

Московский государственный технический университет им. Н.Э.

Баумана

Кафедра «Автоматизированные системы обработки информации и управления»



Курсовой проект

по дисциплине «Параллельные Вычисления (Супер ЭВМ)»

***Автоматизированная информационная система «Учета
пассажиропотоков Московского метро»***

Расчетно-пояснительная записка

(вид документа)

писчая бумага 27

(вид носителя) (количество листов)

ИСПОЛНИТЕЛЬ:

Авдеев Ю. _____

Группа ИУ5-64 " __ " _____ 2018 г.

ПРЕПОДАВАТЕЛЬ:

Калистратов А.П _____

" __ " _____ 2018 г.

Оглавление

Задание	3
1. Описание предметной области	4
2. Архитектура сети	5
2.1. Задание	5
2.2. Теоретический материал	5
2.3. Реализация	6
3. Выбор оборудования	7
3.1. Задание	7
3.2. Параллельные вычисления - теоретический материал	7
3.3. Практическое выполнение	9
3.4. Реализация функции посредством параллельных вычислений	10
4. Выбор операционной системы	12
5. Модель бизнес процесса	13
5.1. Словесное описание	13
5.2. Теоретический материал	13
5.3. Модель процесса в нотации BPMN	14
6. Создание БД и визуализация данных	16
6.1. Задание	16
6.2. Теоретический материал	16
7. Выбор режимов архивации и восстановления	12
7.1. Определение объема базы данных	21
7.2. Интенсивность наполнения и работы с БД	21
7.3. Способы архивации данных	21
7.4. Выбор режима архивации	25
8. Регламент выполнения задач КП	31
9. Требования к предоставляемой информации КП	31
10. Заключение	32
11. Список литературы	32

Задание

Для автоматизированной информационной системы необходимо создать проект сети:

- по архитектуре клиент-сервер;
- по структуре 5-7 офисов, один из которых главный.

Выбрать оборудование и операционные системы для офисов из линейки Z/OS. Для центрального офиса выбрать супер ЭВМ типа мэйнфрейм и советующую операционную систему.

Выбор произвести, используя методику анализа прототипов и аналогов (не менее 7-10-х аналогов по 5-7 критериям). Задаться количественными характеристиками запросов для АИС.

Разработать модель бизнес процесса АИС в нотации BPMN в виде диаграмм. Диаграмма моделирования бизнес процесса в нотации BPMN должна содержать следующие элементы:

- объекты потока управления;
- соединяющие объекты;
- роли (не менее 2 пулов или 1 пул и 2 дорожки).

Привести словесное описание бизнес процесса.

Выбрать в качестве СУБД DB2 для использования с «1С: Предприятие 8».

Обосновать выбор режимов архивации и восстановления информационной БД.

Создание БД и визуализация данных

Построение даталогической модели.

1. Описание предметной области

АИС «учета пассажиропотоков Московского метро» предназначена для того, чтобы обеспечить сотрудникам возможность оперативно вести статистику, документацию, реорганизовывать работу подземного транспорта, а так же в виде базы данных хранить необходимую информацию.

Визуализация данных в АИС «учета пассажиропотоков Московского метро» позволяет получить в первую очередь наглядное представление о загруженности той или ной станции или ветки метрополитена. Возможно посмотреть количество входящих или выходящих пассаижиров, загруженность вагонов в то или иное время дня. Помимо всего прочего доступен список наиболее популярных, загруженных станций.

Работа системы основывается на сервере(или нескольких серверах) и его базе данных, в которую заносится информация обо всех входах или выходах на той или иной станции.

2. Архитектура сети

2.1. Задание

Для автоматизированной информационной системы необходимо спроектировать сеть:

- по архитектуре клиент-сервер;
- по структуре из 5-7 офисов, один из которых – главный.

2.2. Теоретический материал

Клиент-сервер (англ. client–server) – вычислительная или сетевая архитектура, в которой задания или сетевая нагрузка распределены между поставщиками услуг, называемыми серверами, и заказчиками услуг, называемыми клиентами. Фактически клиент и сервер – это программное обеспечение. Обычно эти программы расположены на разных вычислительных машинах и взаимодействуют между собой через вычислительную сеть посредством сетевых протоколов.

VPN (англ. *Virtual Private Network* – виртуальная частная сеть) – обобщённое название технологий, позволяющих обеспечить одно или несколько сетевых соединений (логическую сеть) поверх другой сети (например, Интернет).

По степени защищённости используемой среды:

- 1) **Защищённые.** Можно создать надёжную и защищённую сеть на основе ненадёжной сети, как правило, Интернета.
- 2) **Доверительные.** Используются в случаях, когда передающую среду можно считать надёжной и необходимо решить лишь задачу создания виртуальной подсети в рамках большей сети.

По способу реализации:

- 1) **Специальное программно-аппаратное обеспечение.** Реализация VPN сети осуществляется при помощи специального комплекса программно-аппаратных средств. Такая реализация обеспечивает высокую производительность и, как правило, высокую степень защищённости.
- 2) **В виде программного решения.** Используют персональный компьютер со специальным программным обеспечением, обеспечивающим функциональность VPN.

По назначению:

1) **Intranet VPN.** Используют для объединения в единую защищённую сеть нескольких распределённых филиалов одной организации.

2) **Remote Access VPN.** Используют для создания защищённого канала между сегментом корпоративной сети и одиночным пользователем.

3) **Extranet VPN.** Используют для сетей, к которым подключаются внешние пользователи. Низкий уровень доверия.

4) **Internet VPN.** Используется для предоставления доступа к интернету провайдерами, обычно если по одному физическому каналу подключаются несколько пользователей.

5) **Client/Server VPN.** Обеспечивает защиту передаваемых данных между двумя узлами корпоративной сети.

2.3 Реализация

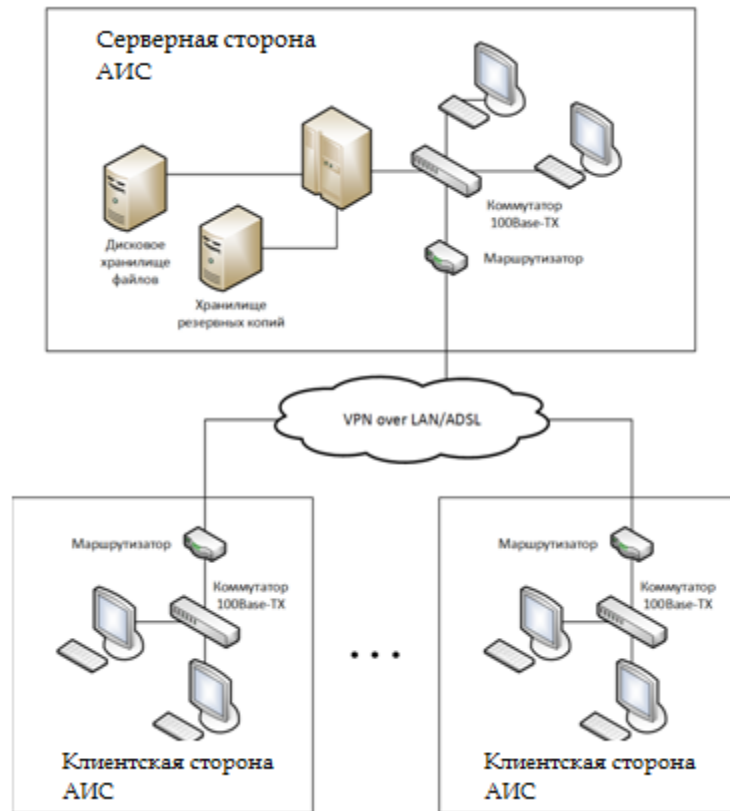


Рисунок 1. Схема сети АИС «учета пассажиропотоков Московского метро»

