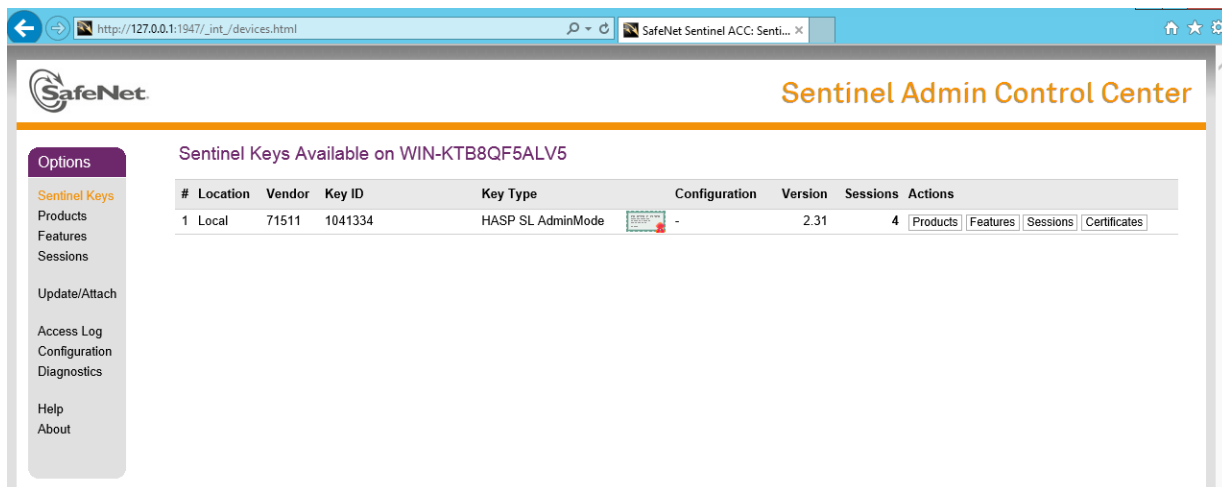


Рисунок 26 – Проверка ключа

5.5.7 Отключение лицензионного ключа

В некоторых случаях, например при замене действующего программного триального HASP-ключа на другой, необходимо отключить старый ключ.



При установке нескольких лицензионных ключей приоритетно работает только первый активированный. При установке несколько ключей лицензии не суммируются.

Через веб-интерфейс определите название удаляемого лицензионного ключа.

В Windows для этого необходимо удалить старый лицензионный файл из каталога **C:\Program Files (x86)\Common Files\SafeNet Sentinel\Sentinel LDK\installed\71511\<KeyId>_<type>.v2c**.

В Linux необходимо удалить старый лицензионный файл из каталога **/var/hasplm/installed/71511/**.

После удаления перезагрузите сервис защиты.

Через веб-интерфейс убедитесь, что старый лицензионный файл был удалён. Произведите установку нового лицензионного файла.

5.5.8 Проверка активации лицензий

Проверьте установленные компоненты в списке программ.

Убедитесь, что сервер лицензий видит все лицензии (рис. 27).

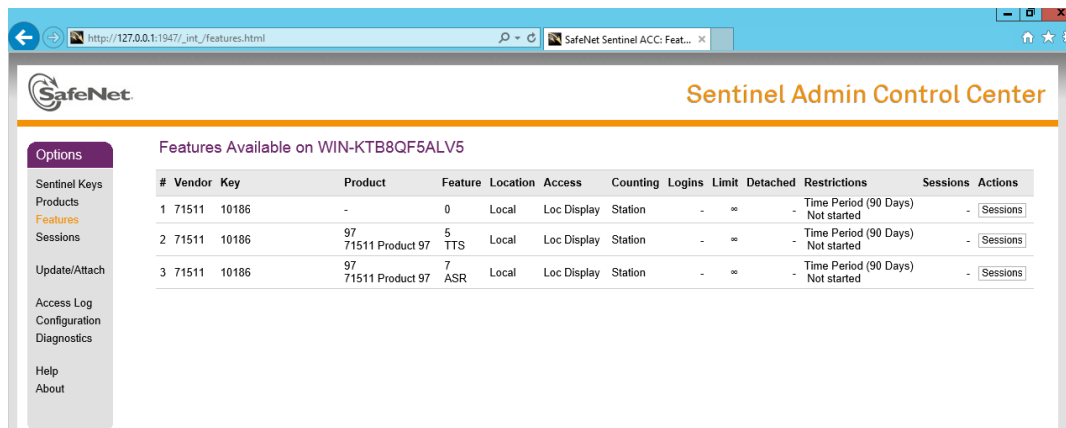


Рисунок 27 – Проверка лицензий. Часть 1

Лицензии активированы по количеству каналов именно с целевого сервера (рис. 28).

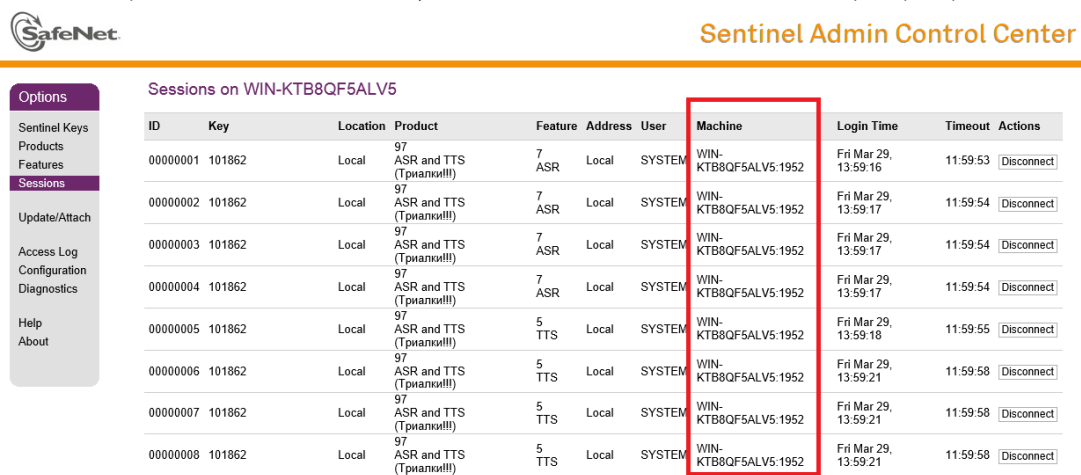


Рисунок 28 – Проверка лицензий. Часть 2

Файловым редактором (например, **Far** позволяет автоматически пересчитывать конец файла) можно смотреть в режиме онлайн количество свободных компонентов в Windows (рис. 29):

```

C:\Program Files\Speech Technology Center\Resource Server\Logs\stats.log      t 1252      70368 Col  0   100%  2:12 F
class ResourceServer::ArtifactStorage::ServiceLink 3/3/0
class ResourceServer::Proxy::ArenaPool 2/2/0
struct ResourceServer::Transport::Message 1/486/485
struct ResourceServer::Proxy::Message 4/245/241
class ResourceServer::ArtifactStorage::StdioFile 0/1206/1206
struct ResourceServer::Model::Message 0/268/268
class ResourceServer::Transport::Link 1/1/0
class ResourceServer::Proxy::Link 1/1/0
class ResourceServer::Proxy::ArenaMessageAllocator 10/10/0
struct ResourceServer::Proxy::Arena 2/2/0
class ResourceServer::Proxy::Formatter 1/1/0
class ResourceServer::Proxy::Stream 1/1/0
class ResourceServer::Service::AsrControl 4/4/0
struct ResourceServer::Media::Message 8/8/0
struct ResourceServer::Service::Message 60/60/0
class ResourceServer::Service::TtsControl 4/4/0
29.03.2019 14:12:36
occupiedAsrEngines=0
freeAsrEngines=4
successfulAsrSessionDuration=0
successfulAsrSessionDurationMin=0
successfulAsrSessionDurationMax=0
asrErrors=0
matchRecognitions=0
noMatchRecognitions=0
noInputRecognitions=0
asrSessions=0
asrRequestsPerSecond=0
mediaDataPerAsrSession=0
mediaDataPerAsrSessionMin=0
mediaDataPerAsrSessionMax=0
occupiedOfflineAsrEngines=0
freeOfflineAsrEngines=0
occupiedTtsEngines=0
freeTtsEngines=4
successfulTtsSessionDuration=0
successfulTtsSessionDurationMin=0
successfulTtsSessionDurationMax=0
ttsErrors=0
ttsSessions=0
ttsRequestsPerSecond=0
mediaDataPerTtsSession=0
mediaDataPerTtsSessionMin=0
mediaDataPerTtsSessionMax=0
```

Рисунок 29 – Просмотр свободных компонентов Windows

В Linux можно просмотреть редактором **mc** (рис. 30).

```

opt/RS/Logs/stats.log 107954/105K 100%
ccupiedOfflineAsrEngines=0
freeOfflineAsrEngines=0
ccupiedTtsEngines=0
freeTtsEngines=4
successfulTtsSessionDuration=0
successfulTtsSessionDurationMin=0
successfulTtsSessionDurationMax=0
tsErrors=0
tsSessions=0
tsRequestsPerSecond=0
mediaDataPerTtsSession=0
mediaDataPerTtsSessionMin=0
mediaDataPerTtsSessionMax=0

alive/created/destroyed
I14ResourceServer15ArtifactStorage11ServiceLinkE 3/3/0
I14ResourceServer5Proxy9ArenaPoolE 2/2/0
I14ResourceServer9Transport7MessageE 1/518/517
I14ResourceServer5Proxy7MessageE 4/261/257
I14ResourceServer9Transport4LinkE 1/1/0
I14ResourceServer5Proxy4LinkE 1/1/0
I14ResourceServer5Proxy21ArenaMessageAllocatorE 10/10/0
I14ResourceServer5Proxy5ArenaE 2/2/0
I14ResourceServer5Proxy9FormatterE 1/1/0
I14ResourceServer5Proxy6StreamE 1/1/0
I14ResourceServer7Service10AsrControlE 4/4/0
I14ResourceServer5Media7MessageE 8/8/0
I14ResourceServer7Service7MessageE 60/60/0
I14ResourceServer15ArtifactStorage9StdioFileE 0/1284/1284
I14ResourceServer5Model7MessageE 0/286/286
I14ResourceServer7Service10TtsControlE 4/4/0
6.04.2019 14:55:31
ccupiedAsrEngines=0
freeAsrEngines=4
successfulAsrSessionDuration=0
successfulAsrSessionDurationMin=0
successfulAsrSessionDurationMax=0
asrErrors=0
matchRecognitions=0
noMatchRecognitions=0
noInputRecognitions=0
asrSessions=0
asrRequestsPerSecond=0
mediaDataPerAsrSession=0
mediaDataPerAsrSessionMin=0
mediaDataPerAsrSessionMax=0
ccupiedOfflineAsrEngines=0
freeOfflineAsrEngines=0
ccupiedTtsEngines=0
freeTtsEngines=4
successfulTtsSessionDuration=0
successfulTtsSessionDurationMin=0
successfulTtsSessionDurationMax=0
tsErrors=0
tsSessions=0
tsRequestsPerSecond=0
mediaDataPerTtsSession=0
mediaDataPerTtsSessionMin=0
mediaDataPerTtsSessionMax=0

```

Рисунок 30 – Просмотр свободных компонентов Linux

6. ГРАММАТИКИ РАСПОЗНАВАНИЯ РЕЧИ

6.1 Распознавание по закрытым грамматикам

Распознавание по закрытым грамматикам – это распознавание одного слова (голосовой команды) из списка слов (базы). Понятие «закрытые грамматики» означает, что в систему заложена определённая конечная база слов, в которой система будет искать произнесённое абонентом слово или выражение. В этом случае система должна поставить вопрос абоненту так, чтобы получить однозначный ответ, состоящий из одного слова.

Пример:

Вопрос системы: «Какой день недели вас интересует?»

Ответ абонента: «Вторник».

В этом примере вопрос поставлен так, что система ожидает совершенно определённый ответ от абонента. База слов в приведённом примере может состоять из следующих вариантов ответа: «понедельник», «вторник», «среда», «четверг», «пятница», «суббота», «воскресенье». Следует предусмотреть и заложить в базу следующие варианты ответов: «не знаю», «все равно», «любой» и т.д. Ответы абонента должны быть предусмотрены и обработаны системой отдельно, согласно заранее заложенным сценариям диалога.

Компоненту распознавания необходимо передать файл с набором закрытых грамматик. Для этого в качестве стандарта оформления файла закрытых грамматик используется стандарт **SRGS** (англ. Speech Recognition Grammar Specification).

Для получения достоверного ответа можно дополнительно задать вопрос (опция настраивается отдельно), например: «Вы выбрали Тариф-1, правильно?».

Логика построения закрытых грамматик представлена на рисунке 31.

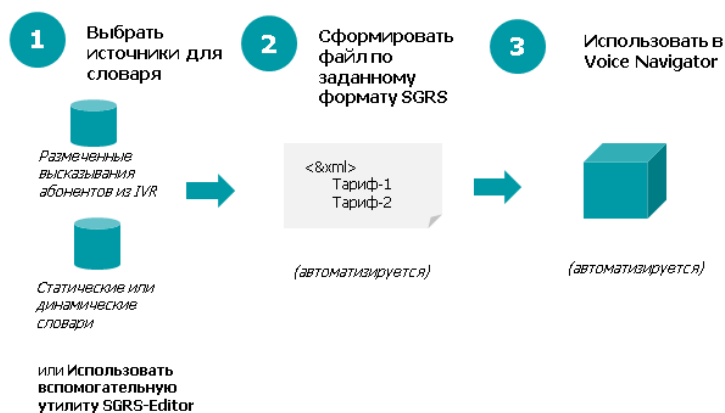


Рисунок 31 – Построение открытых грамматик

Формирование закрытых грамматик можно осуществлять «на лету» по заданному формату в текстовом файле. Далее представлен пример закрытой грамматики **SRGS**.

```
<?xml version="1.0" encoding="UTF-8"?>
<grammar type='application/srgs+xml'
version='1.0'
mode='voice'
root='main'
xml:lang='ru-RU'
xmlns='http://www.w3.org/2001/06/grammar'>
```

```

tag-format='semantics/1.0'>

  <rule id="main" scope="public">
<item repeat='0-1'><ruleref uri="#prefix"/></item>
<one-of>
<item><ruleref uri="#balance"/><tag>out.cxtag='balance'</tag></item>
<item><rulerefuri="#min_payment"/><tag>out.cxtag='min_payment'</tag></item>
<item><ruleref uri="#arrears"/><tag>out.cxtag='arrears'</tag></item>
</one-of>
  </rule>

  <rule id="balance" scope="public">
<one-of>
<item>баланс</item>
<item>по балансу</item>
<item>о балансе</item>
<item>текущий баланс</item>
<item>текущий остаток</item>
<item>остаток на карте</item>
<item>доступный остаток</item>
<item>сколько у меня осталось</item>
<item>остаток на счёте</item>
<item>сколько на карточке</item>
<item>денег на карточке</item>
<item>остаток по карте</item>
<item>уточнить остаток</item>
<item>остаток денег</item>
<item>доступный остаток</item>
<item>счёт на карточке</item>
</one-of>
  </rule>

  <rule id="min_payment" scope="public">
<one-of>
<item>минимальный платёж</item>
<item>по минимальному платежу</item>
<item>о минимальном платеже</item>
<item>оплата</item>
<item>оплата кредита</item>
<item>сколько надо заплатить</item>
<item>меня интересует оплата кредита</item>
<item>минимальный платёж</item>
<item>размер минимального платежа</item>
<item>остаток минимального платежа</item>
<item>сумму минимального платежа</item>
<item>какого числа мне платить</item>
<item>какую сумму платить</item>
<item>мне какую сумму внести</item>
<item>какую сумму мне внести</item>
<item>платёж когда</item>
<item>платёж у меня когда</item>
<item>когда у меня платёж</item>
</one-of>
  </rule>

  <rule id="arrears" scope="public">
<one-of>
<item>общая задолженность</item>
<item>по общей задолженности</item>
<item>об общей задолженности</item>
<item>сколько полностью нужно оплатить</item>
<item>полный расчёт</item>

```

```

<item>общая сумма задолженности</item>
<item>полностью оплатить кредит</item>
<item>какой долг</item>
</one-of>
</rule>

<rule id="prefix" scope="public">
<one-of>
<item>скажите</item>
<item>скажите пожалуйста</item>
<item>а вот</item>
<item>мне узнать</item>
<item>мне бы узнать</item>
<item>хочу узнать</item>
<item>нужно узнать</item>
<item>я хочу узнать</item>
<item>мне нужно узнать</item>
</one-of>
</rule>
</grammar>

```



Описание общего стандарта на грамматики **SRGS** доступно по адресу <http://www.w3.org/TR/speech-grammar/>.

Файл грамматики **SRGS** должен использоваться в кодировке UTF-8.



Не рекомендуется создавать закрытые грамматики с более чем 1000 записей.

Файлы закрытых грамматик занимают небольшой объем на диске и обычно передаются через **MRCP**.

6.2 Распознавание по открытым грамматикам

Распознавание по открытым грамматикам – это распознавание всей произнесённой абонентом фразы целиком. Это позволяет системе задать абоненту открытый вопрос и получить ответ, сформулированный в свободной форме. Понятие «открытые грамматики» означает, что система ожидает услышать от абонента не конкретное слово/команду, а все смысловое предложение целиком, в котором систему будет интересовать каждое слово.

Пример:

Вопрос системы: «Что вас интересует?»

Ответ абонента: «Какие документы нужны для кредита?».

В этом примере распознаётся каждое слово в ответе абонента и выявляется общий смысл сказанного. На основании распознанных ключевых слов и понятий в предложении формируется запрос в базу данных и собирается ответ абоненту, а именно предоставляется справочная информация.

