Streaming to humans:

Can open source hack it?

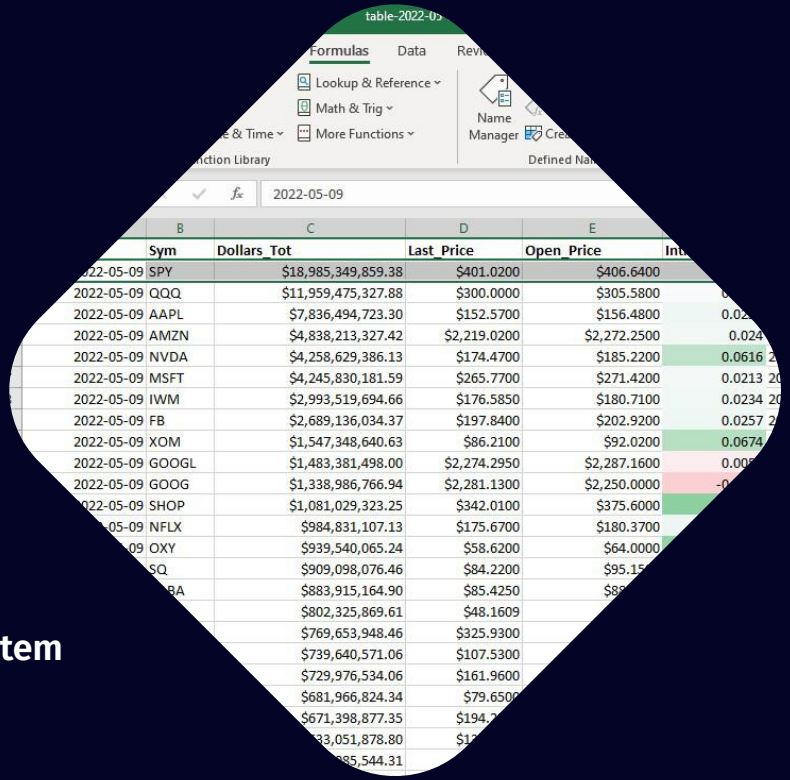
Deephaven

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN →  WALLEYE CAPITAL → Deephaven


Getting data to
people matters

What do people want?

1. Receive data in real-time
2. Consume & produce data
3. See tables (& widgets)
4. Be first class citizens in the data system



	Sym	Dollars_Tot	Last_Price	Open_Price
022-05-09	SPY	$18,985,349,859.38	$401.0200	$406.6400
2022-05-09	QQQ	$11,959,475,327.88	$300.0000	$305.5800
2022-05-09	AAPL	$7,836,494,723.30	$152.5700	$156.4800
2022-05-09	AMZN	$4,838,213,327.42	$2,219.0200	$2,272.2500
2022-05-09	NVDA	$4,258,629,386.13	$174.4700	$185.2200
2022-05-09	MSFT	$4,245,830,181.59	$265.7700	$271.4200
2022-05-09	IWM	$2,993,519,694.66	$176.5850	$180.7100
2022-05-09	FB	$2,689,136,034.37	$197.8400	$202.9200
2022-05-09	XOM	$1,547,348,640.63	$86.2100	$92.0200
2022-05-09	GOOGL	$1,483,381,498.00	$2,274.2950	$2,287.1600
2022-05-09	GOOG	$1,338,986,766.94	$2,281.1300	$2,250.0000
022-05-09	SHOP	$1,081,029,323.25	$342.0100	$375.6000
-05-09	NFLX	$984,831,107.13	$175.6700	$180.3700
09	OXY	$939,540,065.24	$58.6200	$64.0000
	SQ	$909,098,076.46	$84.2200	$95.15
	BA	$883,915,164.90	$85.4250	$8
		$802,325,869.61	$48.1609	
		$769,653,948.46	$325.9300	
		$739,640,571.06	$107.5300	
		$729,976,534.06	$161.9600	
		$681,966,824.34	$79.650	
		$671,398,877.35	$194.7	
		33,051,878.80	$1	
		85,544.31		

Those needs catalyze today's discussion

- Distill requirements
- Review OSS options
- Itemize unsolved problems
- Describe solutions

Technical needs

1. **Browser** compatibility
2. **Tables** that don’t suck
3. A ubiquitous **backplane**
4. Efficient data packaging

Browser Compatibility

- Easy-access
- Platform-independent
- Resource-light
- Mobile-friendly

Tables that don’t suck...

