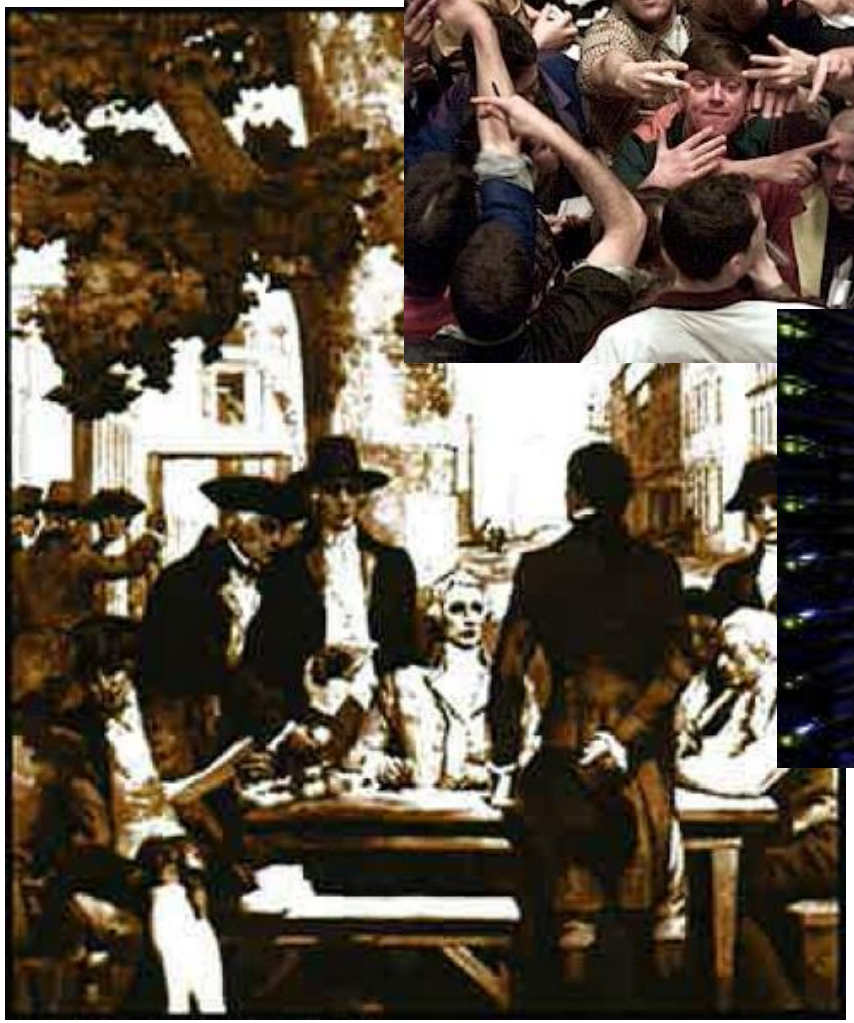
Fast Data & Big Data
Challenges in Electronic Trading

High Speed, Efficient Scalability, and Low TCO
for a Higher Performance Market Data Infrastructure

Nothing new under the sun…

- Shares of Dutch East India Company begin trading in Amsterdam in 1602
- Derivatives (futures) begin trading in 1607
- Market-making (secondary market) begins in 1630s

From trees to pits to data centers



Lexicon: bids, offers, and books

- Trading venues accept orders to buy (bid) and sell (ask or offer) financial instruments
- Trading venues maintain sorted listings of orders (order books) for each financial instrument
- Buy and sell orders that meet at a common price result in an execution (trade)
- Orders at common prices are aggregated into “price books”
- Best bid and offer (BBO) price points comprise a “quote”
- Trading venues distribute order, price, quote, and trade data on “feeds”

