

Big Data Challenges at the CFTC

By Joanne Morrison

Every day the Commodity Futures Trading Commission monitors some 300 million data records covering a range of market information such as time and sales transactions, futures and swaps positions, intraday trades and swap events. This is just a portion of the multi-terabytes of data the CFTC collects each day and the amount is growing as the agency expands the range and scope of its market surveillance. Added to that are problems with inconsistencies in the data as well as limited staff, experience and resources to manage, store and analyze the information coming in.

In June, the CFTC's Technology Advisory Committee held a meeting to discuss the data challenges facing the agency. CFTC officials outlined the scope of data collected and monitored as part of the agency's daily surveillance operations and discussed issues with data inconsistencies and resource constraints in a post Dodd-Frank world.

CFTC officials also invited comment from industry participants represented on the TAC, many of whom are facing similar challenges in their trading and risk management functions. While many of the industry participants agreed there are challenges in harmonizing data, they urged regulators to look to the surveillance that is already conducted within the industry and avoid duplicating what is already collected and monitored.

CFTC's Data Landscape

Much of the growing data demands at the CFTC are the result of new reporting requirements under the Dodd-Frank Act. In particular, the CFTC must now collect data on over-the-counter swaps from multiple industry sources across multiple markets and in a variety of formats.

Jorge Herrada, the associate director of the CFTC's office of data and technology, explained at the TAC meeting that the amount of data coming into the agency is rising "significantly" but staffing levels have remained relatively flat. "We've got the raw materials for a great surveillance program, but there is still, obviously, lots of work to do," said Herrada, emphasizing that the CFTC is in its "infancy" in terms of ingesting swaps data.

Each time the CFTC adds a new data set, a number of things must come into play, Herrada explained. He compared it to setting up a new production line in a factory. Each new data set requires data standardization, data architecture, custom loaders, data quality checks and monitoring of feedback, he said. "Any time we put new parts online, we have to do a major ramp-up to do that."

Data standardization and harmonization are critical for meaningful analysis, Herrada asserted, adding that even basic data points such as execution time stamps must be harmonized. "The challenge is taking these data sets and trying to merge them together in a meaningful way," he said. He added that the development of standard identifications such as the so-called

legal entity identifier as well as the new ownership and control reporting requirements have helped in viewing across asset classes.

Market Surveillance in Play

Officials from CME Group and Intercontinental Exchange also spoke at the TAC meeting and described the systems they use to maintain and analyze data down to the millisecond.

“The increased innovation and the speed of electronic trading has challenged us to ensure that our markets operate with integrity and are fair and open to all customers,” said Bryan Durkin, CME’s chief operating officer. “We maintain a complete and comprehensive audit trail of every single message, order and trade down to the message ID of every participant in our markets.”

Andrew Vrabel, executive director global investigations at CME, described in detail the exchange’s surveillance capabilities. He cautioned that an effective surveillance operation requires highly experienced investigative staff, systems that manage highly granular and enriched messaging, and significant IT support.

He noted that CME’s investigative staff works very closely with the IT department and other technical business systems analysts to refine regulatory systems and said the development of high-speed trading has changed the CME’s approach to investigations. “It’s different than it was in the past, where we were looking at trading infractions that occurred during the course of 15-minute brackets. Our investigators today are looking at fractions upon fractions of seconds,” he said.

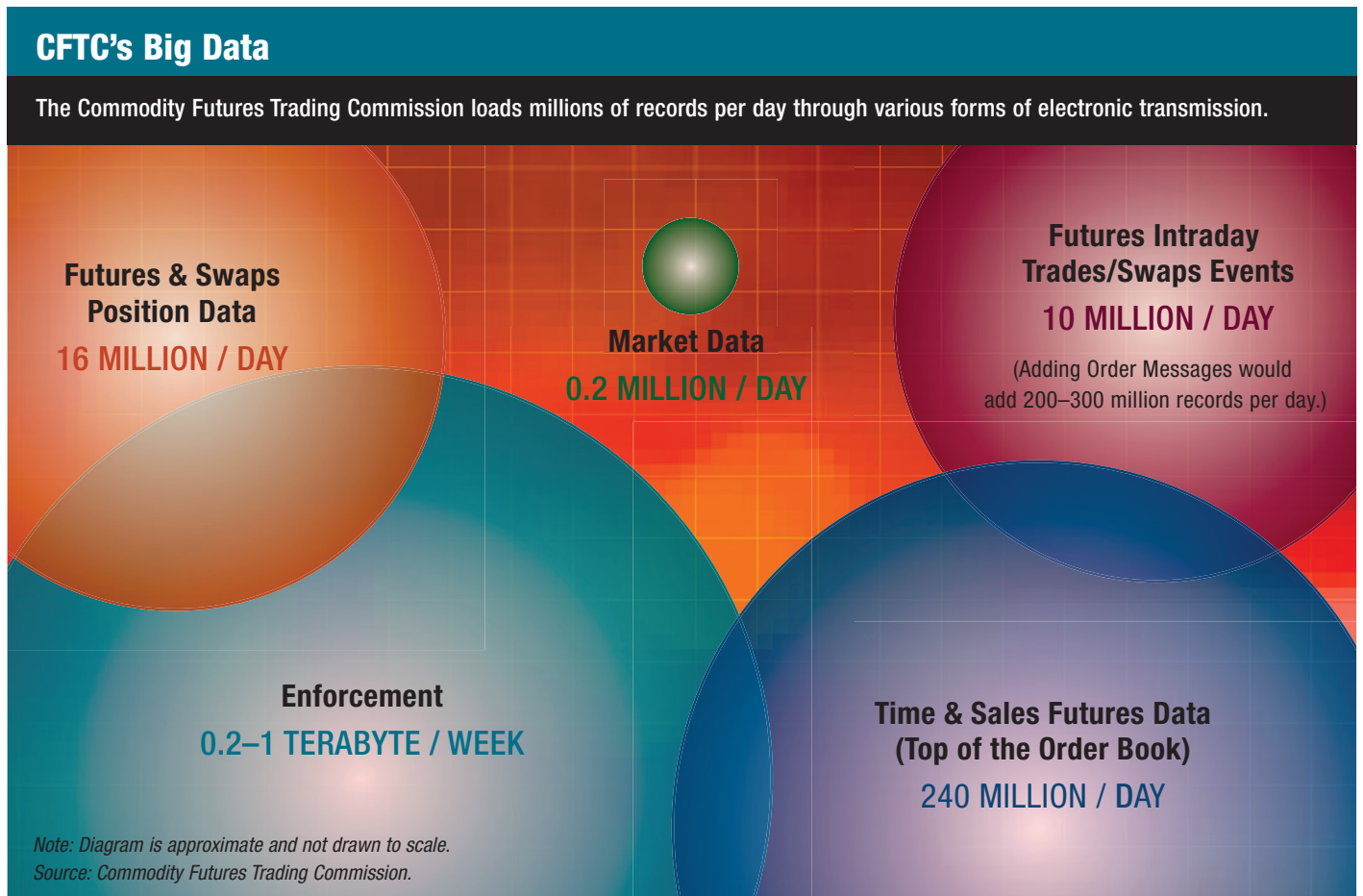
Trabue Bland, vice president of regulation at ICE, highlighted the importance of knowing what specific data is needed and knowing what to look for. “We can build hundreds and hundreds of reports to look for discrete things,” he said, adding that it boils down to two things – is the market performing and is the market healthy.

