Федеральное агентство связи

федеральное государственное бюджетное образовательное учреждение

высшего образования

«Сибирский государственный университет телекоммуникаций

и информатики»

(СибГУТИ)

Кафедра БиУТ

Допустить к защите зав. кафедрой

/С.Н. Новиков /

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА СПЕЦИАЛИСТА

Разработка проекта защищенного DWH (Data Warehouse)

Пояснительная записка

Студент _____ / Д.А. Федоров /

Факультет АЭС Группа АБ-56

Руководитель _____ / О.И. Солонская /

Консультанты:

– по экономическому обоснованию

_____/_____/

– по безопасности жизнедеятельности

_____/ _____/

Рецензент: _____/ _____/

Новосибирск 2021

Федеральное агентство связи

федеральное государственное бюджетное образовательное учреждение

высшего образования

«Сибирский государственный университет телекоммуникаций и информатики»

(СибГУТИ)

КАФЕДРА

Безопасность и управление в телекоммуникациях

ЗАДАНИЕ

НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ СПЕЦИАЛИСТА

СТУДЕНТА Д.А. Федорова ГРУППЫ АБ-56

«УТВЕРЖДАЮ»

« 28 » июля 2020 г.

Зав. кафедрой БиУТ

_____/ С.Н. Новиков /

Новосибирск 2020

1. Тема выпускной квалификационной работы специалиста:

Разработка проекта защищенного DWH (Data Warehouse)

утверждена приказом по университету от « 28 » июля 2020 г. № 4/1011о-20

2. Срок сдачи студентом законченной работы « 15 » января 2021 г.

3. Исходные данные по проекту (эксплуатационно-технические данные, техническое задание):

Концептуальная модель DWH

Схема построения базы данных

Методика определения угроз безопасности информации в информационных системах ФСТЭК России

Заказчик DWH: Авиакомпания

4. Содержание расчетно-пояснительной записки (перечень подлежащих разработке вопросов)	Сроки выполнения по разделам
Введение	13.09.2020 г.
1. Анализ подходов к проектированию Data warehouse	11.10.2020 г.
2. Анализ программного обеспечения для проектирования Data warehouse	08.11.2020 г.
3. Разработка проекта Data warehouse	06.12.2020 г.
4. Разработка проекта обеспечения информационной безопасности Data warehouse	13.12.2020 г.
5. Безопасность жизнедеятельности	20.12.2020 г.
6. Технико-экономическое обоснование работы	27.12.2020 г.
7. Заключение	07.01.2021 г.
8. Список литературы	09.01.2021 г.
9. Приложения	12.01.2021 г.

Консультанты по ВКР (с указанием относящихся к ним разделов):

1. Раздел по технико-экономическому обоснованию

_____

_____

_____

_____/_____/

2. Раздел по безопасности жизнедеятельности

_____

_____

_____

_____

_____/ _____/

Дата выдачи задания	Задание принял к исполнению
« 01 » сентября 2020 г.	« 01 » сентября 2020 г.
_____/ О.И. Солонская /	_____/ Д.А. Федоров /
(подпись, Ф.И.О. руководителя)	(подпись, Ф.И.О. студента)

Федеральное агентство связи

федеральное государственное бюджетное образовательное учреждение

высшего образования

«Сибирский государственный университет телекоммуникаций и информатики»

(СибГУТИ)

РЕЦЕНЗИЯ

на выпускную квалификационную работу студента Д.А. Федорова

по теме «Разработка проекта защищенного DWH (Data Warehouse)»

доц. каф. БиУТ Шевнина И. Е.

« 18 » января 2021 г.

С Рецензией ознакомлен /Д.А. Федоров/

« 18 » января 2021 г.

Федеральное агентство связи

федеральное государственное бюджетное образовательное учреждение

высшего образования

«Сибирский государственный университет телекоммуникаций и информатики»

(СибГУТИ)

ОТЗЫВ

о работе студента Д.А. Федорова в период подготовки выпускной квалификационной работы по теме «Разработка проекта защищенного DWH (Data Warehouse)»

Работа имеет практическую ценность		Тема предложена предприятием	
Работа внедрена		Тема предложена студентом	
Рекомендую работу к внедрению		Тема является фундаментальной	
Рекомендую работу к опубликованию		Рекомендую студента в магистратуру	
Работа выполнена с применением ЭВМ		Рекомендую студента в аспирантуру	

Руководитель выпускной квалификационной работы специалиста

Доц. каф. БиУТ, к.т.н. Солонская Оксана Игоревна

« 15 » января 2021 г.

С Отзывом ознакомлен /Д.А. Федоров/

« 15 » января 2021 г.

Приложение к Отзыву

Уровень сформированности компетенций у студента

Д.А. Федорова

Компетенции		Уровень сформированности компетенций		
		высокий	средний	низкий
1		2	3	4
Профессиональные	ПК-1 - способностью осуществлять анализ научно-технической информации, нормативных и методических материалов по методам обеспечения информационной безопасности телекоммуникационных систем			
	ПК-5 - способностью проектировать защищенные телекоммуникационные системы и их элементы, проводить анализ проектных решений по обеспечению заданного уровня безопасности и требуемого качества обслуживания, разрабатывать необходимую техническую документацию с учетом действующих нормативных и методических документов			
	ПК-7 - способностью осуществлять рациональный выбор средств обеспечения информационной безопасности телекоммуникационных систем с учетом предъявляемых к ним требований качества обслуживания и качества функционирования			
	ПК-12 - способностью выполнять технико-экономические обоснования, оценивать затраты и результаты деятельности организации в области обеспечения информационной безопасности			

АННОТАЦИЯ

Выпускной квалификационной работа студента Д.А. Федорова

по теме Разработка проекта защищенного DWH (Data WareHouse)

Объём работы – 92 страницы, на которых размещены 11 рисунков и 15 таблиц. При написании работы использовалось 34 источников.

Ключевые слова: Data warehouse, система управления базой данных, защита данных, ETL, политика безопасности.

Работа выполнена на: кафедре БиУТ СибГУТИ

Руководитель: доц. каф. БиУТ Солонская О.И.

Целью работы: Разработка проекта защищенного DWH (Data WareHouse)

Решаемые задачи: анализ подходов к проектированию Data warehouse, анализ программного обеспечения для проектирования Data warehouse, разработка проекта Data warehouse, разработка проекта обеспечения информационной безопасности Data warehouse, безопасность жизнедеятельности, технико-экономическое обоснование работы.

Основные результаты: спроектирован защищенный Data warehouse.

Graduation thesis abstract

of D.A. Fedorov on the theme Development of a secure DWH (Data Warehouse) project

The paper consists of 92 pages, with 11 figures and 15 tables/charts/diagrams. While writing the thesis 34 reference sources were used.

Keywords: Data warehouse, database management system, data protection, ETL, security policy.

The thesis was written at BIUT department SibSUTIS
(name of organization or department)
Scientific supervisor associate professor of the BiUT Solonskaya Oxana

The goal/subject of the paper is development of a secure DWH (Data Warehouse) project.

Tasks: analysis of approaches to designing a Data warehouse, analysis of software for designing a Data warehouse, development of a Data warehouse project, development of a project for ensuring information security of a Data warehouse, life safety, feasibility study of work.

Results: secured Data warehouse designed.

ОГЛАВЛЕНИЕ

Введение

Интеллектуальное хранение данных и их дальнейшее использование в современном мире уже стали неотъемлемой частью успешно развивающегося бизнеса. Для этой цели существуют корпоративные хранилища данных.

Поддержка и развитие хранилищ данных это сложная, требующая длительного времени и подверженная ошибкам задача. Основная причина заключается в том, что среда хранилища данных постоянно меняется, а оно само должно обеспечивать стабильный и согласованный интерфейс для доступа к информации, охватывающей различные периоды времени, в том числе 24/7.

Целью выпускной квалификационной работы является:

— анализ подходов к проектированию Data warehouse;

— анализ программного обеспечения для проектирования Data warehouse;

— разработка проекта Data warehouse;

— разработка проекта обеспечения информационной безопасности Data warehouse;

— безопасность жизнедеятельности;

— технико-экономическое обоснование работы.

В данной работе будет рассмотрено и спроектировано хранилище данных для авиакомпании с информацией, содержащей коммерческую тайну и персональные данные. Поскольку у авиакомпаний заведомо много источников данных за счет большого количества возможностей покупки билета и предоставления услуг, то главными критериями выбора для проектирования станут:

— скорость обработки данных;

— простота построения базы данных;

— наименьшее занимаемое дисковое пространство;

— возможность гибкого изменения структуры.

1 Анализ подходов к проектированию Data warehouse

1.1 Постановка задачи

Перед тем как начать проектирование Data warehouse, нужно проанализировать:

— содержание термина и применение Data warehouse;

— концептуальную модель Data warehouse;

— модель угроз;

— модель нарушителя.

1.2 Содержание термина и применение Data warehouse

Хранилище данных (Data warehouse) — это центральный репозиторий информации, который можно анализировать для принятия более обоснованных решений. Данные поступают в хранилище из транзакционных систем, реляционных баз данных и других источников, как правило с определенной периодичностью. Эти данные предварительно обрабатываются и загружаются в хранилище в ходе процессов извлечения, преобразования и загрузки, называемых ETL (Extract, Transform, Load). Бизнес-аналитики, специалисты по работе с данными и лица, ответственные за принятие решений, получают доступ к данным с помощью инструментов бизнес-аналитики, SQL(Structured query language) -клиентов и других приложений для аналитики[1].

Хранилище данных может содержать несколько баз данных. В каждой базе данных хранятся данные, упорядоченные по таблицам и столбцам. В каждом столбце можно определить описание данных: целые числа, поле данных, строка и т. д. Таблицы можно структурировать в схемы, которые во многом похожи на папки с файлами. После поступления данные хранятся в различных таблицах, описанных в этой схеме. С ее помощью инструменты запросов определяют, к каким таблицам данных следует обратиться для анализа[1].

Использование Data warehouse позволяет получить [1]:

— возможность принимать обоснованные решения;

— консолидация данных из множества источников;

— исторический анализ данных;

— высокое качество, непротиворечивость и точность данных;

— изолирование операций аналитики от транзакционных баз данных для повышения производительности обеих систем.

1.3 Концептуальная модель Data warehouse

Data warehouse имеет структуру состоящую из нескольких слоев или уровней, их можно воспринимать как отдельные компоненты системы – со своими задачами, зоной ответственности, правилами работы.

Уровневая архитектура – это средство борьбы со сложностью системы. Каждый последующий уровень абстрагирован от сложностей внутренней реализации предыдущего. Такой подход позволяет выделять однотипные задачи и решать их единообразным образом[2].

Схематично концептуальная архитектурная схема представлена на рисунке 1.1. Это упрощенная схема, которая отражает лишь ключевую идею.

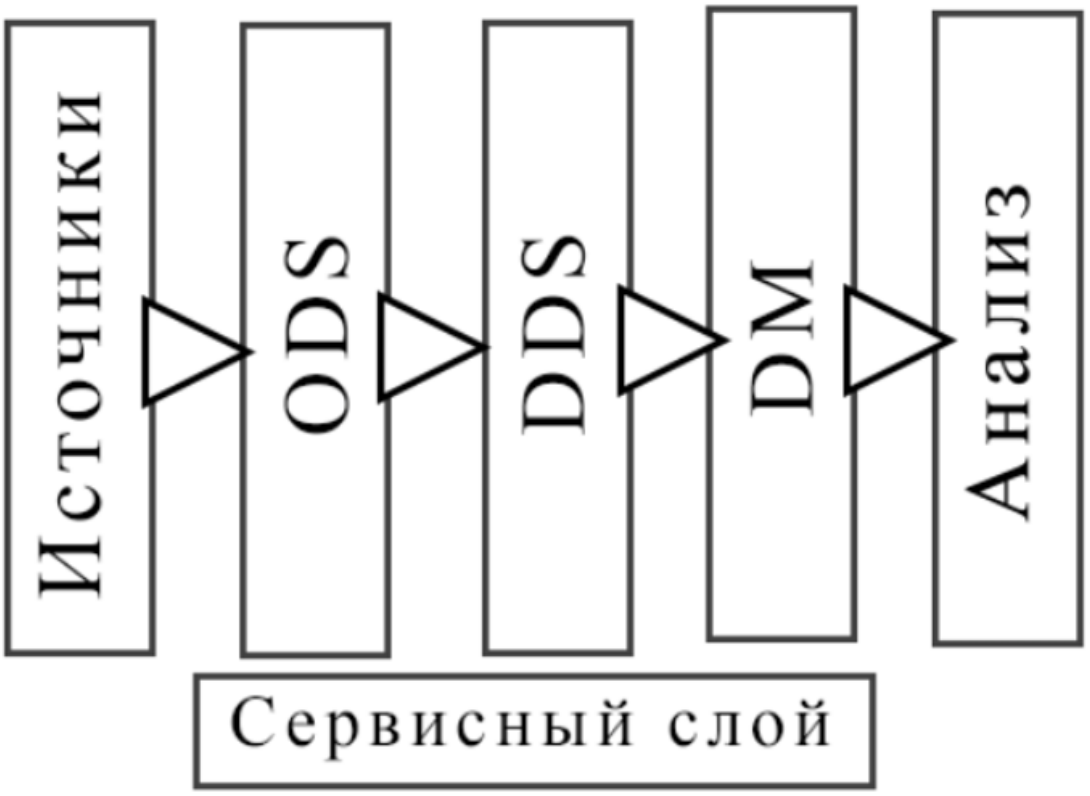

Рисунок 1.1 – Концептуальная архитектурная схема

Источники данных — это то, откуда система получает исходные данные, например бухгалтерские, биллинговые, банковские и тому подобные системы[3].

ODS (Operational Data Store) — область хранения операционных данных, где данные загружаются в первозданном виде. Зачастую, в крупных проектах, ODS выступает как отдельная система. В нее загружаются данные со всех существующих систем источников компании, где появляется возможность работать с ними в единой среде, для анализа и нахождения различного рода зависимостей между источниками. Это позволяет быстрее и точнее выявлять закономерности в данных, которые потом можно структурированно преобразовать в хранилище, выстраивать на этих данных отчеты и принимать управленческие решения[2].

DDS (Detail Data Store) — ядро хранилища, центральный компонент системы, который отличает хранилище от «большой свалки данных», поскольку его основная роль – это консолидация данных из разных источников, приведение к единым структурам, ключам. Именно при загрузке в ядро осуществляется основная работа

с качеством данных и общие трансформации, которые могут быть достаточно сложны.

Задача данного слоя — отделить своих потребителей от особенностей логического устройства источников данных и необходимости сопоставлять данные из различных систем, обеспечить целостность и качество данных[2].

Data Mart Layer — презентационный слой хранилища данных, алгоритмы работающие на этом слоем реализуют соединение очищенных и обогащенных данных детального слоя, представляя данные в натуральном денормализованном виде. Совокупность полученных данных называют витринами данных, на основе данных DM (Data Mart) в BI (Business intelligence) системе реализуются агрегированные отчеты, на основе которых происходит управление различной деятельностью компании[2].

Service Layer (Сервисный слой) — специальный компонент (или набор компонентов), который оказывает сервисные функции для всех слоев. Он содержит область метаданных, область для работы с качеством данных, а также имеет возможность содержать в себе и другие структуры в зависимости от возложенных на него функций[2].

Такое четкое разделение системы на отдельные компоненты существенно повышает управляемость развития системы[2]:

— снижается сложность задачи, которая ставится разработчику функционала того, или иного компонента поскольку в таком случае он не должен одновременно решать и вопросы интеграции с внешними системами, и продумывать процедуры очистки данных, и думать об оптимальном представлении данных для потребителей;

— можно подключать к работе различных исполнителей, команд или подрядчиков, так как такой подход позволяет эффективно распараллеливать задачи, снижая их взаимное влияние друг на друга;

— наличие персистентной структуры позволяет быстро подключить источники данных, не проектируя целиком ядро, либо витрины для всей предметной области, а далее постепенно достраивать остальные слои согласно приоритетам, при

этом данные будут уже в хранилище – доступные системным аналитикам, что существенно облегчит задачи последующего развития хранилища;

— наличие ядра позволяет всю работу с качеством данных, а также, возможные промахи и ошибки, скрыть от витрин и от конечного пользователя, а главное используя этот компонент как единый источник данных для витрин, можно избежать проблем со сходимостью данных в силу реализации общих алгоритмов в одном месте;

— наличие сервисного слоя позволяет выполнять сквозной анализ данных, использовать унифицированные средства аудита данных, общие подходы к выделению дельты изменений, работе с качеством данных, управления загрузкой, средства мониторинга и диагностики ошибок, ускоряет разрешение проблем.

Кроме вышеперечисленных слоев, зачастую в Data warehouse добавляют дополнительные вспомогательные слои для определённых бизнес-задач или же облегчения последующей поддержки системы.

ETL (Extract, Transform, Load) – это совокупность процессов управления хранилищами данных, включая[4]:

— извлечение данных из внешних источников таких как таблицы баз данных, файлы;

— преобразование и очистка данных согласно бизнес-потребностям;

— загрузка обработанной информации в корпоративное хранилище данных.

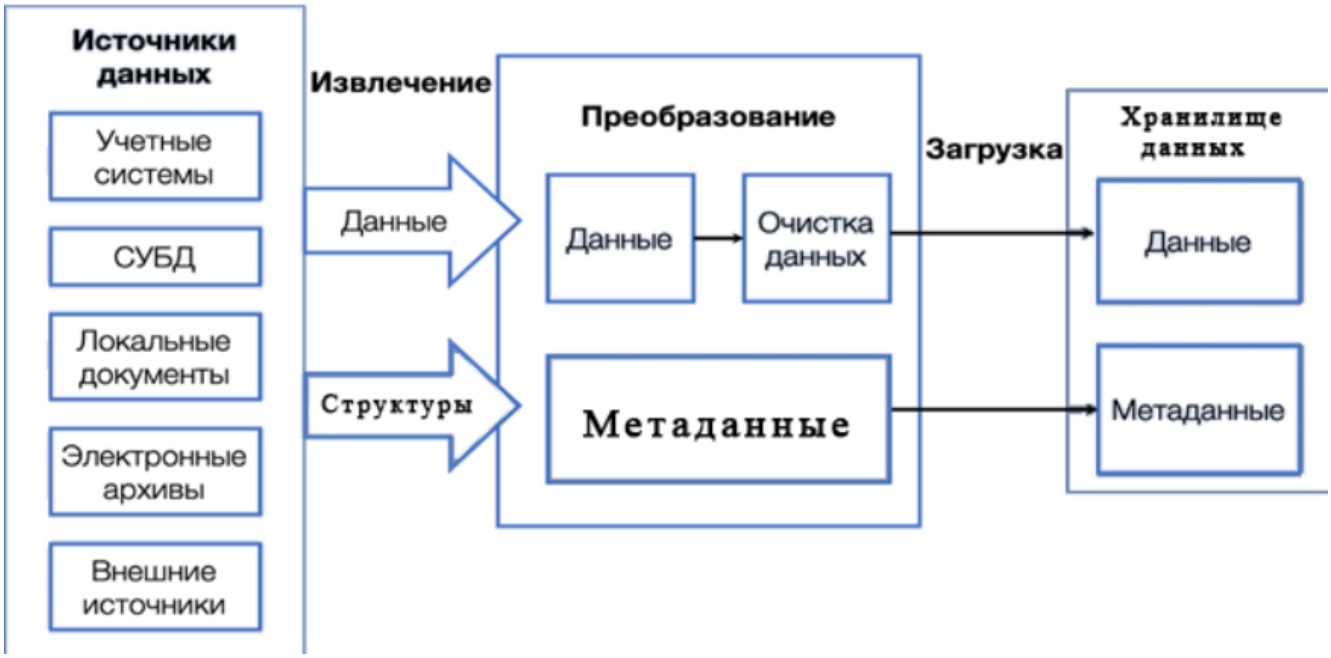

Рисунок 1.2 Работа ETL процесса[4]

Таким образом, ETL-процесс представляет собой перемещение информации (поток данных) от источника к получателю через промежуточную область, содержащую вспомогательные таблицы, которые создаются временно и исключительно для организации процесса выгрузки. Требования к организации потока данных описывает аналитик. Поэтому ETL – это не только процесс переноса данных из одного приложения в другое, но и инструмент подготовки данных к анализу. На рисунке 1.2 показана работа ETL процесса[4].

