Network Design Considerations for Trading Systems

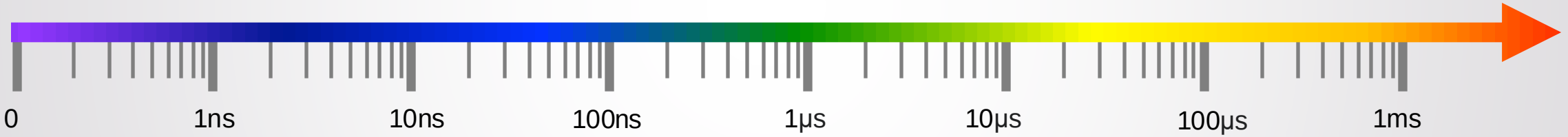
Andy Myers Brian Nigito Nate Foster

Latency is key to competitive trading

Competitive trading strategies: a continuum

Distance light travels...

Edge of podium [30cm]	First row of seats [3m]	Back row of seats [30m]	Parking lot [300m]	Pacific Ocean [3km]	Long Beach [30km]	San Jose [300km]



High-Frequency Trading Systems

Simple algorithms that mostly react to patterns and shifts in market conditions

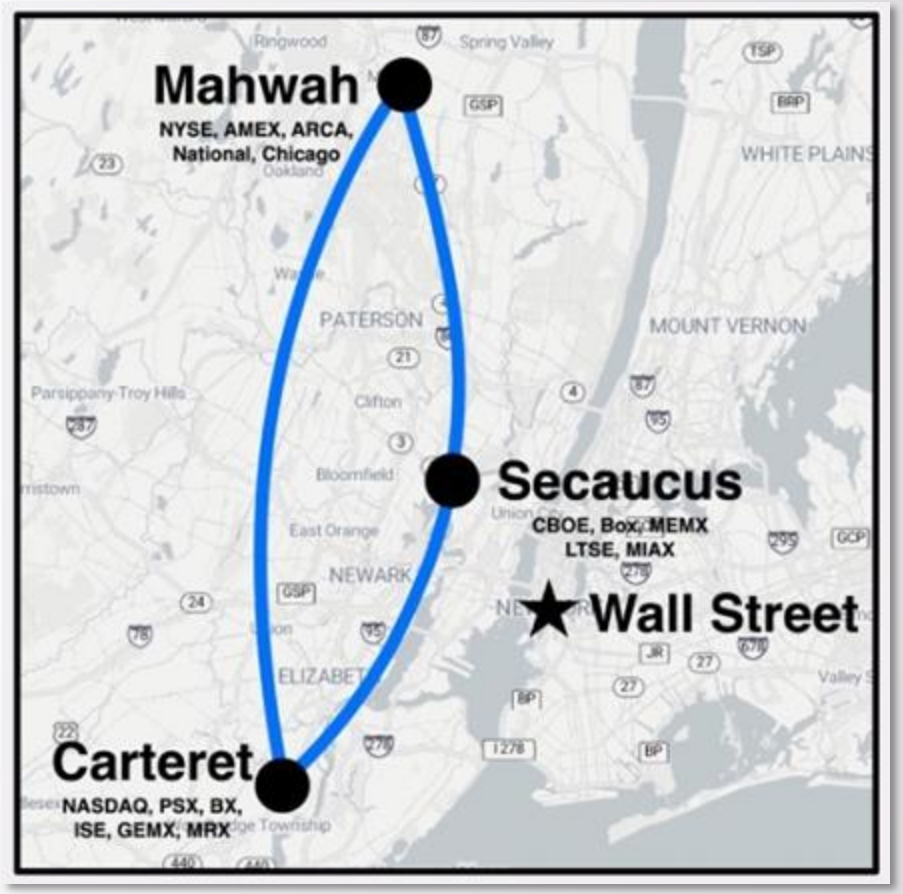
“Scaled up” Trading Systems

More sophisticated algorithms that consume data from multiple sources, perform more complex computations, etc.