The surveillance picture is different in the equities market, according to Steven Joachim, executive vice president at the Financial Industry Regulatory Authority, the self-regulatory organization that oversees the equities markets. He told panelists at the TAC meeting that unlike the futures markets, the equities market is more fragmented. Trades in one asset and related products can be conducted across a myriad of platforms in a variety of ways, making surveillance a challenge. “Frankly today, our regulatory community does not treat the marketplace on a broad scale in one view,” he said.

Buried in Data

Panelists at the TAC meeting cautioned the CFTC against collecting more data than is needed and encouraged the agency to concentrate its resources in areas where it would not be duplicating systems already developed by self regulatory organizations.

“You want to make sure that you have a clear process of how you’re going to analyze this and the questions you want to answer,” ICE’s Bland told CFTC officials. “The one recommendation I have for the CFTC is just to make sure you don’t get lost in the data.”



In addition, industry officials urged the CFTC to look to the surveillance systems and data that are already collected in the industry when developing its own surveillance systems. “In the interest of time and cost, I would urge the Commission not to reinvent the wheel,” said Richard Gorelick, chief executive officer of RGM Advisors, a Texas-based trading firm.

Rob Creamer, president and chief executive officer of Geneva Trading and chairman of FIA Principal Traders Group, suggested that the CFTC should take advantage of the model adopted by the Securities and Exchange Commission, which contracted with a technology firm to develop its MIDAS platform for capturing market data. “This should serve as a template or road map for the development of a solution to address market data for futures and swaps,” he said.

Creamer also suggested that the CFTC should concentrate on harmonizing data across trading venues and developing tools for identifying and analyzing patterns in that data. This should be done in a manner that allows for “rigorous, efficient and flexible analysis” but also respects the “confidentiality of the [market] participants whose activity is represented in this data,” he said.

One key challenge in coordinating data is making sure time stamps for transactions are in sync. Clocks “drift” off their times, according to Keith Fishe, managing partner at TradeForecaster Global Markets, a Chicago-based trading firm.

Scott O’Malia, the CFTC commissioner who chaired the TAC meeting, agreed that the agency needs to do more to “leverage” the existing resources. O’Malia, who has since left the CFTC to serve as CEO of the International Swaps and Derivatives Association, also cautioned that very little work has been done to coordinate with international regulators and markets.

“We’re not moving quickly enough [in] setting up international agreements to share that data,” said O’Malia. “There’s no agreement today if we wanted to begin harmonization efforts across jurisdictions,” he said.

At the end of the TAC meeting, CFTC officials requested that panelists submit comments and recommendations on how to improve its surveillance. In August, FIA and FIA Principal Traders Group responded by suggesting that the CFTC should consider five principles as it examines ways to reform its market surveillance and oversight in a “technologically adept way.” The five principles were:

- Continue to delegate “front-line surveillance to exchanges
- Increase the analytical expertise of the CFTC staff
- Use existing reports to enhance cross-market surveillance
- Avoid the duplication of existing systems built or commissioned by exchanges, and
- Maintain data privacy and cybersecurity.

“The Commission already collects a range of valuable data points, and it has access to even more surveillance data through the exchanges,” said Walt Lukken, president and CEO of FIA. “Rather than creating redundant systems, we believe the CFTC would more effectively enhance its surveillance efforts by increasing its analytical capacity and using existing data to test for cross-exchange abuses.” ■■

.....

Joanne Morrison is deputy editor for *Futures Industry* magazine.



Bryan Durkin, chief operating officer of CME Group, discusses the surveillance in place and cautions against duplicating.



Keith Fishe, managing partner at TradeForecaster Global Markets, highlights problems with time stamps.



Rob Creamer, president and chief executive officer of Geneva Trading and chairman of FIA Principal Traders Group, urges regulators to leverage surveillance from existing systems.



CFTC officials hear the challenges in managing data at an all-day meeting of the Technology Advisory Committee.