... support real-time data

... support new tables & schema on-the-fly

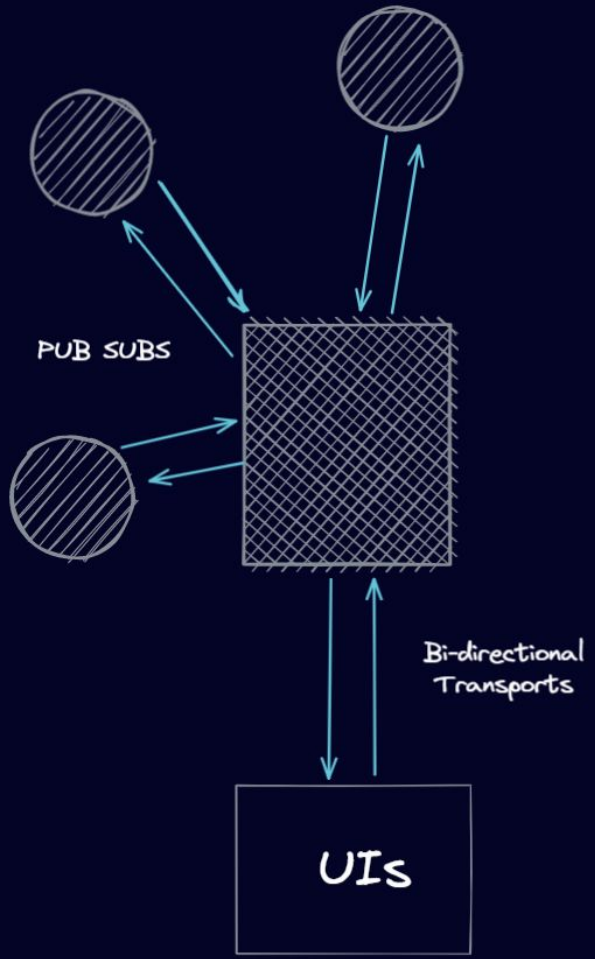
... support widgets

Tables that don't suck...

... support real-time data

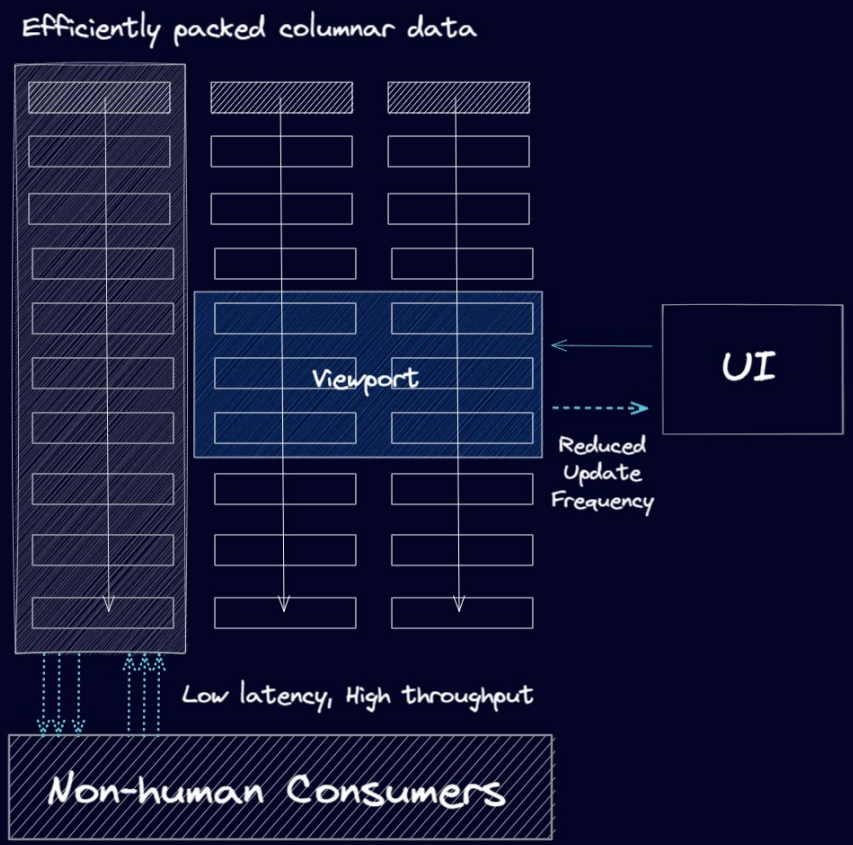
... support new tables & schema on-the-fly

... support widgets



A ubiquitous backplane

- Reduce costs
- Accelerate interoperability
- Satisfy client apps
- Support bi-directionality