Логика построения графа для распознавания слитной речи представлена на рисунке 32.

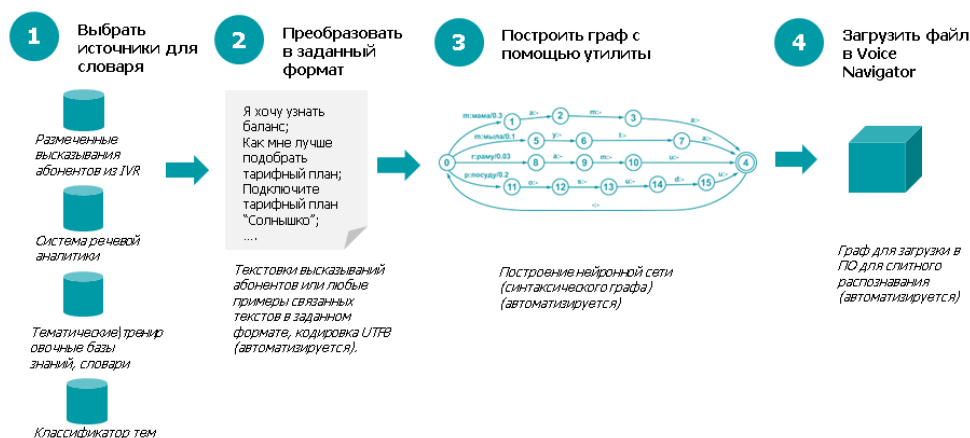


Рисунок 32 – Построение открытых грамматик

Распознавание слитной речи даёт системе больше возможностей для автоматизации процесса диалога с абонентом, а также возрастает скорость и удобство пользования системой со стороны абонента. Однако такие системы сложнее в реализации, так как помимо распознавания требуется классификация (приведение фразы к односложной теме, на основе которой надо принять решение в сценарии).

Граф слитного распознавания представляет собой набор файлов, представленных на рисунке 33.

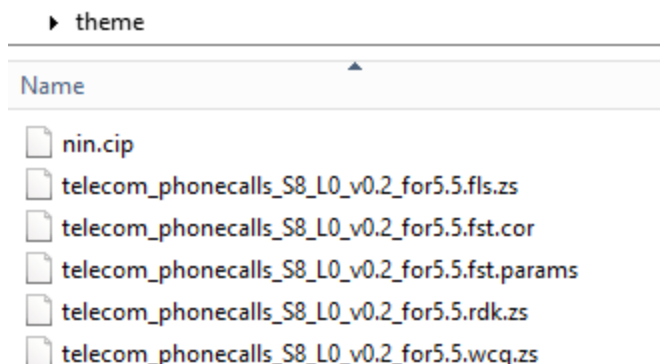


Рисунок 33 – Граф слитного распознавания



Набор файлов и структура могут отличаться.

Пример потока вызова в общем виде представлен на рисунке 34.

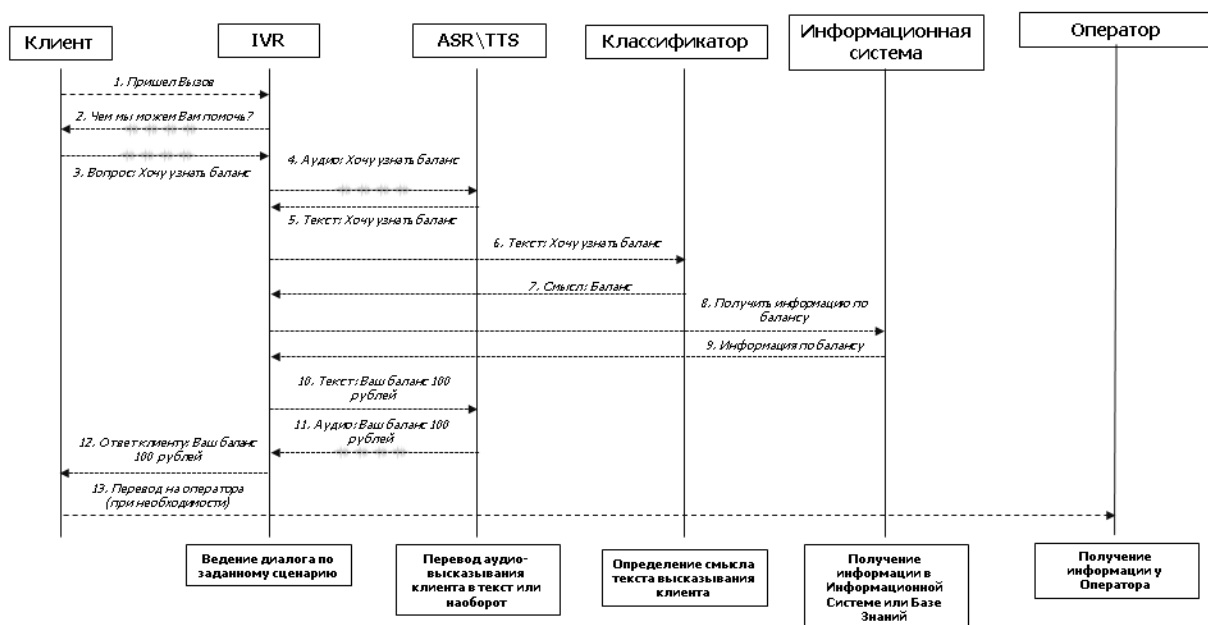


Рисунок 34 – Поток вызова

6.3 Пакеты для слитного распознавания

Пакет представляет собой набор файлов, необходимых для запуска определённого сценария в **ASR SDK**. В состав пакета обязательно входит акустическая модель с набором параметров (профиль), а также опционально тема, а именно граф распознавания с настройками декодера. Состав пакета может быть расширен моделями для дополнительных компонентов: предиктор, препинатор, грамматики свёртки и т.д.



Для возможности создания графов распознавания под каждую версию пакета существуют свои версии технологических утилит, поставляемых по запросу с инструкцией разработчика грамматик.

Все пакеты делятся на три части:

- Серийные пакеты являются основными сценариями и языками (обязательны к релизу).
- Специальные пакеты под конкретную задачу и/или заказчика – разовые сборки, которые могут не выходить со следующими релизами **ASR SDK**.
- Пакеты для тестирования, которые собираются для нужд тестирования. В такие пакеты, к примеру, могут включаться старые модели для регрессионных тестов. В релиз такие пакеты не включаются.

Ниже приведены обозначения сценариев, для которых выпускаются пакеты для работы с **VoiceNavigator**:

- **IVR** – распознавание команд по таймаутам в телефонном канале, но бывают и исключения (основной пакет).
- **OffRec** – многопроходное распознавание речи по большому графу в офлайн режиме (дополнительный пакет).
- **Respeaking** – распознавание речи в режиме «диктовка» (дополнительный пакет).



Дополнительные пакеты устанавливаются по запросу заказчика.

6.4 Обучение системы распознавания слитной речи

При поставке ПО **VoiceNavigator** по запросу заказчика предоставляется готовый пакет с графом слитного распознавания для распознавания общеупотребительного набора слов и фраз по заданной тематике. Однако из-за уникальности каждой компании и бизнес-процессов набор слов и фраз может не охватывать все варианты использования. Для этого необходимо провести дополнительное обучение существующей модели графа или создать собственный граф, который будет удалять неиспользуемые формы, что значительно повысит качество распознавания.

Основой распознавания слитной речи является статистика, поэтому любой проект по внедрению распознавания слитной речи начинается со сбора статистики.

Сбор информации начинается с прослушивания и анализа реальных записей разговоров абонентов. Затем происходит транскрибирование собранных записей. Каждая запись прослушивается специалистом, который переводит их в текст. Транскрибированные фразы не должны содержать пунктуационные знаки и специальные символы. Все слова и аббревиатуры должны быть прописаны полностью так, как их произносит клиент. Данная работа занимает достаточно много времени, но не требует специфических знаний, поэтому к работе обычно одновременно привлекается сразу много сотрудников.

Транскрибированные файлы служат для построения тренировочного файла (словарь слов и выражений, конфигурационные параметры), который представляет собой xml-документ. Чем объёмней будет тренировочный файл по предметной области, тем более качественным будет распознавание. Тренировочный файл позволяет сформировать языковую модель (SLM), которая является основой распознавания слитной речи. Для этого тренировочный файл загружается в специальную утилиту **ASR Конструктор**, которая строит языковую модель. Затем языковая модель загружается в программное обеспечение **VoiceNavigator**.

Распознавание речи представляет собой технологию, позволяющую переводить звучащую речь в текстовое представление (последовательность слов). При этом последовательность слов, получаемая в процессе распознавания, может отличаться от последовательности слов, которые были произнесены из-за ошибок, совершаемых системой распознавания. Причинами ошибок могут быть следующие факторы:

- наличие в речи слов, отсутствующих в словаре системы распознавания (англ. out-of-vocabulary words, OOV);
- особенности стиля речи (высокий/низкий темп, «проглатывание» окончаний и т.п.), дикции (картавость, шепелявость и т.п.) и произношения (растягивание гласных, оканье и т.п.);
- омонимы (слова, имеющие разное написание, но одинаковое произношение);
- внешние помехи и окружающий шум, в том числе фоновая или одновременная речь;
- несоответствие тематики разговора модели языка, используемой системой распознавания.

При обучении могут быть достигнуты следующие показатели:

№	Количество транскрибированных аудиоданных в часах	Качество целевого распознавания речи
1	100	Отлично, выше 87%
2	50	Хорошо, до 76-86%
3	25	Удовлетворительно, до 66-75%

4	10	Допустимо, до 55-65%
---	----	----------------------



В некоторых случаях, например, со специфической акустической обстановкой или другого языка, требуется переобучение моделей вендором с количеством аудиоданных более 1000 часов.

Для задач NLP (понимания естественной речи) в распознанном списке слов необходимо выявить тему обращения абонента (классификация) и выделить слотовую информацию (сущность).



Дополнительная информация может быть получена только в ходе обучения.

6.5 Установка и обновление графа слитного распознавания

Граф для слитного распознавания содержит миллионы словоформ и занимает большой объём на жёстком диске, поэтому тема должна быть предварительно размещена в специальном каталоге Resource-сервера, путь к которому настраивается в конфигурационных файлах Resource-сервера.

6.6 Установка и обновление графа слитного распознавания в Windows

Создайте свой или получите граф слитного распознавания и, при необходимости, разархивируйте его.

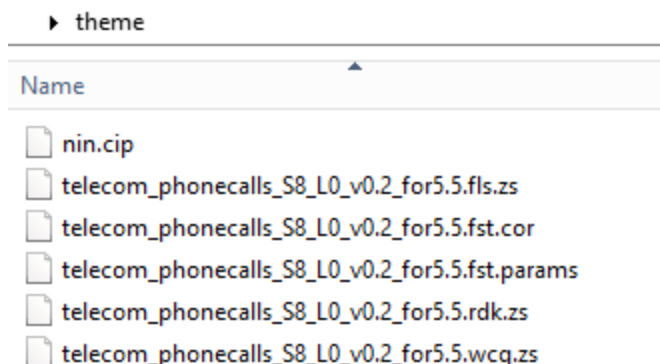


Рисунок 35 – Созданный граф слитного распознавания

Установите или обновите граф в каталоге **%allusersprofile%\Speech Technology Center\ASR_Themes\<theme_name>** (рис. 36).

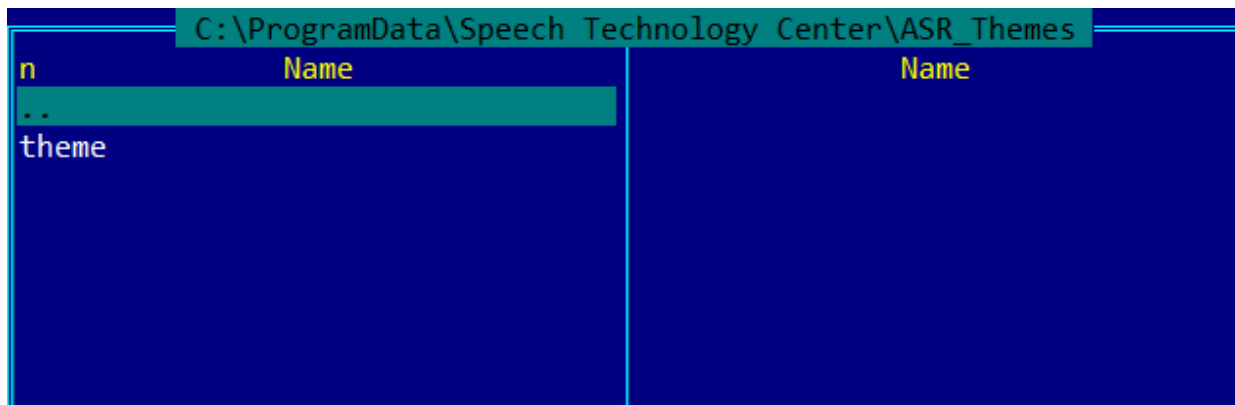


Рисунок 36 – Создание/обновление графа

Чтобы принять изменения (активные сессии не прерываются), создайте файл **reload.txt** (рис.37).

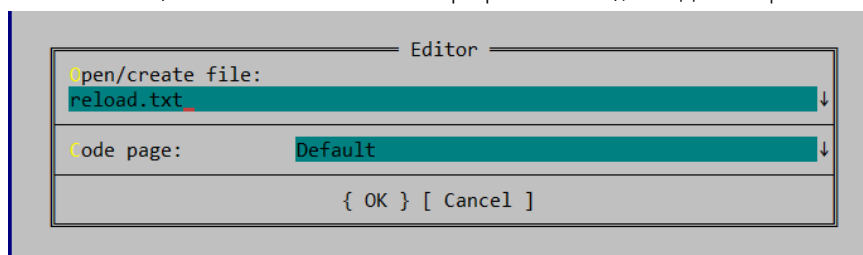


Рисунок 37 – Создание файла

Дождитесь, пока файл исчезнет. Процесс может занять некоторое время. Запустите log-файл **STC_ResourceServer** (рис. 38).

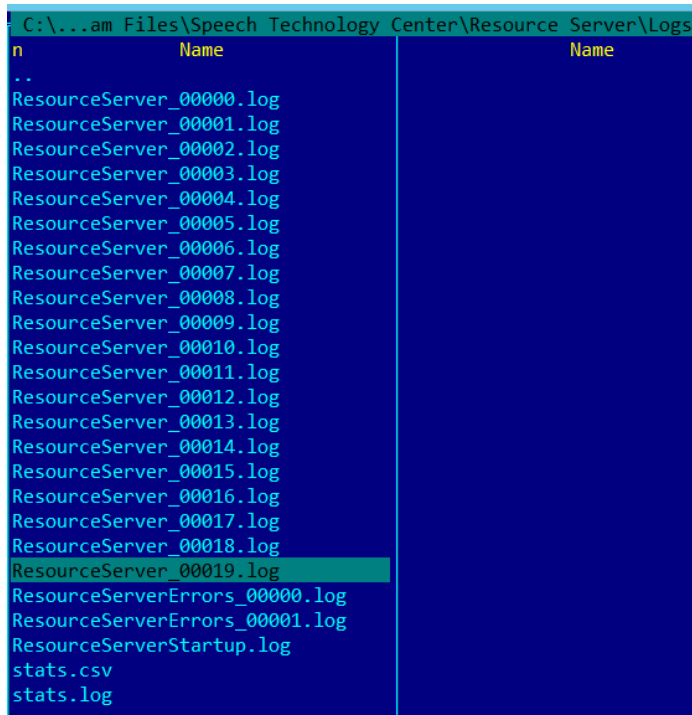


Рисунок 38 – Выбор файла

Проверьте последние записи log-файла **STC_ResourceServer** (рис. 39).

```

C:\Program Files\Speech Technology Center\Resource Server\Logs\ResourceServer_00019.log  t 1252 754877 Col 0 100% 1:51 P
04.04.19 13:50:26.570345 0x00000494 0x0000056c INFO: Runner-1: Responded in 21.959 ms
04.04.19 13:50:27.068509 0x00000494 0x0000056c INFO: TcpLink-105: Sync sending 12 bytes to 0.0.0.0:0
04.04.19 13:50:27.070509 0x00000494 0x0000056c INFO: TcpLink-105: Received 12 bytes
04.04.19 13:50:27.070509 0x00000494 0x0000056c INFO: ProxyLink-105: PingRsp 0 received
04.04.19 13:50:29.567745 0x00000494 0x0000056c INFO: Runner-1: Responded in 18.007 ms
04.04.19 13:50:32.070188 0x00000494 0x0000056c INFO: TcpLink-105: Sync sending 12 bytes to 0.0.0.0:0
04.04.19 13:50:32.072184 0x00000494 0x0000056c INFO: TcpLink-105: Received 12 bytes
04.04.19 13:50:32.072184 0x00000494 0x0000056c INFO: ProxyLink-105: PingRsp 0 received
04.04.19 13:50:32.565798 0x00000494 0x0000056c INFO: Runner-1: Responded in 16 ms
04.04.19 13:50:34.904607 0x00000494 0x00000530 INFO: AsrGrammar: 0000001E3A99E2C0 created: 2 objects
04.04.19 13:50:34.904607 0x00000494 0x00000530 INFO: Created grammar from source 0000001E3A8642C0
04.04.19 13:50:34.904607 0x00000494 0x00000530 INFO: GrammarManager-45: Theme C:\ProgramData\Speech Technology Center\ASR_Themes\theme loaded
04.04.19 13:50:35.564649 0x00000494 0x0000056c INFO: Runner-1: Responded in 13.996 ms
04.04.19 13:50:35.916668 0x00000494 0x0000056c INFO: TcpLink-105: Received 15 bytes

```

Рисунок 39 – Проверка записей log-файла

Убедитесь, что тема успешно загрузилась.



