



# Going with the flow: Are financial firms ready for the data capture challenge?

*In this article, Mike O'Hara and Adam Cox of The Realization Group look at the challenges facing the financial industry as firms prepare for new regulation requiring them to capture, store and provide far more trade-related data than ever before. No longer will transaction data be enough. Regulators want granular data on quotes and timestamps, presenting substantial technological issues in terms of scale and complexity. The new regulatory environment represents a major juncture for an industry that has already dealt with several waves of new rules for trading and regulatory reporting.*

*What's more, the changes will apply to a range of financial market participants, ranging from banks to hedge funds to asset managers. Addressing all of this will involve more than technology; firms will need to think about their operating models, vendor relationships and how much flexibility to build into their systems. Nonetheless, despite the formidable challenges, it's not all doom and gloom. There are cost-effective solutions and benefits to be had as firms put in place new data capture systems. Mike and Adam hear from **Rick Truitt** of Napatech, **Clive Posselt** at Instrumentix, **Stephen Taylor** of Stream Financial, **David Grocott** from Financial Technology Advisers, **Tony Woodhams** at Deloitte and **William Garner** of Charles Russell Speechlys.*



## Introduction

Think about every electronic trade that takes place in world markets. We're not just talking about cash equities or highly liquid exchange-traded futures, but virtually all market segments. Now think about how much data that represents. It's a very large number. Storing and providing all of that data for compliance and regulatory needs is already a gargantuan task. But that's only scratching the surface. Now think about how many quotes may have been made by all the firms in the markets before each of those transactions took place. The number just got exponentially bigger.

Capturing all of the quote data – potentially in real time – with a high degree of granularity is the task set for market participants as MiFID II and MiFIR (the Markets in Financial Instruments Directive II & Regulation) come on stream. Regulators have realised that to combat suspicious trading they need to see far more than the deals that take place. Now they want to see pretty much everything. But monitoring and capturing the data is only part of the problem. Firms will need to make it accessible and be able to deliver it on demand, either for their own forensic analysis or to satisfy regulatory investigations.

The changes are as big as anything trading firms have faced. But advances in technology, particularly in the area of network packet capture and data retrieval, offer companies the means to address the new regulatory requirements. This is not to say that most firms are ahead of the curve. Many are not. But even for those that have yet to start the necessary planning, there is cause for optimism.

### The surveillance era

As if a global crisis had not been bad enough, the financial industry in recent years has been grappling with the fallout from a series of scandals that have generated endless headlines, stiffened the spines of regulators and led to billions of dollars' worth of fines. Welcome to the era of trading surveillance.

*“There is a fundamental change of emphasis in terms of how firms should be operating”*

William Garner, Charles Russell Speechlys



In such an environment, firms have had little choice but to invest in beefing up their network and trade monitoring systems. **Rick Truitt**, Vice President, Business Development at network monitoring provider Napatech, says his firm's larger clients have estimated some two-thirds of their IT budgets were being taken up with monitoring solutions. Developing scalable solutions for monitoring, storing and analysing trade-related data has become a major priority for these firms.

The rise in spend, however, has not always been matched with a rise in the quality of the solutions. Truitt says it all starts with the monitoring part. “If you have a situation where a lot of your data is not being captured and you end up dropping the packets, then it doesn't matter how sophisticated your analytics are,” he says. “You're only as good as the data you bring in.”

*“If you have a situation where a lot of your data is not being captured and you end up dropping the packets, then it doesn't matter how sophisticated your analytics are. You're only as good as the data you bring in”*

Rick Truitt, Napatech



Making the issue of surveillance that much more critical has been a shift on the part of regulators. **Tony Woodhams**, a capital markets special advisor at Deloitte and experienced former trader, says MiFID II and other regulatory measures show that authorities are committed to building a stronger picture of what is happening in markets. “The regulators over the last 5-10 years who were doing some of these investigations recognised the fact that the warning signs aren't just in the trade executions. They're actually also in the orders going in the market that may not be executed. You can read a lot into some of those orders, especially if you triangulate analysis across trades, orders, trader communications and P&L fluctuations.”