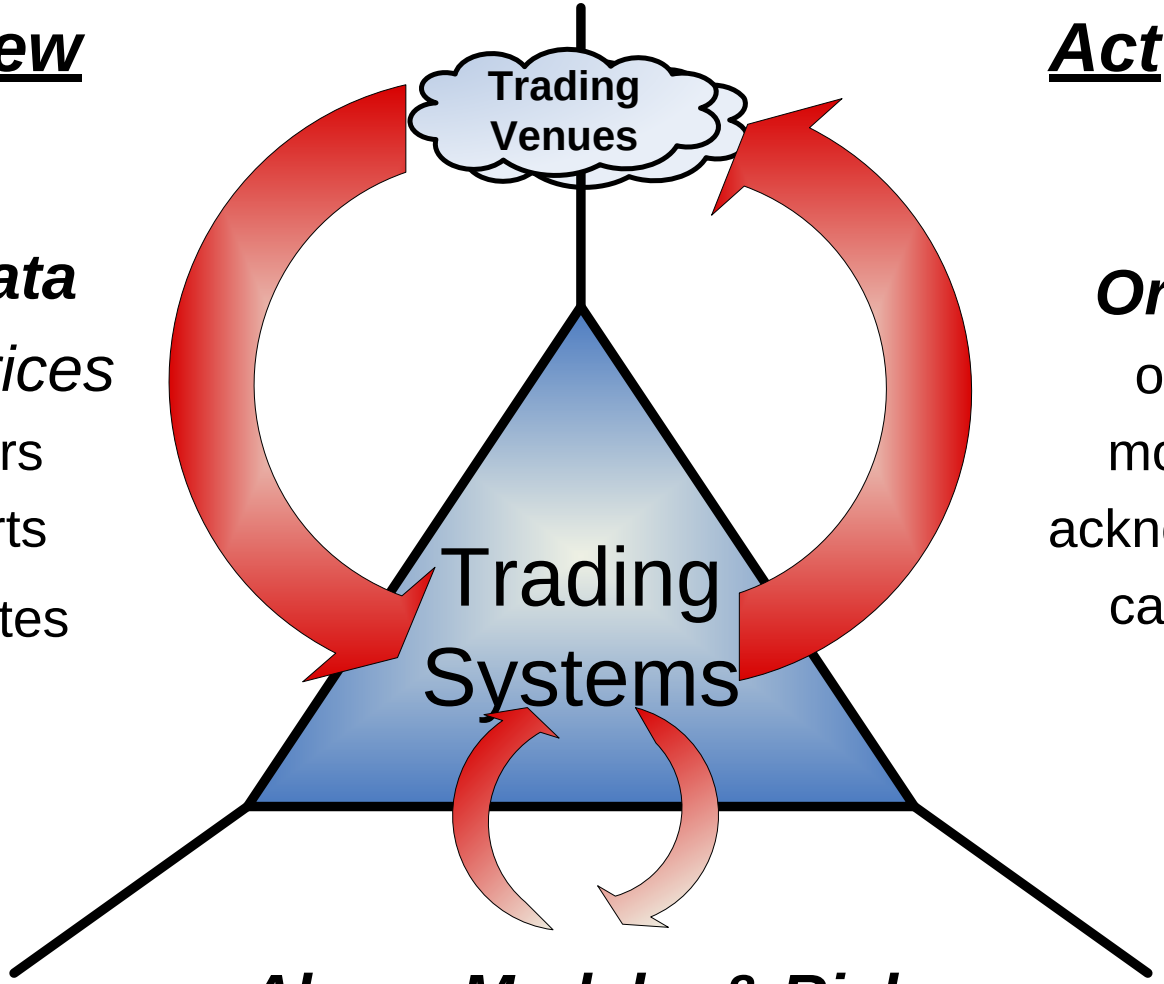
Regional Price Book

Bid		Ask	
54.01	12	54.02	70
53.99	20	54.03	90
53.98	10	54.04	5
53.97	70	54.05	21
53.95	16	54.07	85

Trading Cycle (simplified)