При необходимости повторите операцию на других серверах.

6.7 Установка и обновление графа слитного распознавания в Linux

Создайте свой или получите граф слитного распознавания, и, при необходимости, разархивируйте его (рис. 40).

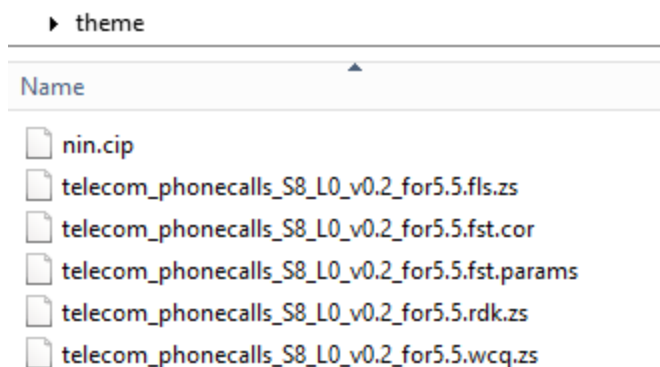


Рисунок 40 – Созданный граф слитного распознавания

Установите или обновите граф в каталоге `/opt/ASR_Themes/<theme_name>` (рис. 41).

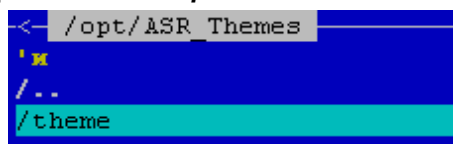


Рисунок 41 – Создание/обновление графа

Чтобы принять изменения (активные сессии не прерываются), создайте файлы **reload.txt** и **touch reload.txt**.

Дождитесь пока файл исчезнет. Процесс может занять некоторое время.

Запустите log-файл **ResourceServer_00003.log** (рис. 42).

```

<- /opt/RS/Logs
'Имя
/..
ResourceServerErrors_00000.log
ResourceServerStartup.log
ResourceServer_00000.log
ResourceServer_00001.log
ResourceServer_00002.log
ResourceServer_00003.log
empty.txt
stats.csv
stats.log

```

Рисунок 42 – Выбор файла

Проверьте последние записи log-файла (рис. 43).

```

mc [root@sn-test-perform-cetos7] /opt/RS/Logs
/opt/RS/Logs/ResourceServer_00003.log 2496844/2439K
18.04.19 17:08:32.818921 00000x2a56 0x00007f48b47fe700 ... Fst graph is loaded
18.04.19 17:08:33.055354 00000x2a56 0x00007f48b47fe700 Graph /opt/ASR_Themes/theme/telecom_phoncalls_S8_L0_v0.2_for5.5.fst.cor loaded in 32.3991 seconds
18.04.19 17:08:33.253970 00000x2a56 0x00007f48b47fe700 Внимание!!! в словаре обнаружено служебное (<...>) слово,незнакомое декодеру: <vtb>

```

Рисунок 43 – Проверка записей log-файла

Убедитесь, что тема успешно загрузилась.



При необходимости повторите операцию на других серверах.

6.8 Удаление графа слитного распознавания в Windows

Перейдите в каталог с установленным графом **%allusersprofile%\Speech Technology Center\ASR_Themes\<theme_name>** и удалите его (рис. 44).

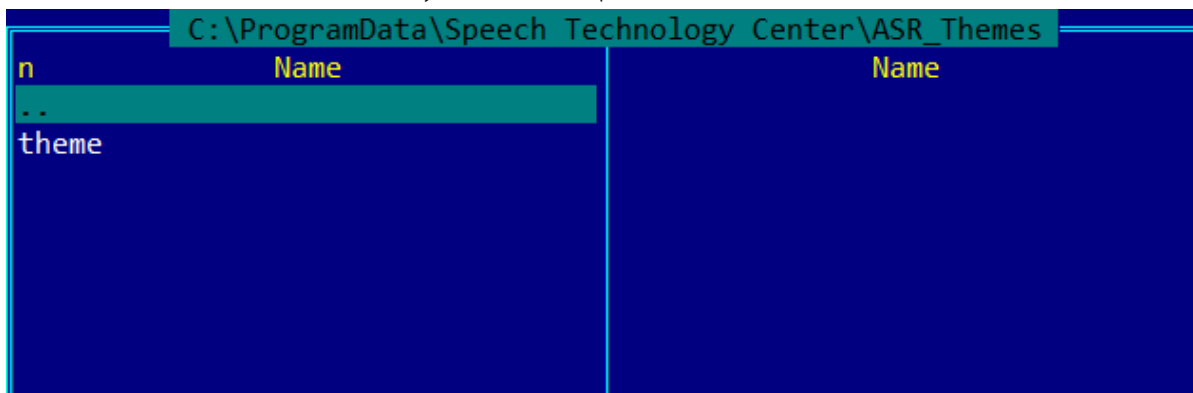


Рисунок 44 – Удаление графа

Чтобы принять изменения (активные сессии не прерываются), создайте файл **reload.txt**. Создание файла представлено на рисунке 45.

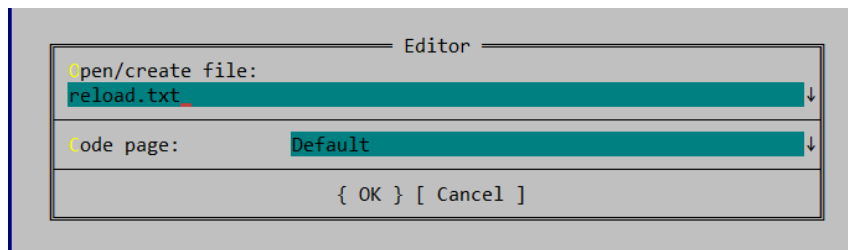


Рисунок 45 – Создание файла

Дождитесь пока файл исчезнет. Процесс может занять некоторое время.



При необходимости повторите операцию на других серверах.

6.9 Удаление графа слитного распознавания в Linux

Перейдите в директорию с установленным графом `/opt/ASR_Themes /<theme_name>` и удалите файл с темой (рис. 46).

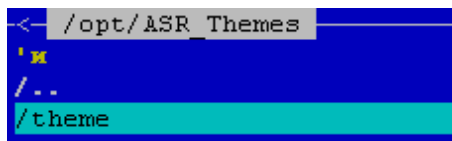


Рисунок 46 – Выбор файла с темой

Чтобы принять изменения (активные сессии не прерываются), создайте файлы **reload.txt** и **touch reload.txt**.

Дождитесь, пока файл исчезнет. Процесс может занять некоторое время.



При необходимости повторите операцию на других серверах.

7. СПОСОБЫ ИНТЕГРАЦИИ

7.1 Интеграция по протоколу MRCP

7.1.1 Общая информация об интеграции через MRCP

Для связи ПО **VoiceNavigator** и голосовой платформы используется протокол MRCP версии 1.0 или версии 2.0. В соответствии с этим могут быть использованы два режима работы, которые основываются на разных наборах протоколов: режим MRCPv1 и режим MRCPv2.



Для ознакомления с протоколом MRCPv1 перейдите по ссылке <https://tools.ietf.org/html/rfc4463>.

Для ознакомления с протоколом MRCPv2 перейдите по ссылке <https://tools.ietf.org/html/rfc6787>.

Для MRCPv1 используется следующий набор протоколов:

- в качестве протокола транспортного уровня для передачи цифрового звукового сигнала используется протокол RTP;
- в качестве протокола управления сессиями используется RTSP.

Для MRCPv2 используется следующий набор протоколов:

- в качестве протокола транспортного уровня для передачи цифрового звукового сигнала используется протокол RTP;
- в качестве протокола сессионного уровня используется SIP,
- в качестве транспортного протокола для которого используется UDP;
- для передачи сообщений прикладного протокола MRCP версии 2.0 открывается отдельный канал с использованием TCP в качестве протокола транспортного уровня.

Голосовая платформа через **STC MRCP Server** запрашивает доступ у **STC ResourceServer** к программному комплексу распознавания речи **STC ASR** и программному комплексу синтеза речи **STC TTS**. В зависимости от этого используются различные схемы взаимодействия, описание которых приведено ниже.

При установке продукта **VoiceNavigator** с интеграцией по протоколу MRCP целевой задачей является создание решений, формируемых при взаимодействии с голосовой платформой (автоматизированная обработка обращений голосовыми сервисами IVR с распознаванием и синтезом речи). На момент публикации данной документации была протестирована совместимость со следующими голосовыми платформами по протоколу MRCP:

- Asterisk (через модуль unimrcp);
- Avaya Portal 5, 6, 7;
- Cisco UCCE 9, 10, 11;
- Cisco UCCX 9, 10, 11;
- Cisco VXML GW ios 12.3 и выше;
- FortelT UIVR;
- Genesys Voice Platform 7, 8;
- Naumen NCC 6.2 и выше.

Количество одновременных компонентов распознавания и синтеза речи рекомендуется рассчитывать на основе количества каналов, используемых голосовой платформой.

Программному комплексу **STC ASR** с аудиоданными может передаваться информация об использовании грамматик (**SRGS** или граф), после обработки приходит ответ в виде NLSML.



SRGS (англ. Speech Recognition Grammar Specification) – стандарт, который описывает структуру грамматики, используемой в распознавании речи. SRGS позволяет задавать слова или словосочетания, которые могут быть распознаны речевым движком. Полное описание стандарта приведено тут: <https://www.w3.org/TR/speech-grammar/>.



NLSML (англ. Natural Language Semantics Markup Language) – язык разметки для представления результатов распознавания. Полное описание стандарта приведено тут: <https://www.w3.org/TR/nl-spec/>.

Программному комплексу **STC TTS** может передаваться текст с информацией об использовании настроек произношения с помощью тегов (**SSML**). По завершении обработки ответ приходит в аудиоформате.



SSML (Speech Synthesis Markup Language) язык разметки, основанный на XML для применения в приложениях синтеза речи. Полное описание стандарта приведено тут: <https://www.w3.org/TR/speech-synthesis/>.

7.1.2 Требования к форматам аудиоданных по MRCP

Формат выходных аудиоданных

Программный комплекс синтеза речи совместно с MRCP-сервером поддерживает следующие форматы выходных аудиоданных, передаваемых по протоколу RTP:

- G.711 a-law, моно, 8000 Гц;
- G.711 μ -law, моно, 8000 Гц.

Формат входных аудиоданных

Программный комплекс распознавания речи совместно с MRCP-сервером поддерживает следующие форматы входных аудиоданных, передаваемых по протоколу RTP:

- G.711 a-law, моно, 8000 Гц;
- G.711 μ -law, моно, 8000 Гц.

Значения идентификатора формата аудиоданных

При передаче аудиоданных по протоколу RTP программные средства **VoiceNavigator** используют следующие значения для идентификатора формата (**Payload type**):

Формат аудиоданных	Значение параметра Payload type
G.711 μ -law	0
G.711 a-law	8

Параметры сигнала, передаваемого на распознавание

Рекомендации к параметрам сигнала, передаваемого на распознавание:

- соотношение сигнал/шум не менее 15 дБ;
- отсутствие искажений канала связи (искажения частотной характеристики; периодические наводки; импульсные помехи; наличие перегрузки);
- высокий уровень речевого сигнала (не менее 10 % от максимума сигнала);
- отсутствие фонового шума (музыка, «речевой коктейль»);
- целевая группа дикторов в возрасте от 18 до 60 лет (без ярко выраженного акцента и дефектов речи).

7.1.3 Взаимодействие STC ASR и голосовой платформы в режиме MRCPv1

Порядок взаимодействия между голосовой платформой и программным комплексом распознавания речи **STC ASR** происходит в последовательности, представленной на рисунке 47.

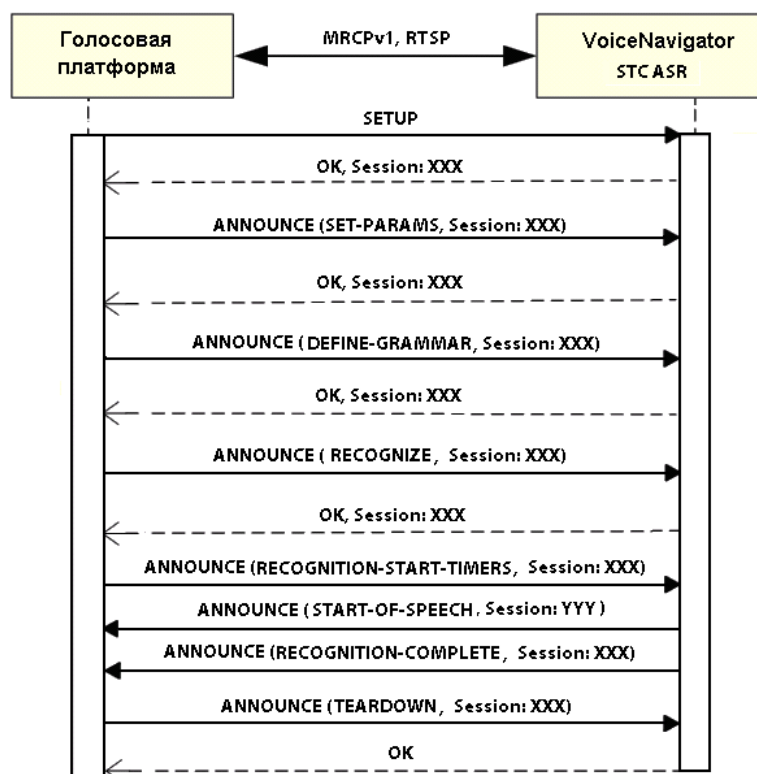


Рисунок 47 – Взаимодействие голосовой платформы и ресурса распознавания речи

При необходимости использования ресурса распознавания речи голосовая платформа устанавливает новое соединение с **STC ASR** с помощью RTSP-команды **SETUP**.

STC ASR создает новую сессию и передаёт ее идентификатор в ответе. Далее голосовая платформа настраивает используемый ресурс, используя MRCP-запросы **SET-PARAMS** (установка параметров распознавания) и **DEFINE-GRAMMAR** (установка грамматики).

Для запуска процесса распознавания голосовая платформа отправляет в **STC ASR** MRCP-запросы **RECOGNIZE** (начало распознавания) и **RECOGNITION-START-TIMERS** (запуск таймеров).

Событие **START-OF-SPEECH** используется при распознавании для индикации того, что ресурс распознавания детектировал речь в передаваемом сигнале.

После завершения процесса распознавания **STC ASR** информирует голосовую платформу об окончании работы ресурса, используя MRCP-запрос **RECOGNITION-COMPLETE** (процесс распознавания завершен).

По завершении работы платформа использует RTSP-команду **TEARDOWN** для прекращения соединения, иначе происходит повторная настройка ресурса и постановка новой задачи на распознавание.

7.1.4 Взаимодействие STC ASR и голосовой платформы в режиме MRCPv2

Порядок взаимодействия между голосовой платформой и программным комплексом распознавания речи **STC ASR** в режиме MRCPv2 происходит в последовательности, представленной на рисунке 48.

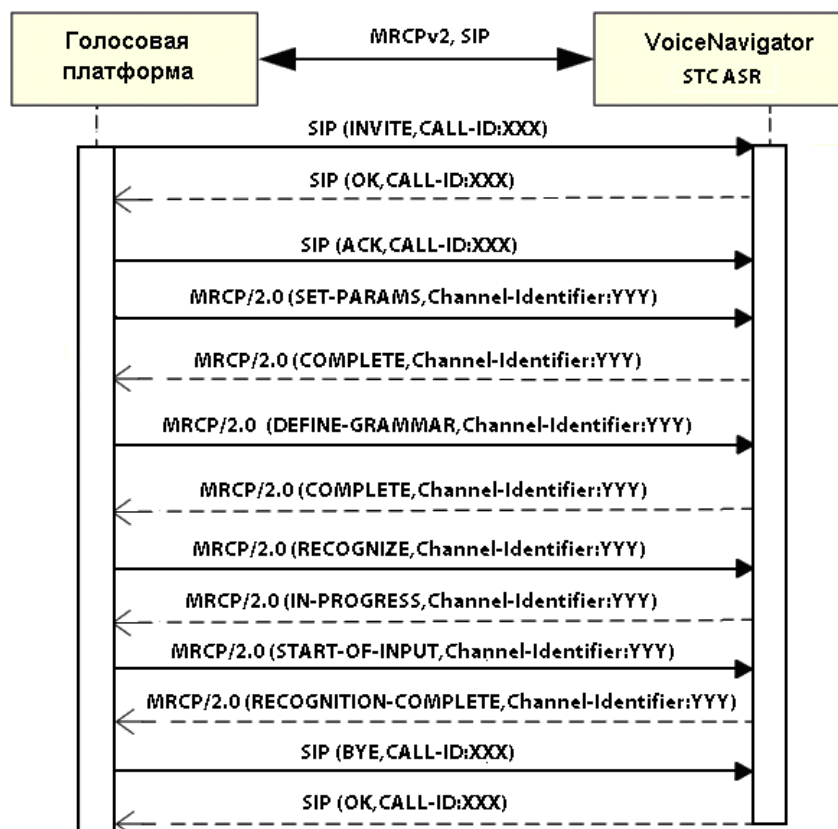


Рисунок 48 – Взаимодействие голосовой платформы и ресурса распознавания речи

При необходимости использования ресурса распознавания речи голосовая платформа устанавливает новое соединение с **STC ASR** с помощью SIP-команды **INVITE**.

STC ASR создаёт новую сессию и передаёт её идентификатор в ответе. После ответа **STC ASR** на запрос, голосовая платформа подтверждает приём ответа с помощью запроса **ACK**.

