Level 3 Order Book Pipelines

P. Daniel Tyreus
Talk Overview

- What is a central limit order book
- Ingesting data from an exchange
- Processing the ingested data with Beam
Crypto Exchanges

- Trusted counterparty
- Exchange cryptocurrencies to USD and back.
- Price-time priority order book to match buyers with sellers.
- FIFO queue of order sizes at each price level.
# BEAM-USD Order Book

<table>
<thead>
<tr>
<th>Asks</th>
<th>Bids</th>
</tr>
</thead>
<tbody>
<tr>
<td>$103</td>
<td>$99</td>
</tr>
<tr>
<td>$102</td>
<td>$98</td>
</tr>
<tr>
<td>$101</td>
<td>$97</td>
</tr>
<tr>
<td>15 BEAM</td>
<td>3 BEAM</td>
</tr>
<tr>
<td>5 BEAM</td>
<td>13 BEAM</td>
</tr>
<tr>
<td>1 BEAM</td>
<td>22 BEAM</td>
</tr>
</tbody>
</table>
Level 3 Order Book Snapshot

```json
{
  "bids": [
    [
      "39733.91",
      "1",
      "95dd1e2c-8bf0-45ee-9cee-a990742817e0"
    ],
    [
      "39733",
      "0.05071409",
      "754d152c-123f-4419-8c05-21591bf61274"
    ],
    [
      "39732.99",
      "0.29",
      "e84582ef-b067-4274-b7ed-8c4155fb982a"
    ],
    [
      "39732.89",
      "1",
      "21013671-32fe-41cb-bd2b-23c9a1fd22c7"
    ],
    [
      "39731.52",
      "0.04585183",
      "32140f85-e679-4906-8c5d-be5491e54bb7"
    ]
  ]
}
```
Websocket Feed

OPEN, DONE, MATCH, RECEIVED

{  
    "type": "received",  
    "time": "2021-08-02T08:19:27.028459Z",  
    "product_id": "BTC-USD",  
    "sequence": 10,  
    "order_id": "d50ec984-77a8-66f114b0de9b",  
    "size": "1.34",  
    "price": "39669.1",  
    "side": "buy",  
    "order_type": "limit"  
}
A Beam Pipeline (GCP Dataflow)

Read from websocket

key: asset name
payload: message JSON

GCP PubSub

FileIO write JSON to Google Storage

Group by asset name

Apply windows

Read from PubSub

FileIO write order books to Google Storage

Window

PTransform: apply message changes

Group by asset name
Scaling up on GCP

- Deployed to Google Cloud Dataflow
- Pipeline is processing over 100 unique keys (asset names) in parallel
- Overall we handle 5-10 billion messages per week
What are we working on?

- Multiple Apache Beam pipelines for trade data, order book data, and derivatives data.
- Custom PTransforms - format for DS

email: daniel@cambrianasset.com
twitter: @tyreus