

Задачи по выбору с 1 по 3

Задача 1: Анализ космических данных с использованием параллельных вычислений

Вы работаете в астрофизической лаборатории, которая получает огромные объемы данных с космического телескопа. Ваша задача - разработать программу для анализа этих данных с использованием параллельных вычислений.

Входные данные - это наборы изображений космических объектов, полученных телескопом. Каждое изображение представляет собой матрицу пикселей. Ваша задача - провести анализ каждого изображения для поиска интересных астрофизических объектов и собрать статистику о них.

Сложность задачи заключается в следующем:

Обработка большого объема данных: У вас есть сотни тысяч изображений, и каждое из них содержит миллионы пикселей. Обработка всех этих данных последовательно займет слишком много времени.

Параллельная обработка: Вам нужно разработать программу, которая может параллельно обрабатывать множество изображений на нескольких ядрах процессора, чтобы ускорить процесс анализа.

Сложный астрофизический анализ: Анализ каждого изображения требует выполнения сложных астрофизических вычислений, таких как выделение объектов, измерение их характеристик и классификация.

Сбор статистики: Для каждого найденного объекта вам нужно собрать статистические данные, такие как его яркость, координаты и другие характеристики.

Управление ресурсами: Важно правильно управлять ресурсами и избегать гонок данных и других конфликтов, которые могут возникнуть при параллельной обработке.

Ваше решение должно включать в себя эффективный механизм разделения задач между потоками/процессами, а также обеспечивать безопасное взаимодействие между ними. В конечном итоге, вы должны получить подробную статистику о найденных астрофизических объектах и предоставить ее исследователям вашей лаборатории.

Задача 2: Распределенная обработка и анализ больших данных

Вы работаете в компании по анализу данных и получили задачу обработать огромный объем текстовых данных из социальных сетей. Ваша задача - разработать программу, которая сначала соберет данные из нескольких источников, затем проведет их предварительную обработку и, наконец, выполнит анализ, чтобы выявить популярные темы и тренды.

Условия задачи:

1. Вам предоставлено 10 различных источников данных (например, VK, telegram, Facebook, Instagram и другие), каждый из которых содержит миллионы текстовых сообщений.
2. Вам нужно создать программу, которая параллельно соберет данные из всех 10 источников. Каждый источник должен быть обработан в отдельном потоке или процессе.
3. После сбора данных, они должны быть предварительно обработаны для удаления стоп-слов, пунктуации и других ненужных символов. Этот этап также должен выполняться параллельно для увеличения производительности.
4. Затем вы должны провести анализ данных с целью выявления ключевых слов, хэштегов и популярных тем, которые обсуждаются в социальных сетях.
5. Результаты анализа должны быть сохранены в базе данных или файле для последующего анализа.

Вызовы:

1. Управление большим объемом данных и обеспечение эффективной параллельной обработки.
2. Эффективное использование ресурсов сервера и управление памятью при работе с большими данными.
3. Разработка алгоритма анализа данных для выявления популярных тем и трендов.
4. Обеспечение надежности и отказоустойчивости программы при сборе и обработке данных.

Эта задача демонстрирует сложности обработки больших данных и требует хороших навыков в параллельном программировании, обработке текста и анализе данных.

Задача 3: Параллельная обработка изображений в многозадачной среде

Вы работаете над проектом по обработке изображений в многозадачной среде. У вас есть набор изображений, которые нужно обработать с помощью нескольких фильтров, таких как фильтр увеличения резкости, фильтр сепии и фильтр уменьшения размера. Ваша задача - создать программу, которая может обрабатывать изображения параллельно с использованием многозадачности.

Условия задачи:

1. У вас есть папка с исходными изображениями, и вы хотите сохранить обработанные изображения в отдельной папке.
2. Каждое изображение должно быть обработано всеми тремя фильтрами.
3. Ваша программа должна поддерживать параллельную обработку с использованием потоков или процессов.
4. При обработке изображений с разными фильтрами не должно возникать конфликтов доступа к данным.
5. Важно обеспечить безопасное сохранение обработанных изображений в выходную папку.

Вызовы:

- Необходимо разработать архитектуру программы, которая позволит эффективно использовать многозадачность для обработки изображений.
- Необходимо учесть, как обрабатывать ошибки и исключения, которые могут возникнуть в процессе обработки изображений.
- Подумайте об эффективном распределении ресурсов (например, процессорного времени) между потоками или процессами, чтобы избежать избыточных накладных расходов.

Эта задача представляет собой сложную задачу параллельной обработки данных, которая требует хорошего понимания концепций многозадачности и обработки изображений, а также умения эффективно управлять ресурсами и избегать конфликтов доступа к данным при параллельной обработке.