Efficient data consumption

- Give client apps control
- Aspire to zero copy
- Package data well

Contenders for transport

Open, popular, modern

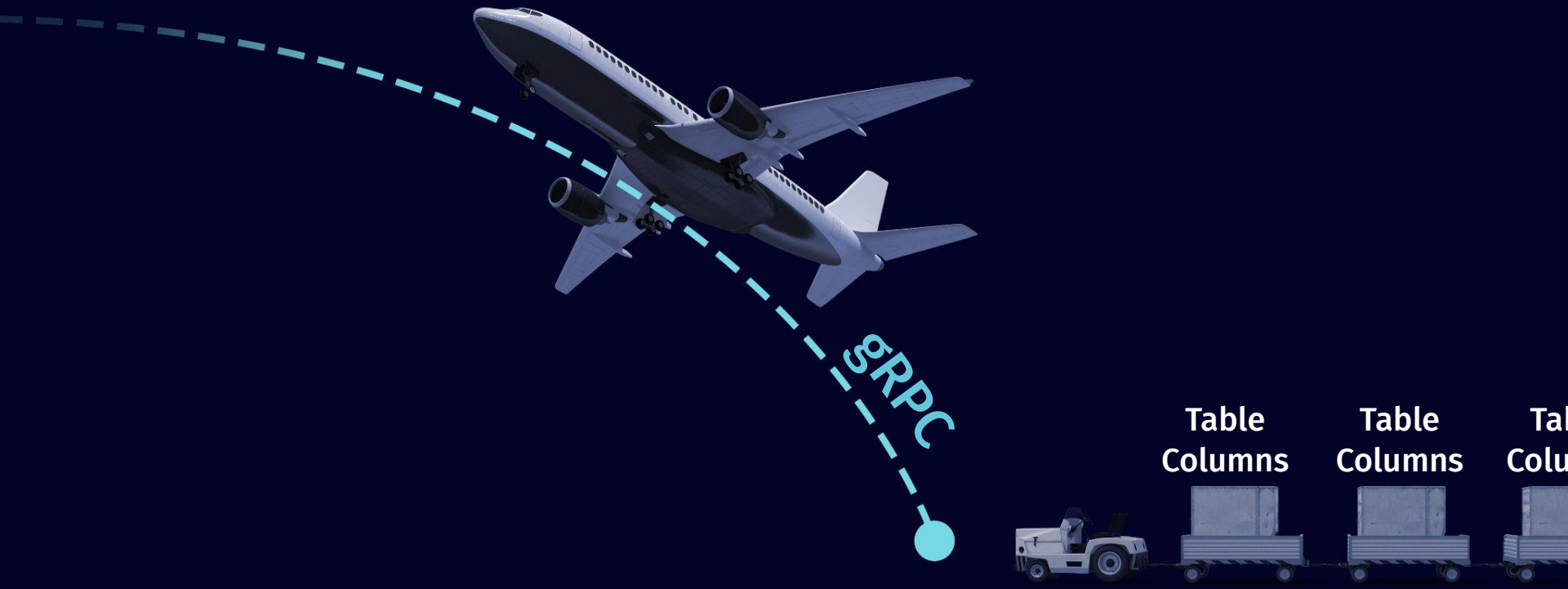
Contenders for transport

Open, popular, modern

Kafka and ZeroMQ

> Reliable & scalable, but...
>
> ... load-ignorant, by design

Arrow Flight *knows* what it's carrying



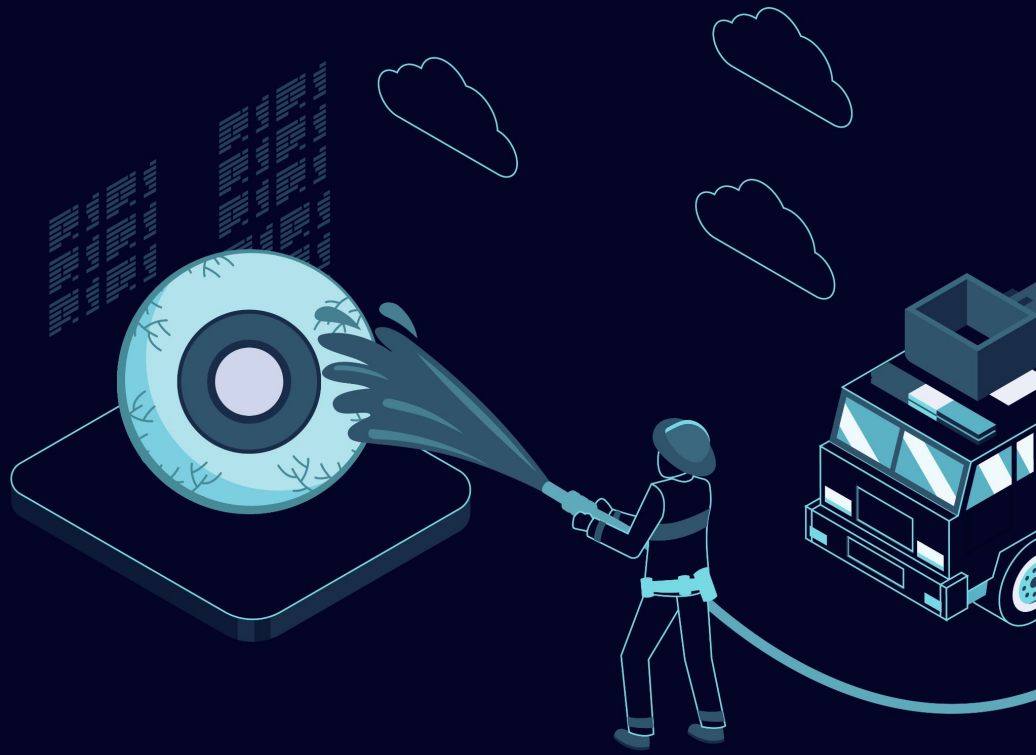
Requirements	ZeroMQ	Kafka	Arrow Flight
Low latency	Can be	Can be	Yes
Variable schema	Yes	Yes	Yes
Can support tables	Yes	Yes	Yes
Efficient for tables	No	No	Yes
Zero Copy	No	No	Yes
Can support real-time tables	Sort of	Sort of	No
Bi-directional streaming	Yes	Yes	No
Client controls throughput and latency	No	No	No
Works in a browser	No	No	No

Kafka and ZeroMQ are non-starters

Blind (and blinding) appending streams

