Refinitiv, an LSEG Business – STAC Summit | Innovation Roundup

Unleashing the value of market data

Pramod Nayak: Director – Product Management, Low Latency, LSEG

.....

Unleashing the Value of Market Data is Difficult & Costly

Data must be accurate in **Data Analysis** order to take action on it Data driven decision making an issue for Understanding data model profitability and compliance requires market data and Vendor costs may be tied market expertise to use cases Data processing consumes • time better spent generating data-driven insights - slower to insights and frustrated staff Analysis building blocks not • available - always starting from scratch Legacy infrastructure hampers data analysis. Move to the cloud has been slow and with hiccups Lack of simplification or agility with current models increases cost and complexity of actionable decisions REFINITIV

Data Sourcing

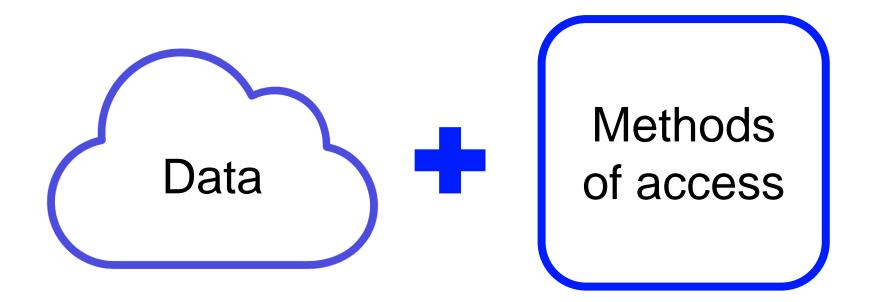
- Tech Costs cost of maintaining colocation and data storage
- Compliance Costs changing and inconsistent licensing terms
- Quality Issues accuracy, precision, completeness

Data Processing

- Frequent specification . changes require ongoing, agile development
- Expertise required to normalize correctly
- Symbology must . accommodate time series analysis and disparate data sources
- Data management costs to maintain consolidated data store include HW, SW, and staff
- Legacy infrastructure hampers efficient data processing
- Support resources lack market data depth

Decision Making

What do we offer ?





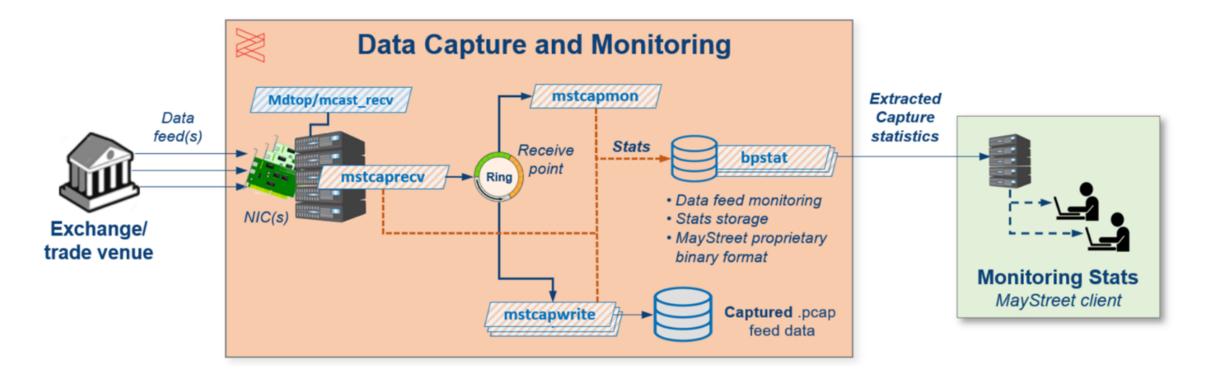
How we make it easy

Programmatic Access to Data **Bellport Enterprise Data Files** Pcaps, Normalized files, **Queryable Services Analytics Workbench** Low latency market data feed handlers Custom data sets Single digit Files available Optimized for Optimized for Orchestration -Consistent Clients access Analytic

microsecond	normalization	through SFTP	MayStreet S3	targeted queries	big data	getting data out	Environment	
latency	across asset		buckets	(HPQ API)	(Parquet in	of Data Lake	with	
	classes				AWS)		Orchestration	
	world wide							



Receiving and storing an incoming data feed, and monitoring the overall capture process





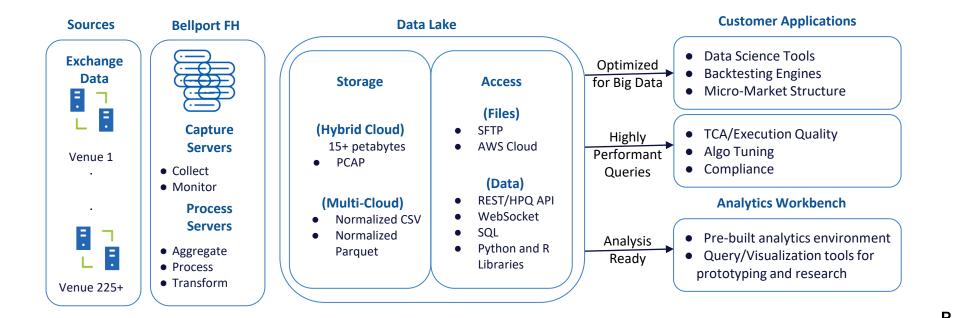
Market Data Lake

MayStreet data store built off of raw PCAP data.