Далее голосовая платформа настраивает используемый ресурс, используя MRCP-запросы **SET-PARAMS** (установка параметров распознавания) и **DEFINE-GRAMMAR** (установка грамматики). **STC ASR** информирует голосовую платформу о выполнении данных запросов с помощью MRCP-запросов **COMPLETE**.

Для запуска процесса распознавания голосовая платформа отправляет в **STC ASR** MRCP-запрос **RECOGNIZE** (начало распознавания).

Событие **IN-PROGRESS** показывает, что происходит процесс распознавания.

Событие **START-OF-INPUT** используется при распознавании для индикации того, что ресурс распознавания обнаружил речь в передаваемом сигнале.

После завершения процесса распознавания **STC ASR** информирует голосовую платформу об окончании работы ресурса, используя MRCP-запрос **RECOGNITION-COMPLETE** (процесс распознавания завершен).

По завершении работы платформа использует SIP-команду **BYE** для прекращения соединения, иначе происходит повторная настройка ресурса и постановка новой задачи на распознавание.

7.1.5 Передача грамматик для распознавания по MRCP

Передача закрытых грамматик по MRCP

Для выполнения распознавания в **IVR** по закрытым грамматикам голосовое приложение должно передать в MRCP-командах **DEFINE-GRAMMAR** или **RECOGNIZE** имя темы в формате **http://<real_server_name>/<name_file>.grxml**.

Например:

http://192.168.20.211/yes_no.grxml.



Проверьте, чтобы ссылки на файлы были доступны для чтения.

Передача грамматик слитного распознавания речи по MRCP

Для выполнения слитного распознавания в **IVR** голосовое приложение должно передать в MRCP-командах **DEFINE-GRAMMAR** или **RECOGNIZE** имя темы в формате **http://<fake_or_real_server_name>/theme:<theme_name>**.

Например:

http://<localhost>/theme:myTheme

В граф необходимо добавлять информацию о языке, на котором собран граф. Для этого в каталог графа добавляется файл **lang.txt**, содержащий код одного из трех поддерживаемых языков:

- ru-RU (Русский язык);
- kk-KZ (Казахский язык);

- en-US (Английский язык).

Каталоги для хранения графов:

- Windows: **%allusersprofile%\Speech Technology Center\ASR_Themes\<theme_name>**;
- Linux: **/opt/ASR_Themes /<theme_name>**.



Проверьте, чтобы ссылки на файлы были доступны для чтения.

Взаимодействие STC TTS и голосовой платформы в режиме MRCPv1

Порядок взаимодействия между голосовой платформой и программным комплексом синтеза речи **STC TTS** происходит в последовательности, представленной на рисунке 49.

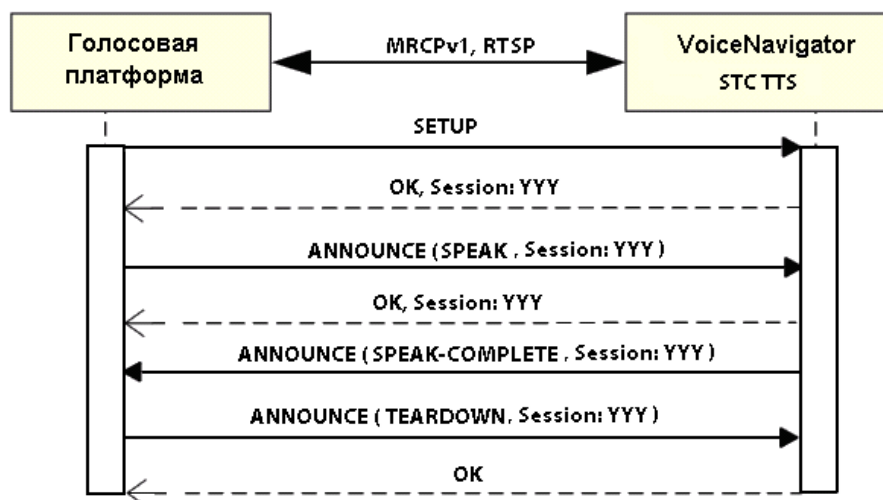


Рисунок 49 – Взаимодействие голосовой платформы и ресурса синтеза речи

Голосовая платформа создает новое соединение с **STC TTS** с помощью RTSP-команды **SETUP** и получает в ответе идентификатор новой сессии.

Для синтеза речи платформа отправляет MRCP-запрос **SPEAK** (начать синтез), в которой содержится синтезируемый текст.

Завершая синтез, **STC TTS** отправляет платформе MRCP-запрос **SPEAK-COMPLETE** (синтез завершен).

По завершении работы платформа использует RTSP-команду **TEARDOWN** для прекращения соединения, иначе происходит повторная настройка ресурса и постановка новой задачи на синтез.

Взаимодействие STC TTS и голосовой платформы в режиме MRCPv2

Порядок взаимодействия между голосовой платформой и программным комплексом синтеза речи **STC TTS** в режиме MRCPv2 происходит в последовательности, представленной на рисунке 50.

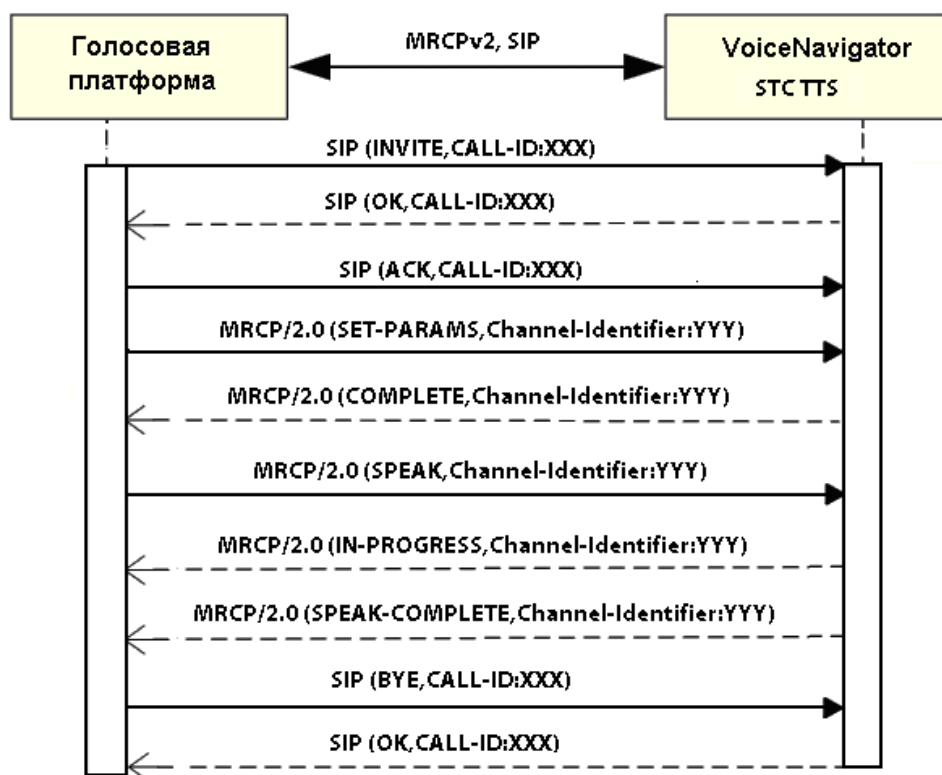


Рисунок 50 – Взаимодействие голосовой платформы и ресурса синтеза речи

Голосовая платформа создает новое соединение с **STC TTS** с помощью SIP-команды **INVITE** и получает в ответе идентификатор новой сессии.

Далее голосовая платформа настраивает используемый ресурс, используя MRCP-запросы **SET-PARAMS** (установка параметров распознавания).

Для синтеза речи платформа отправляет MRCP-запрос **SPEAK** (начать синтез), в которой содержится синтезируемый текст.

Завершая синтез, **STC TTS** отправляет платформе MRCP-запрос **SPEAK-COMPLETE** (синтез завершен).

По завершении работы платформа использует SIP-команду **BYE** для прекращения соединения, иначе происходит повторная настройка ресурса и постановка новой задачи на синтез.

7.2 Интеграция по протоколу WEB API

7.2.1 Общая информация по интеграции через Web API

Для связи ПО **VoiceNavigator**, голосовой платформы и других программных средств, которые не поддерживают протокол MRCP, используется протокол Web API.

Возможности использования функционала Web API заключаются в следующем:

- распознавания ранее записанных аудиофайлов;
- синтез аудиофайлов для последующего прослушивания;
- распознавание речи в режиме реального времени;
- синтез речи в режиме реального времени;

Рекомендуется рассчитывать количество одновременных компонентов распознавания и синтеза речи, основываясь на требуемой производительности.

Выбор конфигурации зависит от количества используемых каналов.

7.2.2 Требования к форматам аудиоданных через WEB API

Формат выходных аудиоданных

Программный комплекс синтеза речи Web API поддерживает следующий формат выходных аудиоданных: audio/x-pcm;bit=16;rate=8000 Linear PCM с частотой дискретизации 8000 Гц и разрядностью квантования 16 бит, моно.

Формат входных аудиоданных

На данный момент Web API распознает речь только в формате: audio/x-pcm;bit=16;rate=8000 Linear PCM с частотой дискретизации 8000 Гц и разрядностью квантования 16 бит, моно.

Значения идентификатора формата аудиоданных

При передаче аудиоданных по протоколу RTP, программные средства **VoiceNavigator** используют следующие значения для идентификатора формата (**Payload type**):

Формат аудиоданных	Значение параметра Payload type
G.711 μ -law	0
G.711 a-law	8

Параметры сигнала, передаваемого на распознавание

Рекомендации к параметрам сигнала, передаваемого на распознавание:

- соотношение сигнал/шум не менее 15 дБ;
- отсутствие искажений канала связи (искажения частотной характеристики; периодические наводки; импульсные помехи; наличие перегрузки);
- высокий уровень речевого сигнала (не менее 10 % от максимума сигнала);
- отсутствие фонового шума (музыка, «речевой коктейль»);
- целевая группа дикторов в возрасте от 18 до 60 лет (без ярко выраженного акцента и дефектов речи).

7.2.3 Настройка взаимодействия через Web API в режиме работы с файлами

Для запуска веб-сервиса необходимо выполнить следующие действия:

- зайти в каталог настройки **/opt/MRCP/Config**;
- открыть файл **MrcpServerConfig.xml** в режиме правки;
- найти раздел `<!-- Web API-->`;
- настроить HTTP-префикс, установив нужный порт, например 8081:

```
<cloud>
  <httpPrefix value="http://*:8081/" />
```



В настройках порт 8081 установлен по умолчанию.

Если порт необходимо закрыть, то требуется закомментировать раздел `<cloud>`:

```
<!-- Web API-->
<cloud>
  <httpPrefix value="http://*:8081/" />
  <!-- Web API Proto -->
  <streaming ip="192.168.20.221" port="39876" />
</cloud> -->
```

или заблокировать его настройками фаервола.



При комментировании только веб-порта, порт службы потоковых сервисов не будет доступен.

Здесь и далее указанные IP-адреса являются примером настроек. Для корректной работы необходимо использовать IP-адреса целевой системы.

Перезагрузите сервис одним из удобных способов, представленных в разделе [ЗАПУСК И ОСТАНОВКА СЛУЖБ](#).

Необходимо убедиться, что порт нормально функционирует:

```
netstat -na
tcp        0      0 0.0.0.0:8081          0.0.0.0:*           LISTEN
```

7.2.4 Настройка взаимодействия через Web API в режиме потока

Сервисы синтеза и распознавания работают поверх протокола TCP. Сообщения кодируются и декодируются с помощью Google ProtoBuffers, а разбиение на сообщения производится с помощью специального маркера.

Данный режим работает как поверх TCP, так и поверх WebSocket.

Для запуска веб-сервиса необходимо выполнить следующие действия:

- зайти в каталог настройки **`/opt/MRCP/Config`**;
- открыть файл **`MrcpServerConfig.xml`** в режиме правки;
- найти раздел `<!-- Web API-->`;
- настроить HTTP-префикс, установив нужный порт, например 39876:

```
<cloud>
  <httpPrefix value="http://*:8081/" />
  <!-- Web API Proto -->
  <streaming ip="192.168.20.221" port="39876" />
</cloud>
```



В настройках порт 39876 установлен по умолчанию.

Если порт необходимо закрыть, то требуется закомментировать раздел `<cloud>`:

```

<!-- Web API-->
<cloud>
  <httpPrefix value="http://*:8081/" />
  <!-- Web API Proto -->
  <streaming ip="192.168.20.221" port="39876" />
</cloud>

```

или заблокировать настройками фаервола.



При комментировании только порта для трансляции, порт службы распознавания файлов будет доступен.

Здесь и далее указанные ip-адреса являются примером настроек. Для корректной работы необходимо использовать ip-адреса целевой системы.

Перезагрузите сервис одним из удобных способов, представленных в разделе [ЗАПУСК И ОСТАНОВКА СЛУЖБ](#).

Необходимо убедиться, что порт нормально функционирует:

```

netstat -na
tcp        0      0 192.168.20.221:39876  0.0.0.0:*        LISTEN

```



Формат разбиения потока на сообщения в TCP: [hex размер сообщения]\n\n[сериализованное protobuf-сообщение].

7.2.5 Распознавание аудиофайлов через Web API

Для распознавания аудиофайлов через WEB API используется HTTP-метод **POST**.

Формат POST-запроса представлен ниже:

```

POST /asr?uuid=<идентификатор пользователя>&key=<API-ключ>&topic=<название темы>&lang=<язык
распознавания>&asr_result=<формат результата> HTTP/1.1

```

```

Host: IP-SN:8081

```

```

Content-Type: audio/x-pcm;bit=16;rate=8000

```

```

... (двоичное содержимое аудиофайла)

```

Описание полей:

- **uuid** – уникальный идентификатор пользователя (не проверяется);
- **key** – API-ключ (не проверяется);
- **topic|srgs - topic** – языковая модель, которую следует использовать при распознавании (graph) | srgs (ссылка на srgs грамматику);
- **lang** – язык распознавания: русский, английский и казахский языки;
- **asr_result** – результат распознавания;
- **nlsmi** – результат распознавания возвращается в виде nlsmi (по умолчанию),
- **srrs** – результат распознавания в виде srrs,
- **words** – результат распознавания с временными задержками по словам;
- **Content-Type** – использование формата audio/x-pcm;bit=16;rate=8000 Linear PCM с частотой дискретизации 8000 Гц и разрядностью квантования 16 бит, моно;
- **Transfer-Encoding** – поле предназначено для одновременного запуска распознавания и передачи аудиоданных в целях ускорения начала распознавания и всего процесса распознавания. Для этого

в HTTP-заголовке **Transfer-Encoding** укажите `chunked` и передавайте данные в теле сообщения частями. Данный метод позволяет серверу обрабатывать аудио ещё до момента поступления всех данных. Результат распознавания формируется практически одновременно с окончанием приёма звука. Поддерживаются любые частоты дискретизации и только моно в форматах (значения заголовка **Content-Type**):

WAV – `audio/pcm;rate=<sample-rate>;`

MP3 – `audio/mp3;`

OGG/VORBIS – `audio/oggvorbis;`

OGG/OPUS – `audio/oggopus;`

g711a – `audio/g711a;`

g711u – `audio/g711u;`

g729 – `audio/g729.`

Пример POST-запроса с использованием **CURL**:

```
curl -X POST -H "Content-Type: audio/x-pcm;bit=16;rate=8000" -H "Transfer-Encoding: chunked" --data-binary "@utterance.wav" "http://
```

Полученный результат:

```
curl -X POST -H "Content-Type: audio/x-pcm;bit=16;rate=8000" -H "Transfer-Encoding: chunked" --data-binary "@utterance.wav" "http://192.168.20.82:8081/asr?uuid=Test-000001&key=test&topic=graph&lang=ru-RU&asr_result=srrs"
```

Примеры результатов распознавания для **asr_result** в виде **srrs**:

```
<?xml version="1.0" encoding="utf-8"?>
<result snr="25" description="success">
  <grammar id="session:bd7d94e2-31c5-4091-a2c3-cf934539de17" confidence="41"
mode="voice" tag-format="semantics/1.0">
    <rule id="WFST">
      <input>
        <input confidence="72" start="2018-09-28 11:58:20.630"
stop="2018-09-28 11:58:21.140">мужская</input>
        <input confidence="73" start="2018-09-28 11:58:21.140"
stop="2018-09-28 11:58:21.620">сборная</input>
        <input confidence="73" start="2018-09-28 11:58:21.620"
stop="2018-09-28 11:58:21.980">России</input>
        <input confidence="73" start="2018-09-28 11:58:21.980"
stop="2018-09-28 11:58:22.130">по</input>
        <input confidence="73" start="2018-09-28 11:58:22.130"
stop="2018-09-28 11:58:22.580">хоккею</input>
        <input confidence="69" start="2018-09-28 11:58:22.580"
stop="2018-09-28 11:58:22.610">с</input>
        <input confidence="70" start="2018-09-28 11:58:22.610"
stop="2018-09-28 11:58:23.090">крупным</input>
        <input confidence="60" start="2018-09-28 11:58:23.090"
stop="2018-09-28 11:58:23.480">счетом</input>
        <input confidence="25" start="2018-09-28 11:58:23.510"
stop="2018-09-28 11:58:24.080">обыграла</input>
        <input confidence="73" start="2018-09-28 11:58:24.080"
stop="2018-09-28 11:58:24.530">команду</input>
        <input confidence="57" start="2018-09-28 11:58:24.530"
stop="2018-09-28 11:58:25.070">Норвегии</input>
      </input>
```

```

        </rule>
    </grammar>
</result>

```

Примеры результатов распознавания для **asr_result** в виде **nlsmi**:

```

<?xml version="1.0" encoding="utf-8"?>
<result>
  <interpretation grammar="session:ccc190c8-69d3-4cbc-8661-2c5496b401c7"
confidence="41">
    <input mode="speech" confidence="41" timestamp-start="2018-09-28T11:59:28.630"
timestamp-end="2018-09-28T11:59:58.450">мужская сборная России по хоккею с крупным счетом
обыграла команду Норвегии </input>
    <instance>мужская сборная России по хоккею с крупным счетом обыграла команду
Норвегии
      <SWI_meaning>мужская сборная России по хоккею с крупным счетом обыграла
команду Норвегии</SWI_meaning>
      <SWI_utteranceSNR>25</SWI_utteranceSNR>
      <STC_wordConfidence>72;73;73;73;73;69;70;60;25;73;57;</STC_wordConfidence>
    </instance>
  </interpretation>
</result>

```

Примеры результатов распознавания для **asr_result** в виде **words**:

```

<words>
  <word begin="630" end="1140">мужская</word>
  <word begin="1140" end="1620">сборная</word>
  <word begin="1620" end="1980">России</word>
  <word begin="1980" end="2130">по</word>
  <word begin="2130" end="2580">хоккею</word>
  <word begin="2580" end="2610">с</word>
  <word begin="2610" end="3090">крупным</word>
  <word begin="3090" end="3480">счетом</word>
  <word begin="3510" end="4080">обыграла</word>
  <word begin="4080" end="4530">команду</word>
  <word begin="4530" end="5070">Норвегии</word>
</words>

```

7.2.6 Синтез речи через Web API

Для синтеза речи через Web API используется HTTP-метод **GET**. Формат запроса представлен ниже:

```

http://IP-SN:8081/tts?uuid=<идентификатор пользователя>&key=<API-ключ>&voice=<Установленный
голос>&lang=<язык распознавания>&text=<Синтезируемый текст>

```

Описание полей:

- **uuid** – уникальный идентификатор пользователя (не проверяется);
- **key** – API-ключ (не проверяется);
- **voice** – установленный голос на сервере, голос должен соответствовать языку **lang**;
- **lang** – язык распознавания: русский, английский и казахский языки;
- **text** – синтезируемый текст в формате UTF-8.