Bad news:

1. **No efficiency for tables**
2. **Pub & sub cannot collaborate**
3. **Workarounds become gross**



Arrow Flight

Table super-powers

Easily extendable

gRPC-based (hello, streaming)

gRPC means http2 (hello, browser)

What we needed to do

1. Put “table changes” in Flight’s payload
2. Use Flight’s DoExchange() to provide real streaming
3. Make a JS client that can connect & stream

Introducing

Barrage

Table changes (not static tables)

- Think: “deltas”
- 4 varieties –> add, remove, modify, shift
- Give this a name: ***streaming_tables***

Barrage: DoExchange() for streaming

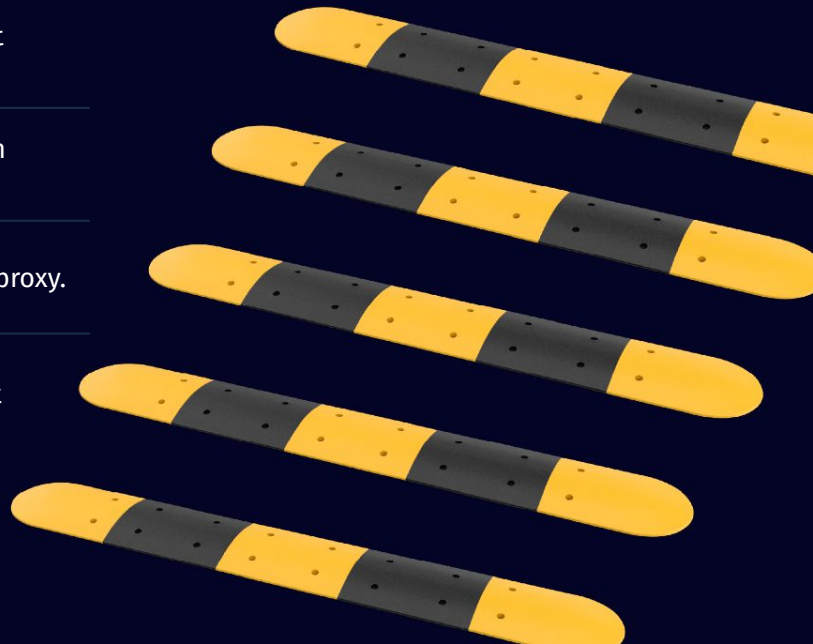
snapshot(): “Tell me about a subset of this table”

subscribe(): “Give me the current contents and push me updates”

publish(): “I want to send streaming_tables too”