View

Act



Market Data

real-time prices

bids & offers

trade reports

status updates

Order Flow

order entry

modifications

acknowledgement

cancellations

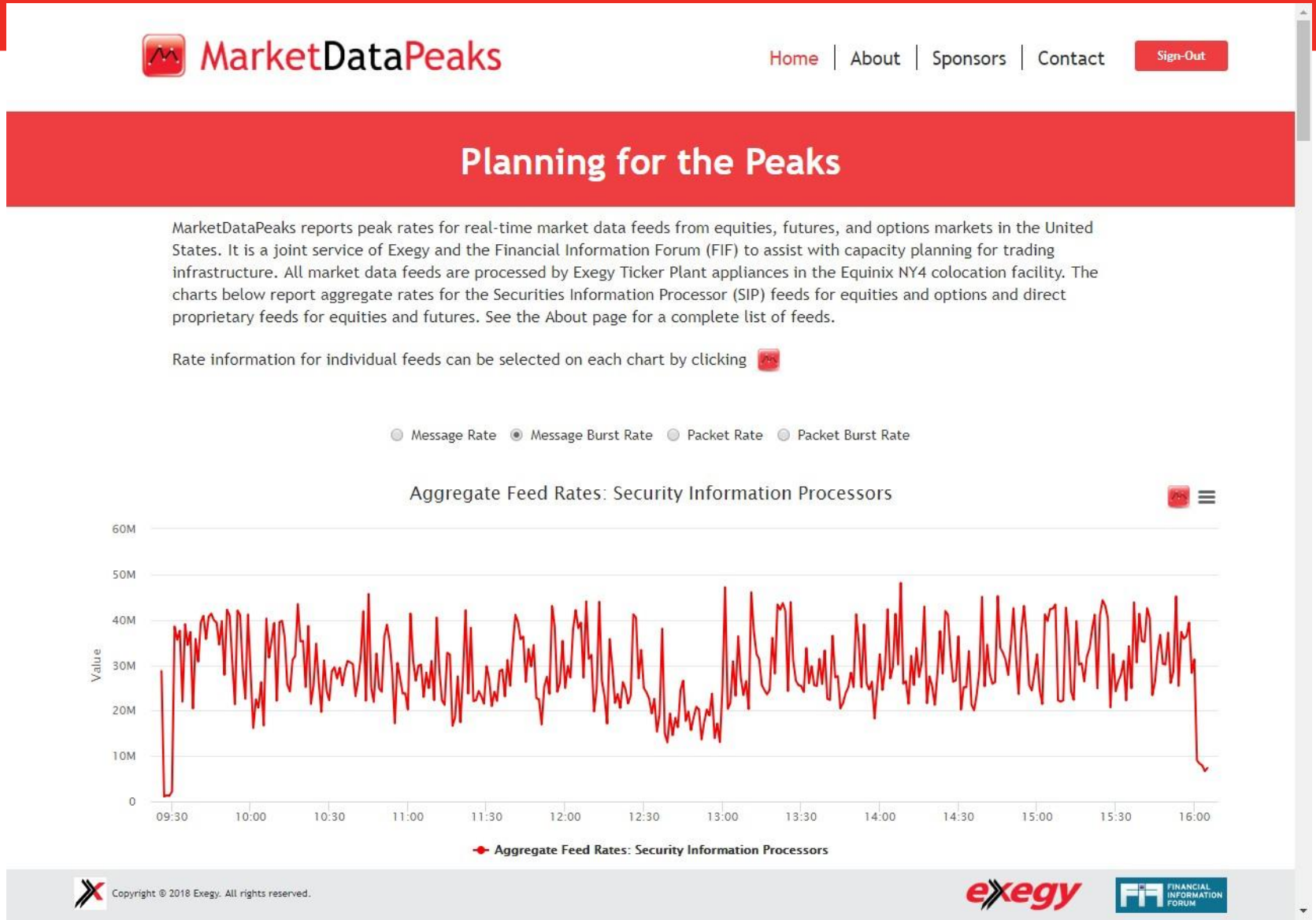
Algos, Models, & Risk

Decide

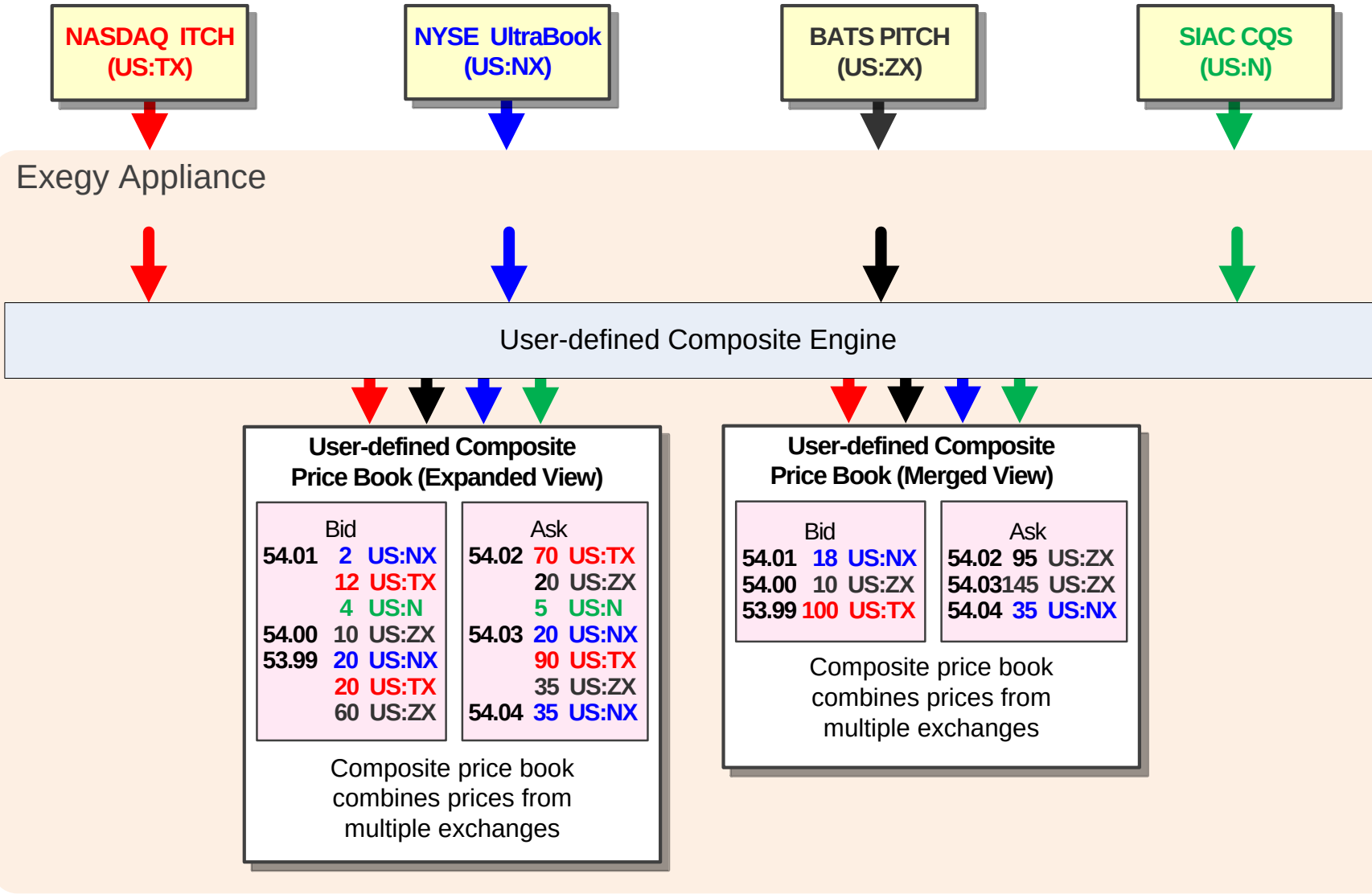
Big Data: Market Scale

- ~8k stocks listed on US exchanges
  - 16+ exchanges → ~132k books to maintain
- ~1M+ stock options listed on US exchanges
  - 14+ exchanges → ~16M books to maintain
- ~1M futures contracts on major commodity markets
- Change in one stock price triggers changes in thousands of derivatives prices
- Rates of dynamic insertions and deletions of orders exceed 40M per second
  - Changes must be reflected in derived pricing views
- Competition among venues precludes standards and fuels constant change
  - Demands maintenance, flexibility, and efficiency

Big Data: Market Data Event Rates



Big Data: Aggregation & Regulation



Big Data: History & Simulation

- Packet capture and storage
  - Terabytes per day of packet capture files
  - Synchronized, high-resolution timestamping