Критерии при проектировании сети

- Низкая стоимость
- Относительно высокая надежность
- Простота прокладки, установки и настройки
- Поддержка любым современным оборудованием
- Расстояние между узлами до 100 м

3. Архитектура сети

3.1 Задание

Выбрать оборудование и операционные системы для офисов из линейки Z/OS. Для центрального офиса выбрать супер ЭВМ типа мэйнфрейм и соответствующую операционную систему.

Выбор произвести, используя методику анализа прототипов и аналогов (не менее 3-4 аналогов и 5-7 критериев). Задаться количественными характеристиками запросов для АИС.

3.2 Параллельные вычисления – теоретический материал

3.2.1. Параллельные вычислительные системы (ПВС)

Параллельные вычислительные системы (ПВС) – физические компьютерные, а также программные системы, реализующие тем или иным способом параллельную обработку данных на многих вычислительных узлах.

3.2.2. Последовательные вычисления

Последовательные вычисления – выполнение линейно упорядоченной последовательности операций над данными, на преобразование входных данных на выходные.

3.2.3. Параллельные вычисления

Параллельные вычисления – способ организации компьютерных вычислений, при котором программы разрабатываются как набор взаимодействующих вычислительных процессов, работающих параллельно, то есть одновременно.

3.2.4. Эффективность параллельных вычислений

Определение эффективности параллельных вычислений возможно двумя методами. По формуле:

$$\delta = \frac{T_{\text{пос}}}{n * T_{\text{пар}}} * 100\%,$$

Ускорением, получаемым при использовании параллельного алгоритма для p процессоров, по сравнению с последовательным вариантом выполнения вычислений. Эффективность использования параллельным алгоритмом процессоров при решении задачи определяется соотношением:

$$Ep(n) = Tl(n) / (p * Tp(n)) = Sp(n) / p, \text{ где } n - \text{размерность задачи.}$$

3.2.5. Общие принципы построения параллельных алгоритмов

В общем плане под параллельными вычислениями понимаются процессы обработки данных, в которых одновременно могут выполняться нескольких машинных операций. Достижение параллелизма возможно только

при выполнимости следующих требований к архитектурным принципам построения вычислительной системы:

1) Независимость функционирования отдельных устройств ЭВМ – данное требование относится в равной степени ко всем основным компонентам вычислительной системы – к устройствам ввода-вывода, к обрабатывающим процессорам и к устройствам памяти;

2) Избыточность элементов вычислительной системы – организация избыточности может осуществляться в следующих основных формах:

– использование специализированных устройств таких, например, как отдельных процессоров для целочисленной и вещественной арифметики, устройств многоуровневой памяти (регистры, кэш);

– дублирование устройств ЭВМ путем использования, например, нескольких однотипных обрабатывающих процессоров или нескольких устройств оперативной памяти.

Дополнительной формой обеспечения параллелизма может служить конвейерная реализация обрабатывающих устройств, при которой выполнение операций в устройствах представляется в виде исполнения последовательности составляющих операцию подкоманд; как результат, при вычислениях на таких устройствах могут находиться на разных стадиях обработки одновременно несколько различных элементов данных.

3.2.6. Функция для реализации в виде параллельных вычислений

Для определения более объективного рейтинга фотографий применяется специальная статистическая техника, известная как байесовская оценка (по имени автора Томаса Байеса). Она призвана взять в расчёт не только среднее арифметическое оценок проголосовавших (средний балл), но и их количество.

$$f = \frac{N}{N + N_{min}} * AvgR + \frac{N_{min}}{N + N_{min}} * 7,2453$$

В этой формуле N – это количество оценок, N_{min} – это необходимый минимум голосов, отданных за данную фотографию, $AvgR$ – среднее арифметическое значение оценок, поставленных фотографии, а константа 7,2453 – некая усреднённая величина, принятая за основу метода. Суть заключается в том, что при небольшом количестве голосов расчётный балл будет близок к 7,2453. По мере увеличения отданных голосов роль среднего балла (среднего арифметического оценок) будет возрастать. Интересной особенностью является то, что при одинаковой средней арифметической голосов больший расчётный балл имеют фотографии с меньшим количеством голосовавших, и это искажение будет исправляться с поступлением новых голосов.

3.3 Практическая реализация

Выбор оборудования осуществим из линейки zSeries методом взвешенной суммы.

Таблица 1

Критерий	H20	H43	H44	H89	HA1	M15	M32	M49	M66	M80
Макс. Число ЦП	20	43	66	89	101	15	32	49	66	80
Макс. Объем ОП (Гб)	704	1392	2272	40	3040	752	1520	2288	3056	3056
Макс. число zPP	10	21	33	44	50	7	16	24	33	40
Кол. Портов Infiniband 1x	64	64	64	64	64	48	48	48	48	48
Версия CryptoExpress	4S	4S	4S	4S	4S	3	3	3	3	3
Цена (тыс \$)	750	1300	1800	2200	2600	500	800	1200	1500	2000

Таблица 1. Исходные значения сравниваемых мейнфреймов

Шаг 1. Числовые характеристики серверов, приведенные в таблице 1, вычисляются посредством деления значения характеристик на коэффициент $K=1.486$ и округляется до целого. Результаты записываем в таблицу 2.

Критерий	H20	H43	H66	H89	HA1	M15	M32	M49	M66	M80	MAX
Макс. Число ЦП	13	29	44	60	68	10	22	33	44	54	68
Макс. Объем ОП (Гб)	474	937	1529	27	2046	506	1023	1540	2057	2057	2057
Макс. число zPP	7	14	22	30	34	5	11	16	22	27	34
Кол. Портов Infiniband 1x	43	43	43	43	43	32	32	32	32	32	43
Версия CryptoExpress	5	5	5	5	5	2	2	2	2	2	5
Цена (тыс \$)	505	875	1211	1480	1750	336	538	808	1009	1346	1750

Таблица 2. Вычисленные значения сравниваемых мейнфреймов

Шаг 2. Проводится нормирование параметров мейнфреймов путем деления значения атрибутов R_{ij} кортежа на максимальное значение R_i .

