Alternative Data Service

J-GATE Timestamp Data

Specification

Version 1.4

October 2023

JPX Market Innovation & Research, Inc.

We shall not be liable to any persons or entities for any disadvantages or problems whatsoever arising from the use of information contained in this document.

The contents of this document are subject to change without notice.

Revision history

#	Date	Chapter	Change description	Remark
1	2021/9/21	-	Initial version.	-
2	2021/10/5	2.1	Added a disclaimer about data provision.	Version 1.1
3	2021/10/5	3.3	Mass cancel entries will not have “msg_itch” and “p6d” column.	Version 1.1
4	2022/10/31	4.2	Added “Internal” to values to be populated in column “liquidity”.	Version 1.2
5	2022/12/5	3.1 3.5 4.1 4.5	Added data for weekend test.	Version 1.3
6	2022/12/5	2.1 3.3.1	Added notes about some error cases.	Version 1.3
7	2023/10/30	2.2.1 2.2.3	Added tuning on latency offset. Added setup of timestamping devices.	Version 1.4

TABLE OF CONTENTS

1 About This Specification

1.1 Introduction

This is a specification document for “J-GATE Timestamp Data”, one of datasets of the Alternative Data Service (“the Service”). It mainly describes information necessary for users of the Service to handle the dataset.

It is prohibited to redistribute this document.

2 Service Details

2.1 Data Overview

J-GATE Timestamp Data is a set of CSV files containing times when each order/trade message of OUCH and OMNetAPI passes each point on network paths constituting arrownet and J-GATE as well as when the message is processed by Matching Engine (“ME”) application.

A timestamp on the network path is issued per ethernet frame using Layer 1 switch (“L1 switch”) after splitting a signal from a physical cable constituting the network using an optical splitter. Then, the frame is captured, protocol-decoded, and used as a timestamp to indicate when the message passes each point. A timestamp on the application is issued when an application logic on ME server starts processing the message.

A file containing order message timestamps is separate from one for execution notice message timestamps.

The files include TOCOM messages.

The data does not guarantee the integrity and consistency with any other data feed from J-GATE (ITCH market data, OMNetAPI broadcast, etc.).

The data might miss some of the entries when the number of orders rapidly increases in a day and the performance of the timestamping data processing system are not enough to process all of the traffic.

The data might be unavailable when there are some malformed packets that do not comply with IP/TCP/UDP and any other application layer protocols in the network traffic between J-GATE and the client systems cause issues on the timestamping data processing system.

As the data has a limitation that it may include records where each timestamp lacks chronological consistency, they need to be treated appropriately such as removed as necessary.

2.2 Timestamp Specification

2.2.1 Timestamping Point

In Figure 2.2.1, T1 through T6 and TA represent points to issue timestamps to each message.

At T1 through T6, it is L1 switches that issue timestamps, while at TA, it is ME servers.

At T1 through T6, timestamps correspond to the time when the exchange device receives or transmits the traffic: i.e., the actual timestamps offset the time elapsed between the optical splitter and the timestamping device.

At each point, timestamps are captured in both directions, i.e. from a client system to the exchange system and from the exchange system to a client system. However, at TA, timestamps are captured in only a direction from a client system to the exchange system, and at T6, only from the exchange system to a client system.

At T3 and T4, timestamps are only captured by primary gateway servers. At T6, which is on ITCH Multicast packet path, timestamps are captured for both ITCH feed 1 and 2.