Пример выполненного запроса представлен ниже:

```
# TTS: wav файлы будут без заголовков!
curl -v 'http://95.213.207.114:81/tts?uuid=uuid&key=key&lang=en-US&voice=Carol8000&text=hello%20world'
>test_en.wav
```

```
# Текст на русском необходимо передавать в кодированном html виде
curl -v 'http://95.213.207.114:81/tts?uuid=uuid&key=key&lang=ru-RU&voice=Julia8000&text=%D1%85%D0%BE%D1%87%D1%83+%D1%83%D0%B7%D0%BD%D0%B0%D1%82%D1%8C+%D0%B1%D0%B0%D0%BB%D0%B0%D0%BD%D1%81' > test_ru.wav
```

В результате запроса должны получить следующие данные:

- синтезированная речь в формате **PCM** без заголовка;
- формат **Linear PCM** с частотой дискретизации 8000 Гц и разрядностью квантования 16 бит, моно;
- такие браузеры как **Chrome** и **Mozilla FireFox** будут нормально передавать русский текст.

7.2.7 Распознавание речи через Web API в потоковом режиме

Распознавание осуществляется в следующих режимах:

Идентификатор сервиса	Режим распознавания	Примечание
asr	Режим одиночного высказывания	
asr_offline	Режим распознавания офлайн	медленное, но качественное распознавание
asr_dictation	Режим диктовки	непрерывное распознавание по теме с обновлением промежуточных результатов

Грамматики устанавливаются в **AdvancedASROptions.Srgs** (грамматика в текстовом формате или ссылка на **SRGS** грамматику).

Результат распознавания в формате строки **NLSML** предоставляется в виде **AsrAddDataResponse.Recognition.Result.Normalized**.

Запрос начала распознавания:

```
AsrConnectionRequest
{
    ServiceName = <Идентификатор сервиса> (см. табл выше)
    Uuid = <Идентификатор пользователя>
    ApiKey = <Ключ пользователя>
    ApplicationName = [Приложение]
    ProtocolVersion = <Версия протокола>
    SpeechkitVersion = <Версия сервиса>
}
```

Ответ:

```
ConnectionResponse.ResponseCode - код завершения
ConnectionResponse.Message - сообщение об ошибке
```

Запрос на передачу звуковых данных:

```
AsrAddData
{
    AudioData = <Порция звуковых данных>
    LastChunk = <Признак конца данных>
}
```

В случае распознавания одиночного высказывания или распознавания в офлайн режиме, высылается одно сообщение с окончательным результатом.

В случае диктовки происходит последовательная выдача промежуточных результатов:

```
AddDataResponse
{
  Recognition
  {
    Result
    {
      Normalized=результат распознавания в формате NLSML
    }
  }
}
```

7.2.8 Синтез речи через Web API в потоковом режиме

Инициализация синтеза речи в потоковом режиме осуществляется с помощью следующего запроса:

```
TtsConnectionRequest
{
  ServiceName = "tts"
  Uuid = <Идентификатор пользователя>
  ApiKey = <Ключ пользователя>
  ApplicationName = [Приложение]
  ProtocolVersion = <Версия протокола>
  SpeechkitVersion = <Версия сервиса>
}
```

Ответ:

```
ConnectionResponse.ResponseCode - код завершения
```

Процесс синтеза речи осуществляется с помощью запроса:

```
GenerateRequest
{
  Application = [Приложение]
  Chunked = true
  Format = Pcm
  Lang = <ru-RU|en-EN|kz-KZ>
  Voice = <Установленный голос>
  Text = <Текст>,
}
```

В результате синтезированной речи поступает частями в сообщении **GenerateResponse**.

Поле **AudioData** содержит речь в формате **PCM**.

Если в поле **Completed** установлено значение **true**, то процесс синтеза речи считается завершённым.

8. НАСТРОЙКА ПРОГРАММНЫХ КОМПЛЕКСОВ

8.1 Настройка MRCP-сервера

8.1.1 Местоположение конфигурационных файлов

Конфигурационные файлы MRCP-сервера расположены в следующих каталогах:

- Windows: **C:\Program Files\Speech Technology Center\MRCP Server\Config**.
- Linux: **/opt/MRCP/Config**.

Название файла/каталога	Описание
MrcpServer.nlog	Конфигурационный файл протоколирования MRCP-сервера
MrcpServerConfig.xml	Конфигурационный файл настройки MRCP-сервера
Grammar	Папка с ABNF-грамматиками, описывающие работу протоколов



Для обеспечения работоспособности ПО выполните настройку брандмауэра или отключите его.

8.1.2 Конфигурационный файл MrcpServerConfig.xml

Все параметры заполняются автоматически при установке и развертывании ресурсов **VoiceNavigator** на одном сервере.

Путь параметра	Описание
common->host	IP-адрес сервера
session->rtspTcpServerAddress	IP-адрес и порт соединения по протоколу MRCPv1 (RTSP (TCP))
session->sipUdpServerAddress	IP-адрес и порт соединения по протоколу MRCPv2 (SIP (UDP))
session->sipTcpServerAddress	IP-адрес и порт соединения по протоколу MRCPv2 (SIP (TCP))
session->rtsp->resources	Список синонимов наименования ресурсов распознавания и синтеза речи, по которым можно обращаться к ресурсам
service->transportTcpServerAddress	IP-адрес и порт соединения по протоколу MRCPv2 (транспорт)
service->cloud	Секция настройки взаимодействия по Web API
service->cloud->httpPrefix	Порт, на котором ожидается соединение по Web API
service->cloud->streaming	IP-адрес и порт для взаимодействия по протоколу protobuf

proxy	Секция настройки взаимодействия с Resource-сервером
proxy->pingTimeout	Интервал времени (в миллисекундах) между проверками доступности Resource-серверов
proxy->establishmentTimeout	Время (в миллисекундах), по истечении которого MRCP-сервер осуществляет повторные попытки соединения с Resource-сервером в случае неудачной предыдущей попытки
proxy->links->server	Сетевые адреса (IP-адрес и порт) Resource-серверов, из расчёта 50 соединений/лицензий на одно соединение. Например, если предполагается 70 одновременных соединений, то нужно прописать 2 resource-сервера (можно одинаковых). Количество соединений к конкретному Resource-серверу должно совпадать с количеством соединений, прописанных в настройках данного Resource-сервера

8.1.3 Конфигурационный файл MrcpServer.nlog

Можно ознакомиться с конфигурационным файлом настройки протоколирования событий работы MRCP-сервера, перейдя по ссылке <https://github.com/nlog/NLog/wiki/File-target>.

8.1.4 Управляемые параметры распознавания речи по MRCP

Программные средства распознавания речи поддерживают дополнительные параметры (на которые можно оказывать влияние), устанавливаемые голосовой платформой в MRCP-запросах SET-PARAMS или RECOGNIZE.



Наивысший приоритет имеют параметры, заданные командой SET-PARAMS, далее идут параметры, заданные в программе настройки MRCP-сервера.

Параметр	Значение
Confidence threshold	Доверительный порог уверенности распознавания. Если Confidence level , возвращаемый при распознавании, меньше Confidence threshold , то результат распознавания no-match (нет соответствия).
Sensitivity level	Уровень чувствительности к несловарным командам (0 – 100). Чем больше значение, тем выше чувствительность к шуму. При значении 100 все данные распознаются как команды, при значении 0 – всегда выдаётся несоответствие команде.
N best list length	Количество возвращаемых результатов распознавания. Возвращается заданное количество результатов распознавания с уровнем достоверности больше Confidence threshold .
No input timeout	Время ожидания аудиоданных, по истечении которого система распознавания может прервать процедуру распознавания, с кодом no-input-timeout (002), в случае, если распознавание началось и не найдено речи.

Параметр	Значение
	Значение задаётся в миллисекундах.
Recognition timeout	Время ожидания результата распознавания, по истечении которого система распознавания может прервать процедуру распознавания с кодом recognition-timeout (003), в случае если распознавание началось или нет ни одного варианта распознавания. Значение задаётся в миллисекундах.
Dtmf interdigit timeout	Время ожидания следующего символа DTMF перед фиксацией введённой последовательности символов DTMF и выдачей результата. Значение задаётся в миллисекундах.
Dtmf term. timeout	Время ожидания после ввода последнего символа DTMF перед выдачей результата, при условии, что текущий результат окончательный. Значение задаётся в миллисекундах.
Dtmf term. char	Символ завершения при распознавании DTMF.
Speech complete timeout	Длительность тишины после фразы пользователя, необходимой для выдачи результата распознавания при наличии гипотезы. Значение задаётся в миллисекундах.
Speech incomplete timeout	Длительность тишины после фразы пользователя, необходимой для выдачи результата распознавания при отсутствии гипотезы (no-match). Значение задаётся в миллисекундах.
Save-Waveform	Флаг включения записи звука, поступающего на распознавание с последующей выдачей URI файла звукозаписи (для правильной работы требуется указание дополнительных параметров). Возможные значения: • true – включить запись звука, • false – отключить запись звука.  При использовании параметра Save-Waveform необходим внешний HTTP или файловый сервер для хранения (в поставку не входит).
Send nomatch	Параметр, с помощью которого устанавливается необходимость включения в ответ гипотез с низким уровнем достоверности (confidence level < confidence threshold). Если параметр отключён, то такие гипотезы будут исключены из ответа.
Speed vs Accuracy	Приоритет между производительностью и качеством распознавания (0 – 100). Чем больше значение, тем выше производительность и ниже качество распознавания.

Параметр	Значение																		
Default SRRS interpreter	Параметр позволяет задавать семантический интерпретатор по умолчанию для тех грамматик, где не задан tag-format. Возможные значения: 0 – не обрабатывается <tag>; 1 – <tag> обрабатывается как ECMAScript; 2 – <tag> обрабатывается как ECMAScript. Подробное описание представлено: http://www.w3.org/TR/speech-grammar/.																		
swirec_secure_context	Флаг принудительного выключения (экранирования) секретных данных в log-файлах и артефактах, работает даже при глобальном включении логирования. <table><tr><th>Заголовок 1</th><th>Заголовок 2</th><th></th></tr><tr><td>Value</td><td>Open</td><td>Значение по умолчанию, записываем все в логи и артефакты</td></tr><tr><td></td><td>Secure</td><td>В логи не пишутся результаты распознавания, в артефакты не пишутся результаты распознавания и wav файлы</td></tr><tr><td>Default</td><td>Open</td><td></td></tr><tr><td>Варианты установки</td><td><property name="swirec_secure_context" value="secure"/></td><td>VXML</td></tr><tr><td></td><td>Использование Vendor-Specific-Parameters заголовка в RECOGNIZE, DEFINE-GRAMMAR, or INTERPRET request (not SET-PARAMS)</td><td>MRCP</td></tr></table>	Заголовок 1	Заголовок 2		Value	Open	Значение по умолчанию, записываем все в логи и артефакты		Secure	В логи не пишутся результаты распознавания, в артефакты не пишутся результаты распознавания и wav файлы	Default	Open		Варианты установки	<property name="swirec_secure_context" value="secure"/>	VXML		Использование Vendor-Specific-Parameters заголовка в RECOGNIZE, DEFINE-GRAMMAR, or INTERPRET request (not SET-PARAMS)	MRCP
Заголовок 1	Заголовок 2																		
Value	Open	Значение по умолчанию, записываем все в логи и артефакты																	
	Secure	В логи не пишутся результаты распознавания, в артефакты не пишутся результаты распознавания и wav файлы																	
Default	Open																		
Варианты установки	<property name="swirec_secure_context" value="secure"/>	VXML																	
	Использование Vendor-Specific-Parameters заголовка в RECOGNIZE, DEFINE-GRAMMAR, or INTERPRET request (not SET-PARAMS)	MRCP																	

8.1.5 Управляемые параметры синтеза речи по MRCP

Программные средства синтеза речи поддерживают дополнительные параметры (на которые можно оказывать влияние), устанавливаемые голосовой платформой в MRCP-запросах SET-PARAMS или RECOGNIZE.

Параметр	Значение
Voice name	Задаёт голос для использования по умолчанию. Возможные значения: Alexander8000, Anna8000, Vladimir8000, Lydia8000, Maria8000, Julia8000, Asel8000, Carol8000, Julia_8000n, Anna_8000n, Vladimir_8000n, Julia_f8000n, Anna_f8000n, Vladimir_f8000n.  Для работы комплекса синтеза речи необходимо обязательно указать в поле <voice language/> голос, который будет использоваться по умолчанию в настройках STCResourceServer.

Параметр	Значение
Prosody volume	Устанавливает относительный уровень громкости синтезируемой речи. Изменяется от 0 до 100.
Prosody rate	Устанавливает темп произнесения речи. Изменяется от 0,5 до 4.

8.2 Настройка Resource-сервера

8.2.1 Местоположение конфигурационных файлов

Конфигурационные файлы Resource-сервера расположены в следующих каталогах:

- Windows: **C:\Program Files\Speech Technology Center\Resource Server\Config**.
- Linux: **/opt/RS/Config**.

Файлы настроек Resource-сервера представлены в таблице:

Название файла\папки	Описание
ResourceServerConfig.xml	Конфигурационный файл настройки Resource-сервера
ArtifactStorageConfig.xml	Конфигурационный файл настройки артефактов и протоколирования событий
tts_config.xml	Конфигурационный файл настройки синтеза речи

Файлы настроек SNMP-сервиса:

Название файла	Описание
ResourceServerSnmpAgent.conf	Конфигурационный файл настройки SNMP сервиса, расположен: • C:\usr\etc\snmp (для Windows), • /usr/etc/snmp (для Linux).
STC-VN-MIB.txt	MIB файл, описание счетчиков, расположен: • C:\usr\share\snmp\mibs (для Windows), • /usr/share/snmp/mibs (для Linux).

В ResourceServerSnmpAgent.conf прописать rocommunity public.

8.2.2 Конфигурационный файл resourceServerConfig.xml

Путь параметра	Описание
common->hostAddress	IP-адрес Resource-сервера

common->nodeId	Имя Resource-сервера (используется в уведомлениях SNMP)
languages	Секция настройки списка поддерживаемых языков и списка тегов для каждого языка, по которым к ним можно обращаться. Первый язык в этом списке – язык по умолчанию (для распознавания речи)
asr	Секция настройки модуля распознавания речи ASR
asr->enginesCount	Максимальное количество одновременно работающих движков онлайн распознавания (задается в атрибуте online) и офлайн распознавания (задается в атрибуте offline) речи
asr->productName	Имя продукта для конфигурирования
asr->profiles->online->profile	Путь к профилю (атрибут path) для установленного языка распознавания (атрибут language). Используется для онлайн распознавания
asr->profiles->offline->profile	Путь к профилю (атрибут path) для установленного языка распознавания (атрибут language). Используется для офлайн распознавания
asr->themes->online	Путь к директории, в которой расположены темы для онлайн распознавания
asr->themes->offline	Путь к директории, в которой расположены темы для офлайн распознавания
tts	Секция настройки модуля синтеза речи TTS
tts->enginesCount	Максимальное количество одновременно работающих движков синтеза речи
tts->params->default->language	Язык по умолчанию
tts->params->default->voice	Голос по умолчанию
proxy	Секция настройки взаимодействия с MRCP-сервером
proxy->listen	Порт (атрибут port), к которому ожидается подключение MRCP-сервера
proxy->resources->client	Подключаемые MRCP-серверы
proxy->resources->client->link	Соединения (link) с указанием количества предоставляемых движков распознавания (asr) и синтеза (tts) речи. Количество соединений должно совпадать с количеством соединений, прописанных в настройках соответствующего MRCP-сервера
logger	Секция настройки параметров протоколирования событий Resource-сервера
logger->fileSizeLimit	Максимальный размер файла протоколирования событий (в байтах)

logger->target	Путь сохранения файлов протоколирования событий
logger->filenamePattern	Маска имени файла протоколирования событий
logger->maxFile	Максимальное количество файлов протоколирования событий
logger->syslog	IP-адрес и порт вывода в Unix syslog
logger->eventLog	Вывод ошибок в Windows EventLog
logger->snmpTraps	Версия SNMP для вывода snmpTraps
logger->statsTimeout	Периодичность (в миллисекундах) вывода статистики использования ресурсов распознавания и синтеза речи. Статистика выводится в файлы stats.csv и stats.log
logger->logLevel	Уровень протоколирования событий

8.2.3 Конфигурационный файл artifactStorageConfig.xml

Конфигурационный файл **ArtifactStorageConfig.xml** перечитывается с периодичностью 30 секунд, поэтому параметры, находящиеся в нем, можно менять в процессе работы **VoiceNavigator** без перезапуска служб.