Критерий	H20	H43	H66	H89	HA1	M15	M32	M49	M66	M80
Макс. Число ЦП	0,191	0,4264	0,64705	0,882	1	0,147	0,3235	0,4852	0,64705	0,79411
Макс. Объем ОП (Гб)	0,230	0,4555	0,7433	0,013	0,994	0,245	0,4973	0,7486	1	1
Макс. число zPP	0,205	0,4117	0,64705	0,882	1	0,147	0,3235	0,4705	0,64705	0,79411
Кол. Портов Infiniband 1x	1	1	1	1	1	0,744	0,7441	0,7441	0,74418	0,74418
Версия CryptoExpress	1	1	1	1	1	0,4	0,4	0,4	0,4	0,4
Цена (тыс \$)	0,288	0,5	0,692	0,845	1	0,192	0,3074	0,4617	0,57657	0,76914

Таблица 3. Нормированные значения сравниваемых мейнфреймов

Шаг 3. Субъективно задается коэффициент α_i и вычисляются весовые коэффициенты Y_i . Проводится вычисление нормированных значений $R_{ij} * Y_i$. параметров серверов и их итоговые значения Y_j .

Критерий	а		H20	H43	H66	H89	HA1	M15	M32	M49	M66	M80
Макс. Число ЦП	6	0,27	0,05	0,12	0,17	0,24	0,27	0,04	0,09	0,13	0,17	0,21
Макс. Объем ОП (Гб)	5	0,23	0,05	0,1	0,17	0	0,23	0,06	0,11	0,17	0,23	0,23
Макс. число zPP	2	0,09	0,02	0,04	0,06	0,08	0,09	0,01	0,03	0,04	0,06	0,07
Кол. Портов Infiniband 1x	1	0,05	0,05	0,05	0,05	0,05	0,05	0,04	0,04	0,04	0,04	0,04
Цена (тыс \$)	2	0,09	0,09	0,09	0,09	0,09	0,09	0,04	0,04	0,04	0,04	0,04
Версия CryptoExpress	2	0,09	0,03	0,05	0,06	0,08	0,09	0,02	0,03	0,04	0,05	0,07
Цена (тыс \$)	4	0,18	0,05	0,025	0,017	0,014	0,0146	0,0028	0,00086	0,00039	0,0002	0,0001
<i>Результат</i>	22	1	0,34	0,475	0,617	0,554	0,8346	0,2128	0,34086	0,46039	0,5902	0,6601

Таблица 4. Взвешенные нормированные значения для сравнения мейнфремов

Итоговый критерий эффективности рассчитывается по формуле $Y_k = \max Y_j$. Из этого следует, что $Y_k = 0,83$, то есть, по итогам сравнения выбранной супер ЭВМ типа мейнфрейм стала модель **НА1** из линейки **ЕС12**.

3.4 Реализация функции посредством параллельных вычислений

3.4.1. Эффективность реализации функции посредством параллельных вычислений

Для вычисления эффективности реализации функции используется формула:

$$\delta = \frac{T_{\text{пос}}}{n * T_{\text{пар}}} * 100\%,$$

где $T_{\text{пос}}$ – время, затраченное на последовательное вычисление;

$T_{\text{пар}}$ – время, затраченное на параллельное вычисление;

n – количество вычислительных потоков.

Эффективность распараллеливания показывает, какой процент времени вычислительные потоки тратят на эффективную работу, на вычисления.

Проведен тест с использованием максимального количества параллельных вычислительных потоков. Для процессора Intel(R) Core(TM) i7-6700HQ CPU @ 2.60 GHz с 3 физическими ядрами и 6 логическими процессорами была вычислена оценка эффективности параллельных вычислений:

$$T_{\text{пос}} = 5.300600 \text{ с};$$

$$T_{\text{пар}} = 0.109058 \text{ с};$$

$$n = 8$$

$$\delta = \frac{5.300600}{6 * 0.109058} * 100 \% \approx 607,5 \%$$

Следовательно, использование параллельного вычисления позволяет получить шестикратный прирост скорости вычисления по сравнению с последовательным вычислением.

4. Выбор операционной системы

Выбранная модель мейнфрейма поддерживает следующие операционные системы:

z/OS – проприетарная 64-битная серверная операционная система, разработанная компанией IBM для мейнфреймов собственного производства. Является дальнейшим развитием операционной системы OS/390, объединяя MVS и системные службы Unix (POSIX-совместимая реализация для Unix, изначально известная как MVS OpenEdition или OpenMVS).

OpenSolaris for System z – дистрибутив на основе операционной системы OpenSolaris, предназначенный для работы на мейнфреймах IBM System z. OpenSolaris for System z доступен для бесплатного скачивания и распространяется на тех же лицензионных условиях, что и OpenSolaris для других платформ. Весь исходный код открыт; нет модулей, существующих только в бинарном виде.

Linux on IBM System z – UNIX-подобный дистрибутив, который наследует в себе все плюсы и минусы ОС данного семейства. Распространяется под лицензией GPL, которая обеспечивает открытость исходных кодов. К преимуществам данной ОС следует отнести высокое быстродействие, защищенность, бесплатность и нетребовательность к ресурсам. Однако обратной стороной является сложность настройки данной ОС.

Исключим z/OS, как устаревшую проприетарную 64-битную серверную операционную систему. Системы OpenSolaris for System z и Linux on IBM System z находятся в бесплатном доступе, исходный код, которых открыт. Но так как **Linux on IBM System z** имеет высокое быстродействие, защищенность и нетребовательность к ресурсам, она будет выбрана в качестве ОС к выбранной модели мейнфрейма, как более распространенная и поддерживаемая широким сообществом

5. Модель бизнес процесса

5.1 Словесное описание

Пользователями АИС «учета пассажиропотоков Московского метро» являются сотрудники метрополитена, ведущие статистический учет загруженности подземки с целью реорганизации внутренней работы для разгрузки станций или целых веток в зависимости от времени суток.

Все пользователи имеют доступ к данным обо всех станциях и ветках, а так же поездах. Данные о входе/выходе пассажира фиксируются автоматически при помощи турникетов.

5.2 Теоретический материал

BPMN (англ. Business Process Model and Notation, нотация и модель бизнес-процессов) – система условных обозначений (нотация) для моделирования бизнес-процессов. Разработана Business Process Management Initiative (BPMI) и поддерживается Object Management Group, после слияния организаций в 2005 году. Предыдущая версия BPMN – 1.2; последняя – 2.0.

Спецификация BPMN описывает условные обозначения для отображения бизнес-процессов в виде диаграмм бизнес-процессов. BPMN ориентирована как на технических специалистов, так и на бизнес-пользователей. Для этого язык использует базовый набор интуитивно понятных элементов, которые позволяют определять сложные семантические конструкции. Кроме того, спецификация BPMN определяет, как диаграммы, описывающие бизнес-процесс, могут быть трансформированы в исполняемые модели на языке BPEL. Спецификация BPMN 2.0 также является исполняемой и переносимой (то есть процесс, нарисованный в одном редакторе от одного производителя может быть исполнен на движке бизнес-процессов совершенно другого производителя, при условии, если они поддерживают BPMN 2.0).