Для более удобного взаимодействия с данными существуют схемы хранения данных, которые основывают логическую структуру базы данных и в корне определяют, каким образом данные могут храниться, организовываться и обрабатываться. Основными схемами являются «звезда» и «снежинка»[5].

Схема типа «звезда» имеет централизованное хранилище данных, которое хранится в таблице фактов. Схема разбивает таблицу фактов на ряд денормализованных таблиц измерений. Таблица фактов содержит агрегированные данные, которые будут использоваться для составления отчетов, а таблица измерений описывает хранимые данные.

Денормализованные проекты менее сложны, потому что данные сгруппированы. Таблица фактов использует только одну ссылку для присоединения к каждой таблице измерений. Более простая конструкция звездообразной схемы значительно упрощает написание сложных запросов.

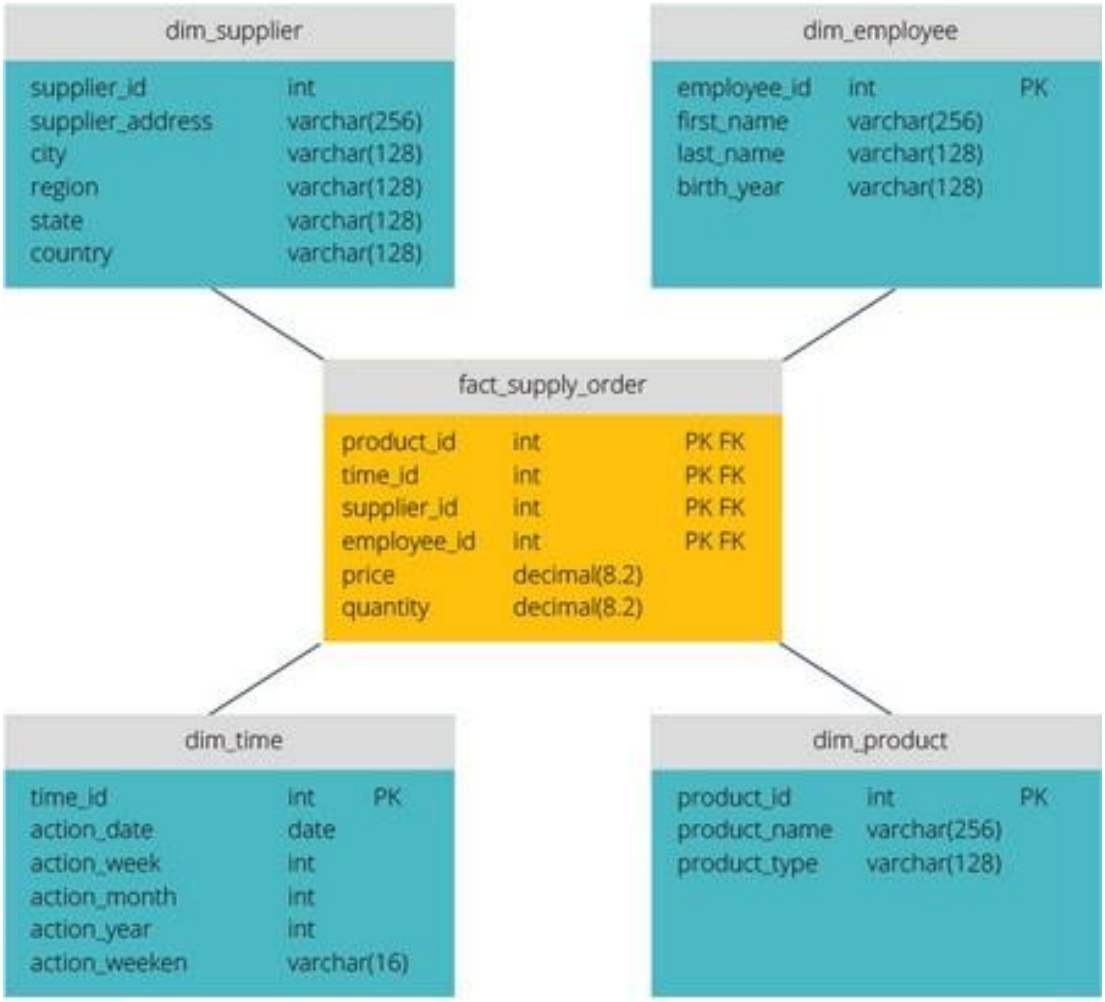

Рисунок 1.3 Схема звезда[5]

Схема типа «снежинка» отличается тем, что использует нормализованные данные. Нормализация означает эффективную организацию данных так, чтобы все зависимости данных были определены, и каждая таблица содержала минимум избыточности.

Схема «снежинки» использует меньше дискового пространства и лучше сохраняет целостность данных. Основным недостатком является сложность запросов,

необходимых для доступа к данным — каждый запрос должен пройти несколько соединений таблиц, чтобы получить соответствующие данные.

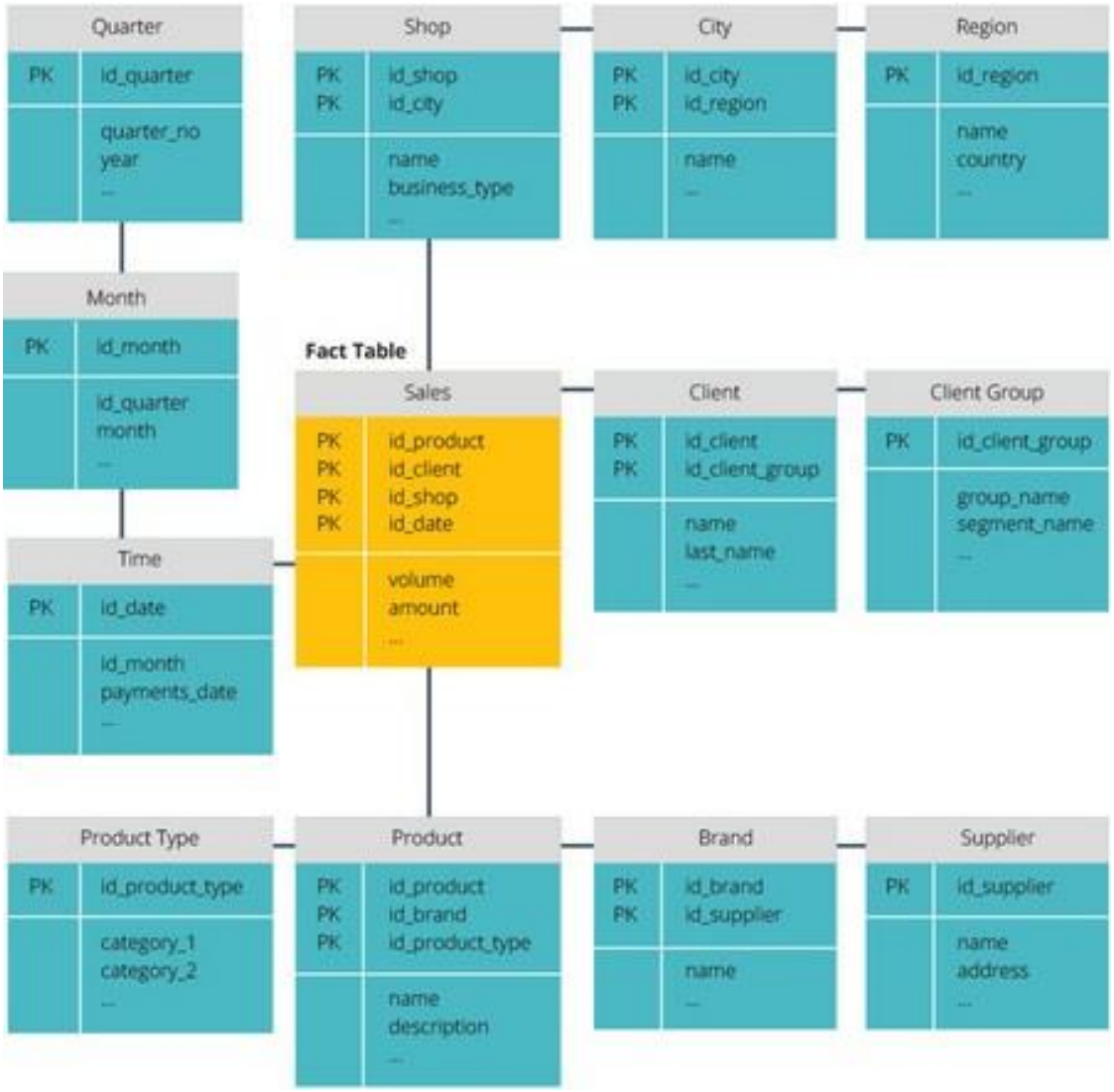

Рисунок 1.4 Схема снежинка[5]

В таблице 1.1 показан сравнительный анализ схемы «звезда» и «снежинка».

Таблица 1.1 – Сравнение схемы «звезда» и «снежинка»

Название	«Звезда»	«Снежинка»
Схема построения	Содержит таблицу фактов, окруженную таблицами измерений	Одна таблица фактов, окруженная таблицей измерений, которая окружена таблицей измерений

Продолжение таблицы 1.1

Соединение	В схеме типа «звезда» только одно соединение создает связь между таблицей фактов и любыми таблицами измерений	Схема снежинки требует много соединений для извлечения данных
Дизайн базы данных	Простой	Сложный
Структура данных	Денормализованная	Нормализованная
Избыточность	Высокая	Низкая
Скорость	Быстрая из-за простоты соединения.	Медленная из-за сложного соединения.

Из данного сравнения видно, что схема «звезда» проще в создании и может быстрее обработать данные чем схема «снежинка», что как раз отвечает выбранным критериям. На основе схемы «снежинка» существует такая схема как Data Vault.

Data Vault — это гибридный подход, объединивший достоинства схемы «звезда» и 3-ей нормальной формы[29]. Впервые эта методология была анонсирована в 2000 году Дэном Линстедтом (Dan Linstedt). Подход был придуман в процессе разработки хранилища данных для Министерства Обороны США и хорошо себя зарекомендовал[6].

Data Vault состоит из трех основных компонентов[6]:

— хаб (Hub),

— ссылка (Link)

— сателлит (Satellite).

Хаб — основное представление сущности (Клиент, Продукт, Заказ) с позиции бизнеса. Хаб содержит одно или несколько полей, отражающих сущность в понятиях бизнеса. Хаб так же содержит метаполя load timestamp и record source, в

которых хранятся время первоначальной загрузки сущности в хранилище и ее (название системы, базы или файла, откуда данные были загружены).

Таблицы Ссылки связывают несколько хабов связью многие-ко-многим. Она содержит те же метаданные, что и Хаб. Ссылка может быть связана с другой Ссылкой, но такой подход создает проблемы при загрузке, так что лучше выделить одну из Ссылок в отдельный Хаб.

Все описательные атрибуты Хаба или Ссылки (контекст) помещаются в таблицы Сателлиты. Помимо контекста Сателлит содержит стандартный набор метаданных load timestamp и record source и один и только один ключ «родителя». В Сателлитах можно без проблем хранить историю изменения контекста, каждый раз добавляя новую запись при обновлении контекста в системе-источнике. Для Хаба или Ссылки может быть сколь угодно Сателлитов, обычно контекст разбивается по частоте обновления. Контекст из разных систем-источников принято класть в отдельные Сателлиты. На рисунке 1.4 показан пример схемы хранения данных Data Vault[6].

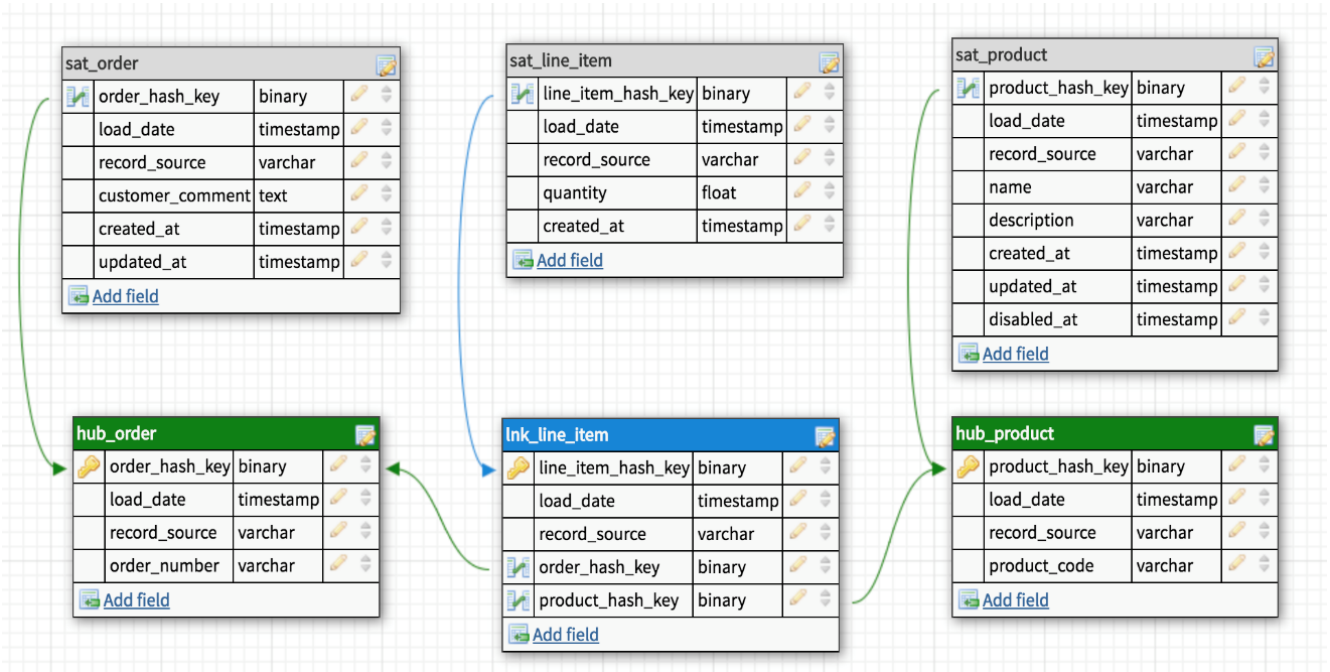

Рисунок 1.4 Схема Data Vault [6]

1.4 Анализ модель угроз

В типовой модели угроз рассматриваются угрозы связанные с случайным, в том числе несанкционированным, доступом в информационных системах с целью копирования, изменения, незаконного распространения данных или разрушающих влияний на элементы информационной системы и обрабатываемых в ней данных с использованием программно-аппаратных и программных средств с целью уничтожения или блокирования данных[7].

Модель угроз представляет собой наступление видов последствий в результате неправомерного или случайного доступа к данным и осуществления угрозы безопасности.

Под угрозами информационной безопасности при ее обработке в системе понимается совокупность всех факторов и условий, создающих реальную или потенциаольно существующую опасность, связанную с утечкой информации и (или) несанкционированными и (или) непреднамеренными воздействиями на нее. Таким образом, угрозы информационной безопасности при ее обработке в системе могут быть связаны как со специально осуществялемыми так и с непреднамеренными действиями персонала или отдельных организций, граждан, а так же иными источниками угроз.

Данная модель состоит из переченя угроз безопасности персональных данных при их обработке в системе, данные угрозы могут быть от разных источников, имеющих стихийный, техногенный, антропогенный характер и воздействующих на уязвимости, характерные для Системы, реализуя тем самым угрозы информационной безопасности.

Данная модель угроз содержит данные по угрозам безопасности, связанным с:

- несанкционированным или случайным доступом в ИС;
- перехватом (съемом) данных по техническим каналам.

Реализация данных угроз может приводить к нарушению заданных характеристик информационной безопасности, а именно таких как[7]:

– перехват или съем данных по техническим каналам может быть произведен с целью неправомерного распространения или копирования, и привести к нарушению конфиденциальности данных;

– несанкционированный доступ к данным может быть произведен с целью копирования, изменения, неправомерного распространения данных или деструктивных воздействий на элементы информационной системы и обрабатываемых в них данных и привести к нарушению доступности, целостности и конфиденциальности обрабатываемых данных.

В информационной системе угрозы несанкционированного доступа к данным подразделяются на[7]:

– угрозы непосредственного доступа;

– угрозы виртуализации.

– угрозы удаленного доступа;

Источниками угрозы безопасности могут быть[7]:

– средства съема сигналов с проводных линий;

– средства перехвата сигналов ПЭМИН;

– закладочные устройства обнаружения и перехвата сигналов;

– средства перехвата информации в каналах передачи данных.

Источниками угроз несанкционированного доступа к информации так же могут быть программно-аппаратные закладки и отчуждаемые носители вредоносных программ.

К угрозам непосредственного доступа относятся[7]:

– угрозы, реализуемые после загрузки операционной среды независимо от того, какая прикладная программа запускается пользователем;

– угрозы, реализуемые в ходе загрузки ОС;

– угрозы, реализуемые после загрузки операционной среды и зависящие от запускаемых прикладных программ (в т. ч. пользователем).

Применительно к информационной системы к угрозам удаленного доступа относятся[7]:

– анализ сетевого трафика;
– внедрение ложного объекта сети;
– угроза выявления пароля;
– сканирование сети;
– подмена доверенного объекта сети;
– отказ в обслуживании;
– навязывание ложного маршрута сети;
– удаленный запуск приложений;
– угрозы внедрения по сети вредоносных программ.

Угроза анализа сетевого трафика реализуется благодаря специальной программе-анализатору пакетов (sniffer), перехватывающей пакеты, которые передаются по сегменту сети, и выделяющей среди них те, в которых передаются пароль и идентификатор пользователя.

Сущность процесса сканирования сети заключается в передаче запросов сетевым службам хостов информационной системы и анализе ответов от них.

Цель реализации выявления пароля состоит в получении неправомерного доступа путем обхода парольной защиты, таким образом злоумышленник может произвести угрозу с помощью целого ряда действий, таких как простой перебор, перебор с использованием специальных словарей, подмена доверенного объекта сети (IP-spoofing), перехват пакетов (sniffing) и установка вредоносной программы для перехвата пароля.

В результате подмены доверенного объекта сети возможно изменение путей прохождения сообщений, несанкционированного доступа к сетевым ресурсам, несанкционированное изменение маршрутноадресных данных, навязывание ложной информации[7].

Угрозы отказа в обслуживании основаны на недостатках сетевого ПО, его уязвимостях, позволяющих нарушителю создавать условия, когда ОС оказывается не в состоянии обрабатывать поступающие пакеты.

Угроза удаленного запуска приложений заключается в стремлении запустить на хосте информационной системы различные предварительно внедренные вредоносные программы, такие как, вирусы, программы-закладки, "сетевые шпионы", основная цель которых это нарушение конфиденциальности, целостности, доступности информации и полный контроль за работой хоста.

Угрозы внедрения по сети вредоносных программ. Основными видами вредоносных программ являются[7]:

– программные закладки;

– классические программные вирусы;

– вредоносные программы, распространяющиеся по сети (сетевые черви);

– другие вредоносные программы, предназначенные для осуществления НСД.

Вредоносные программы могут быть внесены (внедрены) как случайно, так и преднамеренно в ПО, используемое в информационной системе, в процессе его сопровождения, разработки, модификации и настройки.

Угроза возможности копирования и удаления виртуальных машин основана на управлении виртуальной средой даже при отсутствии доступа к данным на самих виртуальных машинах.

В результате изменения настроек виртуальной среды возможен несанкционированный доступ к ресурсам виртуальных машин, а так же нарушение доступности ИС.