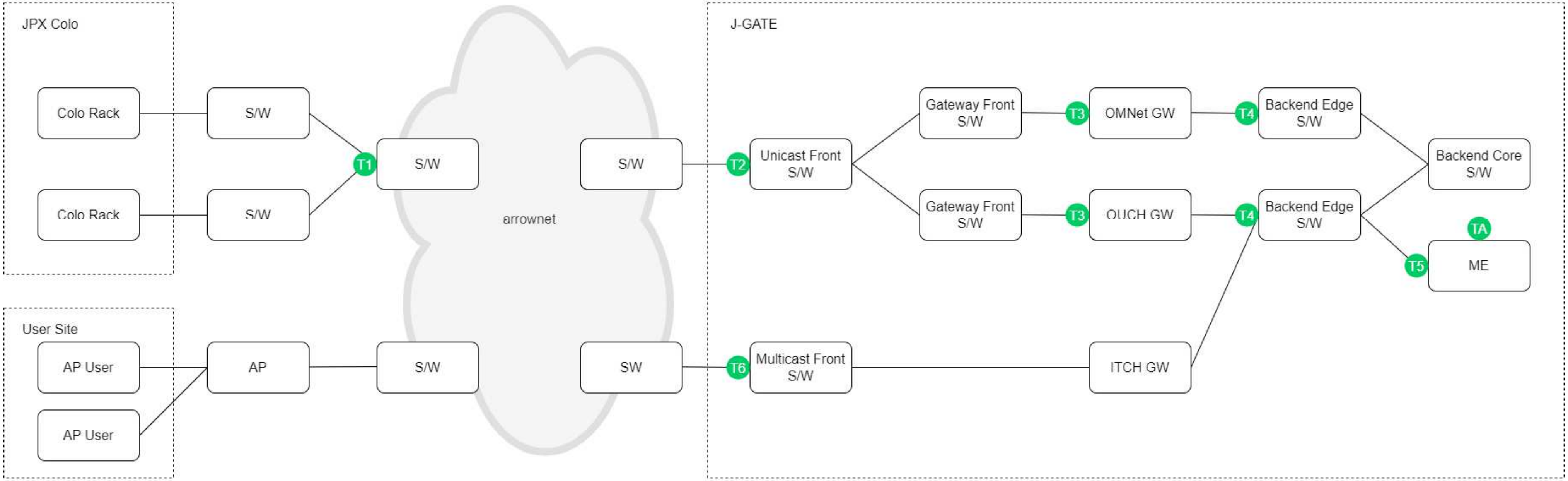


Figure 2.2.1 System architecture and timestamping points

2.2.2 Time Synchronization Precision

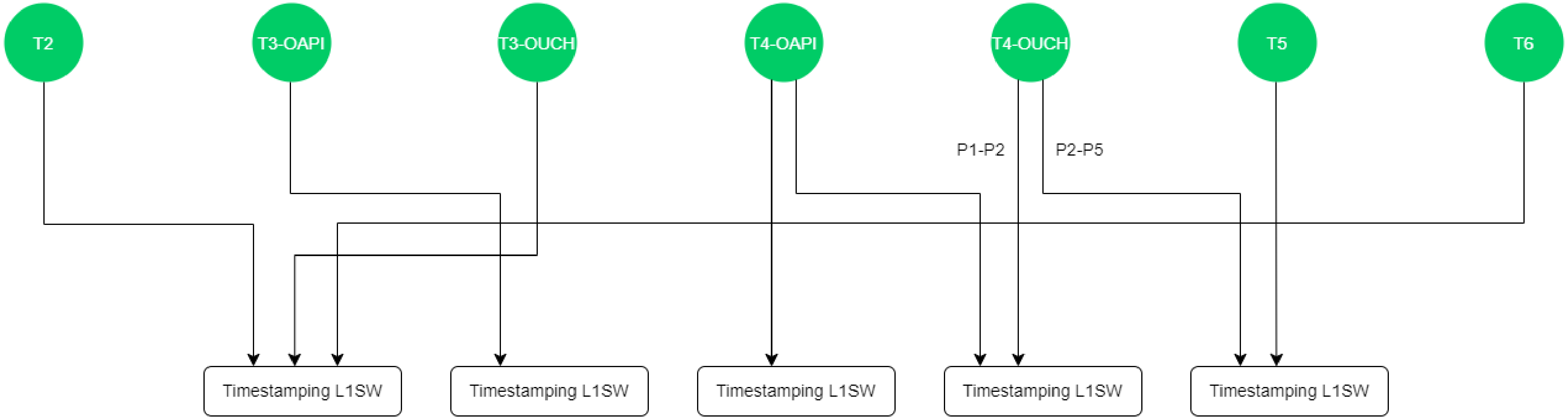
The table below shows time synchronization precision (in nanoseconds) of the L1 switches and ME servers responsible for timestamping. These devices are time-synchronized by PTP and use hardware timestamp.

