

Тестовое задание [Huobi global]

Восстановление лимитной книги заявок

1. Введение

Книга заявок (limit order book / "биржевой стакан") - это перечень актуальных заявок трейдеров (ордеров) на покупку (bid) и продажу (ask) какого либо актива (акции / валюты / криптовалюты).

Данные о текущем стакане - это очень ценная информация, на основе которой строится основная логика высокочастотных торговых роботов.

Price (USDT)	Amount (BTC)	Sum (BTC)
11701.29	0.089242	8.315285
11701.01	0.500000	8.226043
11700.65	4.000000	7.726043
11700.49	0.500000	3.726043
11700.40	0.128341	3.226043
11700.18	0.420000	3.097702
11700.17	0.675000	2.677702
11699.87	0.351210	2.002702
11699.84	0.130000	1.651492
11699.69	0.100000	1.521492
11699.31	0.300000	1.421492
11699.08	0.226600	1.121492
11699.00	0.140000	0.894892
11698.99	0.199800	0.754892
11696.96	0.060991	0.555092
11695.26	0.494101	0.494101
11695.25 ≈ 874469.3 RUB		More
11695.25	8.232211	8.232211
11695.07	0.003271	8.235482
11695.02	0.008072	8.243554
11694.99	0.140000	8.383554
11694.79	0.199500	8.583054
11694.78	0.008072	8.591126
11694.64	0.064090	8.655216
11694.55	0.008072	8.663288
11694.35	0.139009	8.802297
11694.32	0.108072	8.910369
11694.25	0.157500	9.067869
11694.08	0.008072	9.075941
11693.85	0.008073	9.084014
11693.71	0.000529	9.084543
11693.67	0.170826	9.255369
11693.64	1.692804	10.948173

Пример стакана на инструменте bitcoin/usd

В примере зеленым обозначена сторона заявок на покупку BTC, а красным сторона заявок на продажу.

Каждый уровень характеризуется ценой и содержит количество предлагаемой валюты.

Уровень цены 11695.25 называется лучшим уровнем покупки (best bid) и означает, что на текущий момент существует покупатель, готовый приобрести 8.232211 BTC по данной цене.

Уровень 11695.26 соответственно называется лучшим уровнем продажи (best ask) и содержит 0.494101 BTC.

Важно заметить, что данные в стакане анонимные и агрегированы по уровням цен, т.е. на одном уровне могут находиться несколько заявок от разных игроков.

Данный вид агрегации называется L2 market data, существует более подробный L3 уровень, но в текущем задании он не рассматривается.

Однако, передавать целые стаканы (order book snapshots) не оптимально, поэтому биржи отправляют данные в виде изменений (level update), в т.ч. добавлений и удалений уровней стакана.

Имея начальный снимок стакана и набор апдейтов можно восстановить целый стакан на момент каждого апдейта.

2. Техническое задание

Дан набор реальных биржевых данных (формат сообщений описан ниже). В первой строчке находится снэпшот стакана, далее следуют сообщения об апдейтах.

Вам необходимо реализовать программу на языке C++, которая восстанавливает снэпшоты стаканов на каждый апдейт. В выходной файл необходимо вывести цены и объемы лучших уровней best bid и best ask с временной меткой. Подумайте над выбором оптимальной структуры данных для данной задачи, учитывая, что максимальное количество ценовых уровней на одной стороне снэпшота равно 20. Подумайте, что нужно будет поменять, если размер одной стороны книги будет больше - 400 уровней. Основное требование к решению - это его производительность.

Дополнительно необходимо провести benchmark на скорость обновления стакана и получения из него best bid и best ask. Проведите сравнение различных структур данных для книги. Бенчмарк должен включать в себя только применения

распаршенного события к ордербуку, не нужно включать в него время парсинга и время дискового чтения/записи.

3. Формат данных

Входной файл

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/50f943ad-5c8b-466e-92a6-db8a14111604/huobi_global_depth.log.gz

В квадратных скобках указана debug информация, `Get Object:` - сырое биржевое сообщение в формате JSON. `Get Object:` в первой строке файла содержит полный снэпшот стакана, в последующих строках записаны апдейты.

`event_time` - локальный timestamp записи события в микросекундах,

Уровни на покупку и продажу лежат в `["bids"]` и `["asks"]` соответственно, в виде массива `[[price, amount], ..]`. Биды отсортированы по убыванию цены, аски - по возрастанию. Если кол-во валюты на уровне равно 0, это значит что ордер был исполнен или отменен, значит необходимо удалить соответствующий ценовой уровень из ордербука.

Остальные поля являются сервисными и для выполнения задания не потребуются.

Выходной файл

В каждой строке необходимо записать best bid и best ask в формате:

```
{event_time}, {best bid price}, {best bid amount}, {best ask price}, {best ask amount}
```

3. Оформление решения

Оформить в виде Github/Gitlab репозитория с кодом решения, CMakeLists, входными данными. Требуется предоставить readme с описанием основных частей, инструкцией по сборке, запуску бенчмарков.

Код должен быть читаемым, аккуратным. Решение должно быть в лучших традициях именно с++, если нет явной мотивации использовать сишное наследие