Угроза сетевой атаки на виртуальные машины основана на получении удаленного доступа к виртуальным машинам, в том числе со стороны других виртуальных машин данной виртуальной среды.

1.5 Анализ модель нарушителя

Для представления модели нарушителя для начала определим кто является нарушителем.

Нарушитель безопасности информации - это физическое лицо или субъект, случайно или преднамеренно совершившее действия, следствием которых является нарушение безопасности информации при ее обработке техническими средствами в информационных системах[7].

Нарушителей различают как внешнего и внутреннего. Под внутренним нарушителем понимают нарушителя, находящегося внутри информационной системы на момент начала реализации угрозы. Под внешним нарушителем понимают нарушителя, находящегося вне информационной системы на момент начала реализации угрозы.

Для реализации угроз в информационной системе внешний нарушитель должен тем или иным способом получить доступ к процессам, проходящим в информационной системе. При этом дальнейшие свои действия внешний нарушитель выполняет от имени созданного им нового или существующего в системе субъекта.

К внутренним нарушителям относят инсайдеров, не смотря на то, что они могут выполнять инструкции лиц, находящихся вне информационной системы[7].

Воспользуемся утвержденным ФСТЭК методическим документом для определения всевозможных моделей нарушителя. Виды нарушителей и их возможные цели или мотивация реализации угроз безопасности информации приведены в таблице 1.2.

Таблица 1.2 – Виды нарушителей и их типы

№ вида	Виды нарушителя	Типы нарушителя	Возможные цели (мотивация) реализации угроз безопасности информации
1	Специальные службы иностранных государств	Внешний, внутренний	Нанесение ущерба государству, отдельным его сферам деятельности или секторам экономики.

Продолжение таблицы 1.2

2	Террористические, экстремистские группировки	Внешний	Нанесение ущерба государству, отдельным его сферам деятельности или секторам экономики. Совершение террористических актов. Идеологические или политические мотивы. Дестабилизация деятельности органов власти, организаций
3	Преступные группы (криминальные структуры)	Внешний	Причинение имущественного ущерба путем мошенничества или иным преступным путем. Выявление уязвимостей с целью их дальнейшей продажи и получения финансовой выгоды
4	Внешние субъекты (физические лица)	Внешний	Идеологические или политические мотивы. Причинение имущественного ущерба путем мошенничества или иным преступным путем. Любопытство или желание самореализации (подтверждение статуса). Выявление уязвимостей с целью их дальнейшей продажи и получения финансовой выгоды
5	Конкурирующие организации	Внешний	Получение конкурентных преимуществ. Причинение имущественного ущерба путем обмана или злоупотребления доверием

Продолжение таблицы 1.2

6	Разработчики, производители, поставщики программных, технических и программно-технических средств	Внешний	Внедрение дополнительных функциональных возможностей в программное обеспечение или программно-технические средства на этапе разработки. Причинение имущественного ущерба путем обмана или злоупотребления доверием. Непреднамеренные, неосторожные или неквалифицированные действия
7	Лица, привлекаемые для установки, наладки, пусконаладочных и иных видов работ	Внутренний	Причинение имущественного ущерба путем обмана или злоупотребления доверием. Непреднамеренные, неосторожные или неквалифицированные действия
8	Лица, обеспечивающие функционирование информационных систем или обслуживающие инфраструктуру оператора (администрация, охрана, уборщики и т.д.)	Внутренний	Причинение имущественного ущерба путем обмана или злоупотребления доверием. Непреднамеренные, неосторожные или неквалифицированные действия

Продолжение таблицы 1.2

9	Пользователи информационной системы	Внутренний	Причинение имущественного ущерба путем мошенничества или иным преступным путем. Любопытство или желание самореализации (подтверждение статуса). Месть за ранее совершенные действия. Непреднамеренные, неосторожные или неквалифицированные действия
10	Администраторы информационной системы и администраторы безопасности	Внутренний	Причинение имущественного ущерба путем мошенничества или иным преступным путем. Любопытство или желание самореализации (подтверждение статуса). Месть за ранее совершенные действия. Выявление уязвимостей с целью их дальнейшей продажи и получения финансовой выгоды. Непреднамеренные, неосторожные или неквалифицированные действия
11	Бывшие работники (пользователи)	Внешний	Причинение имущественного ущерба путем мошенничества или иным преступным путем. Месть за ранее совершенные действия

В зависимости от потенциала, требуемого для угрозы безопасности информации, нарушителей можно разделить на:

— нарушителей, которые обладают высоким потенциалом нападения при реализации угроз безопасности информации в информационной системе;

— нарушителей, которые обладают средним потенциалом нападения при реализации угроз безопасности информации в информационной системе;

— нарушителей, которые обладают базовым потенциалом нападения при реализации угроз безопасности информации в информационной системе.

К нарушителям с базовым потенциалом относятся внешние субъекты:

— лица, обеспечивающие функционирование информационных систем;

— пользователи информационной системы;

— бывшие работники;

— лица, привлекаемые для установки, наладки, монтажа, пусконаладочных и иных работ.

Они имеют возможность получить информацию об уязвимостях отдельных компонентов информационной системы, которую можно найти в общедоступных источниках, а также возможность получить информацию о методах и средствах реализации угроз безопасности информации.

Нарушители с базовым средним потенциалом это террористические, экстремистские группировки, конкурирующие организации, преступные группы, производители, разработчики, поставщики программных, технических и программнотехнических средств, администраторы информационной системы и администраторы безопасности они имеют осведомленность о мерах защиты информации, применяемых в информационной системе данного типа, так же имеют возможность получить информацию об уязвимостях отдельных компонент информационной системы путем проведения, с использованием имеющихся в свободном доступе программных средств, анализа кода прикладного программного обеспечения и отдельных программных компонент общесистемного программного обеспечения. А также доступ к сведениям о структурно-функциональных характеристиках и особенностях функционирования информационной системы.

Нарушители с высоким потенциалом, это специальные службы иностранных государств, которые обладают всеми возможностями нарушителей с базовым и средним потенциалами. Они имеют возможность осуществлять неправомерный доступ из выделенных сетей связи, к которым возможен физический доступ, и имеют возможность получить доступ к программному обеспечению чипсетов, системному и прикладному программному обеспечению, телекоммуникационному

оборудованию и другим программно-техническим средствам информационной системы для преднамеренного внесения в них уязвимостей или программных закладок. Имеют возможность создания и применения специальных технических средств для получения информации, распространяющейся в виде физических полей или явлений.

В данной работе будет рассмотрено противодействие нарушителям внутреннего и внешнего типа с средним и низким потенциалами, поскольку они больше всего подходят под данный проект.

1.6 Выводы по разделу

В данном разделе рассмотрена концептуальная модель построения Data warehouse, проанализированы функции каждой компоненты.

Представлены модели нарушителя и угроз для хранилища данных авиакомпании.

Для последующего проектирования Data warehouse была выбрана архитектура Data Vault, поскольку схема «звезда» обладает ограниченной гибкостью. Data Vault же обладает:

— простотой масштабирования хранилища данных;

— хорошей поддержкой «историчности».

2 Анализ программного обеспечения для проектирования Data warehouse

2.1 Постановка задачи

В данном разделе необходимо провести анализ программного обеспечения необходимого для создания Data warehouse.

2.1 Выбор системы управления базой данных

Система управления базами данных (СУБД) – это комплекс программно-языковых средств, позволяющих создать базы данных и управлять данными. Иными словами, СУБД — это набор программ, позволяющий организовывать, контролировать и администрировать базы данных. Большинство сайтов не могут функционировать без базы данных, поэтому СУБД используется практически повсеместно[9].

Oracle RDBMS (Oracle Database) эта система часто выбирается разработчиками поскольку она проста в использовании, у нее понятная документация, поддержка длинных наименований, JSON, улучшенный тег списка и Oracle Cloud, написана на Assembly, C, C++[10].

Особенности[10]:

— обрабатывает большие данные;

— поддерживает SQL, к нему можно получить доступ из реляционных БД Oracle;

— Oracle NoSQL Database с Java/C API для чтения и записи данных.

MySQL работает на Linux, Windows, OSX, FreeBSD и Solaris. Можно начать работать с бесплатным сервером, а затем перейти на коммерческую версию. Лицензия GPL с открытым исходным кодом позволяет модифицировать ПО MySQL.

Эта СУБД использует стандартную форму SQL. Утилиты для проектирования таблиц имеют интуитивно понятный интерфейс. MySQL поддерживает до 50 миллионов строк в таблице. Предельный размер файла для таблицы по умолчанию

4 ГБ, но его можно увеличить. Поддерживает секционирование и репликацию, а также Xpath и хранимые процедуры, триггеры и представления. Разработана Oracle Corporation и написана на C, C++[10].

Особенности[10]:

— масштабируемость;

— лёгкость использования;

— безопасность;

— поддержка Novell Cluster;

— скорость;

— поддержка многих операционных систем.

Microsoft SQL Server самая используемая коммерческая СУБД. Она имеет жесткую привязку к Windows. Поддерживает SQL, непроцедурные, нечувствительные к регистру и общие языки баз данных. Разработана Microsoft Corporation, написана на C, C++[10].

Особенности[10]:

— высокая производительность;

— зависимость от платформы;

— возможность установить разные версии на одном компьютере;

— генерация скриптов для перемещения данных.

PostgreSQL масштабируемая объектно-реляционная база данных, работающая на Linux, Windows, OSX и некоторых других системах. В PostgreSQL есть такие функции, как логическая репликация, декларативное разбиение таблиц, улучшенные параллельные запросы, более безопасная аутентификация по паролю на основе SCRAM-SHA-256. Разработана PostgreSQL Global Development Group, написана на C[10].

Особенности[10]:

— поддержка табличных пространств, а также хранимых процедур, объединений, представлений и триггеров;

— восстановление на момент времени;

— асинхронная репликация;

— NoSQL-базы данных.

MongoDB самая используемая NoSQL система управления базами данных. Лучше всего подходит для динамических запросов и определения индексов. Гибкая структура, которую можно модифицировать и расширять. Поддерживает Linux, OSX и Windows, но размер БД ограничен 2,5 ГБ в 32-битных системах. Использует платформы хранения MMAPv1 и WiredTiger. Разработана MongoDB Inc., написана на C++[10].

Особенности[10]:

— высокая производительность;

— автоматическая фрагментация;

— работа на нескольких серверах;

— поддержка репликации Master-Slave;

— данные хранятся в форме документов JSON;

— возможность индексировать все поля в документе;

— поддержка поиска по регулярным выражениям.

DB2 работает на Linux, UNIX, Windows и мейнфреймах. Эта СУБД хорошо подходит для хост-сред IBM. Версию DB2 Express-C нельзя использовать в средах высокой доступности (при репликации, кластеризации типа active-passive и при работе с синхронизируемым доступом к разделяемым данным). Разработана IBM, написана на C, C++, Assembly[10].

Особенности[10]:

— улучшенное встроенное шифрование;

— упрощённая установка и развёртывание.

Microsoft Access система управления базами данных от Microsoft, которая сочетает в себе реляционное ядро БД Microsoft Jet с графическим интерфейсом пользователя и инструментами разработки ПО.

Отлично подходит для начала работы с данными, но производительность не рассчитана на большие проекты. В MS Access можно использовать C, C#, C++, Java,

VBA и Visual Rudimental.NET. Access хранит все таблицы БД, запросы, формы, отчёты, макросы и модули в базе данных Access Jet в виде одного файла. Разработана Microsoft Corporation[10].

Особенности[10]:

— можно использовать VBA для создания многофункциональных решений с расширенными возможностями управления данными и пользовательским контролем;

— импорт и экспорт в форматы Excel, Outlook, ASCII, dBase, Paradox, FoxPro, SQL Server и Oracle;

— формат базы данных Jet.

Cassandra СУБД активно используется в банковском деле, финансах, а также в Facebook и Twitter. Поддерживает Windows, Linux и OSX. Для запросов к БД Cassandra используется SQL-подобный язык — Cassandra Query Language (CQL). Разработана Apache Software Foundation, написана на Java[10].

Особенности[10]:

— линейная масштабируемость;

— быстрое время отклика;

— поддержка MapReduce и Apache Hadoop;

— максимальная гибкость;

— P2P архитектура.

Redis или Remote Dictionary Server — СУБД с открытым исходным кодом, которая снабжена механизмами журналирования и снимков. Поддерживаются списки, строки, хэши, наборы. Все операции в Redis атомарные. Система написана на языке C и поддерживается практически всеми языками программирования. Разработчик Salvatore Sanfilippo[10].

Особенности[10]:

— автоматическая обработка отказа;

— транзакции;

— сценарии LUA;

— вытеснение LRU-ключей;

— поддержка Publish/Subscribe.

Elasticsearch легко масштабируемая поисковая система корпоративного уровня с открытым исходным кодом. Благодаря обширному и продуманному API обеспечивает чрезвычайно быстрый поиск, работает в том числе с приложениями для обнаружения данных. Используется такими компаниями, как Википедия, The Guardian, StackOverflow, GitHub. ElasticSearch позволяет создавать копии индексов и сегментов. Разработана Elastic NV, написана на Java[10].

Особенности[10]:

— масштабируемость вплоть до нескольких петабайт структурированных и неструктурированных данных;

— многопользовательская поддержка;

— масштабируемый поиск, поиск в режиме реального времени.

В таблице 2.1 приведено сравнение систем управления базами данных.

Таблица 2.1 – сравнение систем управления базами данных[10]

СУБД	Разработчик	Лицензия	Написана на
Oracle	Oracle Corporation	Проприетарная	Assembly, C, C++
MySQL	Oracle Corporation	GPL v2 или проприетарная	C, C++
Microsoft SQL Server	Microsoft Corporation	Проприетарная	C, C++
PostgreSQL	PostgreSQL Global Development Group	Лицензия PostgreSQL (бесплатное ПО с открытым исходным кодом, либеральная лицензия)	C

Продолжение таблицы 2.1

MongoDB	MongoDB Inc.	Различные варианты лицензирования	C++, C, JavaScript
DB2	IBM	Проприетарная EULA	Assembly, C, C++
Microsoft Access	Microsoft Corporation	Проприетарное	Нет данных
Redis	Salvatore Sanfilippo	Лицензия BSD	ANSI C
Cassandra	Apache Software Foundation	Нет данных	Java
Elasticsearch	Elastic NV	Нет данных	Java

2.2 Анализ ETL инструментов

Для осуществления ETL - процесса допустимо использовать почти любой современный язык программирования. Однако, если требуется не разовая конвертация, а постоянно выполнять интеграцию данных, то целесообразно рассмотреть специализированные инструменты. Самыми распространенными ETL инструментами являются Informatica PowerCenter и IBM InfoSphere DataStage[11].

Informatica PowerCenter позволяет интегрировать данные любого формата из виртуальных и бизнес систем и распространяет их на всю компанию с достаточной скоростью для улучшения операционной эффективности.

Является основой для любых интеграционных задач организации, включая построение хранилищ данных, управление данными, миграцию данных, построение сервис ориентированной архитектуры (SOA), обмен неструктурированными данными между партнёрами и управление справочными данными (MDM)[12].

Informatica унифицированная платформа интеграции корпоративных данных. Она обеспечивает доступ, извлечение, трансформацию данных из любой системы или бизнес приложения, в любом формате, и доставку этих данных в корпоративном масштабе по требованию.

Является легко масштабируемым, высокопроизводительным

интеграционным программным продуктом корпоративного уровня. Она позволяет извлекать и интегрировать данные из любых виртуальных и реальных бизнес систем и в любых форматах, доставляя данные в любую точку организации.

Так же Informatica имеет возможность масштабирования для поддержки работы с большими объемами данных в режиме параллельной работы нескольких сессий[12].

Соответствуя требованиям безопасности за счет маскирования данных и производительности, данное программное обеспечение может быть применено для решения любых интеграционных задач организации, таких как построение Data warehouse и миграция данных[13].

IBM InfoSphere DataStage позволяет интегрировать большие объемы данных между многочисленными источниками данных и целевыми приложениями. IBM InfoSphere DataStage представляет собой главный модуль платформы интеграции данных InfoSphere Information Server, которая поддерживает универсальный доступ к данным локальной сети, облачной инфраструктуре, мобильной сети, а также к структурированным и неструктурированным данным. Использование решения Huawei FusionInsight, реализованного на базе Hadoop, Oozie, и других решений для работы с большими данными, совместно с IBM InfoSphere DataStage помогает предприятиям максимизировать ценность бизнеса с помощью технологии больших данных[14].

Продукт IBM InfoSphere DataStage поддерживает извлечение, преобразование и передачу больших объемов данных, которые могут иметь как простую, так и весьма сложную структуру. DataStage способен работать как с данными, поступающими в текущий момент времени, так и с данными, поступившими ранее на регулярной или плановой основе. DataStage позволяет компаниям решать крупномасштабные бизнес - задачи, благодаря возможности высокопроизводительной обработки больших объемов данных [14].

С учетом необходимости обработки непрерывно растущих объемов данных, обработки в жестких условиях реального времени и пакетной обработки при постоянно сокращающихся временных окнах, DataStage использует возможности

параллельного выполнения заданий и с легкостью масштабируется в широком диапазоне аппаратных платформ: от систем с симметричной мультипроцессорной обработкой (SMP) и SMP кластеров до серверов с сотнями процессоров и поддержкой архитектуры вычислений с массовым параллелизмом (MPP). При увеличении количества процессоров или серверов, перекомпилировать задания не требуется, при этом изменяются только несколько строк в файле конфигурации.

2.3 Выводы по разделу

В данном разделе рассмотрены СУБД, ETL инструменты и средства хранения информации.

Поскольку DWH в среднем имеет объем на носителе от 5 Тб, то для корректной работы потребуется СУБД способная обрабатывать достаточно больше объемы данных и при этом делать это с оптимальной скоростью. Oracle Database подходит для этих целей, этот продукт проверен временем и имеет отличную поддержку со стороны разработчика.

И Datastage и Informatica - мощные инструменты ETL. Оба инструмента делают почти одно и то же, почти одинаково. Производительность, ремонтопригодность, кривая обучения схожи и сопоставимы. Но Informatica более подходящий инструмент благодаря возможности работы с любыми форматами данных.

3 Разработка проекта Data warehouse

3.1 Постановка задачи

В рамках проектирования Data warehouse, необходимо, используя как основу концептуальную модель (рисунок 1.1), разработать собственную, добавив новые слои с описанием, а также ETL процесс для загрузки данных. На рисунке 3.1 изображена упрощенная схема сети.