Основная цель BPMN – создание стандартного набора условных обозначений, понятных всем бизнес-пользователям. Бизнес-пользователи включают в себя бизнес-аналитиков, создающих и улучшающих процессы, технических разработчиков, ответственных за реализацию процессов и менеджеров, следящих за процессами и управляющих ими. Следовательно, BPMN призвана служить связующим звеном между фазой дизайна бизнес-процесса и фазой его реализации.

В настоящий момент существует несколько конкурирующих стандартов для моделирования бизнес-процессов. Распространение BPMN поможет унифицировать способы представления базовых концепций бизнес-процессов (например, открытые и частные бизнес-процессы, хореографии), а также более сложные концепции (например, обработка исключительных ситуаций, компенсация транзакций).

BPMN поддерживает лишь набор концепций, необходимых для моделирования бизнес процессов. Моделирование иных аспектов, помимо бизнес-процессов, находится вне зоны внимания BPMN. Например, моделирование следующих аспектов не описывается в BPMN:

- Модель данных
- Организационная структура

Несмотря на то, что BPMN позволяет моделировать потоки данных и потоки сообщений, а также ассоциировать данные с действиями, она не является схемой информационных потоков.

5.2.1. Элементы BPMN

Моделирование в BPMN осуществляется посредством диаграмм с небольшим числом графических элементов. Это помогает пользователям быстро понимать логику процесса. Выделяют четыре основные категории элементов:

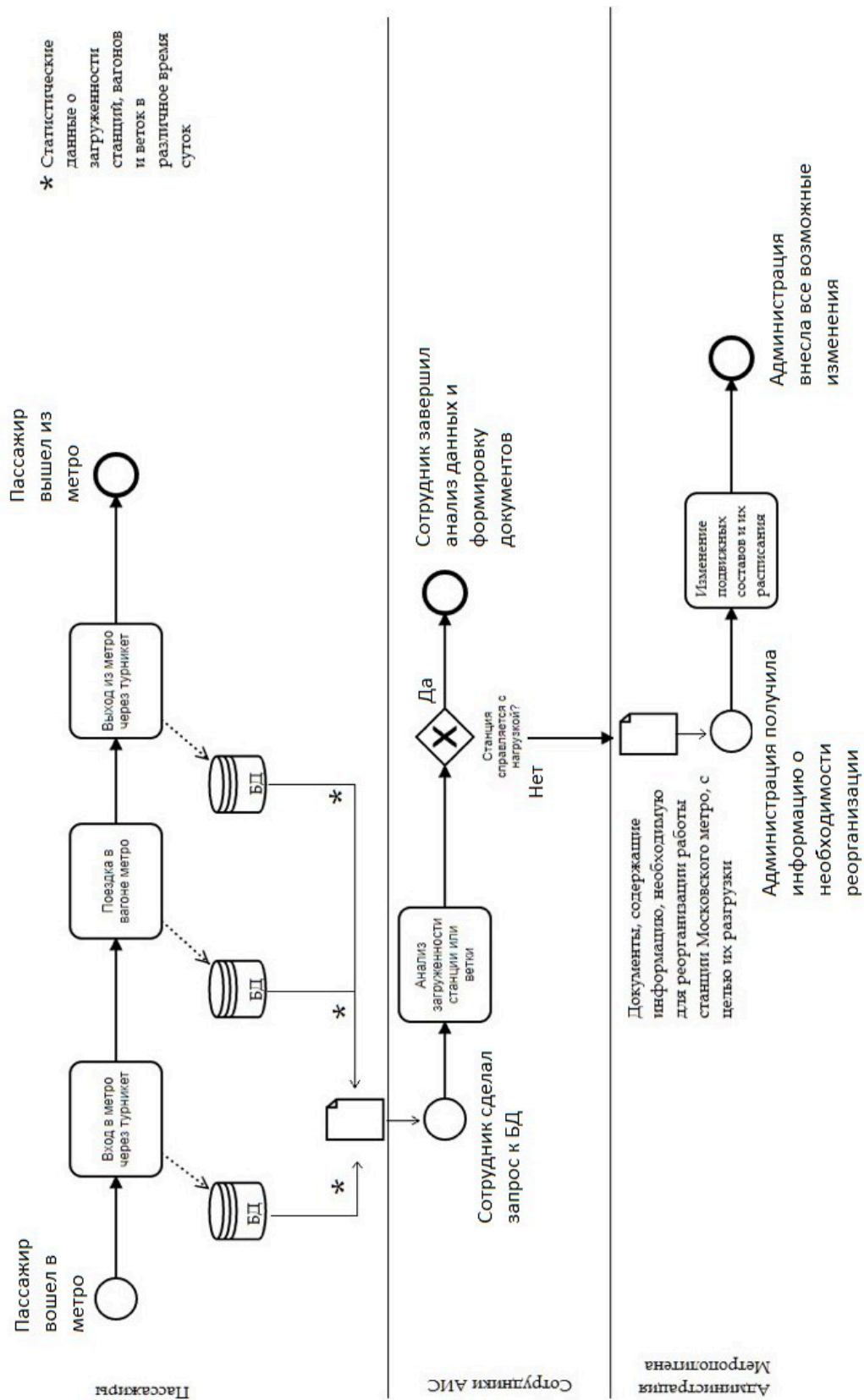
- Объекты потока управления: события, действия и логические операторы
- Соединяющие объекты: поток управления, поток сообщений и ассоциации
- Роли: пулы и дорожки
- Arteфакты: данные, группы и текстовые аннотации.

Элементы этих четырёх категорий позволяют строить простейшие диаграммы бизнес-процессов. Для повышения выразительности модели спецификация разрешает создавать новые типы объектов потока управления и артефактов.

5.3 Модель процесса в нотации BPMN

Диаграмма бизнес-процесса в нотации BPMN приведена в приложении.

Рисунок 3. Схема работы АИС «учета пассажиропотоков Московского метро» в нотации BPMN



6. Создание БД и визуализация данных

БД создается средствами PostgreSQL, что увеличивает скорость разработки базы, а так же скорость ее работы.

Количество записей – кортежей в БД не менее 200, для накопления статистики и построения диаграмм. Для заполнения базы данных тестовыми значениями используется скрипт, написанный на языке Java.