If the challenge starts with the data, it certainly does not end there. **William Garner**, a partner at advisory firm Charles Russell Speechlys, says the work financial firms will need to carry out should not be underestimated. “I think there is a fundamental change of emphasis in terms of how firms should be operating,” he says, noting that the change covers both the data policies and the procedures firms have in place.

It all adds up to pressure on financial firms to make sure they build the right systems and adopt the right operating models.

### A plethora of requirements

The size, speed and diversity of financial markets make trade monitoring a potentially tricky proposition. Two immediate factors that complicate the task are the sheer amount of quotes that a market participant makes and the way that a firm is organised.

“It throws an enormous challenge up,” says **Clive Posselt**, Commercial Director at Instrumentix, a firm that works with Napatech to decode and index data that has been captured on networks. “A lot of the analysis that I've seen over my time has been siloed.”

*“There is, certainly from a business perspective, quite a large hurdle to be able to produce something that monitors end-to-end across the whole platform in real time, and provides the functionality that the various RTSs in MiFID II or MiFIR are asking for”*



Clive Posselt, Instrumentix

To get some of the data they need, many firms still employ a technique known as log file scraping, but that will often involve different solutions. That still leaves the question of whether the results will meet regulatory technical standards (RTSs). “There is, certainly from a business perspective, quite a large hurdle to be able to produce something that monitors end-to-end across the whole platform in real time, and provides the functionality that the various RTSs in MiFID II or MiFIR are actually asking for,” Posselt says.

Further complicating the situation is the diversity of markets. **David Grocott**, Director at Financial Technology Advisers (FTA), notes that different markets and asset classes will have varying protocols. That creates a whole new layer of problems with respect to decoding and classifying data once it has been captured. Historically different asset class operations would often have their own technical teams and might even have segregated parts of a network.

Add to that the fact that sizeable amounts of quote-making is done by voice – all of which still needs to be captured – and the requirements grow even more thorny. Some brokers, Deloitte’s Woodhams points out, could be interacting with hundreds of different clients on any given day. The number of quotes could run into the thousands, and each one of them somehow needs to be logged so that it can later be retrieved and analysed. “Clearly that is potentially a big change in terms of the way the firms operate, but that has to be done, because otherwise you’re non-compliant with some of the major regulations.”

The question of data retrieval presents additional issues. Woodhams notes that many firms have now created a post called Chief Data Officer (CDO). “CDOs can never get all of the data they need into one place because of the large numbers of different systems, access issues and the fact some activity isn’t yet apparent in the data stores”, he says. That means they need to prioritise what data to collect and consolidate and consider effective workarounds. They create what is often now termed a ‘data lake’. What makes it problematic is that different compliance tasks will involve different information coming out of the data lake in different structures.

For instance, a firm may need to do reporting for a variety of dissimilar issues, from trader misconduct to financial crime to money laundering.

*“Chief Data Officers can never get all of the data they need into one place because of the large numbers of different systems, access issues and the fact some activity isn’t yet apparent in the data stores”*



Tony Woodhams, Deloitte

On top of all that, some CDOs will have a background in finance but many may come from other industries. “The point is they need a lot of guidance in terms of what data they’re collecting, in respect of what problems (use cases) they’re looking to solve for, whilst consolidating data and originating these data lakes,” he says.

### A three-step process

Given such diverse and complicated requirements, how should firms approach the task of building robust solutions that can sort and store so much information? “How you do that, and how you tie all of the information into a readily-usable format is obviously the key to this, but that can be slightly mind blowing in terms of what has to be done,” says Garner of Charles Russell Speechlys.

Truitt of Napatech says firms can try to build instead of buy, but the results may not make the grade. “When you individualise the components and try to put them together, that can present firms with some real challenges,” he says.

The solution his firm recommends involves a three-step process. The first step is to capture data on the network. Grocott of FTA notes that realistically that will involve multiple devices in different locations.

Such network packet capture was not borne out of regulatory requirements but was based on the industry’s desire to get better information when monitoring their flow. For instance, firms engaged in the latency “race to zero” needed to understand exactly where their latencies were coming from. In order to do so, a new discipline – Flow Monitoring – arose, which made use of the network capture technology. Apart from revealing the causes of latency, such Flow Monitoring systems also provided a level of visibility that trading desks demanded.