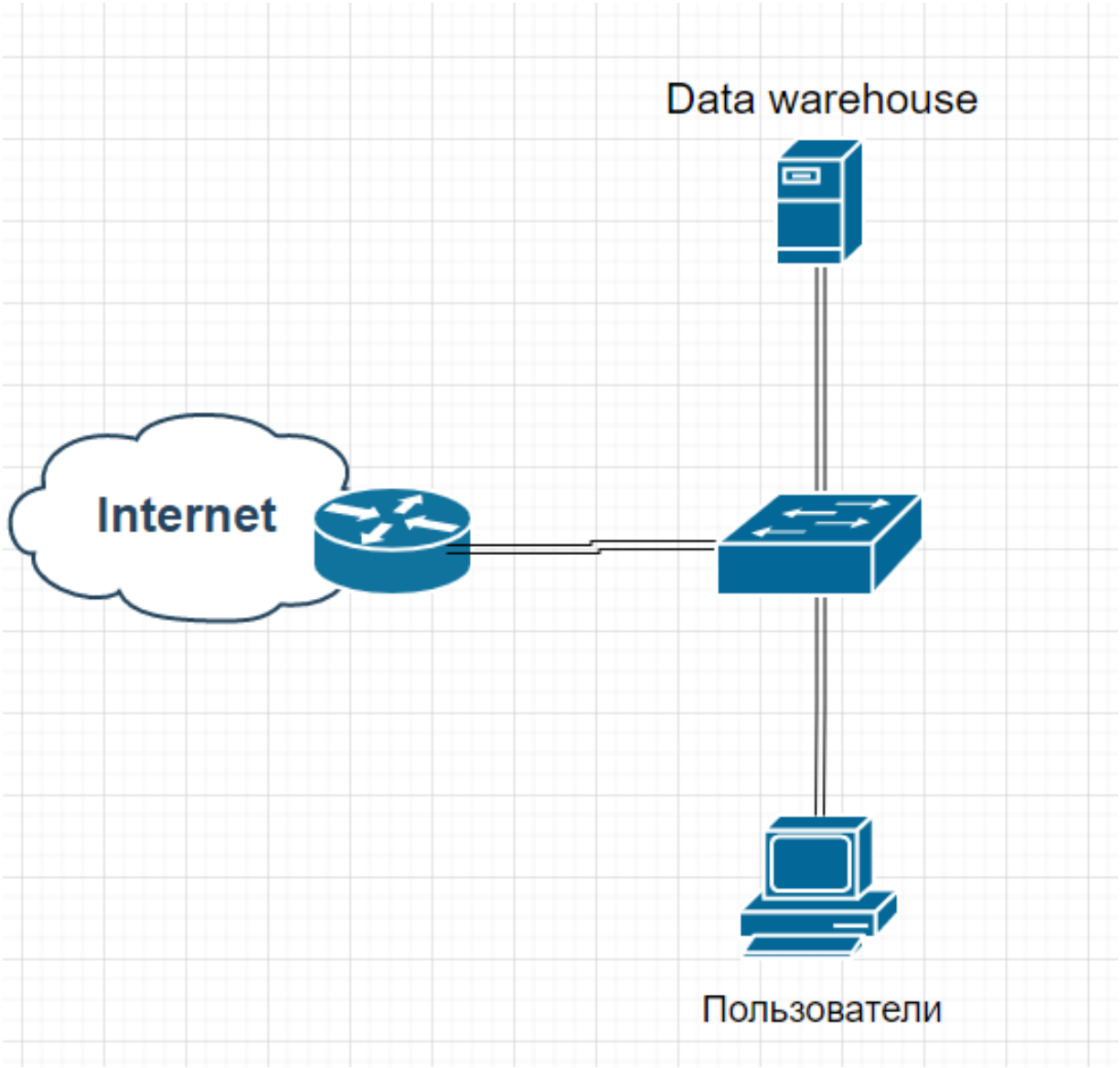

Рисунок 3.1 – Упрощенная схема сети

3.2 Разработка слоев Data warehouse

Взяв за основу концептуальную модель Data warehouse (рисунок 1.1), построим собственную (рисунок 3.1).

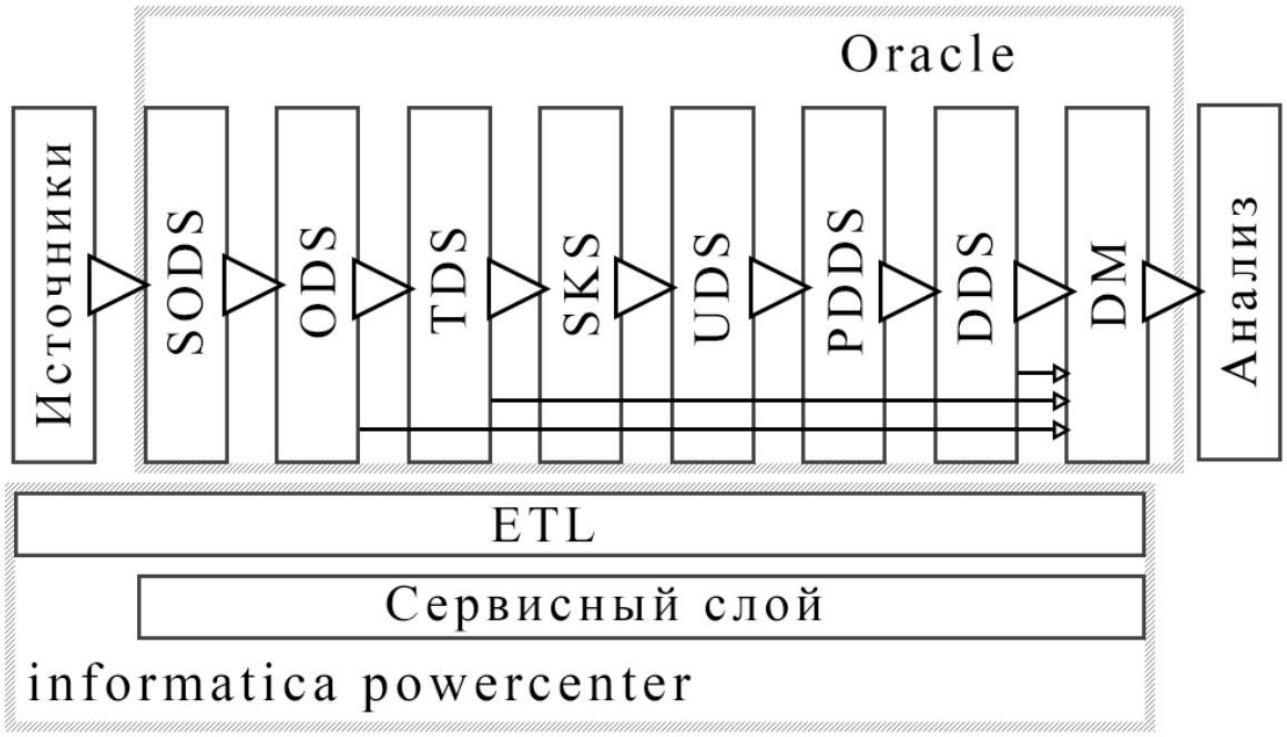

Рисунок 3.2 – Разработанная модель Data warehouse

В отличии от концептуальной модели, в разработанной модели добавлены дополнительные слои для улучшения гибкости системы в случае изменения каких-либо данных.

В качестве источников могут выступать банковские системы, системы бронирования, общие информационные справочники и т.д.

SODS (Staging Operational Data Store) – область извлечения данных, содержит срезы инкрементальных данных из источника. В таблице 3.1 приведено описание технических атрибутов SODS.

Таблица 3.1 – Описание технических атрибутов SODS

Название атрибута	Тип данных	Описание
NCD	NUMBER	Числовой первичный ключ
CCD	VARCHAR2	Символьный первичный ключ
DMLTS	TIMESTAMP(6)	Серверное дата/время загрузки данных в SODS
DML_TYPE	VARCHAR2	Вид DML операции, произведенной над записью в учетной системе. Значения атрибута: I/D - Insert, Delete
PROCESSED_DTTM	DATE	Дата/время загрузки данных
DWH_JOB_ID	NUMBER	Уникальный номер запуска процесса
NCD	NUMBER	Числовой первичный ключ
CCD	VARCHAR2	Символьный первичный ключ
DMLTS	TIMESTAMP(6)	Серверное дата/время загрузки данных в SODS
DML_TYPE	VARCHAR2	Вид DML операции, произведенной над записью в учетной системе. Значения атрибута: I/D - Insert, Delete
PROCESSED_DTTM	DATE	Дата/время загрузки данных

Особенностью схемы SODS является то, что таблицы секционированы по DWH_JOB_ID для более быстрого извлечения данных при загрузке в ODS и быстрой очистки не актуальной истории в SODS.

В рамках разрабатываемого хранилища ODS (Operational Data Store) представляет собой схему базы данных (БД), где осуществляется сбор данных с первоисточников. Такая реализация необходима для более удобной обработки данных в последующих слоях, а также для возможности реализации версионного хранения данных.

При формировании слоя ODS закладывается набор технических атрибутов, необходимых для стандартизации работы ETL, разбора ретроспективных данных, контроля согласованности данных в дальнейших слоях. В таблице 3.2 приведено описание технических атрибутов ODS.

Таблица 3.2 – Описание технических атрибутов ODS

Название атрибута	Тип данных	Описание
NCD	NUMBER	Числовой первичный ключ
CCD	VARCHAR2	Символьный первичный ключ
DML_TS_UTC	TIMESTAMP(6)	UTC дата/время загрузки данных в SODS
DML_TS	TIMESTAMP(6)	Серверное дата/время загрузки данных в SODS
INS_DWH_JOB_ID	NUMBER	Номер запуска процесса первичной вставки
UPD_DWH_JOB_ID	NUMBER	Номер запуска процесса при обновлении
DWH_JOB_ID	NUMBER	Уникальный номер запуска процесса
DELETED_FLAG	VARCHAR2	Флаг удаления записи
PROCESSED_DTTM	DATE	Дата/время загрузки данных
VALID_FROM_DTTM	DATE	Дата/время начала действия записи
VALID_TO_DTTM	DATE	Дата/время окончания действия записи

Для стандартного обозначения первичного ключа таблицы задаются атрибуты NCD (первичным ключом является числовой атрибут таблицы системы источника), CCD (первичным ключом является один символьный или несколько атрибутов таблицы). В качестве промежуточного слоя загрузки в ODS используется слой SODS.

TDS (Technical Data Store) – область хранения инкрементального набора данных ODS. Технический слой нацелен на обработку данных из разных источников для загрузки в DDS. Основными функциональными обязанностями TDS являются:

— дедубликация;

— очистка;

— унификация.

Необходимость выгрузки данных в TDS определяется разработчиком. В самых простых случаях (минимум преобразования исходных данных) допустима загрузка из ODS в DDS без использования TDS.

SKS (Surrogate Key Store) – фабрика суррогатных ключей. В процессе загрузки данных в DDS все сущности обогащаются суррогатными целочисленными ключами. Формирование ключа происходит на основе:

— номера системы источника;

— натурального представления ключа из системы источника.

Суррогатный ключ Data warehouse устойчив, то есть для некоторого натурального ключа системы источника не изменяется значение соответствующего ему суррогата.

Для обеспечения устойчивости суррогатных ключей, их связь с натуральными ключами и натуральными идентификаторами систем источников генерируются один раз и хранятся в SKS - таблицах, причем для каждой таблицы DDS существует отдельная SKS - таблица. В таблице 3.3 приведено описание технических атрибутов SKS.

Таблица 3.3 – Шаблон таблиц схемы SKS

Название атрибута	Тип данных	Описание
NK_CCD	VARCHAR2	Натуральный ключ записи из системы источника
SOURCE_SYSTEM_CCD	VARCHAR2	Номер источника
RK	NUMBER	Суррогатный ключ
PROCESSED_DTTM	DATE	Дата и время вставки/обновления записи
ORPHAN_FLG	VARCHAR2	Признак формирования записи через орфану

Для обеспечения устойчивости даты, входящей в ключ, вводится дополнительная MERGE таблица, в которой хранятся интервалы действия дат, в пределах которых изменение даты не влечет изменения суррогатного ключа. В таблице 3.4 приведен шаблон таблиц MEGRE схемы SKS.

Таблица 3.4 – Шаблон таблиц MEGRE схемы SKS

Название атрибута	Тип данных	Описание
NK_CCD	VARCHAR2	Натуральный ключ записи без даты из системы источника
SOURCE_SYSTEM_CCD	VARCHAR2	Номер источника
RK	NUMBER	Суррогатный ключ
PROCESSED_DTTM	DATE	Дата и время вставки/обновления записи
DATE_FROM	DATE	Интервал действия записи от
DATE_TO	DATE	Интервал действия записи до

UDS (Unified Data Store) – историчный слой хранения преобразованных в единый вид данных. Предназначен для подготовки обогащения данных в детальном слое. Таблицы UDS соответствуют определенные таблицы области DDS и имеют следующие особенности:

- название таблицы UDS = <название_таблицы_TDS>_UNION;
- ключ таблицы UDS = (<ключ_таблицы_dds>, source_system_ccd);
- содержание таблицы – полные актуальные срезы данных о сущности для каждой из систем источников;
- данные приведены к ключам Data warehouse (RK).

Технические атрибуты слоя UDS аналогичны таблицам TDS.*_UNION за исключением добавления DWH_JOB_ID, INS_DWH_JOB_ID и UPD_DWH_JOB_ID аналогичных ODS.

PDDS (Predetail Data Store) – инкрементальный слой, где при помощи алгоритмов приоритезации определяется приоритет источника для каждого атрибута. Источником данных для PDDS является слой UDS. Атрибутный состав и названия таблиц PDDS аналогичен слою UDS, за исключением добавления атрибута SOURCE_SYSTEM_ORDER, он содержит конкатенацию номеров источников, из которых собиралось единое значение.

DDS (Detail Data Store) – область хранения детальных данных. Сущности DDS используется как для формирования витрин данных. Все сущности детального слоя по умолчанию являются версионными. Это означает, что помимо

атрибутного состава PDDS, в сущности добавляются временные интервалы действия записи, аналогично версионному ODS (атрибуты VALID_FROM_DTTM, VALID_TO_DTTM).

DM (Data Mart) – презентационный слой хранилища данных, алгоритмы работающие на этом слоем реализуют соединение очищенных и обогащенных данных детального слоя, представляя данные в натуральном денормализованном виде. Совокупность полученных данных называют витринами данных, на основе данных DM в аналитической системе реализуются агрегированные отчеты, на основе которых происходит управление различной деятельностью компании.

Сервисный слой – отвечает за реализацию общих (сервисных) функций, которые могут использоваться для обработки данных в различных слоях хранилища – управление загрузкой, управление качеством данных, диагностика проблем и средства мониторинга и т.п.

Наличие данного уровня обеспечивает прозрачность и структурированность потоков данных в хранилище.

3.3 Разработка ETL для загрузки данных

Начальным этапом процесса ETL является процедура извлечения записи из источников данных и подготовка их к процессу преобразования. При разработке процедуры извлечения данных в первую очередь необходимо определить частоту выгрузки данных из источников. Выгрузка данных занимает определённое время, которое называется окном выгрузки.

Все ETL процессы хранилища, как и схемы БД разнесены по слоям. Каждый слой в ETL инструменте означает свою папку, где хранятся относящиеся к ней потоки. В таблице 3.5 описаны директории ETL инструмента, в которых должны быть реализованы потоки для загрузки данных в хранилище.

Таблица 3.5 – Описание директорий в ETL инструменте

Название директории	Описание директории
EDW_SODS_<номер источника>	Директории для потоков загружающих данные в SODS
EDW_ODS_<номер источника>	Директории для потоков загружающих данные в ODS
EDW_TDS	Директория для потоков загружающих данные в TDS
EDW_PDDS	Директория для потоков загружающих данные в PDDS
EDW_DDS	Директория для потоков загружающих данные в DDS
EDW_DM	Директория для потоков загружающих данные в DM

Карта ETL процесса для всех сущностей строится на основе общего заранее продуманного разработчиком шаблона. Наименование карты выполняется на основе маски, например для SODS эта маска может быть m_SODS_<номер источника>_<имя таблицы>.

Внутри карты происходят те процессы, которые требуются для конкретного слоя Data warehouse, для загрузки данных с источника на SODS это могут быть:

— обращение к источнику для сбора данных;

— сортировка данных;

— объединение данных с предыдущими записями ;

— добавление технических атрибутов.

Для реализации алгоритмов необходимо для каждой карты создавать поток который будет инициализировать прогрузку данных. Наименование потока происходит на основе маски, для карты с маской m_SODS_<номер источника>_<имя таблицы> поток будет иметь маску wf_SODS_<номер источника>_<имя таблицы>.

3.4 Выводы по разделу

В данном разделе была разработана модель Data warehouse в основе которой лежит концептуальная модель. В процессе разработки были добавлены новые слои для Data warehouse такие как:

— SODS (Staging Operational Data Store);

— TDS (Technical Data Store);

— SKS (Surrogate Key Store);

— UDS (Unified Data Store);

— PDDS (Predetail Data Store).

Так же был разработан ETL процесс для загрузки данных в слои и последующего преобразования данных.

4 Разработка проекта обеспечения информационной безопасности Data warehouse

4.1 Постановка задачи

В рамках этой главы для обеспечения информационной безопасности Data warehouse нужно рассмотреть:

— информационную безопасность на физическом уровне;

— информационную безопасность на техническом уровне;

— информационную безопасность на административном уровне.

4.2 Обеспечение информационной безопасности физического уровня

Меры безопасности на физическом уровне предусматривают защиту инфраструктуры и данных от физического неправомерного доступа и могут включать[8]:

— использование систем доступа на базе биометрии или смарт-карт и турникетов с защитой от проникновения нескольких лиц одновременно и обратного хода, которые разрешают проходить только одному человеку после аутентификации;

— мониторинг внутреннего пространства при помощи датчиков температуры и дыма;

— использование альтернативных источников питания (например, запасного генератора).

Для обеспечения безопасности на физическом уровне от несанкционированного доступа к хранилищу можно использовать СКУД (система контроля и управления доступом)

Обычно система контроля и управления доступом состоит из ряда компонентов начиная с тех, которые идентифицируют пользователей, и заканчивая теми, что принимают решение о предоставлении доступа[15].

Возможные компоненты СКУД[15]:

— идентификаторы;

— считыватели;

— ограждения;

— турникеты;

— электрозамки;

— автономные контроллеры;

— программное обеспечение (ПО).

В таблице 4.1 приведены минимальные цены на компоненты от самых часто используемых компаний по данным Яндекс Вордстат, цены указаны в рублях[15].

Таблица 4.1 – Цены на компоненты СКУД[15].

Компания	Контроллеры	ПО	Турникеты	Ограждения	Замки	Считыватели
PERco	6 569	Беспл.	68 622	18 997	11 984	4 527
Parsec	14 308	Беспл.	Нет	Нет	Нет	7 448
IronLogic	2 600	Беспл.	Нет	Нет	2 430	1 080
Sigur	12 510	Беспл.	Нет	Нет	Нет	5 890
RusGuard	11 890	Беспл.	Нет	Нет	Нет	1 950
Suprema	24 860	Беспл.	Нет	Нет	Нет	13 255
ZKTeco	4 825	Беспл.	38 000	Нет	5 630	950

PERco - единственный производитель в России который производит почти все элементы систем контроля и управления доступом, что можно увидеть из таблицы 4.1

4.3 Обеспечение информационной безопасности технического уровня

Меры безопасности хранилища данных на техническом уровне включают в себя многие процедуры, такие как, защита сетевого периметра, системы

обнаружения вторжений, фаерволы, антивирусы, аутентификация пользователей и контроль доступа.

В контексте аутентификации пользователей и контроля доступа рекомендуется предпринять следующие меры[8]:

— изменение всех стандартных учетных записей;

— избегание совместного использования учетных записей, отследить которые сложно или невозможно;

— назначение ровно таких прав, которые нужны для выполнения роли;

— изменение или снятие прав при увольнении или смены роли пользователя.

Привилегии для всех пользователей в Data warehouse регулируются при помощи ролей. Каждая роль имеет свой уникальный набор привилегий, предназначенный для возможности выполнения задач той или иной группой пользователей. Всего в хранилище существует 3 пользовательские роли и 1 системная.

ANALYST – роль аналитика Data warehouse, аналитик имеет право только на чтение данные всех пользовательских схем, под пользовательскими схемами понимаются все основные слои хранилища данных, в которых происходит загрузка. Аналитик не имеет право на чтение объектов схем ETL и на любые операции манипулирования данными;

DEVELOPER – роль разработчика Data warehouse, разработчик по умолчанию имеет право на чтение любых пользовательских и ядерных объектов в базе данных, а так же на чтение некоторых системных объектов, однако, как и аналитик, роль не имеет права на операции, связанные с манипулированием данных;

DMREADER – роль внешнего пользователя Data warehouse, выдается пользователям, которые не имеют отношение к разработке продукта, а только используют полученным данные для целей бизнеса, такими пользователями могут являться заказчики, либо BI системы. Для группы таких пользователей выдаются только права на чтение презентационного слоя хранилища, т.е. только представлений в схеме DM.

Единственной системной ролью является роль DBA, это стандартная роль

СУБД Oracle. DBA имеет привилегии выполнять любые операции в БД. Данная роль дана только администраторам базы данных.