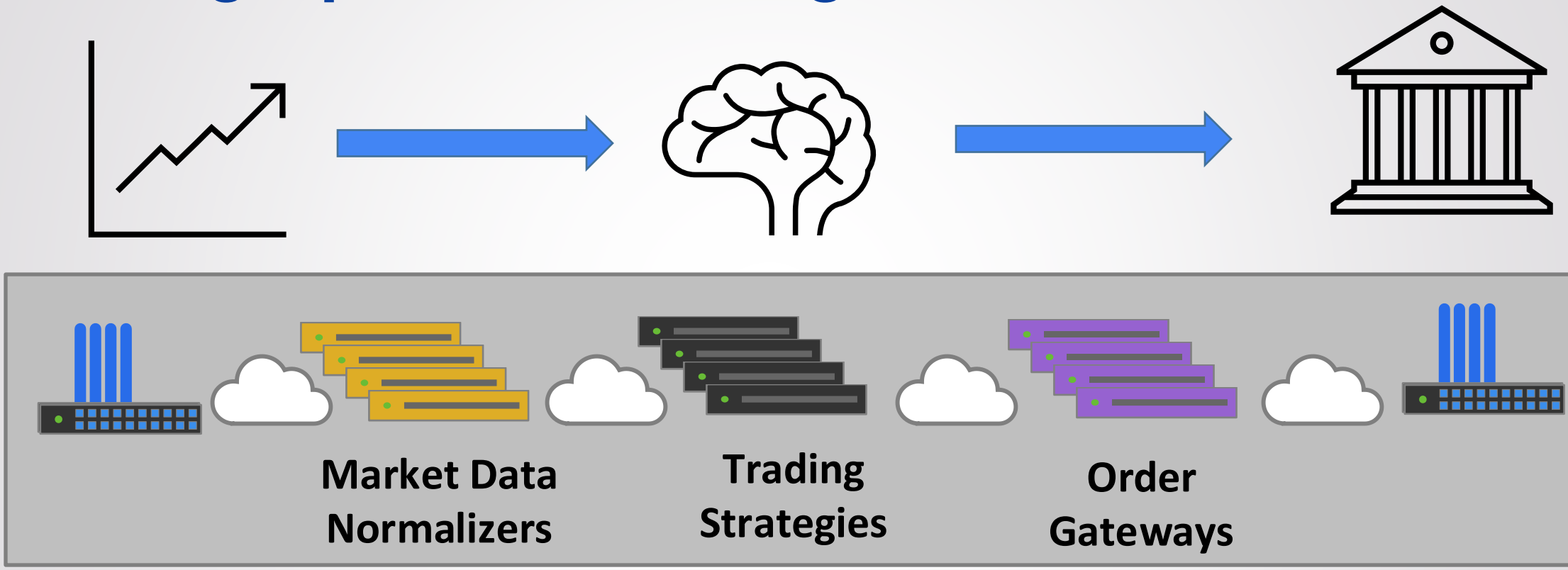
Scaling Up: Across Multiple Sites

Idea: get more sophisticated strategy by processing data from multiple exchanges

- **Co-locate servers with each exchange**
  100s to 1000s of servers per site
- **Connect colos with low-latency links**
  Trading firms sometimes use exotic links (e.g. shortwave, microwave)
- **Need packet capture and timestamps**
  For research, tick to order, etc.



