

После того как кеширование было внедрено в систему, появляются новые задачи, связанные с его управлением. Основные из них — это **инвалидация кеша и стратегия вытеснения данных**.

Инвалидация кеша

Инвалидация кеша — это процесс удаления или обновления данных в кеше, которые стали устаревшими или недействительными.

Когда речь идет о инвалидации кеша, различные стратегии могут быть применены для обновления или удаления данных в кеше и источнике данных (обычно базе данных). Вот некоторые из наиболее распространенных методов:

Сквозная запись (Write-Through)

1. При обновлении данных новые значения одновременно сохраняются и в источнике, и в кеше.

2. **Плюсы:**

- Гарантированная консистентность данных между кешем и источником.
- Меньший риск потери данных.

3. **Минусы:**

- Скорость записи может снизиться, так как необходимо выполнить две операции: одну в источнике и одну в кеше.

Запись в обход (Write-Around)

1. Данные сохраняются непосредственно в источнике, минуя кеш.

2. **Плюсы:**

- Экономия ресурсов кеша, так как кешируются только часто запрашиваемые данные.

3. **Минусы:**

- Если недавно обновленные данные сразу же запрашиваются, кеш не сможет предоставить их, и придется обращаться к источнику.

Реверсивная запись (Write-Back или Write-Behind)

1. Новые или измененные данные сначала сохраняются в кеше, а затем асинхронно копируются в источник данных.

2. **Плюсы:**

- Запись в системе происходит очень быстро, так как основное действие происходит в кеше.
- Увеличивается общая пропускная способность системы.

3. **Минусы:**

- Существует риск потери данных, если они не были асинхронно скопированы в источник до какого-либо сбоя.

Эти стратегии инвалидации кеша могут быть использованы в различных комбинациях или адаптированы в зависимости от специфических требований вашего приложения и шаблонов использования данных. Выбор правильной стратегии зависит от баланса между требованиями к производительности и согласованности данных. Самое главное - вы должны помнить цель внедрения кеша - это уменьшение задержки. Если этого не произошло после внедрения, значит вы что-то сделали не так (кешируете не на том уровне или выбрали не ту стратегию инвалидации).

Стратегии вытеснения данных

Стратегия вытеснения определяет, какие данные должны быть удалены из кеша, когда он заполняется.

Виды стратегий:

1. **First In, First Out (FIFO)**

- Удаляются самые давно добавленные записи.
- Может привести к удалению важных или часто используемых данных.

2. **Last In, First Out (LIFO)**

- Удаляются самые последние добавленные записи.
- Не учитывает частоту или актуальность данных.

3. Least Recently Used (LRU)

- Удаляются данные, к которым дольше всего не обращались.
- Хорошо работает, если недавно использованные данные будут использованы снова в ближайшем будущем.

4. Most Recently Used (MRU)

- Удаляются недавно использованные данные.
- Может быть полезно для сценариев, где старые данные используются чаще.

5. Least Frequently Used (LFU)

- Удаляются наименее часто используемые данные.
- Требуется подсчет частоты обращений, что может быть ресурсоемким.

6. Random Replacement (RR)

- Удаляются случайно выбранные записи.
- Просто в реализации, но может удалить важные данные.

Принцип Парето 80-20:

В соответствии с правилом 80-20, вы можете хранить в кеше лишь 20% самых важных данных. Эти данные, как правило, обеспечивают наибольшую производительность (дают 80% результата). Стратегии вытеснения данных могут быть настроены так, чтобы учитывать это правило, исключая менее важные или редко используемые данные.

Правильное управление кешем требует тщательного планирования и непрерывного мониторинга. Не забывайте анализировать метрики задержки чтобы делать выводы о тех или иных принятых технических решениях.