Item	L1 switch (T1-T6)	ME server (TA)
Mean	9.953488	16.911720
Standard Deviation	6.821222	9.653347
Min	0.000000	0.276000
Median	9.000000	15.403000
90%ile	19.000000	30.402000
95%ile	22.600000	35.536000
Max	33.000000	62.464000

*The values above were observed in our internal test in March 2021 and are presented here for a reference purpose. These levels are not always guaranteed in the Service.

2.2.3 Timestamping Device Setup

Below diagram shows which timestamping devices are responsible for which capture points.



Timestamping L1SW #1	T2: OUCH/OMnetAPI (All traffic) T3: OUCH (All traffic) T6: ITCH (All traffic)
Timestamping L1SW #2	T3: OMnetAPI (All traffic)
Timestamping L1SW #3	T4: OMnetAPI (A part of traffic)
Timestamping L1SW #4	T4: OMnetAPI (A part of traffic) T4: OUCH (All traffic on partitipn 1 and a part of traffic on partition 2)
Timestamping L1SW #5	T4: OUCH (All traffic on partition 3, partition 4 partition 5 and a part of traffic on partition 2) T5: OUCH/OMnetAPI (All traffic)

*At T1, traffic is distributed across multiple timestamping devices. Details are not disclosed due to some security reasons.

*At all capture points from T1 through T6, traffic in both directions, i.e., from a client system to the exchange system and from the exchange system to a client system, are captured by the same timestamping device.

*The setup above is subject to change without any prior notice in future.

3 Order Timestamp File

3.1 File Specification

File name	Format	Character code	Line feed code	Delimiter	Header	Remark
OrderTimestamp_yyyymmdd.csv	CSV	UTF-8-	LF	, (comma)	Present	
OrderTimestamp_yyyymmdd_WeekendTest.csv	CSV	UTF-8-	LF	, (comma)	Present	Data for weekend test

One file records data from order acceptance at 8:00 a.m. of the day to night session close of the following day.

For weekend test data, one file records data from market open to market close of the weekend test.

Messages of all users are recorded in one file. It has no information that can identify the user who sent each individual message.

One line in a CSV file represents one request message sent from a client system.

3.2 List of Columns

#	Item name	Header name	Example value	Description
1	Protocol	protocol	OAPI	Order protocol. OAPI or OUCH.

2	Message type (inbound)	msg_up	MO33	Message type sent from a client system to the exchange system.
				OAPI = MO31, MO33, MO4
				OUCH = O, U, X, Y, M, P

3	Message type (outbound)	msg_down	Ack	Message type sent from the exchange system to a client system.
				OAPI = Ack
				OUCH = A, U, C, M

4	Message type (ITCH)	msg_itch	A	Message type of ITCH message derived from the transaction. A or D.
5	Transaction ID	trans_id	66EE2883000167E0	OAPI -> A value of txid stored in an 8-byte buffer pointer area when calling omniapi_tx_ex() function.
6	Order ID	order_id	66EE2883000167DD	OAPI -> order_number_u OUCH -> OrderID

7	Partition	partition	3	Matching Engine partition. Any of 1 through 5.
8	Timestamp (T1 inbound)	p1u	1618995821717767012	Time when passing the tapping point T1 in the direction from a client system to the exchange system.
9	Timestamp (T2 inbound)	p2u	1618995821717767098	Time when passing the tapping point T2 in the direction from a client system to the exchange system.
10	Timestamp (T3 inbound)	p3u	1618995821717783571	Time when passing the tapping point T3 in the direction from a client system to the exchange system.
11	Timestamp (T4 inbound)	p4u	1618995821717784001	Time when passing the tapping point T4 in the direction from a client system to the exchange system.
12	Timestamp (T5 inbound)	p5u	1618995821715518954	Time when passing the tapping point T5 in the direction from a client system to the exchange system.
13	Timestamp (ME)	t_me	1618995821718042721	Time when passing the tapping point TA in the direction from a client system to the exchange system.
14	Timestamp (T6 outbound #1)	p6d_1	1618995821717994021	Time when passing the tapping point T6 (ITCH feed 1) in the direction from the exchange system to a client system.
15	Timestamp (T6 outbound #2)	p6d_2	1618995821717994054	Time when passing the tapping point T6 (ITCH feed 2) in the direction from the exchange system to a client system.
16	Timestamp (T5 outbound)	p5d	1618995821717976788	Time when passing the tapping point T5 in the direction from the exchange system to a client system.
17	Timestamp (T4 outbound)	p4d	1618995821717977243	Time when passing the tapping point T4 in the direction from the exchange system to a client system.
18	Timestamp (T3 outbound)	p3d	1618995821717994119	Time when passing the tapping point T3 in the direction from the exchange system to a client system.
19	Timestamp (T2 outbound)	p2d	1618995821717995777	Time when passing the tapping point T2 in the direction from the exchange system to a client system.
20	Timestamp (T1 outbound)	p1d	1618995821717998521	Time when passing the tapping point T1 in the direction from the exchange system to a client system.