Scaling Up: Within a Single Site

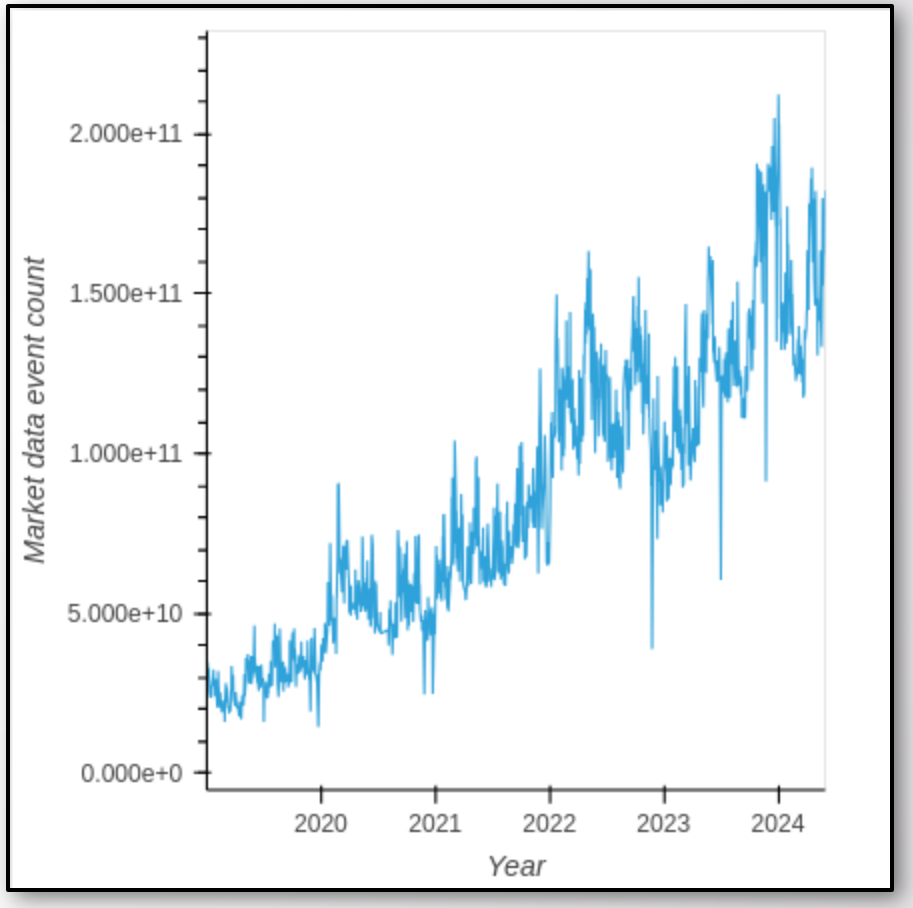


Non-standard network requirement: heavy use of multicast to deliver (normalized) market data to relevant trading strategies

Workload Trends

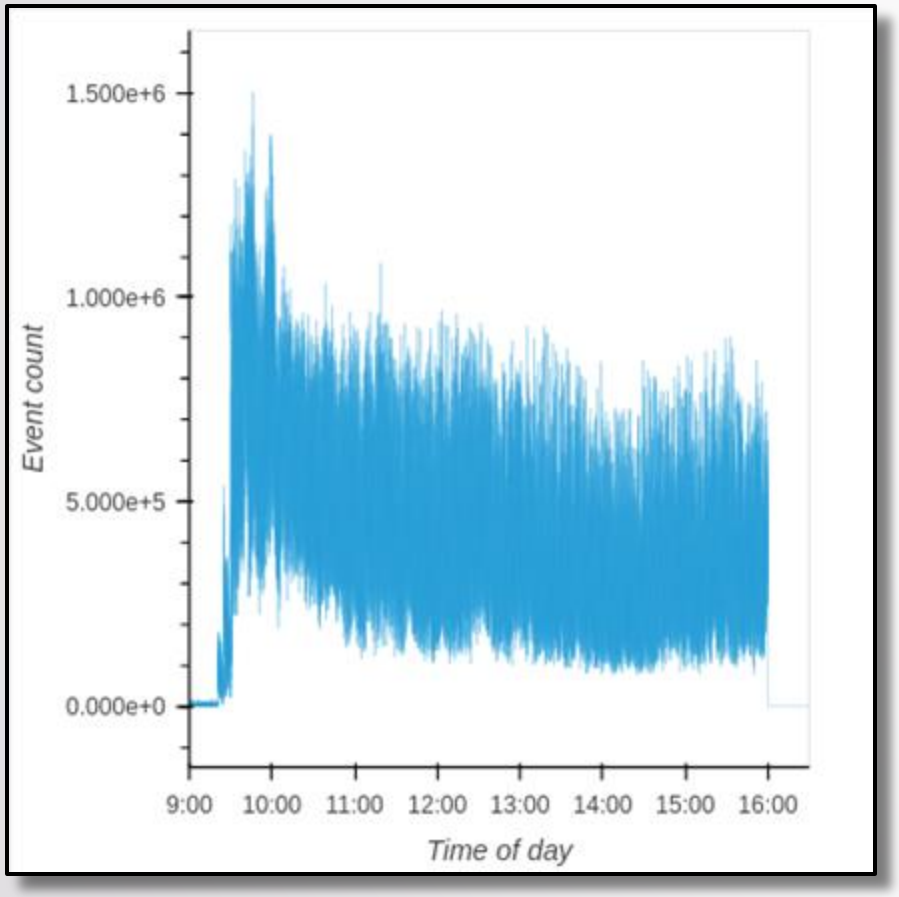
Growth of Market Data

- **Market Data**
  - Steadily growing over time
  - 200B / day in 2024
- **Key contributors**
  - New exchanges
  - Growth in option markets
- **Impact for trading systems**
  - *Must* spread processing across multiple servers

