Seizing cloud opportunities
The consolidated audit trail
Background

The consolidated audit trail (CAT) is here. And, now is the time to act. Several compliance deadlines are on the horizon, and on November 15, 2018, large US broker-dealers will be required to start reporting to the central repository. Given the level of effort and costs involved in standing up a new non-financial regulatory reporting program, individual firm timelines are even shorter. There are important implications for front- to back-office operations, controls, technology strategy, data architecture, and analytics capabilities.

Historically, competing priorities and large data requirements led many firms to build separate data stores for each repository resulting in increased costs and inefficiencies. Unless firms start planning now, costly and inefficient practices will likely continue as experienced during previous regulatory mandates, such as the swaps data repository (SDR) from the Commodity Futures Trading Commission (CFTC) in the US and the trade repository (TR) from the European Market Infrastructure Regulation (EMIR) framework in Europe. The fundamental choices integral to implementing a solution for CAT will likely have a lasting impact on a firm’s capabilities, such as utilizing more sophisticated trading analytics and streamlining execution validation. Creating a strategic technology plan now can help you lead in your industry, navigate compliance risks, seize broader opportunities presented by CAT, and disrupt through innovation.

The creation of the CAT concept was in response to the events of the “Flash Crash” in May of 2010. The Securities and Exchange Commission (SEC) wanted to enhance their capabilities to reconstruct the trading events of the day in order to determine the cause or causes of market disruptions. In that effort, the SEC passed Rule 613: Consolidated Audit Trail, of which the basic requirements for US broker-dealers is to report every equity and options trade lifecycle event. When completed, the CAT will be the world’s largest data repository of securities transactions. All US broker-dealers are required to implement CAT-compliant reporting protocols into a new and/or existing reporting program by November 2018 for large broker-dealers and November 2019 for small broker-dealers.

The creation of the CAT concept was in response to the events of the “Flash Crash” in May of 2010. The Securities and Exchange Commission (SEC) wanted to enhance their capabilities to reconstruct the trading events of the day in order to determine the cause or causes of market disruptions. In that effort, the SEC passed Rule 613: Consolidated Audit Trail, of which the basic requirements for US broker-dealers is to report every equity and options trade lifecycle event. When completed, the CAT will be the world’s largest data repository of securities transactions. All US broker-dealers are required to implement CAT-compliant reporting protocols into a new and/or existing reporting program by November 2018 for large broker-dealers and November 2019 for small broker-dealers.

The securities industry has seen increases in the volume and complexity of data, data storage, and data management. Many firms are working to take advantage of the insights available from these data sets using advanced analytics tools. However, the costs and complexity of managing the enabling technologies have been difficult for internal resources and infrastructure. Cloud technologies, such as Infrastructure as a Service (IaaS) and Platform as a Service (PaaS), enable companies to take advantage of the cost efficiency, scale, elasticity, automation, and orchestration that both private and public cloud infrastructure platforms provide. With the massive data requirements of equity and options reporting to CAT, firms should consider embracing new opportunities in technology such as data management practices and cloud-based infrastructure and platform provisioning, which can allow the strategic use of data for competitive advantages.
Designing for the future: Data management, cloud infrastructure, and CAT

When establishing the next-generation solution architecture to support CAT requirements, the following architecture and solution design principles are fundamental to an effective solution.

Agility and scalability

- Create agility and scalability for compute and storage demands by utilizing cloud-based infrastructure platforms, whether on premise or publicly cloud-hosted. As the pace of regulatory change is increasing, it's important for firms to be proactive. Agility at an organizational level is going to be an important component of successful implementations. Current and future demands for CAT and other regulatory reporting will require an ability to scale up or scale down the technology stack as needed to anticipate market conditions and peak volumes.
- Enhancing resiliency through application-aware platform scaling and performance monitoring can offer a solution to improve availability and responsiveness as data and usage grows.
- Review both private and public cloud service platforms and consider their consistency with the technology and data architecture of CAT requirements for automation, standardization of data feeds, and information security demands.

Flexibility

- Establish a flexible data and overall technology architecture to meet evolving demands of business and external regulatory entities, including functional and data attribute requirements. As we’ve seen with other non-financial trading programs, requirements change over time, and it’s critical that the system architecture and data repositories are designed to anticipate change.

Cost management

- Reduce sunk infrastructure costs by utilizing cloud’s metered cost model to scale up/scale down data storage requirements and computing power.
- Reduce IT personnel costs associated with on-premise infrastructure management.

Security

- Incorporate security and protection in the reporting solution by encrypting data in motion and at rest from existing systems of record through to the aggregation and report generation layer.
- Integrate security into the foundation of the solution rather than relying solely on perimeter security around the application and data environment.