*All timestamps are provided in the format of Epoch time + nanosecond.

3.3 List of Messages to Be Recorded

The CSV file combines a request message from a client system to the exchange system, a response message from the exchange system to a client system, and an ITCH message, and prints it in a single line. The table below lists response message types and ITCH message types to be combined with each request message type. All the other messages are not captured.

Protocol	Request message type	Response message type	ITCH message type

OMnetAPI	MO31 (Order Entry Transaction)	OMnetAPI Ack *1	A (New Order)
	MO33 (Order Alteration Transaction)	OMnetAPI Ack *1	D & A (Deleted Order + New Order)
	MO4 (Order Deletion Transaction)	OMnetAPI Ack *1	D (Deleted Order) *2
OUCH	O (Enter Order)	A (Order Accepted)	A (New Order)
	U (Replace Order)	U (Order Replaced)	D & A (Deleted Order + New Order)
	X (Cancel Order)	C (Order Canceled)	D (Deleted Order)
	Y (Cancel by Order ID)	C (Order Canceled)	D (Deleted Order)
	M (Mass Cancel)	M (Mass Cancel Accepted)	None *2
	P (Enter MM Order)	A (Order Accepted)	A (New Order)

*1: It refers to a packet containing cstatus, txstat, ordered, etc. that is returned when calling omniapi_tx_ex function of OMnetAPI.

*2: The file does not combine the ITCH messages get disseminated upon mass cancel request. msg_itch, p6d_1 and p6d_2 columns will be blank for the entry.

3.3.1 Error Order

In principle, if a request message from a client system provokes an error, the message is not printed in the dataset. The error here means a case where in OMnetAPI, return values (cstatus and txstat) to omniapi_tx_ex function are negative, and a case where in OUCH, J (Order rejected) is returned in Sequenced packet. However, in some cases, certain error orders may be printed.

3.3.2 MO33/MO4 Missing Target

If a client system sends in MO33 or MO4 to cancel or amend an order, and the order specified in the message doesn’t exist on the orderbook, which results in return values (cstatus and txstat) to omniapi_tx_ex function being 0, the message is not printed in the dataset.

3.3.3 Immediately Traded Order and IOC Order

When an ITCH message does not disseminate A-tag but only disseminates E-tag due to an immediately traded order or IOC order, the message is printed in the CSV file with only ITCH message timestamp (msg_itch column and p6d column) being blank.

3.4 How to Identify Individual Message

For a user to identify their own message in the dataset, they need to map a value populated in each OMnetAPI / OUCH message with a certain column in the CSV file. Listed below are fields to be used as keys in each protocol.

Protocol	Column name	Corresponding message field
OMnetAPI	trans_id	A value of txid stored in an 8-byte buffer pointer area when calling omniapi_tx_ex() function.
OUCH	t_me	Sequenced packet -> Timestamp Nanoseconds
	order_id	Sequenced packet -> OrderID

3.5 Caveat to Weekend Test Data

As ME server is not PTP time synchronized during a weekend test, timestamps in the t_me column considerably deviate from the actual time.