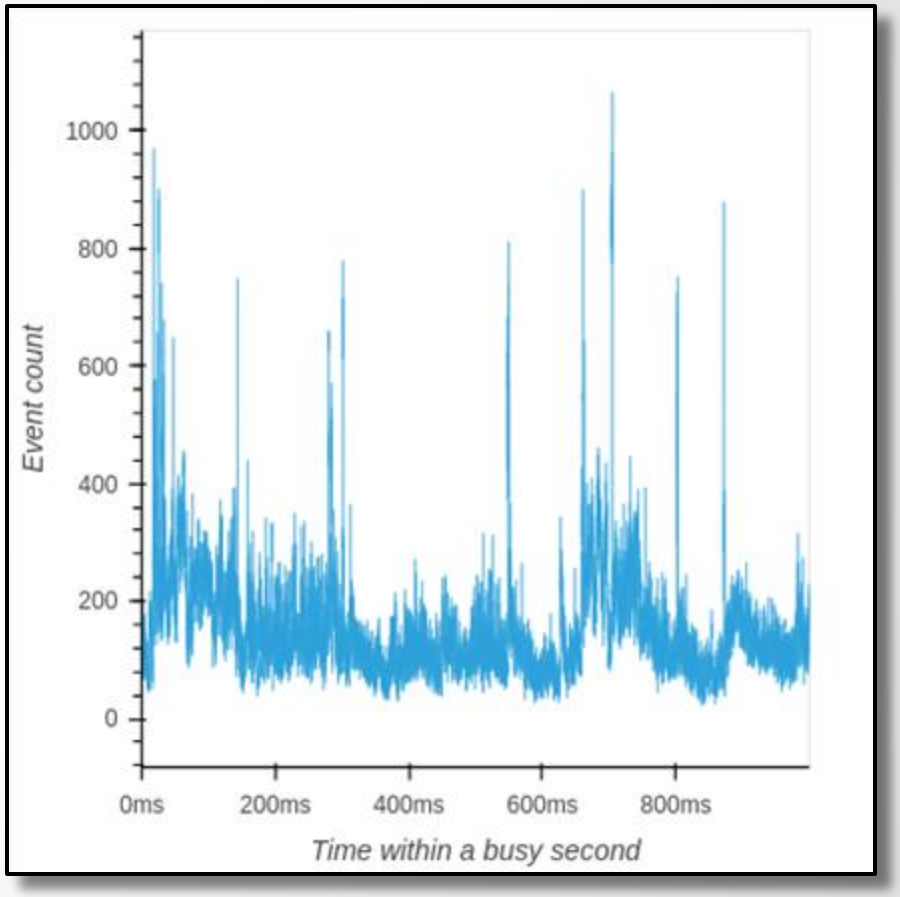


Events per day in US markets

Zooming In: Options Contracts for One Stock

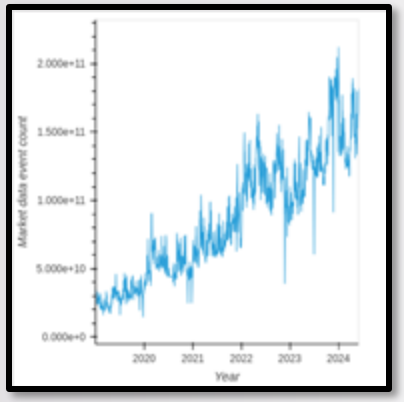


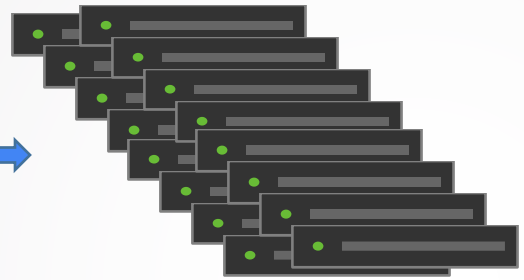
Over entire trading day
Single CPU budget ≈ 600ns / event