But the prospect of comprehensive flow monitoring at a cross-market level and being able to create a holistic view is a relatively new phenomenon. Previously, only a limited number of people outside of the network team, for example Exchange Connectivity specialists, would work with packet capture, whereas other parts of an organisation might work with log files and the trading desk would work with whatever data it had. All of that information then later would be stitched together as well as possible. “It wasn’t real-time monitoring and it was really painful to do,” Grocott says.

*“ Without the advances in capture technology, we could never have given rise to this whole new discipline of these types of monitoring systems and surveillance systems, because they couldn’t exist”*



David Grocott, Financial Technology Advisers

This is where technological progress became absolutely critical. “Without the advances in capture technology, we could never have given rise to this whole new discipline of these types of monitoring systems and surveillance systems, because they couldn’t exist,” says Grocott. “For the first time, with the use of flow monitoring, you were able to do this kind of thing in real time.”

The next step, in order to get to that holistic real-time view, involves processing the network traffic, in real time or using captures, to translate the traffic into usable data. This process relies on the use of “decoders”, that understand the myriad protocols being spoken by the systems involved in the trading lifecycle. These decoders are, more often than not, implemented as part of a Flow Monitoring system and usually represent the border between network capture and Flow Monitoring. Once the traffic has been decoded to expose the business objects such as orders, quotes and ticks, these objects must be correlated to provide any true value, for example to deliver latency measurement. These related processes of decode and correlation together provide the raw data necessary for performance and behavioural analysis to take place.

Finally, there is the question of analysing the data. **Stephen Taylor**, CEO of Stream Financial, says that provided a firm has zero packet loss and can process the network traffic, it should be in a position to do almost any kind of data analysis to satisfy compliance needs. “This means you can run queries across multiple trading venues and join it with data from other systems. These could be key systems like KYC systems, compliance systems, HR systems, and just about anything else you like,” he says.

The advent of federated data query technology takes some of the sting out of the regulatory requirements because, according to Stream's Taylor, it provides the flexibility and performance firms need. "I don't think it's been possible in the past to make the federated solutions perform well enough, whereas it is now. And it's our belief that a federated solution is the only thing that can deal with the rate of change around the organisation."

*"I don't think it's been possible in the past to make the federated solutions perform well enough, whereas it is now. And it's our belief that a federated solution is the only thing that can deal with the rate of change around the organisation"*



Stephen Taylor, Stream Financial

To arrive at the scenario that Taylor describes, where reports can be generated based on data across a wide spectrum of sources, requires more than technology. What firms need to do is ensure their different departments get out of a silo-based mentality. "Banks need to be joining the dots much more effectively across these different yet inter-related departments," says Woodhams of Deloitte.

This is beginning to happen, he adds. "We're seeing some firms starting to join dots across areas of the bank that were historically disconnected".

For instance, alerts could be triggered from email and instant messenger surveillance and reporting systems. The Deloitte consultant sums it up by voicing the questions trading firms may then start to ask themselves: "What trades are being placed in what markets, what does the pattern of that trading look like, and are we worried about that when we see it as a whole? In many cases, if the question is, 'Are there any traders at this bank, or related to this bank, or clients of this bank that we should be worried about?', you only get the right answer if you start joining these things together and reading across related signals. And this has been where many firms have repeatedly failed in the view of regulators.."

### **Next steps**

The net that regulators have cast in terms of the new requirements is wide, encompassing some market participants that are not used to addressing issues such as trade surveillance. That has also led, according to some of the experts, to a degree of uncertainty as to where responsibilities lie. While Woodhams says that firms that have begun to prepare for the new requirements are making some progress, Garner cautions. "I suspect that there is not going to be a uniform approach to how you go about capturing all of this data and ensuring that it's easily and quickly accessible," he says. Posselt adds that the buy side in particular is likely to be playing catch-up.

But the Instrumentix executive says that investing in the appropriate technology will be worth it. "If you need to really prove that you've taken all sufficient steps, I would very much argue that you need to be able to press the button and say, 'That's exactly what's going on in my platform at any one point in time', in order to do the correct analysis." Taylor of Stream Financial also stresses the benefits that come with being able to make ad-hoc queries based on whatever threats or issues become apparent.

Grocott of FTA likens such a solution to an extra level of security. "You never know if somebody's by-passed pre-trade risk checks in some way or if you've got some sort of systems malfunction that you haven't spotted. Having a consolidated approach to network data capture and processing gives you another check on the accuracy of the data coming in and out of the bank in one place."