6.1 Построение даталогической модели

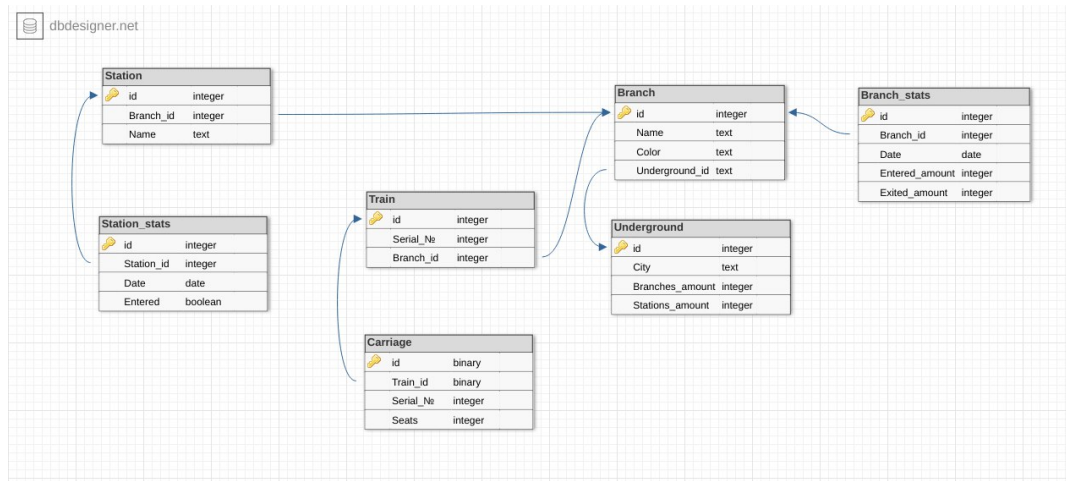


Рисунок 4. Даталогическая модель АИС «учета пассажиропотоков Московского метро»

6.2 Построение модели данных в PostgreSQL

```
CREATE TABLE
Underground(
id SERIAL PRIMARY KEY,
city TEXT,
branches_amount INTEGER,
stations_amount INTEGER
);
```

```
CREATE TABLE Branch(
id SERIAL PRIMARY KEY ,
name TEXT,
color TEXT,
underground_id INTEGER REFERENCES
Underground(id)
);
```

```
CREATE TABLE Branch_stats(
```

```

    id SERIAL PRIMARY KEY ,
    branch_id INTEGER REFERENCES Branch(id),
    date TIMESTAMP ,
    entered_amount INTEGER,
    exited_amount INTEGER
);

CREATE TABLE Station(
    id SERIAL PRIMARY KEY ,
    name TEXT,
    branch_id INTEGER REFERENCES Branch(id)
);

CREATE TABLE Station_stats(
    id SERIAL PRIMARY KEY ,
    date TIMESTAMP ,
    station_id INTEGER REFERENCES Station(id),
    entered BOOLEAN
);

CREATE TABLE Train(
    id SERIAL PRIMARY KEY ,
    serialNo INTEGER,
    branch_id INTEGER REFERENCES Branch(id)
);

CREATE TABLE Carriage(
    id SERIAL PRIMARY KEY ,
    serialNo INTEGER,
    seats INTEGER,
    train_id INTEGER REFERENCES Train(id)
);

```

6.2.3. Заполнение таблиц

Для заполнения таблиц был написан скрипт на языке Java с использованием SpringBoot и JDBC для подключения к базе. Тестовые данные генерируются случайным образом. Текст скрипта приведен в приложении

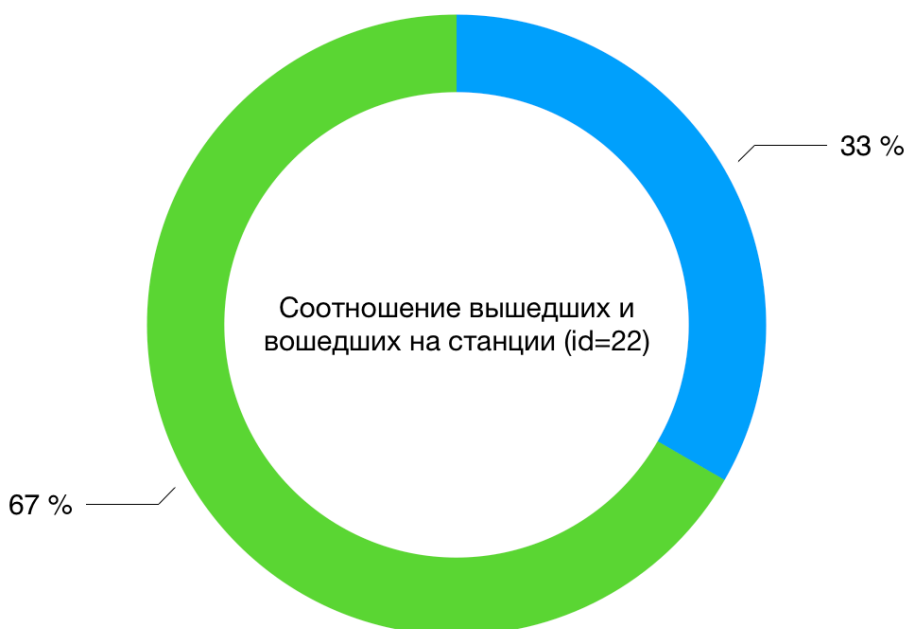
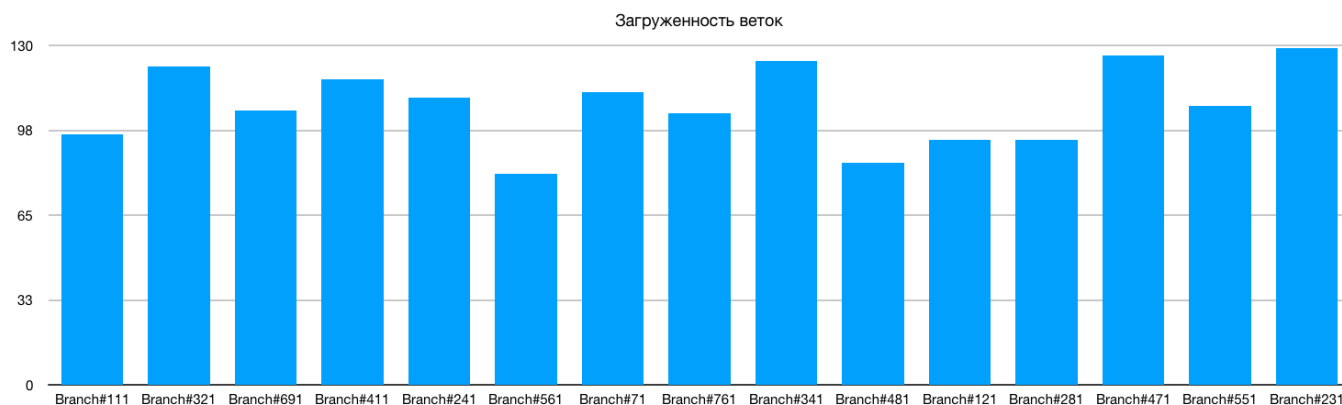
6.2.4. Загрузка из СУБД