Все пользователи хранилища получают привилегии только через присваивание ролей, т.е. непосредственно у самих пользователей нет никаких прав кроме тех, которые определены ролями, предназначенными для них. Единственным исключением является пользователь ETL, при помощи которого происходит выполнение всех алгоритмических операций в хранилище. Данный пользователь обладает возможностью чтения, изменения всех данных и структур объектов в пользовательских и ядерных схемах, однако из соображений безопасности ему не доступны операции над системными объектами. Для оперативного реагирования на инциденты, происходящих в хранилище, доступ к пользователю ETL доступен ведущим разработчикам.

Анализ трафика это одна из наиболее действенных мер в контексте безопасности хранилищ. Отслеживание и детектирование аномальной или подозрительной активности для последующего более тщательного исследования. Эта задача решается при помощи приложений для поведенческого анализа (user and entity behavior analytics, UEBA), систем предотвращения утечек информации (Data Leak Prevention, DLP), сетевых систем обнаружения и предотвращения вторжений(Intrusion Detection/Prevention System, IDS/IPS)[8].

UEBA — это класс систем, которые позволяют на основе массивов данных о пользователях и IT-сущностях (конечных станциях, серверах, коммутаторах и т. д.) с помощью алгоритмов машинного обучения и статистического анализа строить модели поведения пользователей и определять отклонения от этих моделей, как в режиме реального времени, так и ретроспективно. На рисунке 4.1 изображен пример использования UEBA системы.

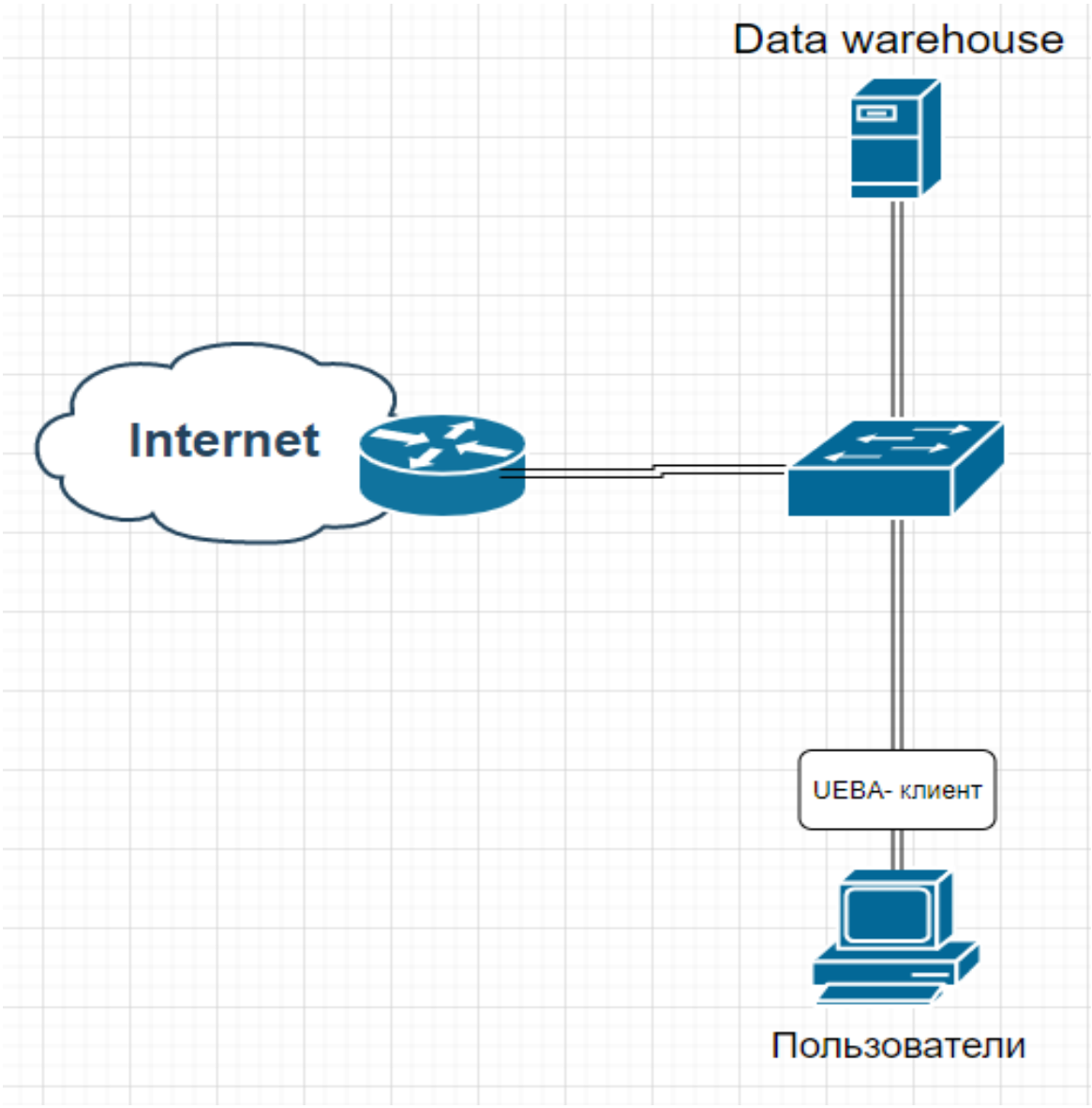

Рисунок 4.1 пример использования UEBA системы

В качестве источников данных могут быть файлы журналов серверных и сетевых компонентов, журналы систем безопасности, локальные журналы с конечных станций, данные из систем аутентификации и даже содержание переписки в социальных сетях, мессенджерах и почтовых сообщениях[16]. В таблице 4.2 приведено сравнение наиболее используемых продуктов в сегменте UEBA.

Таблица 4.2 - Сравнение UEBA продуктов[17].

	Splunk	Securonix	Exabeam Advanced Analytics	Micro Focus Security ArcSight
Локальное программное обеспечение	+	+	+	+
Реагирование на инциденты	+	+	-	+
Машинное обучение	+	+	+	-
Оповещения в реальном времени	+	+	+	+
Настраиваемое уведомление	-	+	-	-
Ролевой доступ к отчетам	-	-	+	-
Интеграция с технологиями	SIEM IAM DLP	SIEM	IAM DLP	SIEM
Логи и пользовательский контекст данных из Active Directory	+	+	+	-
Сетевой поток/Пакетные данные	+	+	+	-
Сбор логов из ОС, приложений, серисов	+	+	+	-
Метаданные электронных сообщений	+	+	-	-
Адаптация системы к динамическим изменениям пользователей	+	+	+	+

DLP — это системы, позволяющие в режиме реального времени производить мониторинг и блокирование входящих и исходящих сообщений сотрудников, отправки файлов на внешние носители, сетевые хранилища информации и веб-ресурсы, а также контроль голосовых и текстовых сообщений, передаваемых по протоколу SIP (Session Initiation Protocol), с целью предотвращения утечки конфиденциальной информации [18].

Системы DLP подразделяются на:

— сетевые;

— агентские;

— гибридные.

Сетевые решения основаны на централизованном мониторинге трафика данных путем подачи его копии на специализированные серверы для обработки согласно заранее настроенным политикам безопасности, преимуществом таких систем является минимальное влияние на существующую инфраструктуру, полное отсутствие какой-либо привязки к рабочим станциям пользователей, относительную простоту внедрения, а также минимизацию рисков несанкционированного доступа к аппаратным компонентам.

Если анализировать только сетевые потоки, будет сложно установить полную картину работы пользователей с конфиденциальной информацией, из за нарастующей популярности мессенджеров, облачных сервисов и других специализированных приложений для обмена информацией[19]. В таком случае может помочь агентское исполнение, которое предполагает установку клиентских программ на все компьютеры пользователей в организации. Эти клиентские программы блокируют несанкционированную передачу конфиденциальных данных, а также контролируют соблюдение политик безопасности и запуск неразрешенных приложений. К тому же агенты собирают максимальное количество сведений о действиях пользователей на корпоративных рабочих станциях и передают информацию в единый центр управления, позволяя специалисту службы безопасности определять инциденты и на основе этих данных строить отчеты[19]. Основным достоинством агентских решений можно считать максимальную «близость» к пользователю, благодаря чему можно контролировать практически любые его действия во всех приложениях.

Гибридный вариант содержит в себе сильные стороны как агентского, так и сетевого решения, в связи с этим переход DLP к гибридной архитектуре выглядит вполне обоснованным (Рисунок 4.2). Многие из существующих систем уже перешли к гибридным решениям[19]. В Приложении А приведено сравнение самых распространенных решений DLP систем[19].

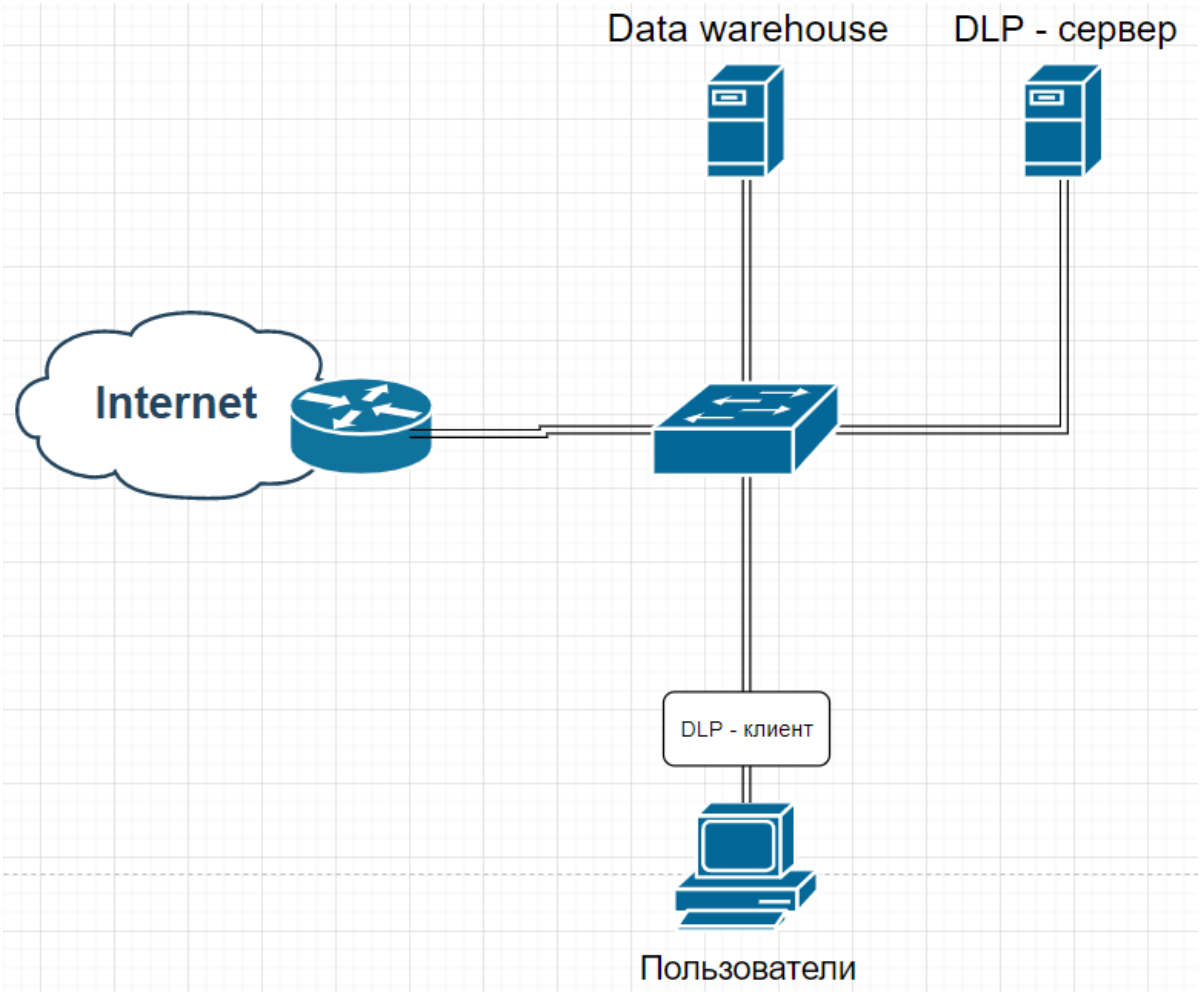

Рисунок 4.2 пример использования DLP системы

Система обнаружения и предотвращения вторжений (IDS/IPS — Intrusion Detection/Prevention System) позволяет реагировать на атаки злоумышленников, использующих известные уязвимости, а также распознавать вредоносную активность внутри сети. IDS/IPS системы предназначены для обнаружения вторжений и защиты сетей компании от атак, неавторизованного проникновения в сеть. Такие решения могут обрывать сомнительные соединения и автоматически настраивать межсетевой экран, который блокирует дальнейшие атаки, а также информируют службу информационной безопасности компании[20].

Чтобы максимально эффективно использовать IDS/IPS, нужно придерживаться следующих рекомендаций[21].

Систему необходимо разворачивать на входе защищаемой сети или подсети и обычно за межсетевым экраном, так как нет смысла контролировать трафик, который будет блокирован. В некоторых случаях датчики устанавливают и внутри сегмента.

Перед активацией функции IPS следует некоторое время погонять систему в режиме, не блокирующем IDS. В дальнейшем потребуется периодически корректировать правила.

Большинство настроек IPS установлены с расчетом на типичные сети. В определенных случаях они могут оказаться неэффективными, поэтому необходимо обязательно указать IP внутренних подсетей и используемые приложения (порты). Это поможет оборудованию лучше понять, с чем она имеет дело.

Если IPS-система устанавливается «в разрыв», необходимо контролировать ее работоспособность, иначе выход устройства из строя может запросто парализовать всю сеть.

Пример использования IDS/IPS изображен на рисунке 4.1.

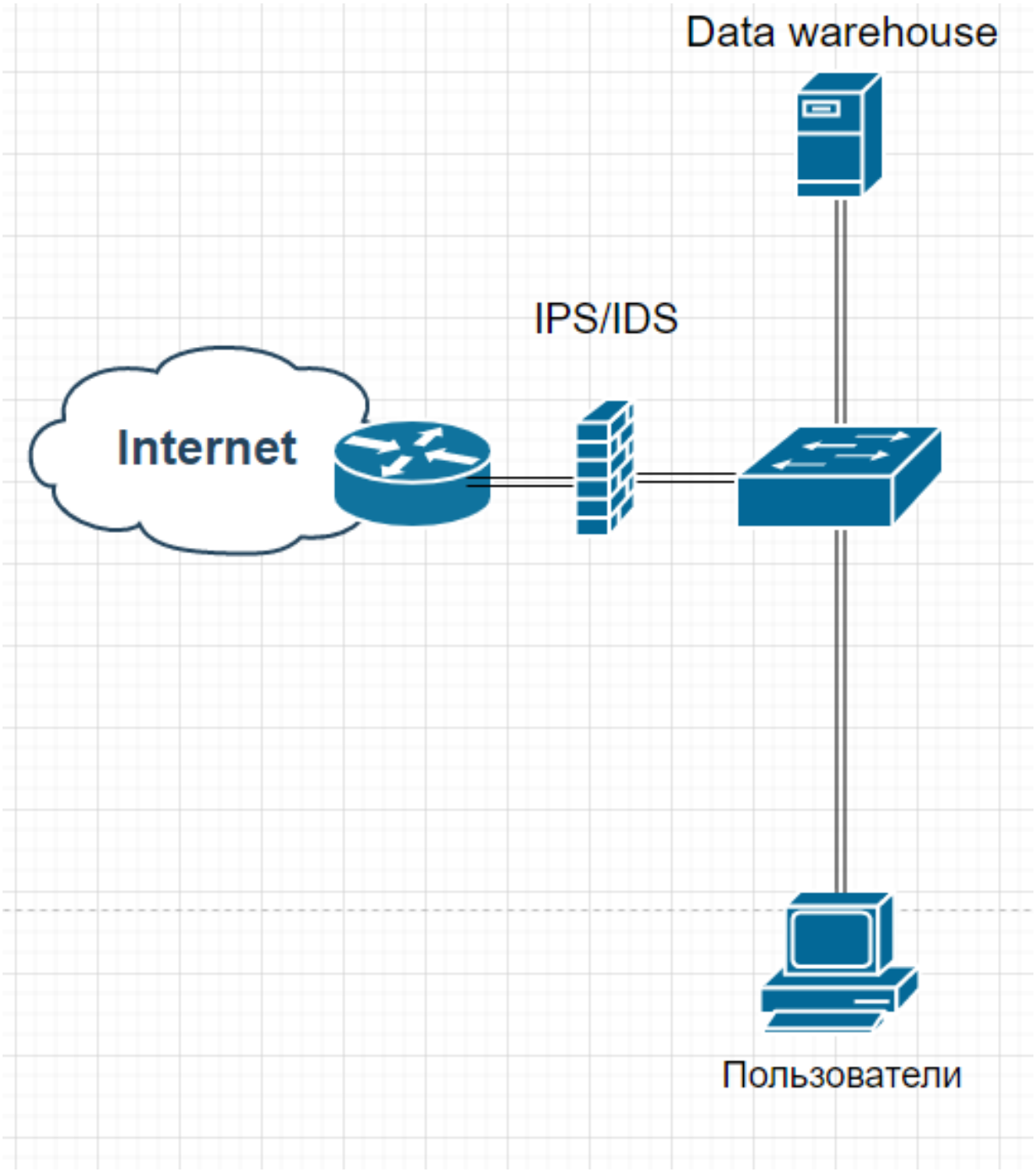

Рисунок 4.3 – Пример использования IDS/IPS[22].

Благодаря использованию IDS/IPS в результате можно получить[22]:

— оперативное обнаружение атак, несанкционированной и запрещённой активности в критических точках сети и на периметре сети;

— уменьшение рисков хакерских атак, проникновения в сеть вирусов, червей, компрометации сетевых ресурсов;

— автоматизация процессов обнаружения и расследования инцидентов информационной безопасности;

— уменьшение потенциального ущерба от возможных инцидентов информационной безопасности;

— соответствие требованиям законодательства, в том числе ФЗ-152.

В России требования к системам обнаружения вторжений появились в 2011 году. ФСТЭК России выделила шесть классов защиты СОВ(Системы обнаружения вторжений). Отличия между ними в уровне информационных систем и самой информации, подлежащей обработке (персональные данные, конфиденциальная информация, гостайна). Соответствие требованиям регулятора важный фактор при выборе решений для защиты от вторжений. Поэтому для гарантированного результата в виде отсутствия санкций относительно выбора ПО стоит обратить внимание на системы обнаружения вторжений, сертифицированные ФСТЭК[23].

Наиболее эффективной идеей защиты инфраструктуры является совместное использование средств IDS и IPS в одном продукте – межсетевом экране, который с помощью глубокого анализа сетевых пакетов, обнаруживает атаки и блокирует их[24].

Межсетевые экраны нового поколения (Next-Generation Firewall, NGFW) — представляют собой интегрированные платформы сетевой безопасности, в которых традиционные брандмауэры сочетаются с другими сетевыми решениями для фильтрации трафика, такими как системы глубокого анализа трафика Deep Packet Inspection (DPI), система обнаружения и предотвращения вторжений (IDS/IPS) и другие.

Решения NGFW производят фильтрацию не просто на уровне портов и протоколов, а на уровне протоколов приложений и функций самих приложений, таким образом заглядывая вглубь транзакций и останавливая активность вредоносного ПО и блокируя сложнейшие методы атак[25].

Согласно определению аналитиков Gartner, межсетевые экраны нового поколения должны гарантированно обеспечивать следующее[25]:

— защиту от непрерывных атак со стороны инфицированных систем;

— стандартные для первого поколения фаерволов возможности;

— сигнатуры определения типов приложений на основе движка IPS;

— полностековое инспектирование трафика, включая приложения, а также детальный и настраиваемый контроль на уровне приложений;

— возможность включать информацию за пределами брандмауэра (например, интеграция с сетевыми каталогами, «белыми» и «черными» списками приложений);

— постоянно обновляемую базу описаний приложений и угроз;

— инспекцию трафика, шифруемого с помощью SSL.