During busiest second
Single CPU budget ≈ 100ns / event

Scaling up to match the growth in workloads





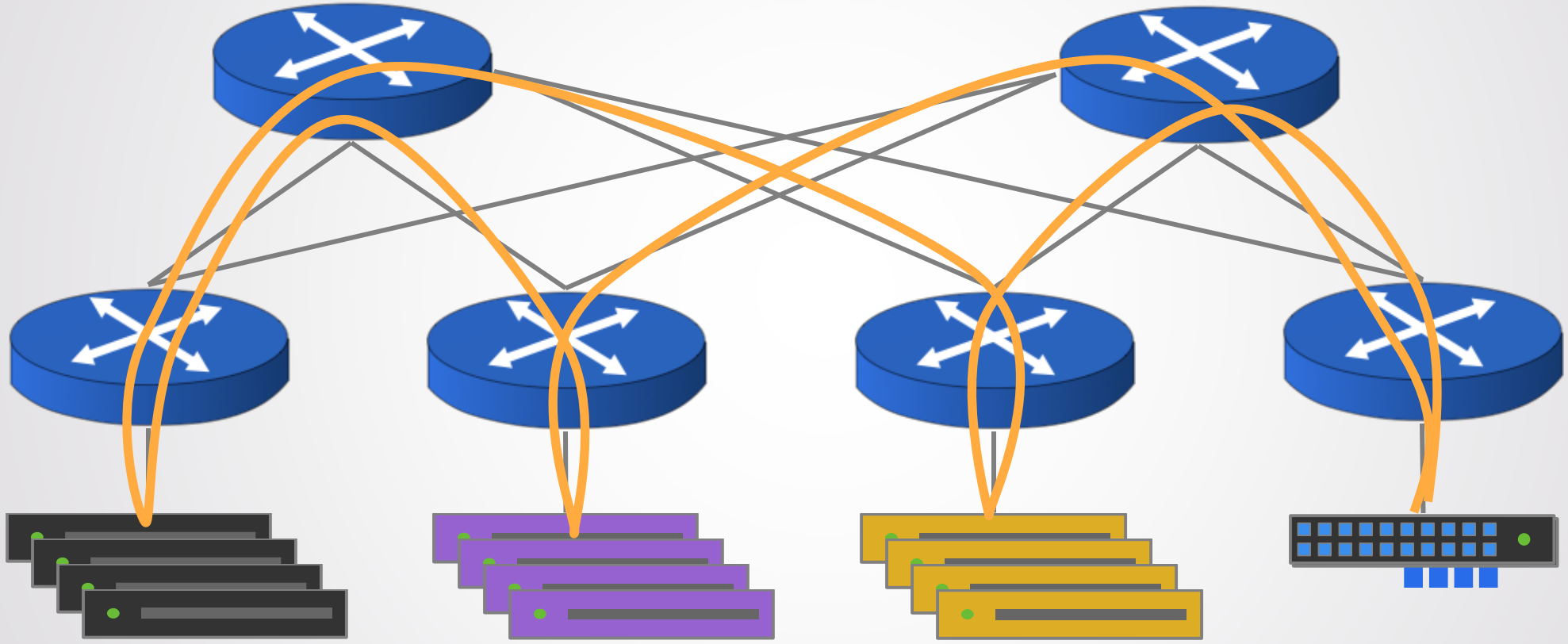
Bigger network
Larger multicast tables
More switch hops

How have commodity switches evolved?