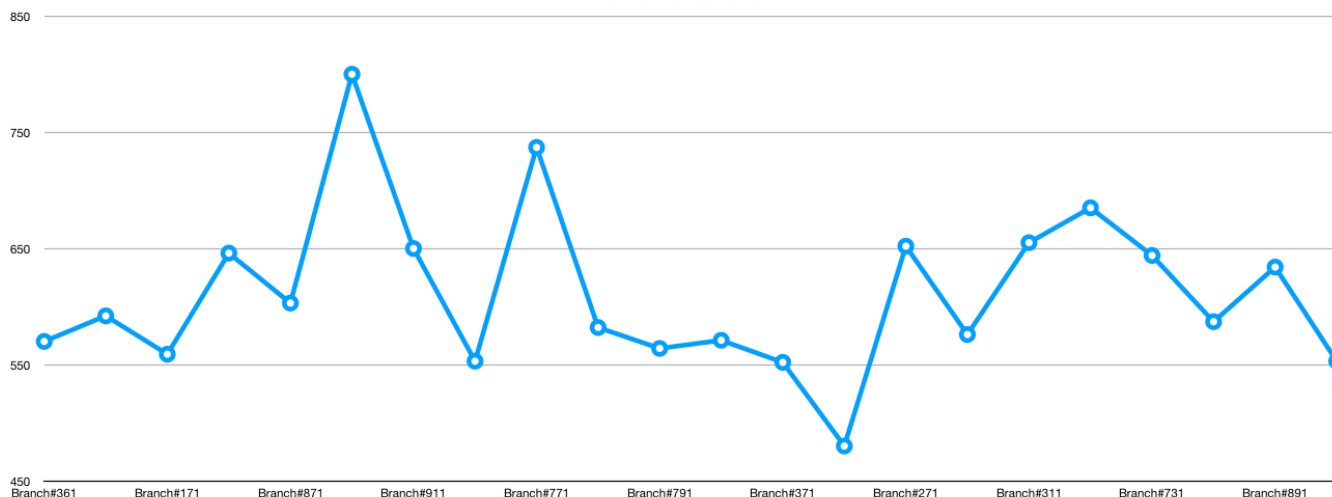
Выгрузка данных из СУБД для визуализации осуществляется стандартными нативными SQL командами, формирующими csv файл, читаемый Excel.

```
COPY (SELECT * FROM your_table) TO 'C:/temp/123.csv' CSV;
```

6.2.5. Визуализация данных

Визуализация данных реализована с помощью следующих диаграмм и графиков:





7. Выбор режимов архивации и восстановления

1. Задание

Выбрать в качестве СУБД DB2 для использования с «1С:Предприятие 8». Обосновать выбор режимов архивации и восстановления информационной БД.

2. Теоретический материал

DB2 – семейство систем управления реляционными базами данных, выпускаемых корпорацией IBM. Чаще всего, ссылаясь на DB2, имеют в виду реляционную систему управления базами данных DB2 Universal Database (DB2 UDB).

DB2 является единственной реляционной СУБД общего назначения, имеющей реализации на аппаратно-программном уровне (система IBM i; также в оборудовании мэйнфреймов IBM System z реализуются средства поддержки DB2).

Ниже представлены наиболее важные функциональные возможности DB2:

1) **Мультиплатформенность.** Даже самая младшая редакция "IBM DB2 Express – C" поддерживает несколько наиболее популярных операционных систем - Windows (включая x64), Linux (серверы POWER и x86), Mac OS X и Solaris x64.

2) **Advanced Copy Services (ACS).** DB2 ACS позволяет использовать технологию быстрого копирования устройства хранения данных для выполнения работы по копированию данных в операциях резервного копирования и восстановления. Возможность копирования данных средствами устройства хранения данных значительно ускоряет операции копированием через мгновенную копию (snapshot backup).

3) **Онлайновая реорганизация таблиц (REORG,** реорганизация "на месте") позволяет пользователю проводить реорганизацию таблицы без прекращения полного доступа к ней.

4) Поддержка High Availability Disaster Recovery (HADR).

Функциональность DB2 HADR обеспечивает поддержку высокой готовности и аварийное переключение для баз данных DB2.

5) Поддержка 64-разрядных экземпляров.

6) Поддержка **Materialized Query Tables** (таблицы материализованных запросов), **Query Parallelism** (параллелизм запросов), **Multidimensional Clustering Tables (MDC)**, многомерная кластеризация таблиц).

7) Поддержка сжатия данных при резервном копировании

8) Поддержка **SQL-репликации, резервного копирования и восстановления**. Резервное копирование с использованием DB2 ACS называется резервным реплицированием.

9) Поддержка Database Partitioning (разбиение баз данных).

Используется для масштабируемых кластеров: прежнее название – DB2 EEE. Данная функциональная возможность позволяет распределить один образ базы данных на несколько физических серверов.

10) **DB2 Text Search**. Функция DB2 Text Search позволяет вести поиск в текстовых столбцах таблиц DB2. Поддержка текстового поиска позволяет использовать встроенные в DB2 функции CONTAINS, SCORE и xmlcolumn-contains для поиска в текстовых индексах, построенных на основе заданных вами аргументов поиска.

11) **IBM Data Studio** – это инструментальная платформа, охватывающая весь жизненный цикл приложений (проектирование, разработка, развертывание, поддержка и управление) для всех реляционных СУБД IBM, с перспективой дальнейшего расширения поддержки. Это означает, что вы имеете не только переносимый SQL API, но и набор инструментальных программ, позволяющий реализовать бизнес-логику в масштабе всего предприятия.

1С:Предприятие – программный продукт компании 1С, предназначенный для автоматизации деятельности на предприятии.

1С:Предприятие – это одновременно и технологическая платформа, и пользовательский режим работы. Технологическая платформа предоставляет объекты (данные и метаданные) и механизмы управления объектами. Объекты описываются в виде конфигураций. При автоматизации какой-либо деятельности составляется своя конфигурация объектов, которая и представляет собой законченное прикладное решение. Конфигурация создаётся в специальном режиме работы программного продукта под названием «Конфигуратор», затем запускается режим работы под названием «1С:Предприятие», в котором пользователь получает доступ к основным функциям, реализованным в данном прикладном решении (конфигурации).

Большинство конфигураций на базе платформы версий 8.0 и 8.1 имеют:

- Возможности общей настройки программы (такой, как задание даты запрета редактирования данных);
- Возможности индивидуальной настройки программы для каждого пользователя (такой, как основной склад для автоматической подстановки в документы);
- Множество переключаемых интерфейсов;
- Множество наборов прав (ролей), причём, часть прав может назначаться пользователями в режиме Предприятия;
- Более гибко настраиваемые отчёты, нежели в версии 7.7;
- Возможность построения произвольных отчётов (так называемая «Консоль отчётов»);
- Встроенный универсальный обмен данными;
- Встроенные возможности обновления через Интернет.

В СУБД IBM DB2 есть несколько способов резервного копирования. DB2 предоставляет способы полного, инкрементного и инкрементного – дельта резервного копирования.

Резервное копирование – процесс создания копии данных на носителе, предназначенном для восстановления данных в оригинальном или новом месте их расположения в случае их повреждения или разрушения. Варианты резервного копирования:

1) **Полное резервное копирование** – обычно затрагивает всю систему и все данные. Ежедневное, ежемесячное и ежеквартальное резервное копирование подразумевает создание полной копии всех данных. Обычно оно выполняется тогда, когда копирование большого объёма данных не влияет на работу организации. Для предотвращения большого объёма использованных ресурсов используют алгоритмы сжатия, а также сочетание этого вида с другими: дифференциальным или резервным копированием журнала транзакций. Полное резервное копирование незаменимо в случае, когда нужно подготовить резервную копию для быстрого восстановления системы с нуля.