Adaptability

- Consider integration requirements from both legacy applications and systems of record (SORs) on premise, as well as net new integration requirements as part of any greenfield CAT build-out and reporting requirements.
- Design a robust abstraction layer capable of decoding highly normalized data from source systems and insulating data flow from changes in the System of Origin (SOO)/System of Record (SOR).

Accuracy, consistency, and completeness

- Ensure accuracy and consistency of information through management and self-service reporting, which could be realized through in-memory architectures and innovative caching to deliver extreme performance against voluminous data.
- Consider how to keep pace with the broader changes in the market from the perspective of completeness of reportable data. CAT reporters will need to keep pace with the increased demands and sophistication of the regulators.
Driving strategic opportunities

CAT requirements provide firms with the opportunity to enhance or transform their non-financial regulatory reporting processes. Since CAT reporters will need to understand the sources and accuracy of their regulatory submissions to CAT, there’s an opportunity to determine how best to incorporate or replace legacy reporting systems. The timing of the retirement of duplicative systems is still a work in progress. It is likely that, in addition to CAT reporting, organizations will be required to operate and manage duplicative regulatory reporting systems such as Order Audit Trail System (OATS), Electronic Blue Sheets (EBS) for some time frame. The overhead of managing multiple disparate reporting repositories while ensuring a high level of data quality, reliability, and consistency across all these reporting systems significantly increases complexity and costs of regulatory reporting.

Figure 1: Conceptual plan processor framework

New generation of reporting
Thesys solution will not be based on existing industry reporting protocols (e.g., OATS)—reporters need to source and report on equities options and customer information.

Tech specs
Thesys will accept a wide variety of existing message formats to minimize new requirements.

Self-support model
Thesys will leverage self-supporting tools, where possible, with functions to assist CAT reporters with testing, onboarding, dashboard for reporting performance and customer support.

Cyber security
Security and protection will be a critical consideration in the reporting solution by encrypting data in flight and at rest. Firms need to protect the aggregation and report generation layer.

Analytics
Thesys has deep experience in big data analytics—this can be expected to flow through into SEC and SRO surveillances.

Thesys partners
Thesys will leverage the IBM Cloud and services to provide CAT Reporters with support services such as help desk.
Seizing cloud opportunities

The consolidated audit trail

The consolidated audit trail

The consolidated audit trail

Seizing cloud opportunities

Thesys Technologies, LLC, a subsidiary of Tradeworx, Inc., was selected by the self-regulatory organizations (SROs) as CAT plan processor to build and operate the CAT repository. As such, the Thesys bid highlighted their experience with big data, trade analytics, and high-performance computing. The SROs will be able to leverage this level of technology experience to enhance their regulatory program, market surveillance, fraud analysis, market structure analysis, and order/trade anomaly detection. More than just duplicative compliance reporting, firms will be pressured to become more data-centric and adept in utilizing large amounts of data to uncover deeper insights into trading behavior. Computing power and efficient data processing are a necessity for the next generation of trading surveillance and fraud detection algorithms that often incorporate a heavy dose of matrix operations, machine learning, and time-series analysis. Effectively applying these advanced techniques requires firms to build computing infrastructure across multiple environments including development, testing, and production. These computing-intensive algorithms rely on an advanced heterogeneous technology stack across a distributed server farm that is foundational for large-scale data processing.

Firms across multiple industries are taking advantage of cloud technologies to reduce their costs related to staffing, training, maintenance, and infrastructure by taking advantage of cloud-based infrastructure and application services. One of the impediments to the use of cloud technologies for broker dealers has been the uncertainty of regulatory approval for storage of data in the cloud. Since the data will reside in the cloud once reported, broker-dealers now have the opportunity to take advantage of cloud-based architectures. Moving data to the cloud creates a tremendous amount of value in terms of tools, infrastructure, and flexibility and allow for the rapid deployment and metered cost structure that make meeting regulatory requirements such as CAT faster and more cost efficient.

In addition to regulatory mandates, trade analytics and surveillance can take advantage of a flexible big compute, leading edge data platform that enables experimentation with and use of more advance algorithms. These techniques usually require cost-prohibitive hardware-based acceleration technologies such as graphics processing units (GPUs), field programmable gate arrays (FPGAs), and distributed architectures. They also rely on the ability to store massive amounts of data, such as current and historical options and equities lifecycle events. A cloud-based, on-demand high-performance architecture helps efficiently transform, store, and analyze all data.