Market data from all major trading venues is captured on performance network hardware with 4-way redundancy. A|B tapes are arbitrated for completeness.

GPS PTP signals ensure highest quality nanosecond-level timestamping.

- Multiple data delivery configurations To meet user need, data is available in real-time, near-time, delayed and/or end-of day scheduling as well as historical data.
- Multiple formats PCAPs, Normalized (Parquet / CSV), Queryable access (HPQ and Workbench)
- Flexible delivery SFTP, Cloud (AWS)



Analytics Workbench

Queryable access to MayStreet and LSEG Data Lake using Python and SQL in an interactive IDE

Interactively analyse market data in Jupyter notebooks

Easily spin up Dask clusters and schedule cloud-based Dask jobs to execute processor-heavy data queries.

Build custom services - Using EngineRoom, build custom services/apps running on user-defined servers that can be accessed outside of the Analytics Workbench desktop.

MayStreet high-integrity data - All Analytics Workbench data is sourced from MayStreet's Market Data Lake datasets, renown for nanosecond accuracy and best-in-class integrity algorithms.

Allows you to query the order book for AAPL on NASDAQ at "09:45:00.458398903" on 2022-10-18

F README.	.mdp 뀩 🛛 🍨 custom-expo	ort.py 📒	Data Lake.session.sql 🏾	🔹 extrac 🔲 🚥	Data Lake: SELECT *	🖯 Data Lake: S	ELECT * ×			
a. Î	Run on active connection SELECT	= Select bloc	k 🗋 Copy python code	<u> Ea</u>	bidprice_1	bidquantity_1	bidnumorders_1	bidprice_2	bidquantity_2	bidnumorders_2
2					a <mark>o</mark> c Filter	₀ <mark>b</mark> ⊂ Filter	ه <mark>ه</mark> د Filter	abc Filter	alle Filter	allic Filter
3	FROM				108.77	51	2	108.6	302	2
4	"prod_lake"."p_n WHERE	nst_data_lak	e".mt_aggregated_pric	e_update	108.77	51	2	108.62	100	1
6	dt = '2022-07-05' AM	1D			108.77	51	2	108.63	100	1
7	<pre>f = 'total_view' AND</pre>)			108.77	51	2	108.6	2	1
8	product = 'AMZN'				108.77	51	2	108.6	2	1
10					108.77	51	2	108.62	100	1
11 12					108.77	51	2	108.62	100	1
13					108.77	51	2	108.6	2	1

Bellport Enterprise

Feed handler library providing low-latency, in-process data ingest, transform and delivery

Written in C++, with Java bindings

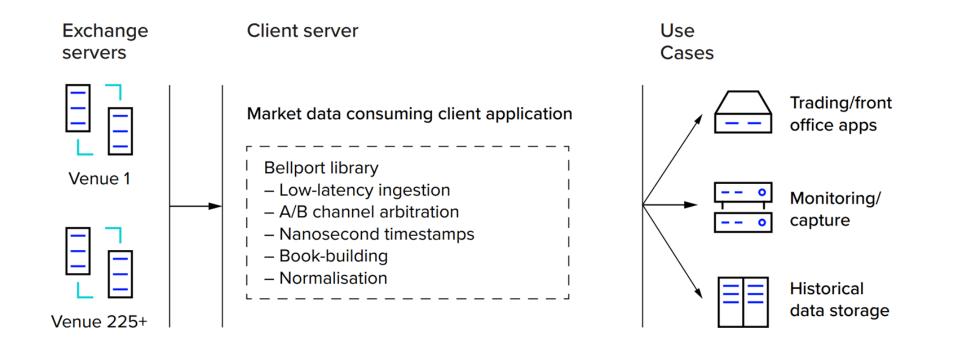
Single high-performance software solution for real-time and historical market data

Clear, transparent, extensible data model

A/B channel arbitration, book building, per-product subscription

Three data access levels: raw, normalized and aggregated

Support for kernel bypass NICs such as Solarflare, Exanic, Mellanox and Myricom. Direct integration with Solarflare EFVI for enhanced performance



MayStreet Makes it Easy

Simplify your data

Unlock value

Content

Enhancements

Innovation

Data Sourcing

 Over 225 feeds already provisioned in Market Data Lake **Data Processing**

•

•

•

•

Developers maintain feed

Data normalized across all

Normalized, Parquet) to meet

Firms can bring their queries

to MayStreet or take all the

Multiple formats (PCAP,

exchange specs

data feeds offered

data analysis needs

data post-processing

handlers in accordance with

- Leverages Bellport Feed Handler for capture. For onprem solutions,
- Meets required quality standards - lossless, nanosecond timestamps,

Data Analysis

 Analytic Workbench provides a window into Market Data Lake with built-in Python libraries and book-building capabilities

• Pre-built analytics in Workbench

- For data science use cases using industry-standard tools, MayStreet supports Parquet file formats.
- Access to pre-built analytics like TWAP and VWAP are available via API. On-demand access to the state of the book at any period of time

Decision-Making

•

- MayStreet QAs its data not just for completeness but also for accuracy within expected parameters
- MayStreet supplies data to regulators and understands compliance use cases
 - MayStreet customers can shift their focus from data processing to value-add activities associated with the analysis and monetization of data



Thank you

REFINITIV

An LSEG Business