В приложении Б приведено сравнение межсетевых экранов[26].

4.4 Обеспечение информационной безопасности административного уровня

К административному уровню информационной безопасности относятся действия общего характера, предпринимаемые руководством организации.

Главная цель мер административного уровня сформировать программу работ в области информационной безопасности и обеспечить ее выполнение, выделяя необходимые ресурсы и контролируя состояние дел.

Основой является политика безопасности, отражающая подход организации к защите своих информационных активов. Руководство каждой организации должно осознать необходимость поддержания режима безопасности и выделения на эти цели значительных ресурсов[27].

Политика безопасности строится на основе анализа рисков, которые признаются реальными для информационной системы организации. Когда риски проанализированы и стратегия защиты определена, составляется программа обеспечения информационной безопасности. Под эту программу выделяются ресурсы, назначаются ответственные, определяется порядок контроля выполнения программы.

Типовой структурой политики безопасности может быть[28]:

1. Общие положения.

1.1. Назначение документа.

1.2. Основания для разработки документа.

1.3. Основные определения.

2. Идентификация системы.

2.1. Идентификатор и имя системы.

2.2. Ответственные подразделения.

2.3. Режим функционирования системы.

2.4. Описание и цели системы.

2.5. Цели и задачи ПБ.

2.6. Системная среда.

2.6.1. Физическая организация системы.

2.6.2. Логическая организация системы.

2.7. Реализованные сервисы системы.

2.8. Общие правила, принятые в системе.

2.9. Общее описание важности информации.

3. Средства управления.

3.1. Оценка рисков и управление.

3.2. Экспертиза СЗИ.

3.3. Правила поведения, должностные обязанности и ответственность.

3.4. Планирование безопасности.

3.5. Разрешение на ввод компонента в строй.

3.6. Порядок подключения подсетей подразделения к сетям общего пользования.

4. Функциональные средства.

4.1. Защита персонала.

4.2. Управление работой и вводом-выводом.

4.3. Планирование непрерывной работы.

4.4. Средства поддержки программных приложений.

4.5. Средства обеспечения целостности информации.

4.6. Документирование.

4.7. Осведомленность и обучение специалистов.

4.8. Ответные действия в случаях возникновения происшествий.

5. Технические средства.

5.1. Требования к процедурам идентификации и аутентификации.

5.2. Требования к системам контроля и разграничения доступа.

5.3. Требования к системам регистрации сетевых событий.

Примерные инструкции по реализации ПБ могут быть, например, следующими:

1. Требования к защите портов и служб.

2. Порядок проведения экспертизы СЗИ.

3. Порядок проведения анализа рисков.

4. Использование автоматизированных систем анализа защищенности.

5. Порядок восстановления автоматизированных систем после аварийных ситуаций.

4.4 Выводы по разделу

В данном разделе были рассмотрены способы обеспечения информационной безопасности на:

— физическом уровне;

— техническом уровне;

— административном уровне.

На рисунке 4.4 изображена структура сети с обеспечением информационной безопасности благодаря интегрированным системам:

— IDS/IPS;

— DLP;

— UEBA.

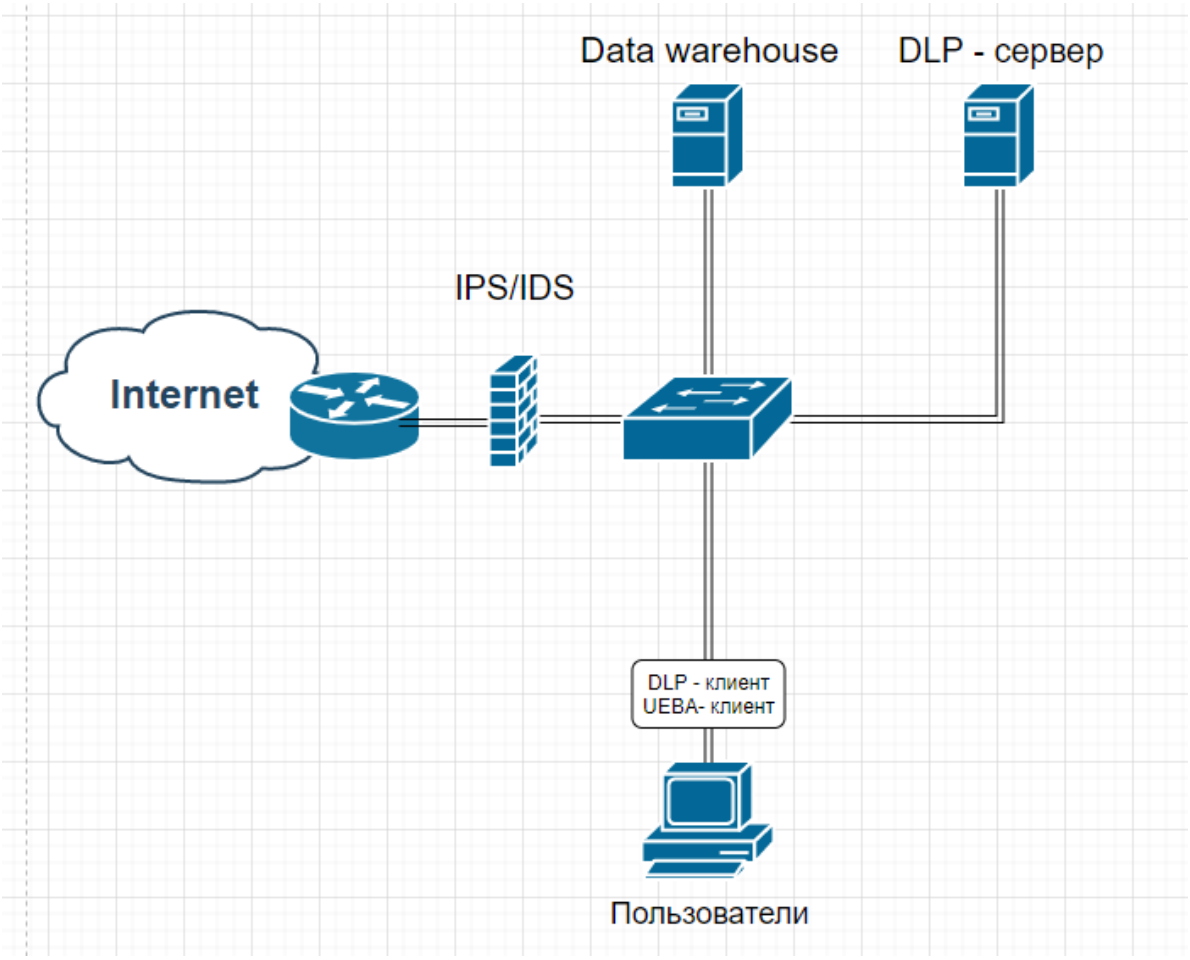

Рисунок 4.4 – структура сети с обеспечением информационной безопасности

Лучшим выбором UEBA системы из предложенных в таблице 4.2 будет Splunk поскольку он имеет возможность реагирования на инциденты, а так же интеграцию в DLP систему.

Из DLP систем больше всего подходит Infowatch Traffic monitor Enterprise поскольку одновременно имеет лицензию как ФСТЭК, так и ФСБ.

В качестве IDS/IPS выступает межсетевой экран с модулем IDS/IPS, наилучшим выбором будет оборудование Usergate, за счет более широкого выбора возможностей.

5 Безопасность жизнедеятельности

5.1 Постановка задачи

В данной главе будут рассмотрены следующие вопросы по безопасности жизнедеятельности:

5.2 Характеристика условий труда при работе с ПК

Условиями труда согласно ст. 209 ТК РФ является совокупность факторов производственной среды и трудового процесса, оказывающих влияние на работоспособность и здоровье работника. Вредным же производственным фактором в силу упомянутой ст. 209 ТК РФ признается фактор, воздействие которого на работника может привести к его заболеванию.

Регулярная работа за компьютером сопровождается постоянным влиянием множества вредных для здоровья факторов. Зачастую специалисты, проводящие больше 12 часов в день за компьютером, со временем начинают страдать от профессиональных заболеваний. Поэтому для работников, которые работают с персональной электронно-вычислительной машиной (ПЭВМ) очень важна правильная организация рабочего места. В таблице 5.1 приведены требования к условиям труда[30].

Таблица 5.1 – Требования к условиям труда[31]

№	Требования
Требования к помещениям для работы с ПЭВМ	
1	Естественное и искусственное освещение должно соответствовать требованиям действующей нормативной документации. Окна в помещениях, где эксплуатируется вычислительная техника, преимущественно должны быть ориентированы на север и северо-восток. Оконные проемы должны быть оборудованы регулируемыми устройствами типа: жалюзи, занавесей, внешних козырьков и др.

Продолжение таблицы 5.1 [31]

2	Площадь на одно рабочее место пользователей ПЭВМ с ВДТ на базе электронно-лучевой трубки (ЭЛТ) должна составлять не менее 6 м2, в помещениях культурно-развлекательных учреждений и с ВДТ на базе плоских дискретных экранов (жидкокристаллические, плазменные) - 4,5 м2.
3	Для внутренней отделки интерьера помещений, где расположены ПЭВМ, должны использоваться диффузно отражающие материалы с коэффициентом отражения для потолка - 0,7 - 0,8; для стен - 0,5 - 0,6; для пола - 0,3 - 0,5.
4	Помещения, где размещаются рабочие места с ПЭВМ, должны быть оборудованы защитным заземлением (занулением) в соответствии с техническими требованиями по эксплуатации.
5	Не следует размещать рабочие места с ПЭВМ вблизи силовых кабелей и вводов, высоковольтных трансформаторов, технологического оборудования, создающего помехи в работе ПЭВМ.
Требования к микроклимату, содержанию аэроионов и вредных химических веществ в воздухе на рабочих местах, оборудованных ПЭВМ	
1	В производственных помещениях, в которых работа с использованием ПЭВМ является основной (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.) и связана с нервно-эмоциональным напряжением, должны обеспечиваться оптимальные параметры микроклимата для категории работ 1а и 1б в соответствии с действующими санитарно-эпидемиологическими нормативами микроклимата производственных помещений. На других рабочих местах следует поддерживать параметры микроклимата на допустимом уровне, соответствующем требованиям указанных выше нормативов. Подробнее в СанПиН 2.2.4.3359-16 Санитарно-эпидемиологические требования к физическим факторам на рабочих местах.
2	В помещениях, оборудованных ПЭВМ, проводится ежедневная влажная уборка и систематическое проветривание после каждого часа работы на ПЭВМ.
3	Уровни положительных и отрицательных аэроионов в воздухе помещений, где расположены ПЭВМ, должны соответствовать действующим санитарно-эпидемиологическим нормативам.
4	Содержание вредных химических веществ в производственных помещениях, в которых работа с использованием ПЭВМ является основной (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.), не должно превышать предельно допустимых концентраций загрязняющих веществ в атмосферном воздухе населенных мест в соответствии с действующими гигиеническими нормативами.
Требования к уровням шума и вибрации на рабочих местах, оборудованных ПЭВМ	
1	В производственных помещениях при выполнении основных или вспомогательных работ с использованием ПЭВМ уровни шума на рабочих местах не должны превышать предельно допустимых значений, установленных для данных видов работ в соответствии с действующими санитарно-эпидемиологическими нормативами. Подробнее в СанПиН 2.2.4.3359-16 Санитарно-эпидемиологические требования к физическим факторам на рабочих местах

Продолжение таблицы 5.1[31]

2	Шумящее оборудование (печатающие устройства, серверы и т.п.), уровни шума которого превышают нормативные, должно размещаться вне помещений с ПЭВМ
Требования к освещению на рабочих местах, оборудованных ПЭВМ	
1	Рабочие столы следует размещать таким образом, чтобы видеодисплейные терминалы были ориентированы боковой стороной к световым проемам, чтобы естественный свет падал преимущественно слева
2	Искусственное освещение в помещениях для эксплуатации ПЭВМ должно осуществляться системой общего равномерного освещения.
3	Освещенность на поверхности стола в зоне размещения рабочего документа должна быть 300 - 500 лк. Освещение не должно создавать бликов на поверхности экрана. Освещенность поверхности экрана не должна быть более 300 лк.
4	Следует ограничивать прямую блесткость от источников освещения, при этом яркость светящихся поверхностей (окна, светильники и др.), находящихся в поле зрения, должна быть не более 200 кд/м2.
5	Следует ограничивать отраженную блесткость на рабочих поверхностях (экран, стол, клавиатура и др.) за счет правильного выбора типов светильников и расположения рабочих мест по отношению к источникам естественного и искусственного освещения, при этом яркость бликов на экране ПЭВМ не должна превышать 40 кд/м2 и яркость потолка не должна превышать 200 кд/м2.
6	В качестве источников света при искусственном освещении следует применять преимущественно люминесцентные лампы типа ЛБ и компактные люминесцентные лампы (КЛЛ). При устройстве отраженного освещения в производственных и административно-общественных помещениях допускается применение металлогалогенных ламп.
7	Применение светильников без рассеивателей и экранирующих решеток не допускается.
8	Коэффициент запаса (Кз) для осветительных установок общего освещения должен приниматься равным 1,4.
9	Коэффициент пульсации не должен превышать 5%.
Требования к уровням электромагнитных полей на рабочих местах, оборудованных ПЭВМ	
1	Временные допустимые уровни ЭМП, создаваемых ПЭВМ на рабочих местах пользователей представлены в таблице 5.2
Требования к визуальным параметрам ВДТ, контролируемым на рабочих местах	
1	Предельно допустимые значения визуальных параметров ВДТ, контролируемые на рабочих местах, представлены в таблице 5.3
Общие требования к организации рабочих мест пользователей ПЭВМ	
1	При размещении рабочих мест с ПЭВМ расстояние между рабочими столами с видеомониторами (в направлении тыла поверхности одного видеомонитора и экрана другого видеомонитора) должно быть не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов - не менее 1,2 м.
2	Рабочие места с ПЭВМ при выполнении творческой работы, требующей значительного умственного напряжения или высокой концентрации внимания, рекомендуется изолировать друг от друга перегородками высотой 1,5 - 2,0 м.
3	Экран видеомонитора должен находиться от глаз пользователя на расстоянии 600 - 700 мм, но не ближе 500 мм с учетом размеров алфавитно-цифровых знаков и символов.

Продолжение таблицы 5.1 [31]

4	Конструкция рабочего стола должна обеспечивать оптимальное размещение на рабочей поверхности используемого оборудования с учетом его количества и конструктивных особенностей, характера выполняемой работы. При этом допускается использование рабочих столов различных конструкций, отвечающих современным требованиям эргономики. Поверхность рабочего стола должна иметь коэффициент отражения 0,5 - 0,7.
5	Конструкция рабочего стула (кресла) должна обеспечивать поддержание рациональной рабочей позы при работе на ПЭВМ, позволять изменять позу с целью снижения статического напряжения мышц шейно-плечевой области и спины для предупреждения развития утомления. Тип рабочего стула (кресла) следует выбирать с учетом роста пользователя, характера и продолжительности работы с ПЭВМ.
6	Поверхность сиденья, спинки и других элементов стула (кресла) должна быть полумягкой, с нескользящим, слабо электризующимся и воздухопроницаемым покрытием, обеспечивающим легкую очистку от загрязнений.

Таблица 5.2 - Временные допустимые уровни ЭМП, создаваемых ПЭВМ на рабочих местах [31]

Наименование параметров		ВДУ
Напряженность электрического поля	в диапазоне частот 5 Гц - 2 кГц	25 В/м
	в диапазоне частот 2 кГц - 400 кГц	2,5 В/м
Плотность магнитного потока	в диапазоне частот 5 Гц - 2 кГц	250 нТл
	в диапазоне частот 2 кГц - 400 кГц	25 нТл
Напряженность электростатического поля		15 кВ/м

Таблица 5.3 - Визуальные параметры ВДТ, контролируемые на рабочих местах[31]

N п/п	Параметры	Допустимые значения
1	Яркость белого поля	Не менее 35 кд/кв. м
2	Неравномерность яркости рабочего поля	Не более +/- 20%
3	Контрастность (для монохромного режима)	Не менее 3:1
4	Временная нестабильность изображения (мелькания)	Не должна фиксироваться

Продолжение таблицы 5.3

5	Пространственная нестабильность изображения (дрожание)	Не более 2 х 1Е(-4L), где L - проектное расстояние наблюдения, мм

5.2 Влияние условий труда на здоровье работников

Условия труда это достаточно сложное явление, характеризующее среду протекания трудового процесса, формирующееся под воздействием взаимосвязанных факторов социально-экономического, технико-организационного и естественно-природного характера и влияющее на здоровье, работоспособность человека, его отношение к труду и степень удовлетворенности трудом, а следовательно, на эффективность труда и другие экономические результаты деятельности[32].

Основными директивными документами, регламентирующими условия труда, являются санитарные нормы проектирования предприятий, Строительные нормы и правила (СНиП), ГОСТы, требования техники безопасности и охраны труда.

В санитарных нормах проектирования промышленных предприятий установлены предельно допустимые концентрации (ПДК) содержания вредных веществ в рабочей зоне. Для обеспечения нормальных условий труда необходимо совершенствование технологии, герметизация и автоматизация оборудования, вентиляция производственных помещений.

Условия труда представляют собой совокупность различных по воздействию на человека элементов, которые можно разделить на четыре группы:

санитарно-гигиенические элементы, образующие предметную внешнюю среду: микроклимат, состояние воздушной среды (запыленность, загазованность), освещение, производственные излучения, шум, вибрация[32];

психологические и физиологические элементы, обусловленные содержанием трудовой деятельности, различными нагрузками на двигательный аппарат, нервную систему и психику человека в процессе труда;

эстетические элементы, которые формируют у человека отношение к среде протекания труда с точки зрения ее художественного восприятия и оказывают большое воздействие на формирование определенного эмоционального состояния;

социально-психологические элементы, характеризующие психологическое состояние работников и коллектива и создающие соответствующий психологический, эмоциональный настрой работника.

Можно выделить четыре группы факторов, влияющих на формирование и изменение условий труда[32]:

социальные и экономические факторы, действие которых обуславливает положение трудящихся в обществе:

нормативно-правовые факторы (законы о труде, правила, нормы, стандарты в области организации, оплаты, условий и охраны труда, режимов труда и отдыха, установления льгот и социальных гарантий отдельным категориям работ-ников, а также система государственного и общественного контроля за их соблюдением);

социально-психологические факторы, характеризующие отношение в обществе к сфере трудовой деятельности и условиям труда, совокупность интересов и ценностных ориентаций работников, состав и особенности персонала, стиль руководства и т.п.;

общественные факторы (общественные организации, движения, за улучшение экологической обстановки, создание благоприятных условий труда и др.);

экономические факторы (система льгот, гарантий и компенсаций работникам, с одной стороны, а с другой – система экономических санкций за нарушение норм, стандартов и проч.);

технические и организационные факторы, оказывающие непосредственное воздействие на формирование материально-вещественных элементов условий труда: средства труда, предметы труда, технологические процессы, организационные формы производства труда и управления, режимы труда и отдыха, формы разделения и кооперации труда, приемы и методы труда, нормирования труда и т.п.;

Естественно-природные факторы, характеризующие воздействие на работников географо-климатических, геологических и биологических особенностей местности, где протекает трудовой процесс;

хозяйственно-бытовые факторы, связанные с организацией питания работников, их санитарного и бытового обслуживания.

Таким образом, условия труда могут рассматриваться в технических, организационных, психофизиологических, социальных, правовых и других аспектах[32].

5.4 Причины и профилактика зрительного утомления

Работа с дисплеем предполагает, прежде всего, визуальное восприятие отображенной на экране монитора информации, поэтому значительной нагрузке подвергается зрительный аппарат работающих с ПК.