Figure 2: Broker-dealers profile from the CAT perspective

Developing Data & Cloud Strategies Based on Firm Characteristics

<table>
<thead>
<tr>
<th>Firm size</th>
<th>Trading systems</th>
<th>Data repositories</th>
<th>Current trade reporting model</th>
<th>Current and emerging infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Option A</strong></td>
<td>• Single trading platform</td>
<td>• 1-2 data repositories</td>
<td>• Out of the box reporting</td>
<td>• Standard infrastructure</td>
</tr>
<tr>
<td></td>
<td>• Limited order Entry</td>
<td>• Vendor-based platform</td>
<td>• Limited or no custom reporting</td>
<td>• Limited internal support</td>
</tr>
<tr>
<td></td>
<td>• Low trade volume</td>
<td>• Very limited proprietary systems</td>
<td>• Limited vendor partnership</td>
<td>• Limited/no use of cloud technologies</td>
</tr>
<tr>
<td><strong>Option B</strong></td>
<td>• 3-5 trading platforms</td>
<td>• 2-3 data repositories</td>
<td>• Modified vendor solution</td>
<td>• Combination of firm-owned and outsourced infrastructure platforms</td>
</tr>
<tr>
<td></td>
<td>• Multiple order entry systems</td>
<td>• Combination of proprietary technology and Vendor provided solutions</td>
<td>• Some customization on existing vendor platforms</td>
<td>• Pilot cloud infrastructure project in flight</td>
</tr>
<tr>
<td></td>
<td>• Moderate trade volume</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Option C</strong></td>
<td>• &gt;5 trading platforms</td>
<td>• 3-5 data repositories</td>
<td>• Bespoke trade reporting</td>
<td>• Complex, global infrastructure</td>
</tr>
<tr>
<td></td>
<td>• Complex order entry architecture</td>
<td>• Well-developed proprietary technology architecture</td>
<td>• Independent trade reporting team</td>
<td>• Well-developed infrastructure strategy</td>
</tr>
<tr>
<td></td>
<td>• High trade volume</td>
<td></td>
<td>• Multiple vendor partnerships</td>
<td>• Active cloud migration program</td>
</tr>
</tbody>
</table>
As depicted in Figure 2 above, capital markets firms are at different maturity and complexity levels with regard to their trade lifecycle and historical market data technologies. Firms should evaluate their environment in order to align their use of data and cloud technologies with their unique environments. For firms of all sizes, operating a highly proficient and motivated technology team has serious implications for cost, keeping current with evolving skills and technology, and adding business value. The maintenance of a high-performance, data-intensive infrastructure requires teamwork between system engineers, network engineers, data engineers, technology architects, and developers. This is not just an operations or a software engineering problem but a serious cost and time-to-market issue. Cloud technologies can help reduce these issues by moving these costs to the cloud provider.

Firms can also take advantage of the convergence of cloud technologies and big data solution stacks. Harnessing big data technology offerings combined in a PaaS cloud model provides a unique set of capabilities for a CAT implementation. These capabilities align well with the suggested CAT implementation design principles. But there are risks associated with the convergence of these two technologies which requires close management. For example, concerns around storing data in cloud need to be addressed. Such concerns are being actively worked on by cloud service providers’ capabilities. More importantly, policies and controls around data management and information security should be instituted by firms to manage the risk around these recent technologies.

CAT will require reporting Social Security numbers as well as other personally identifiable information (PII). Processing and storing PII data in the cloud has been a significant industry concern. It requires proper handling and care and will require the use of encryption in flight and data at rest. Data masking and user access control are also very important considerations: Cloud providers are leading the way in research and development in data privacy and data protection algorithms like multi-party computation (MPC) which allows firms to perform computational tasks without ever having to decrypt the data or managing keys. Because the reputational risk can destroy their business model, securing data is a primary objective for most cloud providers.

**Analytics and surveillance**

CAT will provide regulators with a complete end-to-end picture of a broker-dealer’s equities and options trading activities for the first time. The plan processor has expertise with advanced analytical functionality allowing the regulators to surveil this activity. CAT provides an opportunity for firms to strengthen their own internal surveillance programs—before their regulators understand their activities better than they do. To avoid falling behind regulators, broker-dealers should embrace this as a call to action. CAT is an opportunity to mature internal surveillance and analytics programs and further enhance current surveillance practices.

**Conclusion**

The onslaught of regulatory requirements in the post-Dodd-Frank era has strained the capacity of broker-dealers to comply with the seemingly never-ending demands. Complying with CAT requirements provides broker-dealers an opportunity to take advantage of cloud and big data technologies to gain strategic value. It’s also important to remember that CAT will provide regulators, for the first time, a complete end-to-end picture of a broker-dealer’s equities and options trading activities. The plan processor has expertise with advanced analytical functionality allowing the regulators to surveil this activity.

Disruptive regulatory burdens and complexity are only going to increase. And in order to lead in their industries, organizations should evaluate both their tactical and strategic plans in navigating the CAT requirements. Those firms that are in the position to use the CAT initiative as an opportunity to cut cost, decrease complexity, and advance their data management and analytics strategy should evaluate the cloud as a means to accomplish these goals.