A final benefit from working with vendors on a system that is asset-class agnostic is cost. "When you don't get into proprietary ways of capturing traffic or analysing traffic, it gives a lot more ability to scale across multiple groups within one company," Napatech's Truitt says. "They can see the data, use the data, munch or crunch the data that they need internally."

For an industry long used to developing proprietary technology, the message here is: go with the flow. For capturing, processing, indexing and retrieving data, there is little to be gained from reinventing the wheel. In other words, save the unique proprietary technology for other parts of the trade cycle.

**Zero loss capture and flow monitoring: Some technology factors to consider**

The development of improved packet capture & processing has taken flow monitoring to a new level. Following are some of the major technological factors that are making it possible to address the new regulatory requirements.

**PCAP** – PCAP is the industry acronym for packet capture technology. Posselt says the critical issue with PCAP technology is not the openness of the technology but the degree to which PCAP APIs are integrated with other vendor technologies, particularly Flow Monitoring systems

**FPGA** – Field programmable gate array technology has been around for a long time but it became critical for some market participants with the boom of algorithmic trading. The key value of FPGA is its ability to process large amounts of data at very high speeds in a deterministic fashion. FPGA technology can help in particular with highly accurate timestamping.

**Open source** – By focusing a trade monitoring system on data capture, processing and standards-based storage and retrieval rather than analysis, firms are free to consider open source analytics. Posselt says that nowadays all investment banks are using open-source technologies. “We’ve seen a change in the way that they’re prepared to deal with open-source technologies. They’re prepared to deploy them and they’re prepared to integrate them.”

**Capture fabrics** – As colocation took off, firms began deploying what Posselt refers to as “capture fabrics”, composed of network taps, aggregation switches and network packet brokers(NPBs). Originally used for network diagnosis, these infrastructure components facilitated easy, asset-class agnostic access to massive amounts of data being gathered from many different points in the network. More recently, however, capture fabrics are being deployed not only in colocation facilities, but also in the more traditional Campus datacentres

**Data Virtualisation** – As data volumes grow, the challenges of managing centralised data copies are demanding a new approach to cross system queries. Soft schema query engines and high performance data compression technologies have created the possibility of federated data governance. “This is not a new idea” says Taylor “but implementing a system with sufficient performance requires solutions to all of the limitations in the legacy estate.”

**For more information on the companies mentioned in this article, visit:**

[www.napatech.com](http://www.napatech.com)  
[www.stream-financial.com](http://www.stream-financial.com)  
[www.deloitte.com](http://www.deloitte.com)

[www.instrumentix.co.uk](http://www.instrumentix.co.uk)  
[www.fintech-advisers.com](http://www.fintech-advisers.com)  
[www.charlesrussellspeechlys.com](http://www.charlesrussellspeechlys.com)



**Financial Markets Insights** from The Realization Group, is a series of interviews with thought leaders in financial and capital markets. The purpose of the series is to provide exclusive insights into industry developments, through in-depth conversations with C-level executives and key experts from banks, exchanges, vendors and other firms within the financial markets ecosystem. For more information, please visit [www.financialmarketsinsights.com](http://www.financialmarketsinsights.com)



**Other topics in the series:**

Market surveillance -

Getting to grips with the new Market Abuse Regulation [DOWNLOAD](#)

Voice collaboration - what is the future in Financial Services? [DOWNLOAD](#)

Margin compression in interest rate derivatives - The big squeeze [DOWNLOAD](#)

New initiatives in compliance technology [DOWNLOAD](#)

Financial Applications Hub - A high-performance app store for markets [DOWNLOAD](#)

Digital transformation - when business & technology go hand in hand [DOWNLOAD](#)

Emerging markets trading hubs -

Gateways to new venues, products and markets [DOWNLOAD](#)

Making waves: Breaking down barriers with ultra-fast networks [DOWNLOAD](#)

Cleared for launch: A new era for OTC derivatives [DOWNLOAD](#)

Risk & regulatory reporting: centralisation is not the answer [DOWNLOAD](#)

Fixing fixed income - taking a leaf out of social media's playbook [DOWNLOAD](#)