Факторами, наиболее сильно влияющими на зрение, являются:

Несовершенство способов создания изображения на экране монитора. Эта группа факторов включает в себя[33]:

— несовместимость параметров монитора и графического адаптера;

— недостаточно высокое разрешение монитора, расфокусировка;

— избыточная или недостаточная яркость изображения.

Непродуманная организация рабочего места является причиной:

— наличия бликов на лицевой панели экрана;

— отсутствия необходимого уровня освещенности рабочих мест;

— несоблюдения расстояния от глаз оператора до экрана.

Блики относятся к факторам, которые очень сильно мешают воспринимать информацию с экрана монитора. Они заставляют напрягать зрение, чтобы прочесть нужную информацию на экране[33].

Блики создает любой пучок света, отраженный экраном дисплея и попавший на оболочку глаза. Их источниками могут быть расположенные напротив монитора яркие поверхности, светлое оборудование, осветительные приборы, незашторенные окна, часто – светлая одежда оператора. Блики тем заметнее и тем сильнее

снижают контрастность изображения, чем выше коэффициент зеркального отражения экрана. В ряде случаев текст может стать фактически нечитаемым.

В процессе работы с ПК приходится иметь дело с изображениями на экране монитора. Считывание текста, таблиц, графиков с экрана отличается от чтения той же информации с листа бумага по нескольким причинам.

При работе с дисплеем пользователь целиком зависит от положения дисплея, тогда как при чтении печатной продукции легко можно найти положения листа для наиболее комфортного восприятия информации.

Экран, являясь источником света, считается прибором активного контраста, в то время как при чтении с листа бумаги мы имеем дело с отраженным текстом, т. е. с пассивным контрастом, который в малой степени зависит от интенсивности освещения и угла падения светового потока на бумагу[33].

Текст на бумаге является неизменным, а текст на экране периодически обновляется в процессе сканирования электронного луча по поверхности экрана. Достаточно низкая частота обновления вызывает мерцание изображений;

Монитор надолго приковывает к себе внимание оператора, что является причиной длительной неподвижности глазных и внутриглазных мышц, в то время как они нуждаются в динамическом режиме работы. Это приводит к их ослаблению.

Работа компьютера сопровождается акустическими шумами, включая ультразвук.

Длительная работа с ПК требует повышенной сосредоточенности, что приводит к большим нагрузкам на зрительную систему пользователя. Развивается зрительное утомление (астенопия), которое способствует возникновению близорукости, головной боли, раздражительности, нервного напряжения и стресса.

Приведенные выше особенности изображений на экране дисплея, а также характера работы оператора в значительной мере влияют на степень утомляемости зрительного аппарата. Поэтому ведущие производители мониторов постоянно ведут работу по совершенствованию характеристик своих изделий не только с цепью достижения новых технических возможностей отображения информации, но и для

создания максимально комфортных и безопасных условий труда пользователей ПК.

В современных мониторах, чтобы уменьшить отражения, темное или тонированное стекло, проводят специальную химическую обработку лицевой поверхности (покрытие двуокисью кремния, обработку травлением); применяют цилиндрические (или вертикально-плоские экраны - ЭЛТ TRINITRON и DIAMONDNRON) и плоские прямоугольные экраны (обладают лучшими антибликовыми свойствами в силу действия обычных законов отражения), а также используют защитные фильтры.

Проблемы снижения зрительного утомления решают с помощью применения специальных защитных средств, правильной организации рабочего места, режимов труда и отдыха, специальных упражнений для снятия утомления.

5.5 Экологические проблемы утилизации офисного оборудования

Компьютерная техника, которая является не рабочей и устаревшей, не может быть выброшена вместе с бытовыми и другими видами отходов.

Законы регулируют необходимость утилизации офисной техники. За несоблюдение данных законов на организации могут быть наложены серьезные штрафные санкции[34].

Самая главная причина, которую должно учитывать руководство организации при решении об утилизации офисной техники, является забота об окружающей среде. При сдаче на переработку техники количество не переработанных опасных отходов снижается.

Еще одна немаловажная причина, по которой требуется законная утилизация отработанной компьютерной и офисной техники, – необходимый учет драгоценных металлов, которые содержатся в данных видах техники.

Почти во всех компьютерах, мониторах и иной оргтехнике в небольшом количестве присутствуют золото, серебро и другие драгоценные металлы. Любая

организация обязана документально оформлять их поступление, движение, инвентаризацию и выбытие.

Российское законодательство предусматривает ведение строгого учета всех драгоценных металлов, которые имеются на предприятии, в том числе тех, что являются элементами различной компьютерной техники.

Списание офисной техники включает в себя[34]:

— определение технического состояния каждой единицы основных средств;

— оформление необходимой документации;

— получение разрешения на списание;

— демонтаж, разборку;

— утилизацию объектов и постановку на учет материалов, полученных от их ликвидации; списание с балансового учета.

Согласно федеральному закону от 24.06.1998 N 89-ФЗ "Об отходах производства и потребления" учреждение вправе:

— самостоятельно обрабатывать (перерабатывать) собранный лом, содержащий драгоценные металлы;

— реализовывать лом, содержащий драгоценные металлы;

— передавать на давальческой основе аффинажным организациям или организациям, осуществляющим деятельность по заготовке лома и отходов, первичной обработке и переработке, для дальнейшего производства и аффинажа.

5.6 Выводы по разделу

В данном разделе были рассмотрены характеристики условий труда при работе с персональным компьютером. Определено влияние условий труда на здоровье человека. Указаны экологические проблемы утилизации офисного оборудования а так же причины и профилактика зрительного утомления.

Выполненная работа, поможет защитить работника от негативных опасностей антропогенного и естественного происхождения и обеспечит комфортные и безопасные условия труда.

6 Технико-экономическое обоснование работы

6.1 Постановка задачи

Темой дипломного проекта является «Разработка проекта защищенного DWH». Поскольку хранилище данных в основе является базой данных, то данная разработка считается объектом интеллектуальной собственности (ОИС), права на который Законом РФ «О правовой охране программ для электронных вычислительных машин и баз данных» и Законом РФ «Об авторском праве и смежных правах».

В данном разделе будут рассмотрены следующие вопросы:

— расчет трудоемкости и длительности работ;

— расчет себестоимости и цены программного продукта.

6.2 Расчет трудоемкости и длительности работ

Процесс разработки DWH разбит на несколько этапов:

1. анализ возможностей DWH;
2. разработка нарушителя модели и угроз;
3. выбор программного обеспечения;
4. разработка DWH;
5. Обеспечение информационной безопасности DWH;

Далее требуется рассчитать трудоемкость и длительность работ. Поскольку трудоемкость этапов и видов работ носит вероятностный характер, то предпочтительным будет использование метода экспертных оценок.

В этом методе для каждого этапа требуется экспертным путем определить три оценки трудоемкости, в днях:

— наименее возможная величина затрат, a_i;

— наиболее вероятная величина затрат, m_i;

— наиболее возможная величина затрат, b_i;

На основании экспертных оценок средняя величина для a_i, m_i и b_i определяется по формуле (6.1):

$$\overline{T} = \frac{3T_{рук} + 2T_{авт}}{5}, \quad (6.1)$$

где $\overline{T}$ – среднее время, полученное на основании экспертных оценок;

$T_{рук}$ – оценка затрат времени, данная руководителем;

$T_{авт}$ – оценка затрат времени, данная автором проекта.

Результаты расчета средней оценки затрат времени на разработку программного продукта приведены в таблице 5.1.

Таблица 6.1 – Время, затраченное на разработку программного продукта

Этапы разработки программного продукта	Наименее возможная величина затрат (a_i), дни			Наиболее вероятная величина затрат (m_i), дни			Наиболее возможная величина затрат (b_i), дни		
	$T_{авт}$	$T_{рук}$	$\overline{T}$	$T_{авт}$	$T_{рук}$	$\overline{T}$	$T_{авт}$	$T_{рук}$	$\overline{T}$
1. Анализ возможностей DWH	3	2	2,4	4	2	2,8	6	4	4,8
2. Разработка нарушителя модели и угроз	4	3	3,4	6	5	5,4	10	8	8,8
3. Выбор програмного обеспечения	6	5	5,4	8	6	6,8	9	7	7,8
4. Разработка DWH	20	18	18,8	22	20	20,8	26	22	23,6
5. Обеспечение информационной безопасности DWH	15	12	13,2	20	18	18,8	22	20	20,8

На основе средних оценок рассчитываются математическое ожидание и отклонение по каждому этапу разработки программного продукта. Формула расчета математического ожидания для i-го этапа:

$$MO_i = \frac{a_i + 4m_i + b_i}{6}, \quad (6.2)$$

где MO_i – математическое ожидание для i-го этапа;

a_i, m_i, b_i – средние значения.

Стандартное отклонение для каждого этапа разработки программного продукта определяется по формуле:

$$G_i = \frac{\mathrm{b}_i - \mathrm{a}_i}{6}, \tag{6.3}$$

где G_i – стандартное отклонение по i-му этапу.

Зная математическое ожидание по каждому этапу, рассчитываем общую величину математического ожидания в целом по программному продукту:

$$MO = \sum MO_i, \tag{6.4}$$

где MO – общая величина математического ожидания.

Стандартное отклонение G в целом по программному продукту рассчитывается по следующей формуле:

$$G = \sqrt{\sum G_i^2}, \tag{6.5}$$

где G –стандартное отклонение;

G_i – стандартное отклонение по i-му этапу.

На основе расчетов математического ожидания (6.4) и стандартного отклонения (6.5) рассчитываем коэффициент вариации – коэффициент согласованности мнения экспертов. Коэффициент вариации рассчитывается по формуле:

$$v_i = \frac{G_i}{MO_i}, \tag{6.6}$$

где v_i – коэффициент вариации по i-му этапу.

Все произведенные расчеты сведены в таблицу 6.2.

Таблица 6.3 – Затраты на разработку программного продукта

Этапы разработки программного продукта	Средняя величина затрат по этапам, дни			Матем. ожидание (MO_i, дни)	Станд. отклонение (G_i, дни)	Коэффициент вариации (v_i)
	Наименее возможная величина затрат (a_i, дни)	Наиболее вероятная величина затрат (m_i, дни)	Наиболее возможная величина затрат (b_i, дни)			
1. Анализ возможностей DWH	2,4	2,8	4,8	3,07	0,40	0,130
2. Разработка нарушителя модели и угроз	3,4	5,4	8,8	5,63	0,90	0,160
3. Выбор програмного обеспечения	5,4	6,8	7,8	6,73	0,40	0,059
4. Разработка DWH	18,8	20,8	23,6	20,93	0,80	0,038
5. Обеспечение информационной безопасности DWH	13,2	18,8	20,8	18,20	1,27	0,070
Итого	43,2	54,6	65,8	54,57	1,84	0,034

В итоге коэффициент вариации равен 0,034 и не превосходит 0,33. Поэтому мнения экспертов считаются согласованными.

6.3 Расчет себестоимости и цены программного продукта

Себестоимость программного продукта – это все виды затрат, понесенные при разработке продукта. Чтобы определить себестоимость разработки применяется метод экспертных оценок.

Себестоимость программного продукта определяется по формуле (6.7):

$$\mathrm{C} = \frac{З}{m} \cdot k \cdot k_{\text{ТЕР}} \cdot k_{\text{ПР}} \cdot (\mathrm{t}_1 + \mathrm{t}_2) \cdot (1 + \mathrm{k}_{\text{Н}}) + 8 \cdot \mathrm{t}_3 \cdot \mathrm{C}_{\text{М}} + 8 \cdot \mathrm{t}_4 \cdot \mathrm{C}_{\text{И}}, \quad (6.7)$$

где $З$ – среднемесячная заработная плата DWH-разработчика, $З = 80000$;

$k_{\text{ТЕР}}$ – территориальный коэффициент, $k_{\text{ТЕР}} = 1{,}2$ (для НСО);

$k_{\text{ПР}}$ – коэффициент премии, $k_{\text{ПР}} = 1$;

k – коэффициент, учитывающий страховые взносы (фонды пенсионного, социального и медицинского страхования), $k = 1{,}3$;

m – количество рабочих дней в месяце, $m = 22$;

$\mathrm{k}_{\text{Н}}$ – коэффициент, учитывающий накладные расходы (отопление, освещение, уборка и т. д.), $\mathrm{k}_{\text{Н}} = 0{,}4$;

t_1 – время, затраченное разработчиком на разработку требований к программе, т.е. подготовительное время, которое необходимо потратить, чтобы преступить к написанию программы и отладки программы, чел./дни;

t_2 – сборка устройства, составление алгоритма в программе, время, затраченное на написание и отладку программы, чел./дни;

t_3 – время, затраченное на разработку программы с использованием машинного времени, чел./дни;

t_4 – время работы в сети интернет, дни;

$\mathrm{C}_{\text{И}}$ – стоимость 1 дня работы в сети интернет, руб. (оценивается через абонентскую плату);

$\mathrm{C}_{\text{М}}$ – стоимость одного часа машинного времени.

Для расчета стоимости одного часа машинного времени, необходимо определить затраты на эксплуатацию ПК за год по следующей формуле:

$$C_{\text{м}} = \frac{З_{\text{эл}} + З_{\text{а}} + З_{\text{компл}} + З_{\text{пр}}}{T_{\text{общ}}}. \quad (6.8)$$

Общее время работы компьютера за год составляет:

$$T_{\text{общ}} = 22 * 12 * 8 = 2112 \text{ (часов)}$$

Затраты на электроэнергию за год работы (на данный момент тариф $C_{\text{эл}}$ составляет 2,68 руб. за кВт/ч):

$$З_{эл} = T_{общ} * C_{эл} * P, \quad (6.9)$$

где P – потребляемая мощность ПК по паспортным данным в час, $P = 500$ Вт/ч.

По (6.9) затраты на электроэнергию за год работы составляют:

$$З_{эл} = 2112 * 2{,}68 * 0{,}500 = 2830{,}1 \text{ (руб.)}$$

Амортизационные отчисления в год определяются как процент отчисления на амортизацию от первоначальной стоимости основных производственных фондов. Процент отчисления на амортизацию, согласно ст. 258 НК РФ, составляет 34-50% от первоначальной стоимости ПК (компьютер относится ко второй группе имущества со сроком полезного использования свыше 2 лет до 3 лет включительно). Затраты на ПК определяются по формуле:

$$З_а = С * П_р, \quad (6.10)$$

где С – стоимость ПК, руб.;

$П_р$ – процент отчисления на амортизацию, $П_р = 40\%$.

Получим:

$$З_а = 60000 * 0{,}4 = 24000 \text{ (руб.)}$$

Затраты на комплектующие материалы составляют:

$$З_{компл} = 5000 \text{ (руб.)}$$

Прочие расходы составляют 5% от общей суммы затрат:

$$З_{пр} = \frac{0{,}05 * (З_{эл} + З_а + З_{компл})}{0{,}95}. \quad (6.11)$$

По (6.11) прочие расходы равны:

$$З_{пр} = \frac{0{,}05 * (2830{,}1 + 24000 + 5000)}{0{,}95} = 1675{,}2 \text{ (руб.)}$$

По формуле 5.8 стоимость одного часа машинного времени равна:

$$C_м = \frac{2830{,}1 + 24000 + 5000 + 1675{,}2}{2112} = 15{,}86 \text{ (руб.)}$$

Тариф на услугу интернет составляет 1200 руб. в месяц, следовательно, стоимость 1 дня работы в сети интернет равен:

$$С_и = \frac{1200}{30} = 40 \text{ (руб.)}$$

Заключительным этапом расчета является распределение ранее рассчитанной трудоемкости (таблица 6.3) по 4 направлениям:

– t_1 включает первые четыре этапа:

$$t_1 = 3{,}07 + 5{,}63 = 8{,}7 \text{ (дней)}$$

– t_2 включает оставшиеся этапы:

$$t_2 = 6{,}73 + 20{,}93 + 18{,}20 = 45{,}86 \text{ (дней)}$$

– t_3 включает время работы ПК для разработки DWH:

$$t_3 = 55 \text{ (дней)}$$

– t_4 включает время использования интернета для разработки программы:

$$t_4 = 55 \text{ (дней)}$$

Наконец, итоговая себестоимость программного продукта составляет:

$$С = \frac{80000}{22} \cdot 1{,}3 \cdot 1{,}2 \cdot 1 \cdot (8{,}7 + 45{,}86) \cdot (1 + 0{,}4) + 8 \cdot 45 \cdot 15{,}86 + 8 \cdot 55 \cdot 40$$
$$= 456615{,}2 \text{ (руб.)}$$

В случае, если программный продукт будет доработан и реализован на рынке, следует рассчитать цену по следующей формуле:

$$Ц = C * (1 + \frac{P}{100}), \quad (6.12)$$

где C – себестоимость разработки программы, руб;

P – рентабельность, руб.

Определим цену программного продукта, при условии, что значение рентабельности равно 20%:

$$Ц = 456615{,}2 \cdot \left(1 + \frac{20}{100}\right) = 547938{,}24 \text{ (руб.)}$$

Цена с учетом налога на добавленную стоимость находится по формуле:

$$Ц_{НДС} = Ц * К_{НДС}, \quad (6.13)$$

где Ц – цена программного продукта;

$К_{НДС}$ – коэффициент, учитывающий ставку налога на добавленную стоимость (НДС), $К_{НДС} = 1{,}20$

Цена с учетом налога на добавленную стоимость составит:

$$Ц_{НДС} = 547938{,}24 * 1{,}20 = 657525{,}9 \text{ (руб.)}$$

6.4 Выводы по разделу

В данном разделе были определены и рассчитаны трудоемкость и длительность работ, а также рассчитаны себестоимость и цена программного продукта. Итоговая цена с учетом налога на добавленную стоимость составила 657525,9 руб.

Поскольку данный продукт не имеет готового решения и делается в зависимости от требований компании, то сравнить цену с аналогами на рынке не является возможным, но так как без правильной обработки данных для последующего анализа компания обречена на потерю прибыли, то разработка данного продукта является целесообразной.

Заключение

В результате выполнения дипломной работы была достигнута поставленная цель путем решения следующих задач:

— анализ подходов к проектированию Data warehouse;

— анализ программного обеспечения для проектирования Data warehouse;

— разработка проекта Data warehouse;

— разработка проекта обеспечения информационной безопасности Data warehouse;

— безопасность жизнедеятельности;

— технико-экономическое обоснование работы.

При разработке проекта Data warehouse была доработана концептуальная модель, благодаря чему хранилище стало более гибким для изменений и последующих доработок. Основываясь на модели нарушетелей и угроз были приняты меры по обеспечению информационной безопасности на физическом, техником и административных уровнях.

Список литературы

1 Хранилище данных: понятия - URL: https://aws.amazon.com/ru/data-warehouse/ (дата обращения: 12.09.20).

2 Антихрупкость архитектуры хранилищ данных - URL: https://habr.com/ru/post/281553/ (дата обращения: 12.09.20).

3 Не Hadoop'ом единым: что такое КХД и как его связать с Big Data - URL: https://www.bigdataschool.ru/blog/lsa-data-warehouse-architecture.html (дата обращения: 12.12.18).