Путь параметра	Описание
enabled	Сохранение артефактов: • true – сохранение включено, • false – сохранение выключено. Значение по умолчанию – true
callRecording	Запись всего диалога с пользователем. • true – запись включена, • false – запись выключена. Значение по умолчанию – false . Поддерживается не на всех голосовых платформах
basePath	Путь для сохранения артефактов и записей диалога
companyName	Название компании. Параметр участвует в формировании пути сохранения артефактов и позволяет разделять артефакты по компаниям. В конфигурационном файле задается значение по умолчанию. Значение можно переопределить в приложении и на голосовой платформе
applicationName	Название приложения. Параметр участвует в формировании пути сохранения артефактов и позволяет разделять артефакты по приложениям внутри компании.

	В конфигурационном файле задается значение по умолчанию. Значение можно переопределить в приложении и на голосовой платформе
ageLimit	Время хранения артефактов. Можно задать значение в минутах (постфикс к числу "min"), часах (постфикс к числу "h"), днях (постфикс "d"), неделях (постфикс "w"), месяцах (постфикс "m") и годах (постфикс "y"). Постфикс не чувствителен к регистру
sizeLimit	Размер хранилища артефактов. Можно задать значение в килобайтах (постфикс "k"), мегабайтах (постфикс "m") и гигабайтах (постфикс "g"). Постфикс не чувствителен к регистру
cleanupPeriod	Периодичность очистки хранилища артефактов. Сначала удаляются старые записи, затем записи под sizeLimit размер. Можно задать значение в минутах (постфикс "min"), часах (постфикс "h"), днях (постфикс "d"), неделях (постфикс "w"), месяцах (постфикс "m") и годах (постфикс - "y"). Постфикс не чувствителен к регистру
logLevel	Уровень протоколирования событий
backendThreads	Количество потоков для записи артефактов. Параметр не используется
serverPath	Префикс, использующийся для указания пути к файлам, получаемым при включенном в приложении параметре saved waveform URL

8.2.4 Конфигурационный файл tts_config.xml

Путь параметра	Описание
acoustic->soundbase->type	Тип аллофонной базы
acoustic->duration->word	Единица измерения длительности слов и пауз, в отсчетах (SMPL) или миллисекундах (MS)
acoustic->modify->writesegfile	Включение (on)/выключения (off) вывода разметки звука в файл
acoustic->modify->use_human_effects	Включение (on)/выключение (off) использования звуковых эффектов дыхания и смыкания губ
acoustic->modify->delete_last_pause	Включение (on)/ выключение (off) пауз в конце синтезируемой фразы
synthesis_mode	Синтезируемая модель
UnitSelection	Включение (on)/выключение (off) алгоритма синтеза UnitSelection
logging	Параметр не используется

8.3 Настройка статистики и мониторинга

8.3.1 Рекомендации по мониторингу

Программный продукт **VoiceNavigator** поддерживает протоколы **SNMP** и **SNMP TRAP** для интеграции с системами мониторинга.

В системе мониторинга рекомендуется отображать следующие графики, которые представлены на рисунке 51.

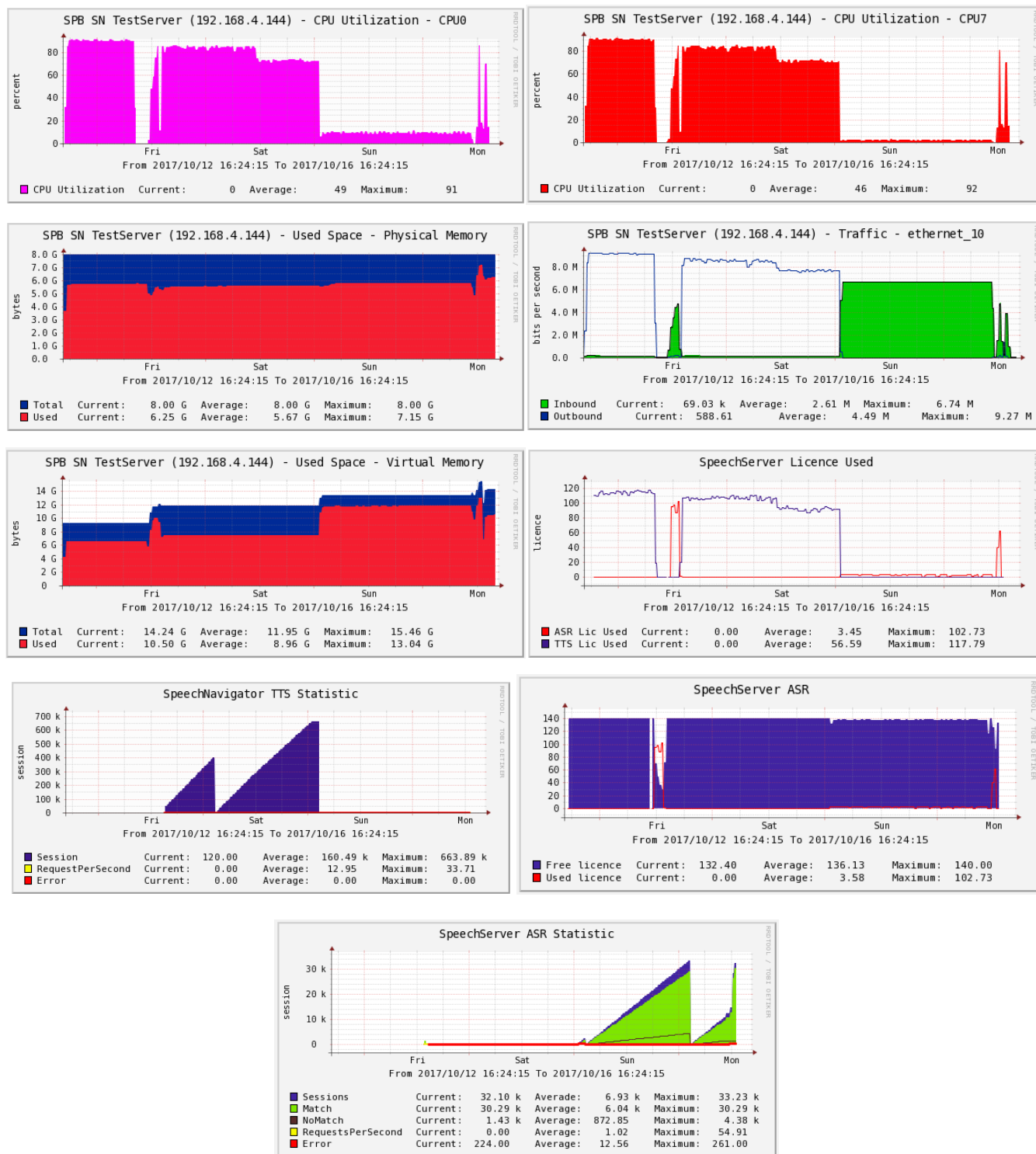


Рисунок 51 – Примеры графиков

8.3.2 Настройка конфигурации SNMP и примеры запросов

SNMP (англ. Simple Network Management Protocol, простой протокол сетевого управления) – стандартный интернет-протокол для управления устройствами в IP-сетях на основе архитектур TCP/UDP. **SNMP** является просто протоколом для сбора и организации информации.

Конфигурационные файлы расположены в следующих каталогах:

- Windows: **C:\usr**.
- Linux: **/usr/local/**.

Название файла	Описание
ResourceServerSnmpAgent.conf	Конфигурационный файл настройки SNMP-сервиса. Расположен в C:\usr\etc\snmp для Windows и в /usr/local/etc/snmp для Linux.
STC-VN-MIB.txt	MIB-файл содержит описание счётчиков. Расположен в C:\usr\share\snmp\mibs для Windows и в /usr/local/share/snmp/mibs для Linux.

Пример запроса всех счётчиков: **snmpwalk -On -c public -v 2c <IP адрес>:56780 .1.3.6.1.4.1.45373.5.1**. Счётчики можно получить тремя способами:

- с помощью текстового файла **stats.log**;
- в виде **CSV** формата **stats.csv**;
- запрос по протоколу **SNMP**.

Статистические данные по этим параметрам сохраняются в файлах **stats.csv** и **stats.log** в следующих каталогах:

- Windows: **C:\Program Files\Speech Technology Center\Resource Server\Logs\stats.log**.
- Linux: **/opt/RS/Logs/stats.log**.

Для оценки общей производительности работы системы **VoiceNavigator** могут быть использованы перечисленные ниже счётчики производительности.

SNMP OID	Значение в stats.log	Движок	Описание
.1.3.6.1.4.1.45373.5.1.1.1.0	occupiedAsrEngines	ASR	Количество занятых движков
.1.3.6.1.4.1.45373.5.1.1.2.0	freeAsrEngines	ASR	Количество свободных движков
.1.3.6.1.4.1.45373.5.1.1.3.0	successfulAsrSessionDuration	ASR	Средняя длина успешной сессии распознавания
.1.3.6.1.4.1.45373.5.1.1.4.0	successfulAsrSessionDurationMin	ASR	Минимальная длина успешной сессии распознавания
.1.3.6.1.4.1.45373.5.1.1.5.0	successfulAsrSessionDurationMax	ASR	Максимальная длина успешной сессии распознавания
.1.3.6.1.4.1.45373.5.1.1.6.0	asrErrors	ASR	Количество ошибок при распознавании
.1.3.6.1.4.1.45373.5.1.1.7.0	matchRecognitions	ASR	Количество успешно распознанных запросов

SNMP OID	Значение в stats.log	Движок	Описание
.1.3.6.1.4.1.45373.5.1.1.8.0	noMatchRecognitions	ASR	Количество нераспознанных запросов
.1.3.6.1.4.1.45373.5.1.1.9.0	noInputRecognitions	ASR	Количество запросов без речи
.1.3.6.1.4.1.45373.5.1.1.10.0	asrSessions	ASR	Общее количество сессий распознавания
.1.3.6.1.4.1.45373.5.1.1.11.0	asrRequestsPerSecond	ASR	Количество запросов в секунду
.1.3.6.1.4.1.45373.5.1.1.12.0	mediaDataPerAsrSession	ASR	Средний размер голосового трафика (байт)
.1.3.6.1.4.1.45373.5.1.1.13.0	mediaDataPerAsrSessionMin	ASR	Минимальный размер голосового трафика (байт)
.1.3.6.1.4.1.45373.5.1.1.14.0	mediaDataPerAsrSessionMax	ASR	Максимальный размер голосового трафика (байт)
Соответствия нет	occupiedOfflineAsrEngines	ASR Offline	Количество занятых движков
Соответствия нет	freeOfflineAsrEngines	ASR Offline	Количество свободных движков
.1.3.6.1.4.1.45373.5.1.1.15.0	occupiedTtsEngines	TTS	Количество занятых движков
.1.3.6.1.4.1.45373.5.1.1.16.0	freeTtsEngines	TTS	Количество свободных движков
.1.3.6.1.4.1.45373.5.1.1.17.0	successfulTtsSessionDuration	TTS	Средняя длина сессии синтезированной речи
.1.3.6.1.4.1.45373.5.1.1.18.0	successfulTtsSessionDurationMin	TTS	Минимальная длина сессии синтезированной речи
.1.3.6.1.4.1.45373.5.1.1.19.0	successfulTtsSessionDurationMax	TTS	Максимальная длина сессии синтезированной речи
.1.3.6.1.4.1.45373.5.1.1.20.0	ttsErrors	TTS	Количество ошибок при синтезе речи
.1.3.6.1.4.1.45373.5.1.1.21.0	ttsSessions	TTS	Общее количество сессий с синтезом речи
.1.3.6.1.4.1.45373.5.1.1.22.0	ttsRequestsPerSecond	TTS	Количество запросов в секунду
.1.3.6.1.4.1.45373.5.1.1.23.0	mediaDataPerTtsSession	TTS	Средний размер голосового трафика (байт)
.1.3.6.1.4.1.45373.5.1.1.24.0	mediaDataPerTtsSessionMin	TTS	Минимальный размер голосового трафика (байт)
.1.3.6.1.4.1.45373.5.1.1.25.0	mediaDataPerTtsSessionMax	TTS	Максимальный размер голосового трафика (байт)



Внимание! Ровно в полночь все счётчики обнуляются.

8.3.3 Настройка конфигурации SNMP Trap

Ловушка SNMP (англ. SNMP-trap) – это особый сигнал, отправляемый устройством с поддержкой протокола SNMP.

Как правило, подобные сигналы отправляются устройствами для оповещения администратора сети о наступлении критических событий.

Например, может быть отправлен SNMP-trap в случае обнаружения ошибки. Подобные ситуации требуют незамедлительного вмешательства обслуживающего персонала, и поэтому устройство/ПО само инициирует отправку сигнала по протоколу SNMP. Параметры передаются в виде критических ошибок.

Отображение данных происходит в:

- Windows Event Logs;
- Unix syslog.

ResourceServerConfig.xml расположен в следующих каталогах:

- Linux: **/opt/RS/Config/ResourceServerConfig.xml**.
- Windows: **C:\Program Files\Speech Technology Center\Resource Server\Config\ResourceServerConfig.xml**.

Секция журналирования **<snmpTraps version="2" />** предназначена для указания версии протокола SNMP-Trap (1,2).

ResourceServerSnmpAgent.conf расположен в следующих каталогах:

- Linux: **/usr/local/etc/snmp/ResourceServerSnmpAgent.conf**.
- Windows: **C:\usr\etc\snmp\ResourceServerSnmpAgent.conf**.

Секция **Trap Destinations** предназначена для указания адреса отправки:

```
#####
# SECTION: Trap Destinations
#
# Here we define who the agent will send traps to.

# trapsink: A SNMPv1 trap receiver
# arguments: host[:portnum] [community]

trapsink 192.168.1.1:162

# trap2sink: A SNMPv2c trap receiver
# arguments: host[:portnum] [community]

trap2sink 192.168.1.1:162 public

# informsink: A SNMPv2c inform (acknowledged trap) receiver
# arguments: host[:portnum] [community]

informsink 192.168.1.1:162
```

9. ЗАПУСК И ОСТАНОВКА СЛУЖБ



Внимание! При перезагрузке сервисов все текущие соединения будут потеряны.

9.1 Запуск и остановка служб в среде ОС Windows



Если после установки всех программных компонентов системы **VoiceNavigator** возникли какие-либо проблемы, перезапустите службы **STC MRCPServer** и **STC ResourceServer**. Служба **STC Srrs2Nlsmi** не является обязательной.

Для запуска служб откройте оснастку ОС Windows **Панель управления > Администрирование > Службы**. Найдите вышеуказанные службы и поочередно выполните для каждой из них команду **Запустить**. Данную команду можно выполнить с помощью контекстного меню (рис. 52) или кнопки **Запуск службы**, расположенной на панели инструментов.

Службы должны запуститься без сообщений об ошибке и остаться в состоянии **Работает**.

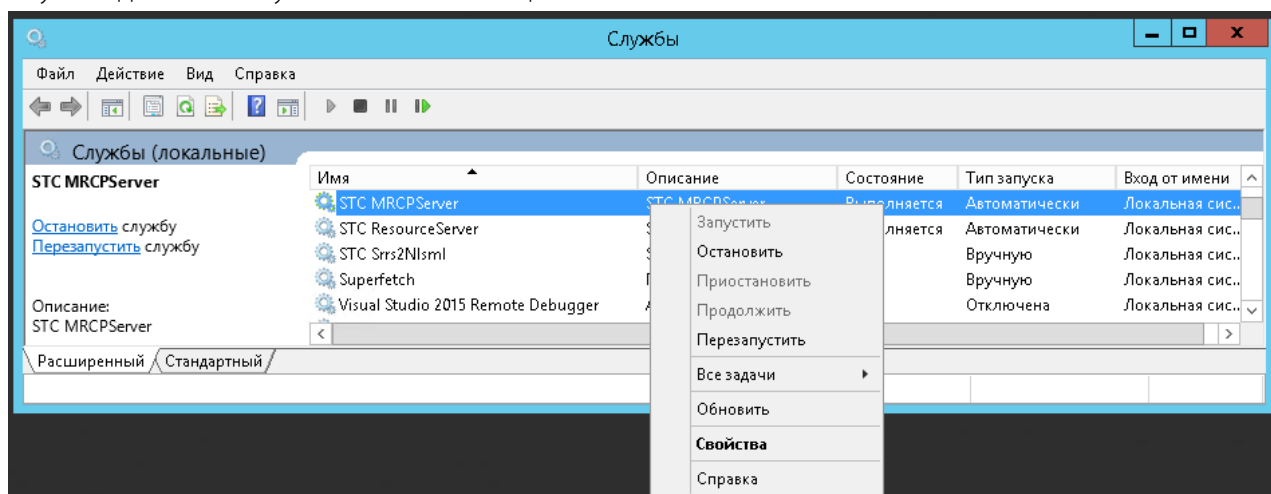


Рисунок 52 – Окно «Службы» операционной системы. Запуск службы

Для остановки служб **STC MRCPServer** и **STC ResourceServer** откройте оснастку ОС Windows **Панель управления > Администрирование > Службы**. Найдите вышеуказанные службы и поочередно выполните для каждой из них команду **Остановить**. Данную команду можно выполнить с помощью контекстного меню (рис. 52) или кнопки **Остановка службы**, расположенной на панели инструментов.