The Javascript client was a hard problem

Problems	Solutions
gRPC doesn't actually work in browsers	gRPC-Web is an almost-gRPC that IS accessible to browsers.
gRPC and gRPC-web are not actually compatible	Put a proxy between server and client (Envoy).
gRPC-web JS client does not support streaming binary data	Use improbable-engineering's custom implementation.
Web browsers cannot stream data back to servers	Split methods and add a server-side proxy.
Browsers require SSL for http2: bad for localhost & secure environments	Improbable-engineering's web socket proxy for gRPC

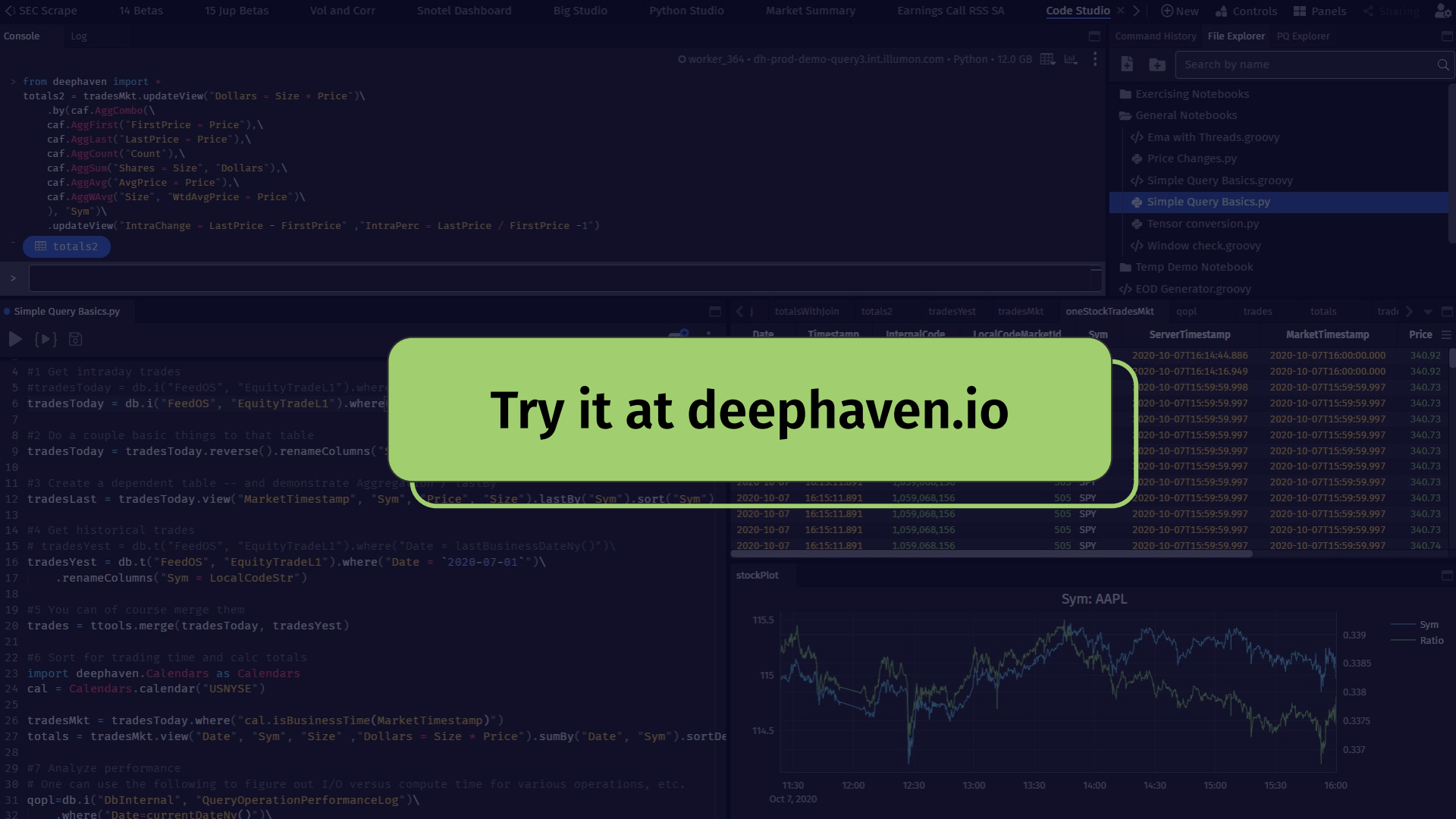


Repos worth exploring

barrage
github.com/deephaven/barrage

deephaven-core
github.com/deephaven/deephaven-core

web-client-ui
github.com/deephaven/web-client-ui


Try it at deephaven.io