4 Trade Timestamp File

4.1 File Specification

File name	Format	Character code	Line feed code	Delimiter	Header	Remark
TradeTimestamp_yyyymmdd.csv	CSV	UTF-8-	LF	, (comma)	Present	
TradeTimestamp_yyyymmdd_WeekendTest.csv	CSV	UTF-8-	LF	, (comma)	Present	Data for weekend test

One file records data from order acceptance at 8:00 a.m. of the day to night session close of the following day.

For weekend test data, one file records data from market open to market close of the weekend test.

Messages of all users are recorded in one file. It has no information that can identify the user who sent each individual message.

One line in a CSV file represents one execution notice message sent from the exchange system.

4.2 List of Columns

#	Item name	Header name	Example value	Description
1	Protocol	protocol	OUCH	Execution notice protocol. OAPI or OUCH.

2	Message type (outbound)	msg_down	E	Message type sent from the exchange system to a client system.
				OAPI = BD6
				OUCH = E

3	Message type (ITCH)	msg_itch	E	Message type of ITCH message derived from the transaction. E or C.
4	Partition	partition	1	Matching Engine partition. Any of 1 through 5.
5	Orderbook ID	orderbook_id	66EE2883000167E0	Orderbook ID populated in each message of ITCH and OUCH.
6	Order ID	order_id	66AEE08102EFBE8D	OAPI -> order_number_u OUCH -> OrderID
7	Match ID	match_id	0106DE4100320FDD	Match ID populated in each message of ITCH and OUCH.

8	Transaction ID	trans_id	66AEE08102EFBE95	OAPI -> A value of txid stored in an 8-byte buffer pointer area when calling omniapi_tx_ex() function.
9	Buy/Sell	buy_sell	Buy	Buy or Sell.
10	Liquidity	liquidity	Passive	Passive, Aggressive or Internal.
11	Deal type	deal_source	1618995821717783891	Normal or Auction.
12	Timestamp (ME)	t_me	1618995821717784001	Time when passing the tapping point TA in the direction from a client system to the exchange system.
13	Timestamp (T6 outbound #1)	p6d_1	1618995821715518954	Time when passing the tapping point T6 (ITCH feed 1) in the direction from the exchange system to a client system.
14	Timestamp (T6 outbound #2)	p6d_2	1618995821718042721	Time when passing the tapping point T6 (ITCH feed 2) in the direction from the exchange system to a client system.
15	Timestamp (T5 outbound)	p5d	1618995821717994021	Time when passing the tapping point T5 in the direction from the exchange system to a client system.
16	Timestamp (T4 outbound)	p4d	1618995821717994054	Time when passing the tapping point T4 in the direction from the exchange system to a client system.
17	Timestamp (T3 outbound)	p3d	1618995821717976788	Time when passing the tapping point T3 in the direction from the exchange system to a client system.
18	Timestamp (T2 outbound)	p2d	1618995821717977243	Time when passing the tapping point T2 in the direction from the exchange system to a client system.
19	Timestamp (T1 outbound)	p1d	1618995821717994119	Time when passing the tapping point T1 in the direction from the exchange system to a client system.

* All timestamps are provided in the format of Epoch time + nanosecond.

4.3 List of Messages to Be Recorded

The CSV file combines an execution notice message from the exchange system to a client system and an ITCH message, and prints it in a single line.

Protocol	Response message type	ITCH message type
OMnetAPI	BD6 (Dedicated Trade Information VIB) *1	E (Execution Notice) C (Execution Notice with Trade Information)
OUCH	E (Order Executed)	E (Execution Notice) C (Execution Notice with Trade Information)

*1: If two or more BD6 messages are disseminated from gateway because multiple users under the same sub-participant subscribe BD6, the file combines the message sent out first.

4.4 How to Identify Individual Message

For a user to identify their own message in the dataset, they need to map a value populated in each message of OMnetAPI and OUCH with a certain column in the CSV file. Listed below are fields to be used as keys in each protocol.

Protocol	Column name	Corresponding message field
OMnetAPI	t_me	BD6->time_spec->tv_nsec
	order_id	BD6->order_number_u
OUCH	t_me	Order Executed->Timestamp Nanoseconds
	match_id	Order Executed->Match ID

4.5 Caveat to Weekend Test Data

As ME server is not PTP time synchronized during a weekend test, timestamps in the t_me column considerably deviate from the actual time.