9.2 Запуск и остановка служб в среде ОС Linux

Запуск служб в среде ОС Linux осуществляется с помощью выполнения следующих команд:

```
systemctl start stcRS (запускается в первую очередь)
```

```
systemctl start stcMRCP (запускается во вторую очередь)
```

Чтобы остановить службы, выполните следующие команды:

```
systemctl stop stcMRCP (остановить в первую очередь)
```

```
systemctl stop stcRS (остановить во вторую очередь)
```

Чтобы перезапустить службы, выполните следующие команды:

```
systemctl restart stcMRCP (перезапустить в первую очередь)
```

```
systemctl restart stcRS (перезапустить во вторую очередь)
```

Для проверки статуса служб выполните команду:

```
systemctl status stcMRCP stcRS
```

10. ПРОТОКОЛИРОВАНИЕ СОБЫТИЙ

10.1 Общие положения

Для настройки параметров и сохранения файлов протоколирования событий, происходящих при функционировании программных средств **VoiceNavigator**, необходимо поставить соответствующий уровень протоколирования в конфигурационных файлах **STC MRCP Server** и **STC Resource Server**.

10.2 Изменение уровней протоколирования событий

Возможны следующие значения уровней протоколирования:

Параметр	Описание
OFF	Протоколирование отключено
FATAL	Протоколируются только критические ошибки
ERROR	Протоколируются все ошибки
WARN	Протоколируются все ошибки и предупреждения
INFO	Протоколируются информационные сообщения
DEBUG	Протоколируется отладочная информация
TRACE	Протоколируется дополнительная отладочная информация

10.3 Конфигурирование хранения log-файлов

Хранение файлов протоколирования событий настраивается в конфигурационных файлах **STC MRCP Server** и **STC Resource Server** (описание приведено в соответствующих разделах).

10.4 Структура пути к артефактам

Путь к артефактам для **STC ASR**:

base/CompanyName/ApplicationName/yyyy/mmMonth/dd/hh/MM/appsessionid/recognizeXXX.

Путь к артефактам для **STC TTS**:

base/CompanyName/ApplicationName/yyyy/mmMonth/dd/hh/MM /appsessionid/synthesisXXX.

где:

- **base** - путь для сохранения артефактов и записей диалога, определяется в конфигурационном файле ArtifactStorageConfig;
- **CompanyName** – имя компании;
- **ApplicationName** – название приложения;
- **yyyy** – год протоколирования;
- **mmMonth** – месяц протоколирования, например, 09September;

- **dd** – день протоколирования;
- **hh** – час протоколирования;
- **MM** – минута протоколирования;
- **appid** – уникальный идентификатор сессии;
- **recognizeXXX** - папка с артефактами распознавания, где XXX - порядковый номер распознавания в сессии;
- **synthesisXXX** - папка с артефактами синтеза, где XXX - порядковый номер синтеза в сессии.

10.5 Формирование в артефактах NLSML-ответа

По умолчанию файлы в виде **nlsml** не сохраняются. Файл формируется при запуске утилиты **MRCP Server\Bin\Srrs2Nlsml.exe** с указанием пути к артефактам.

Формат:

```
"C:\Program Files\Speech Technology Center\MRCP Server\Bin\Srrs2Nlsml.exe"  
srrs2nlsml path=storage_path|endpoint=listen_endpoint [usesingleinput=1][sendnomatch=1]  
[writeutterancesnr=1][forcenbestlength=0]  
path      starting point of scanning for srrs files to convert  
endpoint  service endpoint to listen in format ip:port
```

Пример для Windows:

```
Srrs2Nlsml.exe path="C:\Program Files\Speech Technology Center\Resource Server\Data\artifacts"
```

Пример для Linux:

```
/usr/bin/mono /usr/lib/mono/4.5/mono-service.exe Srrs2Nlsml.exe path=/opt/RS/Data/Artifacts/
```

11. ТЕХНИЧЕСКАЯ ПОДДЕРЖКА

11.1 Порядок обработки обращений в техническую поддержку вендора

Порядок обработки обращений в техническую поддержку регулируется договором о технической поддержке.

Сервисное обслуживание направлено на оказание содействия по поддержанию работоспособности, либо, в случае возникновения отказов, оказание содействия по восстановлению работоспособности.

Сервисное обслуживание предоставляется следующими способами:

- Телефонный номер «горячей линии» службы сервиса для обращений круглосуточно, включая выходные и праздничные дни: **+7(812)325-88-48, доб.2.**
- Адрес электронной почты «горячей линии» службы сервиса: **support@speechpro.com.**
- Интернет адрес электронной системы автоматического приема запросов на техническую поддержку «HelpDesk»: **<http://jira.speechpro.com/>.**

Каждому запросу, направленному исполнителю, присваивается уникальный номер, соответствующий приоритет и назначается специалист, который занимается проблемой на приоритетной основе и решает её в указанные ниже сроки. Приоритет запроса заказчиком определяется самостоятельно в соответствии со степенью важности возникшей проблемы. Параметры предоставления сервисного обслуживания оцениваются в зависимости от приоритета запроса.

11.2 Передача данных для поддержки VoiceNavigator из Windows

Для получения данных о поддержке программных средств **VoiceNavigator** откройте компонент операционной системы **Программы и компоненты**.

В окне **Программы и компоненты** (рис. 53) выберите программу из состава программных средств **VoiceNavigator**.

В нижней части окна отобразится логотип ООО «ЦРТ», номер версии программы и адрес в интернете.

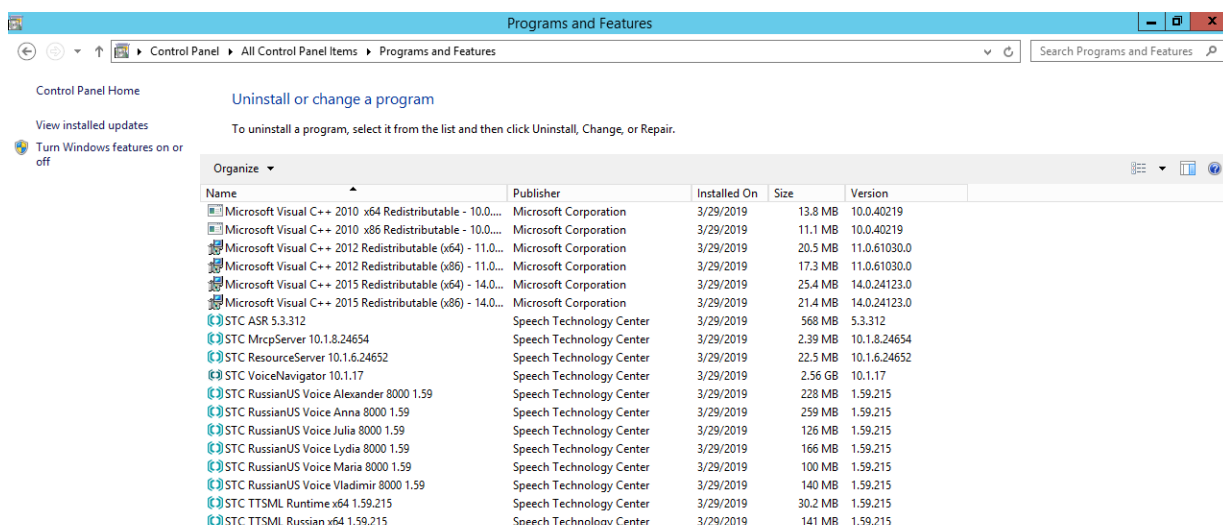


Рисунок 53 – Окно получения данных о программе

Для получения данных об операционной системе, типе компьютера и его конфигурации откройте информацию о системе на контрольной панели (рис. 54).

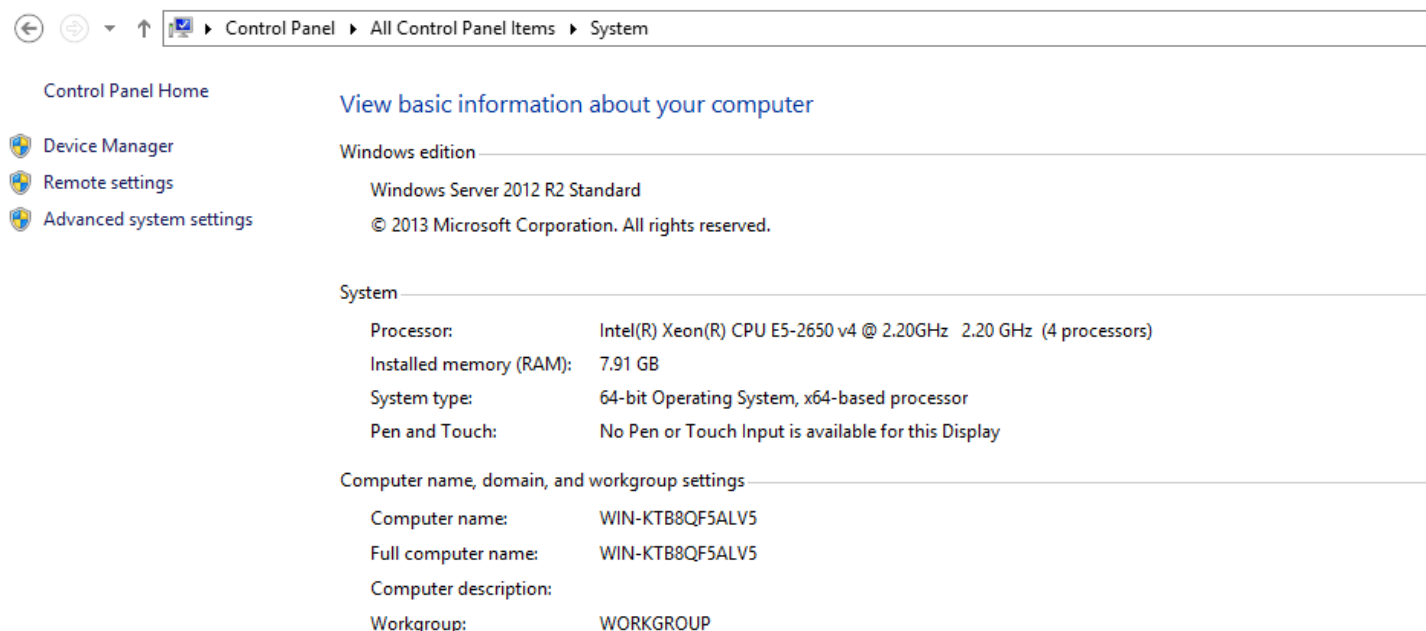


Рисунок 54 – Информация об ОС

В качестве дополнительной информации уточните:

- вид сервера (физический или виртуальный);
- объем жёсткого диска, в том числе свободного.

Для получения log-файлов заархивируйте данные из каталогов:

C:\Program Files\Speech Technology Center\Resource Server

- Logs\ *
- Config\ *
- Version.txt.

C:\Program Files\Speech Technology Center\MRCP Server

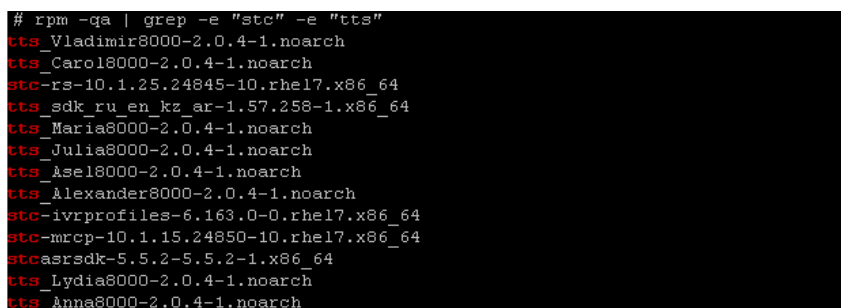
- Logs\ *

- Config*
- Version.txt.

11.3 Передача данных для поддержки VoiceNavigator из Linux

Для получения данных о поддержке программных средств **VoiceNavigator** используйте команду (рис. 55):

```
rpm -qa | grep -e "stc" -e "tts"
```



```
# rpm -qa | grep -e "stc" -e "tts"
tts_Vladimir8000-2.0.4-1.noarch
tts_Carol8000-2.0.4-1.noarch
stc_rs-10.1.25.24845-10.rhel7.x86_64
tts_sdk_ru_en_kz_ar-1.57.258-1.x86_64
tts_Maria8000-2.0.4-1.noarch
tts_Julia8000-2.0.4-1.noarch
tts_Asel8000-2.0.4-1.noarch
tts_Alexander8000-2.0.4-1.noarch
stc_ivrprofiles-6.163.0-0.rhel7.x86_64
stc_mrccp-10.1.15.24850-10.rhel7.x86_64
stcasrsdk-5.5.2-5.5.2-1.x86_64
tts_Lydia8000-2.0.4-1.noarch
tts_Anna8000-2.0.4-1.noarch
```

Рисунок 55 – Получение данных о поддержке программных средств

Для получения данных об операционной системе, типе компьютера и его конфигурации используйте следующие команды:

- lscpu – информация по процессорам;
- lshw – информативное описание компонентов вашего сервера, в том числе CPU, памяти, USB/NIC, аудио и прочих (при необходимости установите пакет);
- uname -r – версия ядра Linux;
- uname -a – полная информация о ядре;
- df -H – информация о пространстве жесткого диска;
- ulimit -a – информация о лимитах открытых файлов;
- cat /etc/profile – информация о настройках профайла;
- echo \$PATH – информация о прописанных путях переменных;
- cat /etc/redhat-release – версия дистрибутива для ОС, основанных на Red Hat Enterprise Linux.

Для получения log-файлов заархивируйте данные из каталогов:

/opt/RS/

- Logs*
- Config*
- Version.txt

/opt/MRCP/

- Logs*
- Config*
- Version.txt

11.4 Требования к персоналу для обслуживания VoiceNavigator

Для обеспечения обслуживания **VoiceNavigator** заказчиком требуются следующие роли:

Роль	Обязанности	Оценка временных затрат
Сотрудник технической поддержки заказчика (предполагается, что уже находится в штате компании)	Мониторинг системы, обеспечение реакции на аварийные события.	Распределяется между всеми проектами равномерно. Сотрудники технической поддержки могут обеспечить 1 и 2 уровень поддержки.
Администратор VoiceNavigator (может выполнять существующий сотрудник, например, администратор голосового портала)	Конфигурирование ПО, выкладывание графов слитного распознавания Профилактические работы, связанные с работой системы распознавания (ОС, ПО VoiceNavigator).	Согласно регламенту, обычно не более 1 часа в неделю.
Разработчик голосовых грамматик (может выполнять сотрудник по разработке голосовых приложений)	Разработка грамматик и их усовершенствование: загрузка артефактов ПО, прослушивание аттерансов, диктование, сборка грамматик, проверка изменения качества работы, передача графов администратору для выкладывания на сервер.	Обычно не учитывается в проектах поддержки, учитывается в проектах разработки и поддержки голосовых сервисов.
Администратор безопасности (предполагается, что уже находится в штате компании)	Профилактические работы по мониторингу безопасности.	Согласно регламенту, обычно не более 1 часа в месяц.

Для работы с **VoiceNavigator** компания ООО «ЦРТ» предлагает следующие курсы:

Наименование	Направленность	Длительность	Примечание
Администрирование ПО VoiceNavigator (установка и обслуживание)	Для инженеров	1 день	Не является обязательным, может выполняться самостоятельно по представленной документации.
Разработка голосовых грамматик	Для разработчиков (разработка и усовершенствование)	2 дня	Не является обязательным, может выполняться самостоятельно по представленной документации и консультациям.

11.5 Рекомендации по обновлению VoiceNavigator

ПО **VoiceNavigator** поддерживает возможность обновления системы на более новую версию.

Смена версии не рекомендуется при выпуске версий, функциональность которых не улучшает качество или функционал имеющегося сервиса. Смена версии рекомендуется при следующих случаях:

- наличие ошибок в ПО;
- наличие ошибок в безопасности ПО;
- улучшение функционала, влияющего на работу сервисов (например, улучшения качества распознавания);
- выпуск мажорной версии (например, обновлении с версии 7 на 8).

Порядок обновления ПО **VoiceNavigator**:

- заказчик производит установку ПО в тестовой зоне (по стандартной или специальной инструкции) и проводит тестирование согласно оговорённой методике;
- при необходимости заново собираются графы слитного распознавания;
- при успешном тестировании заказчик начинает обновление на одном из регионов, для этого:
 - заказчик в час наименьшей нагрузки производит обновление ПО (по стандартной или специальной инструкции);
 - обновленный сервер подключается к голосовому portalу, проверяется прохождение трафика;
 - в случае успеха по такой же схеме обновляются другие сервера и регионы;
 - при сбое происходит откат на предыдущую версию ПО, а log-файлы с описанием отправляются исполнителю.

11.6 Рекомендации по проведению сервисных и регламентных работ

Оптимальный график проведения сервисных и регламентных работ представлен в виде перечня регламентных работ и технологической карты для проведения сервисных работ.

Сборник технологических карт содержит пакет документации, являющийся основой для технической эксплуатации решения.

Перечень регламентных работ включает:

- Обслуживание серверов для поставляемых систем (предполагается, что обеспечивается заказчиком).
- Обслуживание системы виртуализации (предполагается, что обеспечивается заказчиком).
- Обслуживание хранилища (предполагается, что обеспечивается заказчиком).
- Обслуживание КЦ (предполагается, что обеспечивается заказчиком).
- Обслуживание Базы данных для ИС (предполагается, что обеспечивается заказчиком).
- Обслуживание Информационных систем (предполагается, что обеспечивается заказчиком).
- Обслуживание Системы распознавания и синтеза речи.
- Обслуживание Системы интеллектуальной маршрутизации\классификации (опционально).