- **Bandwidth:** *increased* 10x 😀
- **Multicast Capacity:** *increased*, but only 80% □
- **Latency:** *increased* by 20% 😱

Possible Ways Forward

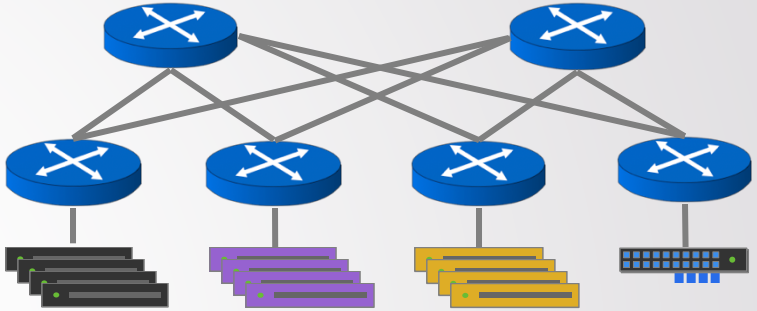
Idea #1: Commodity Switches



Issue: network latency becomes significant (say, at ~500ns per hop)

Idea #2: Cloud-Hosted Trading Systems

Idea: if commodity switches are too slow, can we somehow build a fair network?



Proposals for cloud-hosted exchanges:

- DBO [SIGCOMM ‘23]
- CME, Nasdaq
- Cloud provider equalizes latency

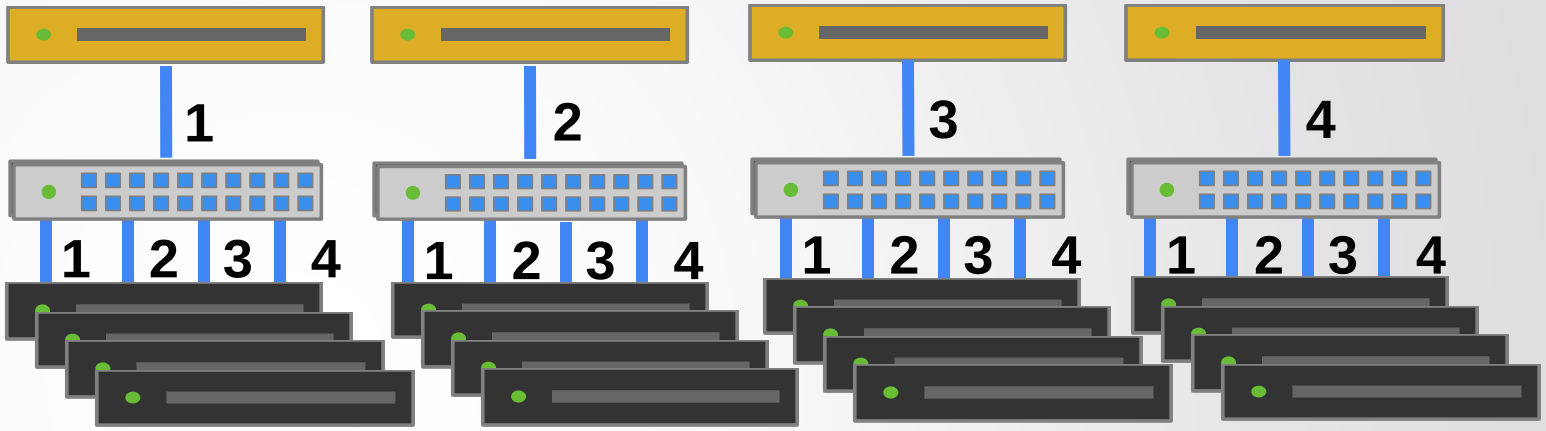
Q: Do these proposals solve the problem?
A: Unfortunately, not yet!

Key Requirements:

- Low-latency connectivity to remote sites
- Broad internal connectivity, including multicast

Idea #3: Layer 1 Switches

Arista 7130
Switch



Layer 1 Switches

- “Crosspoint” switches that can realize arbitrary topologies (including multicast)
- Achieves incredible latency: single-digit nanoseconds! 👍
- But cannot filter, split, merge traffic, so we’d need *lots* of NICs 👎

A Research Agenda for Low-Latency Networks

Hot Take: Over the past few decades, networking research has largely been driven by the needs of hyperscalers (and now AI).

Question: What would we do differently if we took low-latency seriously?

- **Hardware:** can we augment Layer 1 switches with just *slightly* more capabilities to make them more useful?
- **Protocols:** can we co-design protocols and hardware to enable flexible processing while keeping latency low?
- **Routing:** what would routing schemes look like if they took multicast more seriously? Is there a space for content-centric approaches?
- **Cluster Management:** can we control placement to optimize for latency without sacrificing flexibility and scalability

Network Design Considerations for Trading Systems

Andy Myers Brian Nigito Nate Foster