  - High-precision synchronization and replay
- Time series databases
  - Historical record of normalized market events
  - Supports queries to simulate bespoke conditions
    - e.g. Volume-Weighted Average Price (VWAP) of GOOG across three equity markets in the past month
- Complex Event Processing (CEP)
  - Define rules for trading strategy actions
  - Drive from real-time and/or historical data sources

Fast Data: Latency & Location

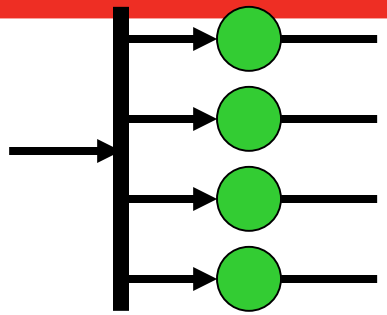
- Tick-to-trade latency goals driving below 10 μs
  - Execute view-decide-act on a single FPGA on the NIC
- Firms co-locate trading infrastructure in same data center as trading venues to minimize speed of light delays
  - Expensive real estate drives computational density
- Venues utilize redundant UDP multicast streams to minimize protocol overhead and maximize fairness
  - Increases bandwidth and receiver computational demands
- Apps require on-demand access to current state of any book (or derived view)
  - Deliver snapshot followed by synchronous stream of real-time updates