Технологические карты устанавливают объемы и порядок выполнения основных работ при техническом обслуживании решения.

В технологических картах приведены данные по периодичности выполнения работ, количеству и квалификации пользователей.

По результатам регламентных работ, выполненных в соответствии с настоящими технологическими картами, сторонами составляется план-график проведения работ.

На работы, выполнение которых не предусмотрено настоящими технологическими картами, могут быть разработаны местные технологические карты, утверждаемые в установленном порядке.

Настоящие технологические карты могут быть использованы для расчета трудоемкости работ, составления производственных заданий отдельным сотрудникам.

Технологическая карта №1:

	Технологическая карта №1	
	Система распознавания и синтеза речи	
Наименование работы	Периодичность	Исполнители
Еженедельная проверка работоспособности системы распознавания	1 раз в неделю	Администратор системы

Необходимые принадлежности: компьютер с доступом в систему.

Выполните следующие действия:

- с помощью логина и пароля зайдите в терминальное окно сервера (по протоколу RDP или SSH);
- убедитесь, что на экране отсутствуют ошибки;
- проверьте корректную установку даты и времени;
- убедитесь через службу сервисов, что сервисы (MRCP, TTS, SNMP, Sentinel) работают;
- убедитесь, что в системных log-файлах отсутствуют ошибки;
- убедитесь, что нагрузка на процессор, память и свободное место не менее 20%;
- убедитесь, что в папках ПО **VoiceNavigator** формируются log-файлы в соответствии с рекомендуемой настройкой и отсутствуют ошибки в работе ПО;
- произведите выход из терминального окна;
- повторите со всеми системами.

Технологическая карта №2:

	Технологическая карта №2	
	Система интеллектуальной маршрутизации	
Наименование работы	Периодичность	Исполнители
Еженедельная проверка работоспособности системы интеллектуальной маршрутизации	1 раз в неделю	Администратор системы

Необходимые принадлежности: компьютер с доступом в систему, телефон.

Выполните следующие действия:

- зайдите в терминальное окно сервера (по протоколу RDP или SSH);
- убедитесь, что на экране отсутствуют ошибки;

- проверьте корректную установку даты и времени;
- убедитесь, что в системных log-файлах отсутствуют ошибки;
- убедитесь, что нагрузка на процессор, память и свободное место не менее 20%;
- убедитесь через службу сервисов, что сервисы (Apache Tomcat, MS SQL, SNMP) работают;
- убедитесь через «Проводник», что в папке ПО Apache Tomcat формируются log-файлы в соответствии с рекомендуемой настройкой и отсутствуют ошибки в работе ПО;
- произведите выход из терминального окна;
- откройте веб-интерфейс интеллектуальной маршрутизации;
- проверьте функциональность работы интерфейса и статистики;
- позвоните по номеру системы;
- проведите тестирование в соответствии с картой тестирования;
- просмотрите наличие записей о вызове;
- произведите выход из веб-интерфейса интеллектуальной маршрутизации;
- проверьте наличие резервированных данных;
- повторите со всеми системами.

Технологическая карта №3:

	Технологическая карта №3	
	Система интеллектуальной маршрутизации	
Наименование работы	Периодичность	Исполнители
Дневная проверка работоспособности системы	2 раз в день	Администратор системы

Необходимые принадлежности: компьютер с доступом в систему.

Выполните следующие действия:

- откройте веб-интерфейс интеллектуальной маршрутизации;
- постройте статистику за предыдущие 6 часов;
- сравните статистику с типовой;
- произведите выход из веб-интерфейса интеллектуальной маршрутизации.

Технологическая карта №4:

	Технологическая карта №4	
	Мониторинг решения	
Наименование работы	Периодичность	Исполнители
Круглосуточный мониторинг решения	24ч	Инженер техподдержки

Необходимые принадлежности: компьютер с доступом в систему мониторинга.

Выполните следующие действия:

- откройте интерфейс мониторинга;
- проверьте записи об ошибках;
- убедитесь, что нагрузка на процессор, память и свободное место каждого сервера не менее 20%.

Технологическая карта №5:

	Технологическая карта №5	
	Тюнинг решения	
Наименование работы	Периодичность	Исполнители
Усовершенствование решения	Раз в полгода	Инженер-разработчик

Необходимые принадлежности: компьютер с доступом в систему.

Выполните следующие действия:

- зайдите в систему интеллектуальной маршрутизации;
- сформируйте отчет за заданный период (1-2 месяца);
- откройте интерфейс списка обращений за заданный период (1-2 месяца);
- прослушайте записи с целью убедиться в корректности работы системы распознавания и определения тематик;
- внесите изменения в настройки тематик;
- внесите изменения в граф слитного распознавания и грамматики;
- проверьте звонком правильность работы;
- подождите несколько часов (4-8 часов);
- сформируйте отчет с целью убедиться в улучшениях.

Технологическая карта №6:

	Технологическая карта №6	
	Обновление ПО	
Наименование работы	Периодичность	Исполнители
Обновление ПО	По выходу нового ПО	Инженер-разработчик

Необходимые принадлежности: компьютер с доступом в систему.

Выполните следующие действия:

- получите новое ПО;
- установите в тестовой зоне и проверьте работоспособность;
- согласуйте план работ;

- действуйте по регламенту.

Рекомендуется следующий график проведения сервисных и регламентных работ:

Наименование	График проведения	Описание
Мониторинг работы серверов системы распознавания	1 час в неделю	Сервисные работы, связанные с работой ОС, загрузкой процессоров, памяти на серверах с ПО VoiceNavigator (системами распознавания и синтеза речи)
Мониторинг работы системы интеллектуальной маршрутизации	1 час в неделю	Сервисные работы, связанные с работой ОС, загрузкой процессоров, памяти на серверах системы интеллектуальной маршрутизации, интерфейсов отчётности и администрирования
Мониторинг работы системы безопасности	Согласно регламенту ИБ	Сервисные работы, связанные с анализом логов ОС и веб-сервисов
Обновление системы распознавания	Раз в полгода или реже при необходимости	Обновление ПО VoiceNavigator согласно регламенту
Обновление системы интеллектуальной маршрутизации	При выпуске обновлений	Обновление системы интеллектуальной маршрутизации согласно регламенту
Регламентная проверка ИБ	Согласно регламенту ИБ	Проверка систем безопасности согласно регламенту
Регламентные работы по усовершенствованию системы распознавания	Первый год: 1 раз в 6 месяцев Далее: 1 раз в 12 месяцев	Проверка качества распознавания. Улучшение системы распознавания
Регламентные работы по тюнингу системы интеллектуальной маршрутизации	Первый год: 1 раз в 6 месяцев Далее: 1 раз в 12 месяцев	Проверка качества диалогов. Совершенствование логики обработки, маршрутизации, диалогов

12. ПРИЛОЖЕНИЕ А. ССЫЛКИ НА ОПИСАНИЯ ПРОТОКОЛОВ И СТАНДАРТОВ

Протокол	Ссылка
MRCP 1.0 – [RFC 4463]	http://tools.ietf.org/html/rfc4463
MRCP 2.0 – [draft-ietf-speechsc-mrcpv2]	http://tools.ietf.org/wg/speechsc/draft-ietf-speechsc-mrcpv2/
RTSP – [RFC 2326]	http://tools.ietf.org/html/rfc2326
SIP – [RFC 3261]	http://tools.ietf.org/html/rfc3261
SDP – [RFC 4566]	http://tools.ietf.org/html/rfc4566
RTP – [RFC 3550]	http://tools.ietf.org/html/rfc3550
VoiceXML 3.0	http://www.w3.org/TR/voicexml30/
SRGS	http://www.w3.org/TR/speech-grammar/
NLSML	http://www.w3.org/TR/nl-spec/
SSML	http://www.w3.org/TR/speech-synthesis/
PLS	http://www.w3.org/TR/pronunciation-lexicon/

13. ПРИЛОЖЕНИЕ Б. ПЕРЕЧЕНЬ ТЕРМИНОВ И СОКРАЩЕНИЙ

В настоящем документе применяют следующие термины и сокращения:

Термин или сокращение	Определение
<i>API</i>	(англ. Application Programming Interface) – интерфейс программирования приложений - набор готовых классов, процедур, функций, структур и констант, предоставляемых приложением (библиотекой, сервисом), для использования во внешних программных продуктах.
<i>ASR</i>	(англ. Automatic Speech Recognition) – автоматическое распознавание речи.
<i>DTMF</i>	(англ. Dual-Tone Multi-Frequency) – двухтональный многочастотный аналоговый сигнал.
<i>HASP</i>	(англ. Hardware Against Software Piracy) – аппаратно-программная система защиты программ и данных от нелегального использования и несанкционированного распространения.
<i>ISDN</i>	(англ. Integrated Services Digital Network) – цифровая сеть с интеграцией обслуживания.
<i>ISUP</i>	(англ. ISDN User Part) – Прикладная часть протокола ISDN.
<i>IVR</i>	(англ. Interactive Voice Response) – система, предоставляющая голосовой интерфейс для диалога (голосовое приложение или сервис) с пользователем в целях выполнения задач маршрутизации звонков или системы голосового самообслуживания (СГС). При этом управление (навигация в меню) может быть посредством распознавания голоса (ASR) или тонального набора. Озвучивание информации осуществляется ранее записанными фразами или с помощью синтеза речи (TTS).
<i>MRCP</i>	(англ. Media Resource Control Protocol) – протокол управления медиаресурсами.

<i>NLP</i>	(англ. Natural Language Processing) – общее направление искусственного интеллекта и математической лингвистики. Оно изучает проблемы компьютерного анализа и синтеза естественных языков. Применительно к искусственному интеллекту анализ означает понимание языка, а синтез – генерацию грамотного текста.
<i>NLSML</i>	(англ. Natural Language Semantics Markup Language) – язык разметки для описания результата распознавания речи.
<i>PBX</i>	(англ. Private Branch Exchange) – телефонная станция, обслуживающая государственные или частные организации и расположенная в зданиях, в которых работает пользователь.
<i>PCI</i>	(англ. Peripheral component interconnect, взаимосвязь периферийных компонентов) – шина ввода/вывода для подключения периферийных устройств к материнской плате компьютера.
<i>PCM</i>	(англ. Pulse Code Modulation) – импульсно-кодовая модуляция сигнала.
<i>PLS</i>	(англ. Pronunciation Lexicon Specification) – язык описания словаря транскрипций для применения в программах распознавания и синтеза речи.
<i>RTF</i>	(англ. Real Time Factor) – показатель для измерения скорости работы системы автоматического распознавания речи.
<i>RTP</i>	(англ. Real Time Protocol) – протокол передачи медиаданных в режиме реального времени.
<i>RTSP</i>	(англ. Real Time Streaming Protocol) – потоковый протокол реального времени.
<i>SDP</i>	(англ. Session Description Protocol) – сетевой протокол, предназначенный для описания сессии передачи потоковых данных.
<i>SIP</i>	(англ. Session Initiation Protocol) – протокол прикладного уровня, установки соединения. Применяется для создания соединений по протоколу MRCP 2.0.
<i>SLA</i>	(англ. Service Level Agreement) – соглашение об уровне предоставления услуг (SL).

<i>SRGS</i>	(англ. Speech Recognition Grammar Specification) – формат описания грамматики для распознавания речи (имеются в виду закрытые грамматики).
<i>SRRS</i>	(англ. Speech Recognition Result Specification) – формат представления результата распознавания речи и тонов DTMF в звуковом сигнале по грамматикам SRGS.
<i>SSML</i>	(англ. Speech Synthesis Markup Language) – язык разметки для описания текста в системах синтеза речи.
<i>TCP/IP</i>	(англ. Transmission Control Protocol/Internet Protocol) – набор сетевых протоколов разных уровней используемых в сетях.
<i>TTS</i>	(англ. Text-to-Speech) – синтез речи из текстовых данных.
<i>UDP</i>	(англ. User Datagram Protocol) – транспортный протокол для передачи данных в сетях IP без установления соединения.
<i>VXML (VoiceXML)</i>	(англ. Voice Extensible Markup Language) – открытый стандартный язык разметки для голосовых приложений. Язык разработан для создания сценариев звуковых диалогов.
<i>Абонент</i>	Клиент центра обслуживания вызовов.
<i>Администратор ИБ</i>	Должностное лицо организации, специалист, отвечающий за информационную безопасность.
<i>Голосовая платформа (VP, Voice Portal\Platform)</i>	Программное обеспечение, обеспечивающее взаимодействие голосовых каналов (обычно sip), серверов приложений (обеспечивают логику сценария, обычно VXML или API), серверов распознавания и синтеза речи (обычно по протоколу MRCP или Web API) в целях автоматизации обращений в голосовом канале.
<i>Директивное высказывание</i>	Высказывание, которое объясняет абоненту, что именно он должен произнести или написать, в зависимости от определённого условия. Пример директивного высказывания: «Если вы хотите купить билет, скажите «билет».

<i>Закрытая грамматика</i>	Метод распознавания речи, основанный на распознавании голосовой команды, состоящей из отдельного слова или фразы. Система, использующая закрытую грамматику, не анализирует реплику целиком, а осуществляет лишь поиск голосовой команды в реплике. Работа такой системы строится на вопросах, требующих от абонента однозначного ответа, в котором должна содержаться голосовая команда.
<i>Информационная система</i>	Любая внешняя (с точки зрения чат-платформы) система, содержащая информацию, которая используется при автоматизации обращений.
<i>Инцидент</i>	Любое событие, не являющееся частью стандартных операций по предоставлению услуги, которое привело или может привести к нарушению или снижению качества этой услуги. В контексте данного документа – событие, не описанное в рамках стандартных сценариев поведения системы.
<i>Канал</i>	Совокупность технических и программных средств необходимых для реализации единичного обращения абонента к центру обслуживания вызовов.
<i>Клиент</i>	Пользователь услуг Заказчика, в том числе по договору об оказании услуг.
<i>Ключевой показатель эффективности</i>	(англ. Key Performance Indicator) – количественно измеримый индикатор фактически достигнутого результата.
<i>Контент-менеджер</i>	Должностное лицо организации, специалист ответственный за наполнение баз данных, контента для голосовых, чат или омниканальных сервисов.
<i>Микрофонный канал</i>	Обращения, в которых общение с системой обеспечивается непосредственно через микрофон устройства (гарнитура, микрофон телефона и пр.). Характеризуется широкой полосой передачи голоса (16 бит).
<i>Оператор</i>	Сотрудник контактного центра заказчика, общающийся с абонентом и осуществляющий его обслуживание.

ОПО	(Общее программное обеспечение) – часть программного обеспечения, разработанного вне связи с созданием системы данного продукта. ОПО представляет собой совокупность программ общего назначения, предназначенных для организации вычислительного процесса и решения часто встречающихся задач обработки информации.
Открытая грамматика	Метод распознавания речи, основанный на обработке всех слов реплики абонента, выраженной в свободной форме речи или письма. Система, использующая открытую грамматику, осуществляет семантический и грамматический разборы каждой реплики абонента для извлечения из неё смысла.
ПО	Программное обеспечение.
Пользователь	Сотрудник компании, участвующий в бизнес-процессах обслуживания клиентов.
Поток звонка	Правила обработки запроса абонента в соответствии с его репликой.
Разработчик	Должностное лицо организации, специалист по разработке приложений. В зависимости от специализации может различаться на разработчика мобильных, веб-, голосовых и прочих приложений, а также грамматики.
Реплика	Составная часть диалога. Ответ на слова собеседника.
СГС	Система голосового самообслуживания (взаимодействия) – система IVR, выполняющая функцию маршрутизации звонков внутри центра обслуживания вызовов, на основе информации, вводимой абонентом с помощью тонального набора или голоса.
Семантическая онтология	Структура, концептуальная схема, описывающая и формализующая значения элементов некоторой предметной области. Онтология состоит из набора терминов и правил описывающих их связи, отношения.
Сервер обработки	Программное обеспечение, обеспечивающее возможности сбора данных, передачи потока речи в реальном режиме от клиентских приложений, сбора и передачи аудиофайлов, организации очередей и балансировки нагрузки.

Сессия	Период между первым обращением пользователя и истечением установленного таймута последнего сообщения. Сессия в чат-платформе не может быть более 24 часов.
Система	Комплекс средств автоматического обслуживания.
Система распознавания и синтеза речи	Программное обеспечение, позволяющее добавлять функционал распознавания и синтеза речи в существующие или новые решения, диалоговые и интерактивные сервисы.
Системный администратор	Должностное лицо организации, специалист по обслуживанию программных и технических средств отвечающий за работу системы в штатном режиме.
СПО	Специальное программное обеспечение.
Телефонный канал	Обращения, в которых общение с системой обеспечивается через телефонные каналы связи. Характеризуется ограниченной полосой передачи голоса (8 бит).
Транскрибирование	Перевод (распознавание) аудиоданных в текст.
ТфОП	Телефонная сеть связи общего пользования.
Центр обслуживания (ЦОВ, КЦ)	Комплекс оборудования информационного обслуживания абонентов (РД 45.191-2001). Является аналогом так называемого «Call Centre» – центра обслуживания телефонных соединений. Подразумевается набор программно-аппаратных средств, предназначенных для обслуживания обращений в различных каналах.
ЦОД	Центр обработки данных. Подразумевается место размещения оборудования в зданиях с выполнением жёстких требований (по безопасности, отказоустойчивости, кондиционированию, электропитанию, каналам связи и пр.).
Широковещательный адрес	Условный (не присвоенный никакому устройству в сети) адрес, который используется для передачи широковещательных пакетов в компьютерных сетях.