## Контрольная работа

Цель работы: изучение организации источников данных с распределенной структурой, изучение организации облачных сервисов в контексте создания цифровых решений для распределенной обработки данных.

**Раздел 1. Секционирование исходных данных для анализа**

Задание

1. Рассмотреть задачу анализа больших данных, поставленную при выполнении лабораторной работы 1 «Характеристика больших данных организации», с точки зрения их распределенной природы. Разбить все данные на несколько секций. Заполнить таблицу:

|  |  |  |
| --- | --- | --- |
| Секция | Данные | Источник |

Источник указывать для исходных данных.

1. Распределить секции по узлам хранения с учетом востребованности данных в узлах. Построить структурную схему распределенной системы.

Методические указания к выполнению

1. Распределенные данные – база данных, составные части которой размещаются в различных узлах компьютерной сети в соответствии с каким-либо критерием. Чаще всего таким критерием является а) место порождения данных, и б) наиболее вероятное место обработки данных. От распределения секций по узлам хранения зависит масштабируемость системы хранения данных.

Распределение секций по узлам хранения обычно осуществляется по результатам истории использования секций. Если ряд секций запрашивается чаще остальных, то система автоматически распределит их по нескольким узлам хранения, распределяя трафик между несколькими серверами.

Хранение данных одной секции вместе позволяет повысить эффективность их обработки и других методов повышения производительности.

2. При построении структурной схемы в качестве обозначений узлов можно использовать прямоугольники с надписями, либо пиктограммы с изображениями ЭВМ (стереотипные элементы). Связи между узлами могут быть обозначены обычными линиями.

Возможно применение линий разных стилей для обозначения каналов связи разных видов.

Возможно применение изображения облака для обозначения виртуальной распределенной вычислительной (облачной) среды.

Грубый пример подобной схемы показан далее.



Перед выполнением работы необходимо ознакомиться с материалом лекций 1 и 2 конспекта лекций по дисциплине.

**Раздел 2. Разработка архитектуры облачной платформы**

Задание

1. Разработать архитектуру цифровой платформы для обслуживания распределенной системы, структура которой разработана при выполнении практической работы 1. В состав платформы включить следующие сервисы типа SaaS, либо типа, производного от SaaS:
* сервис для сбора и предобработки данных;
* сервис для решения задачи анализа;
* сервис для представления результатов анализа.

Возможно более детальное распределение задач между сервисами. Желательно, чтобы данные одной секции обрабатывались одним сервисов.

Заполнить таблицу:

|  |  |  |
| --- | --- | --- |
| Наименование сервиса | Тип сервиса | Связанные секции данных |

1. С учетом структуры распределенной системы и необходимости обработки данных, расположенных на разных узлах, спланировать процедуры репликации данных для обеспечения надежной и максимально быстрой работы системы. Заполнить таблицу репликации:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Узел-источник | Секция данных  | Условие выбора данных | Узел-приемник | Параметры репликации |
| Частота | Длительность |

Методические указания к выполнению

Перед выполнением работы необходимо ознакомиться с материалом лекций 3 и 4 конспекта лекций по дисциплине.

**Раздел 3. Оптимизация ИТ-персонала организации в условиях эксплуатации облачного решения**

Задание

Разработать решения по перераспределению компетенций ИТ-персонала организации в условиях перехода на облачные сервисы. Представить (можно примерно) штатное расписание и организационную структуру ИТ-отдела AS-IS (как есть) и TO-BE (как должно быть).

Методические указания к выполнению

Перед выполнением работы необходимо ознакомиться с материалом лекций 5 и 6 конспекта лекций по дисциплине.

Пример штатного расписания и организационной структуры ИТ-отдела в компании «Яндекс.Сообщения»

**As-Is:**

|  |  |  |  |
| --- | --- | --- | --- |
| Должность | Оклад | Количество | Итог |
| Руководитель отдела | 300000 руб. | 1 | 300000 руб. |
| Руководитель команды | 150000 руб. | 5 | 750000 руб. |
| Программист | 90000 руб. | 60 | 3600000 руб. |
| Тестировщик | 45000 руб. | 15 | 675000 руб. |
| Главный системный администратор | 150000 руб. | 1 | 150000 руб. |
| Системный администратор | 70000 руб. | 10 | 210000 руб. |
| Уборщик серверных помещений | 35000 руб. | 3 | 105000 руб. |
| Технический писатель | 45000 руб. | 3 | 135000 руб |
| Специалист технической поддержки | 35000 руб. | 20 | 700000 руб. |



**To-Be:**

|  |  |  |  |
| --- | --- | --- | --- |
| Должность | Оклад | Количество | Итог |
| Руководитель отдела | 300000 руб. | 1 | 300000 руб. |
| Руководитель команды | 150000 руб. | 5 | 750000 руб. |
| Программист | 90000 руб. | 60 | 3600000 руб. |
| Тестировщик | 45000 руб. | 15 | 675000 руб. |
| DevOps | 90000 руб. | 3 | 270000 руб. |
| Технический писатель | 45000 руб. | 3 | 135000 руб. |
| Специалист технической поддержки | 35000 руб. | 20 | 700000 руб. |



Перед выполнением работы необходимо ознакомиться с материалом лекций 5 и 6 конспекта лекций по дисциплине.

Контрольные вопросы и задания

1. Что такое распределенные данные?
2. По каким критериям распределяются данные в распределенной системе?
3. Что такое секционирование данных и каково его назначение?
4. Сколько секций необходимо для оптимальной производительности распределенной системы?
5. Перечислите критерии распределения данных в вашей индивидуальной задаче.
6. Как связаны между собой сервисы типа Saas и PaaS?
7. Что такое репликация данных?
8. Перечислите критерии эффективности репликации данных.
9. Как внедрение распределенных технологий обработки данных отражается на структуре персонала предприятия?