Fast Data: Pigeons to Microwaves

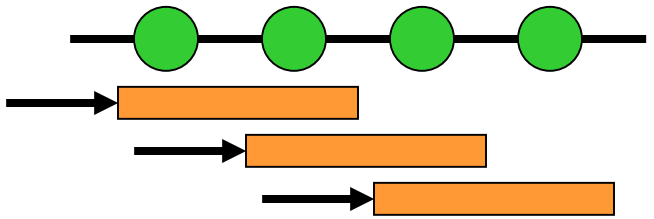
- Paul Reuter uses carrier pigeons to transport stock quotations between London and Paris in the 1850s
- DRW Trading announced plans for 1,050 ft microwave tower in 2016

Purpose-built Machines

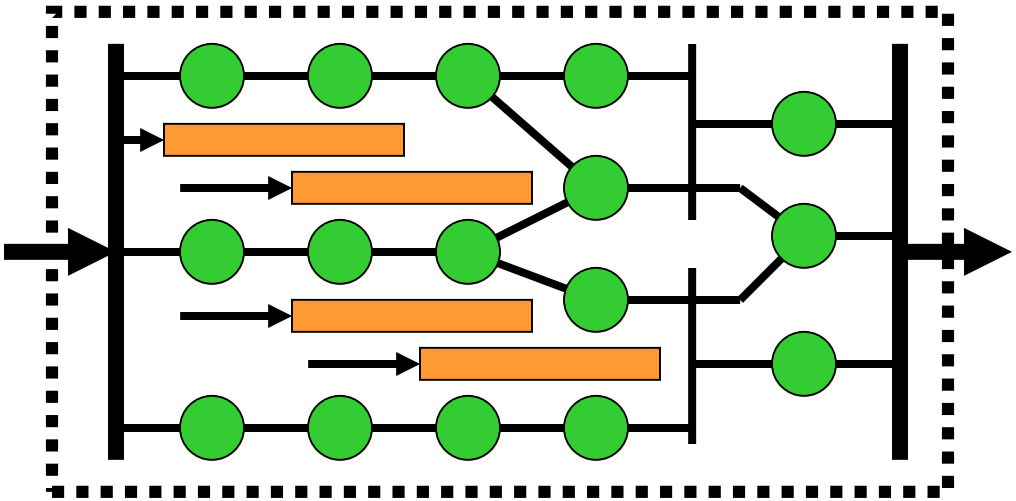
- Data Parallelism
  - Parallelism in the *space* dimension
  - Process independent data streams concurrently



- Functional Parallelism
  - Parallelism in the *time* dimension
  - Process sub-tasks in pipeline stages (assembly line)



- Combine on a single, *configurable* appliance!



Scalable by parallel pipelines of cores and FPGAs

Opportunities to Participate

- Sell-side Investment Banks
  - Morgan Stanley, Goldman Sachs, UBS, RBC
- Buy-side Asset Managers (Hedge Funds)
  - AQR, Tudor, Renaissance
- Retail Funds and Brokerages
  - Vanguard, Fidelity, Scottrade
- Proprietary Trading
  - Two Sigma, DRW Trading, Citadel
- Exchanges
  - Cboe (Bats), ICE (NYSE), NASDAQ, Deutsche Borse
- Vendors
  - Reuters, Bloomberg, FactSet, Exegy

Skills Required for Participation

- “It’s what you know”
  - Computer architecture
  - Networks (all layers, NICs, APIs)
  - High-performance software (C, drivers, kernels)
  - Hardware description languages, FPGA tool flows
- AND “it’s who you know”
  - Invest in your communication skills (writing, speaking, rhetoric)
  - Invest in peer, professor, and professional relationships
  - Integrity and trust matter
  - Leverage your relationships to identify opportunities

Exegy

- Founded in 2003 by four WUSTL faculty and a trustee
- Launched first product in financial services in June 2006 – Exegy Ticker Plant
- 120 employees, most at Saint Louis HQ
  - Sales & services offices in New York, London, Hong Kong
  - BS, MS, and DSc/PhD graduates midwest universities
  - Domain experts from Bridge/Thomson Reuters/Refinitiv
  - ***Significant investment in internship program***
- Serve leading banks, asset managers, traders, exchanges
- Provide managed services, products, and surveilance
- 24x7 operations center manages global estate of 100s of appliances