4 ETL: что такое, зачем и для кого - URL: https://chernobrovov.ru/articles/etl-chto-takoe-zachem-i-dlya-kogo.html(дата обращения: 12.09.20)

5 Архитектура хранилищ данных: традиционная и облачная - URL: https://habr.com/ru/post/441538/ (дата обращения: 12.09.20)

6 Введение в Data Vault - URL: https://habr.com/ru/post/348188/ (дата обращения: 24.10.20)

7 Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных - URL: https://fstec.ru/component/attachments/download/289 (дата обращения: 14.11.20)

8 Лучшие практики по безопасности хранилищ данных - URL: https://www.securitylab.ru/analytics/502751.php (дата обращения: 14.11.20)

9 Что такое СУБД - URL: https://www.nic.ru/help/chto-takoe-subd_8580.html (дата обращения: 14.11.20)

10 Системы управления базами данных - URL: proglib.io/p/databases-2019?comment=5bb34165-a79c-4310-a21c-28e75c1d3f54 (дата обращения: 12.09.20)

11 Особенности ETL инструментов - URL: https://elibrary.ru/item.asp?id=37624966 (дата обращения: 14.11.19)

12 Informatica PowerCenter - URL: https://www.tadviser.ru/index.php/%D0%9F%D1%80%D0%BE%D0%B4%D1%83%D0%BA%D1%82:Informatica_PowerCenter (дата обращения: 14.11.20)

13 Магический квадрант гартнера. Технологии маскировки данных - URL: https://www.dataarmor.ru/quadrant/ (дата обращения: 12.12.20)

14 IBM InfoSphere DataStage - URL: https://www.tadviser.ru/index.php/%D0%9F%D1%80%D0%BE%D0%B4%D1%83%D0%BA%D1%82:IBM_InfoSphere_DataStage (дата обращения: 12.12.20)

15 Как выбрать систему контроля и управления доступом (СКУД)? - URL: https://securityrussia.com/blog/vibrat_skud.html (дата обращения: 12.12.20)

16 Dlp, siem, ngfw и другие средства защиты - URL: https://ideco.ru/company/news/kto-est-kto-v-mire-ib-dlp,-siem,-ngfw-i-drugie-sredstva-zashhityi (дата обращения: 12.12.20)

17 Как UEBA помогает повышать уровень кибербезопасности - URL: https://habr.com/ru/company/roi4cio/blog/436082/ (дата обращения: 12.12.20)

18 Системы защиты от утечек конфиденциальной информации (DLP) - URL: https://www.anti-malware.ru/security/data-leak-protection (дата обращения: 12.09.20)

19 Сравнительный обзор средств предотвращения утечек данных (DLP) URL: https://safe-surf.ru/specialists/article/5233/609990/ (дата обращения: 22.12.20)

20 Обнаружение и предотвращение атак и вторжений (IDS/IPS) - URL: https://tchk.net/usergate/ids-ips/ (дата обращения: 22.12.20)

21Примерение IDS/IPS - URL: https://xakep.ru/2012/10/29/ids-ips/ (дата обращения: 22.12.20)

22 Сетевые системы предотвращения вторжений (Network IPS) - URL: https://www.fgts.ru/page/network-ips (дата обращения: 22.12.20)

23 Сетевые системы обнаружения атак — принцип действия - URL:https://gardatech.ru/articles/smi/setevye-sistemy-obnaruzheniya-atak-printsip-deystviya/ (дата обращения: 22.12.20)

24 Общие понятия о системах обнаружения и предотвращения вторжений - URL: https://habr.com/ru/company/otus/blog/479584/ (дата обращения: 08.01.21)

25 Next Generation Firewall - URL: https://www.tadviser.ru/index.php/Next-Generation_Firewall (дата обращения: 10.01.21)

26 Сравнение универсальных шлюзов безопасности USG (NGFW) - URL: https://www.anti-malware.ru/compare/USG-NGFW (дата обращения: 12.01.21)

27 Лекция 6: Административный уровень информационной безопасности - URL: https://intuit.ru/studies/courses/10/10/lecture/306 (дата обращения: 12.01.21)

28 Политика безопасности - URL: https://npo-echelon.ru/doc/politics_development.php (дата обращения: 12.01.21)

29 Третья нормальная форма - URL: https://metanit.com/sql/tutorial/2.4.php (дата обращения: 12.01.21)

30 Требования к организации и оборудованию рабочих мест с ВДТ и ПЭВМ – URL: https://zakonbase.ru/content/part/130299?print=1 (дата обращения: 12.01.21)

31 Часто задаваемые вопросы по организации работы компьютерного класса – URL: https://mosmetod.ru/metodicheskoe-prostranstvo/srednyaya-i-starshaya-shkola/informatika-ikt/kabinet/chasto-zadavaemye-voprosy-po-organizatsii-raboty-kompyuternogo-klassa.html (дата обращения: 12.01.21)

32 Условия труда на предприятии – URL: https://studref.com/445985/ekonomika/usloviya_truda_predpriyatii (дата обращения: 12.01.21)

33 Факторы воздействия компьютера на человека – URL: https://laws.studio/knigi-deloproizvodstvo_860/faktoryi-vozdeystviya-kompyutera.html (дата обращения: 12.01.21)

34 Причины утилизации компьютерной техники – URL: https://greenologia.ru/utilizaciya-texniki/ofisnaya/kompyutery (дата обращения: 12.01.21)

Приложение А

Таблица А.1 – Сравнительный анализ DLP систем[19].

	Гарда Предприятие	Infowatch Traffic monitor Enterprise	Solar Dozor
Лицензиия ФСБ России	Нет	Да	Нет
Лицензия ФСТЭК России	Да	Да	Да
Сертификаты на продукт ФСТЭК России	Да	Да	Да
Исполнение подсистемы контроля (агенты+сетевая часть, только агенты)	Агенты рабочих мест и сетевая часть (анализатор)	Агенты (поставляется в виде отдельных компонентов - Device Monitor и Person Monitor) и сетевая часть	Агенты и сетевая часть
Работа в режиме мониторинга	Да	Да	Да
Работа в режиме блокировки	Да	Да	Да
Поддержка IPv6	Да	Да	Да
Максимальная пропускная способность подсистемы перехвата сетевого трафика	до 40 Гбит/с на каждый модуль, размеры кластера не ограничены	400 Мбит/с, по 200 Мбит/с на одно плечо кластера	10 Гбит/с в кластере
Интерфейс управления системой	веб-интерфейс, всегда единое "окно" управления	веб-интерфейс, консоли управления разные для разных компонентов, отдельно поставляется модуль визуализации Vision	веб-интерфейс, единое "окно" управления

Поддерживаемые устройства для управления	Любое устройство с установленным браузером	Любое устройство, рекомендуется установленный браузер Google Chrome	Любое устройство с установленным браузером
Адаптации интерфейса под мобильные устройства	Да	Да	Да
Возможность ролевого доступа к системе	Да	Да	Да
Хранение перехваченных данных	Гибридное хранилище собственной разработки	Oracle DB, PostgreSQL, MS SQL, MySQL	Oracle, PostgreSQL
Поддержка каталогов LDAP	Active Directory и иные LDAP-каталоги	Active Directory, Domino Directory, Novell eDirectory, Astra Linux Directory	Active Directory и любые другие LDAP каталоги
Интеграция с почтовыми серверами	Microsoft Exchange, IBM Lotus, любые другие SMTP, IMAP серверы	MS Exchange, MDaemon, IBM Lotus Domino и другими SMTP-, IMAP-серверами	Microsoft Exchange, IBM Lotus Notes, CommuniGate
Электронная почта	SMTP POP3 IMAP MAPI NNTP S/MIME Контроль веб-почты	SMTP POP3 IMAP MAPI S/MIME Контроль веб-почты	SMTP POP3 IMAP Контроль веб-почты
Системы мгновенного обмена сообщений	OSCAR (ICQ, QIP) MMP (любые клиенты, поддерживающие этот протокол, например, Mail.Ru Агент) MSN (Windows Live Messenger)	OSCAR (ICQ, QIP) MMP (любые клиенты, поддерживающие этот протокол, например, Mail.Ru Агент) XMPP (Google Talk, Jabber)	OSCAR - ICQ, QIP, MSN - Windows Live Messenger и прочие, XMPP - Google Talk, Jabber и прочие, IRC, Yahoo messenger, Skype, mail.ru агент и веб-почта

	XMPP (Google Talk, Jabber) YMSG (любые клиенты, поддерживающие этот протокол, например, Yahoo Messenger Protocol) HTTPIM (обмен сообщениями в социальных сетях) Microsoft Lync Skype Telegram Viber	YMSG (любые клиенты, поддерживающие этот протокол, например, Yahoo Messenger Protocol) HTTPIM (обмен сообщениями в социальных сетях) Microsoft Lync Skype Telegram Viber WhatsApp	
Контроль IP-телефонии	Да, контроль SIP, SDP, H.323, T.38, MGCP, SKINNY и др., включая видеотелефонию	Нет	Нет
Запись голоса в VoIP-телефонии	Да, каждый сеанс VoIP-телефонии может быть представлен в виде полного диалога или только отдельные каналы (как входящие, так и исходящие)	Нет	Нет
Контроль HTTP	Только на сети, либо по ICAP	На сети, на агенте	На сети
FTP и его модификации	Да	Да	Да
Возможность установки агента на ОС, отличные от семейства Windows	Нет	Astra Linux, функционал ограничен	Astra Linux, GosLinux, CentOS, функционал ограничен
Сканирование рабочих станций	Да	Да	Да
Снимки экрана	Да	Да	Да

Просмотр рабочего стола в режиме реального времени	Да, по задаваемым условиям	Нет	Нет
Запись фото/видео через веб-камеру	Нет	Да	Нет
Контроль и журналирование использования приложений	Да	Да	Да
Кейлоггер	Да	Да	Нет
Учет рабочего времени пользователя	Да	Да	Нет
Возможность запретить или разрешить использование устройства	Да	Да	Да
Контроль печати	Да	Да	Да
Контроль буфера обмена	Да, в момент операции "вставить"	Да, настраивается по приложению-источнику или приемнику	Да
Уведомление администратора безопасности	Да, по электронной почте	Да, по электронной почте	Да, по электронной почте
Блокировка соединения	Да	Да	Да
Блокировка передачи файлов	Да	Да	Да
Гибко настраиваемые интерактивные дашборды в интерфейсе	Да	Да	Нет
Досье и карточки сотрудников	Да, автозаполнение	Да	Да
Выявление стеганографических контейнеров	Нет	Нет	Нет

Приложение Б

Таблица Б.1 – сравнительный анализ DLP систем[26].

	Cisco	Huawei	Код Безопасности	Usergate
Полное название системы	Межсетевые экраны нового поколения Cisco Firepower	Межсетевые экраны нового поколения Huawei	"Континент" 4.0	Универсальный шлюз безопасности "UserGate"
Сравниваемая линейка продуктов (модели, версии ОС)	Серия Cisco Firepower (1010, 1100, 2100, 4100, 9300) на FTD OS 6.4	Huawei USG v5 (6320, 6330, 6350, 6360, 6370, 6380, 6390, 6620, 6630, 6650, 6660, 6670, 6680, 9560, 9580)	Континент 4.0.3 (IPC-10, IPC-25, IPC-50, IPC-100, IPC-500, IPC-500F, IPC-600, IPC-800F, IPC-1000F, IPC-3000F, IPC-3000FC, IPC-1000NF2, IPC-3000NF2)	UserGate на UGOS 5.0.6 (модели C, D, E, F, X)
Целевой сегмент	Малый, средний и крупный бизнес, государственный и коммерческий сектор	Малый, средний и крупный бизнес, государственный и коммерческий сектор	Малый, средний и крупный бизнес, государственный и коммерческий сектор	Малый, средний и крупный бизнес, государственный и коммерческий сектор
Сертификаты	Сертификат ФСТЭК России №3973 со сроком действия до 25.07.2021 на межсетевой экран серии Cisco ASA 5500-X (ASA 5506-X, ASA 5508-X, ASA 5516-X) с установленным программным обеспечением Cisco ASA	Сертификат ФСТЭК №4083, срок действия до 04.02.2024, на версию V500, профиль защиты МЭ по новым требованиям (А четвертого класса защиты ИТ.МЭ.А4.ПЗ, Б четвертого класса защиты ИТ.МЭ.Б4.ПЗ).	Планируется получение сертификатов: ФСТЭК по требованиям к межсетевым экранам тип "А" 4-го класса и СОВ уровня сети 4-го класса, ФСБ на СКЗИ класса КС3	Сертификат ФСТЭК №3905, срок действия до 26 марта 2021 года на “Универсальный шлюз безопасности “UserGate”. Сертификация была пройдена по требованиям к Межсетевым Экранам (4-й класс, профили А и Б) и по требованиям к Системам Обнаружения Вторжений (4-й класс) для

	версии 9.x , профиль защиты МЭ по новым требованиям (А шестого класса защиты. ИТ.МЭ.А6.ПЗ, Б шестого класса защиты. ИТ.МЭ.Б6.ПЗ). В процессе сертификации - Firepower 2100, Firepower 4100, профиль защиты МЭ (А шестого класса защиты ИТ.МЭ.А6.ПЗ и Б шестого класса защиты ИТ.МЭ.Б6.ПЗ)			программно-аппаратных (модели UserGate C, D, D+, E, E+, F, X1) и виртуальных платформ UserGate.
Локальное производство в РФ	Нет	Нет	Устройства проходят выходной контроль в Москве. Ожидается получение статуса "Телекоммуникационного оборудования российского происхождения" (ТОРП) от Минпромторга.	Сборка в Новосибирске на основе аппаратных платформ Lanner
Соответствие требованиям ФЗ-187 "О безопасности критической информационной инфраструктуры РФ"	Удовлетворяет требованиям приказа ФСТЭК №235 в части используемых средств защиты объектов КИИ и выполняет часть мер,	Сертифицированный МЭ, обеспечивающий безопасность информационных систем, информационно-телекоммуникационных сетей КИИ. Удовлетворяет требованиям	Удовлетворяет требованиям приказа ФСТЭК №235 в части используемых средств защиты объектов КИИ, выполняет часть мер, перечисленных в приказе ФСТЭК №239	Сертифицированное средство, обеспечивающее безопасность информационных систем, информационно-телекоммуникационных сетей, автоматизированных систем управления (АСУ) субъектов

	перечисленных в приказе ФСТЭК №239	приказа ФСТЭК №235 в части используемых средств защиты объектов КИИ и выполняет часть мер, перечисленных в приказе ФСТЭК №239		КИИ, сбор и хранение информации о произошедших на объекте инцидентах безопасности для последующей передачи этих данных в ГосСОПКА
Поддерживаемые варианты исполнения	Аппаратное, виртуальное	Аппаратное, виртуальное, контейнерное	Аппаратное	Аппаратное, виртуальное
Статическая трансляция сетевых адресов SNAT (Static Network Address Translation)	Да	Да	Да	Да
Динамическая трансляция сетевых адресов DNAT (Dynamic Network Address Translation)	Да	Да	Да	Да
Варианты поддержки трансляции сетевых адресов	Static NAT, dynamic NAT, PAT, NAT64, PAT64, policy-based NAT/PAT, Carrier Grade NAT, dual NAT и т.п.	Source IP address based NAT, Destination IP address based NAT, NAT No-PAT, NAPT, Easy IP, Smart NAT, Bidirectional NAT, NAT ALG, NAT444, DS-Lite, NAT64, NAT66(2019.7), IPv4 over IPv6, IPv6 over IPv4	Source NAT, Destination DNAT, Hide NAT, Dynamic NAT	Трансляция портов NAT с PAT, Persistent NAT, NAT64, двойной NAT, PBR
Многоадресная передача (Multicast)	Да, IGMP v2 и v3, PIM-SM, PIM	Да, IGMP v2 и v3, PIM-SM, PIM-SSM, PIM-DM	Нет	Нет

	Boostrap Router, Stub Multicast Routing			
Качество обслуживания (QoS)	Да	Да	Да	Да
Поддержка глубокого пакетного анализа (Deep packet inspection, DPI)	Да	Да	Да	Да
Поддержка расшифрования входящего и исходящего трафика по протоколу SSL/TLS	Да, HTTPS Inspection	Да, MITM	Да, MITM	Да, MITM. Применение согласно политикам дешифрования
Поддержка формирования исключений из инспекции трафика SSL/TLS	Да, по IP, URL, категориям и т.п.	Да, по категориям, по IPv4 и IPv6-адресам, по FQDN, по URL	Да, по IP	Да, по категориям, по URL, по IP, по пользователям и группам
Возможность расшифрования протокола SSH	Нет	Да	Нет	Нет
Поддержка поведенческого анализа	Да, в модуле AVC	Да, в модулях IPS, AV, APT	Да, модуль поведенческого анализа	Да, сценарии для автоматизированной реакции системы на нестандартное поведение
Встроенная корреляция событий	Да, Firepower Management Center	Да, Huawei Monitoring System	Нет	Да, через сценарии реагирования

Система обнаружения / предотвращения вторжений (IDS/IPS)	Да, модуль IPS/IDS	Да, модуль IPS/IDS	Да, модуль IPS с собственными сигнатурами	Да, модуль IPS с собственными сигнатурами
Поддержка формирования исключений для сигнатур	Да	Да	Да, реализован список правил, позволяющий использовать / не использовать IPS в каждом конкретном правиле фильтрации. Для каждого шлюза безопасности может устанавливаться индивидуальный профиль сигнатур	Да, через список правил, позволяющий указывать тип трафика для проверки и применяемый к нему профиль IPS/IDPS
Поддержка возможности индивидуализации сигнатур	Да	Да	Да	Нет
Поддержка возможности импорта сторонних сигнатур	Да, в формате SNORT	Да, в формате SNORT	Да, в формате Suricata	Да, через службу технической поддержки
Поддержка автоматического сбора дампа трафика при срабатывании сигнатуры для последующего анализа	Да	Да	Да	Нет
Поддержка автоматического применения новых сигнатур после обновления	Да	Да	Да	Да
Поддержка уведомлений, отправка отчетов	Да	Да	Да	Да

Противодействие известным методам обхода сигнатурного анализа	Да, препроцессоры обработки транспортных и прикладных протоколов, нормализация, декодирование и т.п.	Да, IDS Avoidance	Да, IP Packet Fragmentation, Tunnel decoding, Stream Segmentation, URL Obfuscation, HTML Obfuscation, Protocol-level Misinterpretation	Да, нормализация HTTP / HTTPS трафика. Возможность блокировки фрагментированных пакетов (включена по умолчанию), защита от VPN over DNS, блокирование туннелей Teredo, IP6-IP4, IP4-IP6
Возможность блокирования нераспознанных приложений	Да	Да	Да	Да
Защита от DDoS-атак	Да, возможность установки модуля vDP Radware для Firepower 4100/9300	Да	Да, обнаружение small packet MTU, DNS mismatch, DNS reply mismatch, SYN flood, SMURF, FIN/RST flood, FRAGGLE attack, LAND-attack	Обнаружение и защита от SYN-, UDP-, ICMP-флуда, защита приложений от превышения сессий
Антивирусная защита (Anti-virus)	Да, Cisco AMP	Да, Huawei AV	Защита от вирусов осуществляется путем запрета доступа к URL-адресам с низкой репутацией (Malicious URL block)	Да, собственный и Kaspersky
Возможность анализа содержимого архивных файлов	Да, ZIP, RAR, TAR и другие, используя автоматизированный анализ или технологию Glovebox	ZIP, RAR, 7Z, JAR, ACE, CAB и др.	Нет	Поддерживаются все основные форматы, более 6 тыс. версий

Поддерживаемые методы детектирования и блокировки бот-зараженных машин	Обнаружение ботов по аналитике Talos, профилю трафика, корреляции событий. Блокирование соединений, в т.ч. на других МСЭ и маршрутизаторах, микросегментация с использованием Cisco ISE и т.п.	По репутации C&C, по сигнатурам в трафике, DNS Trap	По репутации C&C	По репутации C&C
Блокировка взаимодействия бот-сети с командными серверами (C&C)	Да	Да	Да	Да
Анализ и подмена вредоносных DNS-запросов к системам управления бот-сетями (фильтрация DNS-запросов)	Да	Да	Нет	Да
Защита почтового трафика (безопасность почты, антиспам)	Нет	Да, модуль Mail Filtering	Нет	Да, модуль защиты почтового трафика
Фильтрация и анализ почтового трафика SMTP, POP3, IMAP	Нет	Да	Нет	Да
Проверка ссылок в теле письма	Нет	Да	Нет	Да

Поддержка проверки вложенных файлов и архивов	Нет	Да	Нет	Да
Поддержка блокировки скачивания по типам файлов	Нет	Да	Нет	Да
Веб-фильтрация	Да, HTTP/HTTPS-фильтрация	Да, модуль URL Filtering	Только запрет доступа к опасным ресурсам на основе данных Kaspersky Feed	Да, модуль контентной фильтрации