2) **Резервное копирование журнала транзакций** – в резервные копии журнала транзакций записываются все изменения базы данных; обычно такое резервное копирование выполняется в периоды между созданием полных резервных копий базы данных. Резервирование журнала транзакций сохраняет все изменения, произведённые с момента

предыдущего резервирования журнала, что позволяет предотвратить потерю данных. Однако данный способ резервного копирования имеет смысл только при наличии полной резервной копии базы данных.

3) **Дифференциальное резервное копирование** – при дифференциальном («разностном») резервном копировании каждый файл, который был изменен с момента последнего полного резервного копирования, копируется каждый раз заново. Дифференциальное копирование ускоряет процесс восстановления. Все копии файлов делаются в определенные моменты времени, что, например, важно при заражении вирусами.

Восстановление – это воссоздание базы данных после какой-либо аварии, например, отказа носителя или системы хранения, перерыва в питании или ошибки в программе. Типы восстановления:

1) **Восстановление после аварии** позволяет не оставлять базу данных в несогласованном или непригодном к использованию состоянии при неожиданном прерывании транзакции.

2) **Восстановление версии** – это возврат к предыдущей версии базы данных с помощью образа, созданного при резервном копировании.

3) **Восстановление с повтором транзакций** можно использовать для повторного применения изменений, внесенных транзакциями, которые были приняты после резервного копирования.

3. Практическое выполнение

3.1. . Определение объема базы данных

Есть несколько способов определить объем базы данных Postgres:

1. Физический размер БД, при помощи команды du.

```
# du -hsx /var/lib/pgsql
```

2. Просмотр размера базы по имени БД .

```
SELECT pg_database_size('имя базы');
```

Размер базы = 15,63 МБ

3. Каков размер таблицы(только данные)? pg_relation_size

```
SELECT DISTINCT
```

```
table_name,
```

```
pg_relation_size(table_name) / 1024
```

FROM information_schema.columns

WHERE table_schema = 'public'

Результат(представлен в КБ):

underground = 8

Branch = 8

train = 440

branch_stats = 56

station_stats = 1160

carriage = 2600

station = 512

Теоретический расчет (представлен в КБ):

Первый множитель - размер одной записи

Первый множитель - количество записей

underground = 8 * 1

Branch = 0.08 * 100

train = 0.0440 * 10000

branch_stats = 0.056 * 1000

station_stats = 0.051169 * 22670

carriage = 0.04(3) * 60000

station = 0.0512 * 10000

Как видно из приведенных выше пунктов, возможно определить как полный объем базы данных со служебными данными, так и исключительно размер чистых данных. Под служебными данными понимаются индексы, схемы, имена, пользователи, различные настройки и т.д

7.3.2 Способы архивации данных

Резервное копирование – процесс создания копии данных на носителе, предназначенном для восстановления данных в оригинальном или новом месте их расположения в случае их повреждения или разрушения.

Существует три основных способа резервного копирования:

1) **Полное резервное копирование** – обычно затрагивает всю систему и все данные. Еженедельное, ежемесячное и ежеквартальное резервное копирование подразумевает создание полной копии всех

данных. Обычно оно выполняется тогда, когда копирование большого объёма данных не влияет на работу организации. Для предотвращения большого объёма использованных ресурсов используют алгоритмы сжатия, а также сочетание этого вида с другими: дифференциальным или резервным копированием журнала транзакций. Полное резервное копирование незаменимо в случае, когда нужно подготовить резервную копию для быстрого восстановления системы с нуля.

2) Резервное копирование журнала транзакций – в резервные копии журнала транзакций записываются все изменения базы данных; обычно такое резервное копирование выполняется в периоды между созданием полных резервных копий базы данных. Резервирование журнала транзакций сохраняет все изменения, произведённые с момента предыдущего резервирования журнала, что позволяет предотвратить потерю данных. Однако данный способ резервного копирования имеет смысл только при наличии полной резервной копии базы данных.

3) Дифференциальное резервное копирование – при дифференциальном («разностном») резервном копировании каждый файл, который был изменен с момента последнего полного резервного копирования, копируется каждый раз заново. Дифференциальное копирование ускоряет процесс восстановления. Все копии файлов делаются в определенные моменты времени, что, например, важно при заражении вирусами.

7.3.5. Выбор режима архивации

Поскольку из-за особенностей предметной области база данных не равномерно наполняется новыми записями в течение дня, утрата которых критична для системы, то наиболее подходящей для БД рассматриваемой АИС «учета пассажиропотоков Московского метро» является смешанная стратегия резервирования:

– еженедельно необходимо производить полное резервное копирование БД; в наибольшей мере для этой цели подходит ночное время с воскресенья на понедельник, поскольку большинство пользователей АИС в этот момент спят перед началом рабочей недели, и потому вероятная востребованность в АИС меньше, чем в другие ночи;

– каждую ночь необходимо производить разностное резервное копирование для повышения сохранности данных; даже если система откажет в середине или в конце недели, наличие разностных резервных копий позволит быстро восстановить её данные; так же, как и полное

резервное копирование, разностное РК удобно будет проводить ночью – в 04:00 утра;

– из-за небольшой интенсивности транзакций в БД, следует также раз в неделю производить резервное копирование журнала транзакций; наиболее подходящее для этого время – ночь с четверга на пятницу – на кануне выходных дней возрастает востребованность в метро.

Схема описанной выше смешанной стратегии резервного копирования представлена на рисунке:

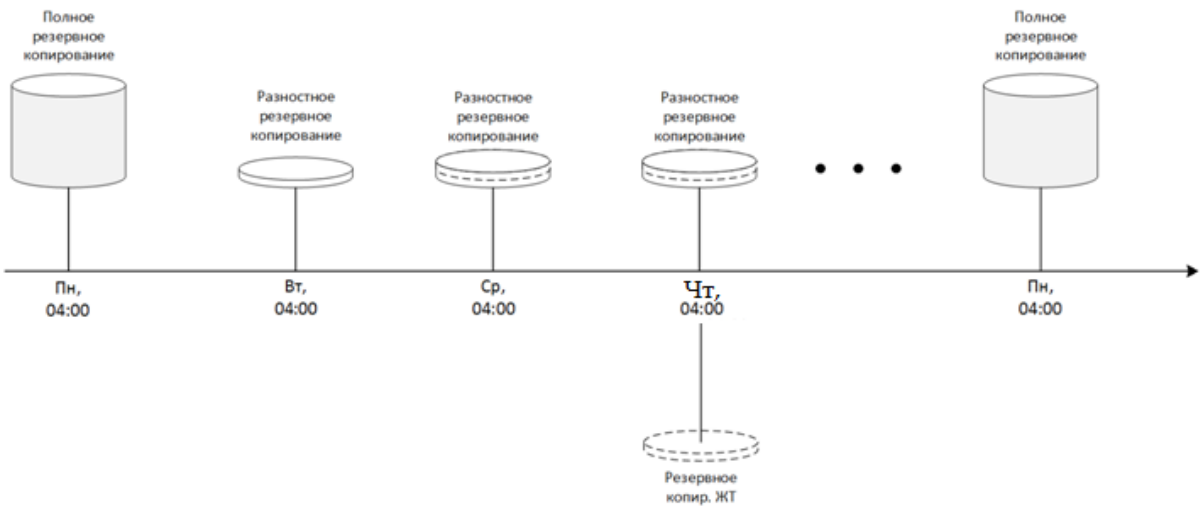


Рисунок 13. Схема резервного копирования

Достоинства этой стратегии резервного копирования базы данных заключаются в высокой надёжности и высокой скорости восстановления в случае сбоя. Разностные резервные копии позволяют восстановить данные быстрее, чем резервные копии ЖТ, и они занимают меньше места, чем полные резервные копии. Однако если разностные копии также будут повреждены, то благодаря наличию резервных копий ЖТ можно будет избежать потери данных ценой более долгого восстановления.

8. Регламент выполнения задач КП

План контроля выполнения задач КП по датам (неделя) приведены в табл. 5.

Таблица 5.

Пункты задач КП	1	2	3	4	5	6	7	9-11
Контролируемая неделя семестра	2	4	6	8	10	12	13	14

Критерии оценки в баллах приведены в табл. 6.

Таблица 6.

Пункты	№	Параметры оценки выполнения курсовой работы	Баллы	
			В срок	Не в срок
1	1	Студент получил и приступил к анализу исходных данных личного варианта курсовой работы, и не предъявил рабочие материалы	3	-1
	2	Студент приступил к выполнению пункта задания, предъявил рабочие материалы, подтверждающие выполненную работу	5	-2
	3	Студент выполнил, частично, пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу	6	-2
	4	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - с замечаниями	7	-2
	5	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - без замечаний	9	-2
Баллы за каждый из пунктов 2 - 6				
2	1	Студент приступил к выполнению пункта задания и не предъявил рабочие материалы , подтверждающие выполненную работу	3	-1
	2	Студент приступил к выполнению пункта задания, предъявил рабочие материалы, подтверждающие выполненную работу	5	-2
	3	Студент выполнил, частично, пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу	6	-2
	4	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - с замечаниями	7	-2
	5	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - без замечаний	9	-2
7	1	Студент приступил к выполнению пункта задания и не предъявил рабочие материалы , подтверждающие выполненную работу	3	-1
	2	Студент приступил к выполнению пункта задания, предъявил рабочие материалы, подтверждающие выполненную работу	5	-2
	3	Студент выполнил, частично, пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу	6	-2
	4	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - с замечаниями	7	-2
	5	Студент выполнил полностью пункт технического задания и предъявил рабочие материалы, подтверждающие выполненную работу - без замечаний	9	-2
Баллы суммарно за пункты 9 - 11				
9-11	1	Студент выполнил полностью все пункты технического задания и предоставил КП - с замечаниями .	5	-2
	2	Студент выполнил полностью все пункты технического задания и предоставил КП - без замечаний .	7	-2

Итоговое значение суммы баллов по КП:

- 1) ниже 57 – неудовлетворительно
- 2) 57–70 – удовлетворительно
- 3) 71–84 – хорошо
- 4) 85–100 – отлично

9. Требования к предоставляемой информации ДЗ студентом

1. Структурная схема клиент серверной архитектуры в формате *.vsd
2. Вычисления, проводимые в табл.1-4 в формате Excel *.xls
3. Диаграмма бизнес процесса BPMN в формате Visio – *.vsd
4. Исходная база данных АИС в PostgreSQL *.db
5. Даталогическая модель исходной базы данных АИС в формате *.JPEG
6. База данных АИС в Excel в формате – *.xls
8. Диаграммы визуализации данных в формате *.JPEG
9. Диаграммы архивации базы данных в формате Visio – *.vsd

10. Заключение

В результате выполнения курсовой работы были достигнуты следующие результаты:

1. Спроектирована сетевая архитектура АИС «учета пассажиропотоков Московского метро»
2. Осуществлён выбор оборудования и операционных систем для серверной стороны АИС.

3. Разработана модель бизнес-процесса АИС «учета пассажиропотоков Московского метро» в нотации BPMN, приведена её графическая схема и словесное описание.

4. Проведён количественный анализ информации, используемой в АИС, на основе предположительных оценок и знаний о хранении данных в ЭВМ.

5. На основе полученной информации выбран режим архивации для базы данных АИС «учета пассажиропотоков Московского метро»

6. Построена даталогическая модель БД, база данных заполнена тестовыми записями в количестве, достаточном для анализа в среде бизнес-аналитики.

7. Данные АИС «учета пассажиропотоков Московского метро» визуализированы средствами Excel в удобном для восприятия и информативном виде.

11. Список литературы

1) Тоноян С.А. Черненький В.М., Балдин А.В., Информационная управляющая система МГТУ им. Н.Э.Баумана «Электронный Университет». Изд. МГТУ им. Н.Э.Баумана 2009. -376 с. 304-325.

2) Введение в современные мейнфреймы: основы zOS/ Эбберс М., О'Брайен У., Огден Б. – М: IBM Redbooks, 2007 г.- 635 стр.

3) От мейнфреймов к облакам[электронный ресурс]/Электрон. дан. – Режим доступа: <http://www.osp.ru/os/2010/06/13003731/> – Загл. с экрана.

4) BPMN Modeling and Reference Guide/ Stephen A. White, Derek Miers - Future Strategies Inc., 2008 г. - 226 стр

5) Тоноян С.А., Балдин А.В., Елисеев Д.В. «Методика модернизации стандартных модулей типовой конфигурации на базе технологической платформы «1С: Предприятие 8» с минимальными доработками». Наука и образование (МГТУ им. Н.Э. Баумана). № 08, август 2012 URL: <http://technomag.edu.ru/>.

6) Тоноян С.А. «Супер ЭВМ - мейнфрейм». Методическое указание по выполнению лабораторных работ. Изд. МГТУ им. Н.Э.Баумана 2013г.

7) . М. Г. Радченко, Е. Ю. Хрусталева 1С: Предприятие 8.2 Практическое пособие разработчика М.: «1С-Пабблишинг»; СПб. :Питер 2009.-613с.

8) Быстрый старт для серверов DB2 [электронный ресурс]/Электрон. дан. – Режим доступа: <ftp://ftp.software.ibm.com/ps/products/db2/info/vr9/pdf/let..> – Загл. с экрана.

9) Начало работы с DB2 Express 9.7/ Рауль Ф. Чон – М.: Москва, 2010г – 269стр.

10) Официальный сайт компании IBM[электронный ресурс]/Электрон. дан. – Режим доступа: IBM.com – Загл. с экрана.