APPENDIX D

CAT NMS Plan Processor Requirements

Appendix D, CAT NMS Plan Processor Requirements, outlines minimum functional and technical requirements established by the Participants of the CAT NMS Plan for the Plan Processor. Given the technical nature of many of these requirements, it is anticipated, as technology evolves, that some may change over time. The Participants recognize that effective oversight of, and a collaborative working relationship with, the Plan Processor will be critical to ensure the CAT achieves its intended purpose, namely enhanced investor protection, in an efficient and cost-effective manner. The Participants also recognize that maintaining the efficiency and cost-effectiveness of the CAT requires flexibility to respond to technological innovations and market changes. For example, these minimum functional and technical requirements allow the Plan Processor flexibility to make certain changes to the Technical Specifications, while limiting others to the Operating Committee, and anticipate agreement between the Operating Committee and the Plan Processor on SLAs relating to, among other things, development, change management, and implementation processes and timelines. Maintaining such flexibility to adapt in these and other areas relating to the development and operation of the CAT is a foundational principle of this Appendix D.

1. Central Repository Requirements

1.1 Technical Architecture Requirements

The Central Repository must be designed and sized to ingest, process, and store large volumes of data. The technical infrastructure needs to be scalable, adaptable to new requirements and operable within a rigorous processing and control environment. As a result, the technical infrastructure will require an environment with significant throughput capabilities, advanced data management services and robust processing architecture.

The technical architecture must be scalable and able to readily expand its capacity to process significant increases in data volumes beyond the baseline capacity. The baseline capacity requirements are defined in this document. Once the CAT NMS Plan is approved, the Operating Committee will define the baseline metrics on an ongoing basis. CAT capacity planning must include SIP, OPRA and exchange capacity and growth forecasts. The initial baseline capacity requirements will be based on twice (2X) the historical peaks for the most recent six years, and the Plan Processor must be prepared to handle peaks in volume that could exceed this baseline for short periods. The SLA(s) will outline details of the technical performance and scalability requirements, and will be specifically targeted to the selected Bidder's solution.

The Central Repository must have the capacity and capability to:

- Ingest and process throughput to meet baseline capacity requirements as well as scalability to meet peak capacity requirements, including staging, loading, speed of processing, and linking of data;

- Accommodate data storage and query compute, such as:

- Scalable for growth data storage and expansion capability, including but not limited to, resizing of database(s), data redistribution across nodes, and resizing of network bandwidth;
- Robust processes to seamlessly add capacity without affecting the online operation and performance of the CAT System; and
- Quantitative methods for measuring, monitoring, and reporting of excess capacity of the solution;

- Satisfy minimum processing standards as described in the CAT RFP and that will be further defined in the SLA(s);

- Adapt to support future technology developments and new requirements (including considerations for anticipated/potential changes to applicable rules and market behavior);

- Handle an extensible architecture that is capable of supporting asset classes beyond the initial scope of NMS Securities and OTC Equity Securities;

- Comply with the clock synchronization standards as set forth in Article VI, Section 6.8; and

- Handle an extensible data model and messaging protocols that are able to support future requirements such as, but not limited to:

  - Expansion of trading hours, including capability and support for 24-hour markets;
  - Sessions for securities;[262] and
  - New asset classes, such as debt securities or derivative instruments.

1.2 Technical Environments

The architecture must include environments for production, development, quality assurance testing, disaster recovery, industry-wide coordinated testing, and individual on-going CAT Reporter testing. The building and introduction of environments available to CAT Reporters may be phased in to align with the following agreed upon implementation milestones:

- Development environment – the development environment must be created to build, develop, and maintain enhancements and new requirements. This environment must be separate from those listed below.

[262] Equity markets currently have morning, primary, and evening sessions. It is possible that over time sessions may cross into the next calendar day.

- Quality assurance environment – a quality assurance (QA) environment must be created to allow simulation and testing of all applications, interfaces, and data integration points contained in the CAT System.

    - The QA environment shall be able to simulate end-to-end production functionality and perform with the same operational characteristics, including processing speed, as the production environment.
    - The QA environment shall support varied types of changes, such as, but not limited to, the following:

        - Application patches;
        - Bug fixes;
        - Operating system upgrades;
        - Introduction of new hardware or software components;
        - New functionality;
        - Network changes;
        - Regression testing of existing functionality;
        - Stress or load testing (simulation of production-level usage); and
        - Recovery and failover.

    - A comprehensive test plan for each build and subsequent releases must be documented.

- Production environment – fully operational environment that supports receipt, ingestion, processing and storage of CAT Data. Backup/disaster recovery components must be included as part of the production environment.

- Industry test environment –

    - The Plan Processor must provide an environment supporting industry testing (test environment) that is functionally equivalent to the production environment, including:

        - End-to-end functionality (e.g., data validation, processing, linkage, error identification, correction and reporting mechanism) from ingestion to output, sized to meet the standards of the production SLA;
        - Performance metrics that mirror the production environment; and
        - Management with the same information security policies applicable to the production environment.

    - The industry test environment must also contain functionality to support industry testing, including:

        - Minimum availability of 24x6;
        - Replica of production data when needed for testing;
        - Data storage sized to meet varying needs, dependent upon scope and test scenarios; and

        - Support of two versions of code (current and pending).

    - The industry test environment must support the following types of industry testing:

        - Technical upgrades made by the Plan Processor;
        - CAT code releases that impact CAT Reporters;
        - Changes to industry data feeds (e.g., SIP, OPRA, etc.);
        - Industry-wide disaster recovery testing;
        - Individual CAT Reporter and Data Submitter testing of their upgrades against CAT interfaces and functionality; and
        - Multiple, simultaneous CAT Reporter testing.

    - The industry test environment must be a discrete environment separate from the production environment.
    - The Plan Processor must provide the linkage processing of data submitted during coordinated, scheduled, industry-wide testing. Results of the linkage processes must be communicated back to Participants as well as to the Operating Committee.
    - Data from industry testing must be saved for three months. Operational metrics associated with industry testing (including but not limited to testing results, firms who participated, and amount of data reported and linked) must be stored for the same duration as the CAT production data.
    - The Plan Processor must provide support for industry testing, including testing procedures, coordination of industry testing, publish notifications, and provide help desk support during industry testing.
    - The Participants and the SEC must have access to industry test data.

1.3 <u>Capacity Requirements</u>

System capacity must have the following characteristics.[263]

The Central Repository must be:

- Designed such that additional capacity can be quickly and seamlessly integrated while maintaining system access and availability requirements;

- Able to efficiently and effectively handle data ingestion on days with peak and above-peak data submission volumes; and

[263] References to data sizing refer to raw, uncompressed data and do not account for benefits of compression, overhead of data storage or indices. Data sizing estimates do not include meta-data and are based on delimited, fixed length data sets. The Plan Processor is responsible for calculating its platform capacity capabilities based on its proposed solution. Three years after the finalization of the CAT NMS Plan, when all CAT Reporters submit their data to the Central Repository, the Central Repository must be sized to receive process and load more than 58 billion records per day.

- Required to maintain and store data for a 6-year sliding window of data. System access and availability requirements must be maintained during the maintenance of the sliding window. It is expected that the Central Repository will grow to more than 29 petabytes of raw, uncompressed data.

The Plan Processor must:

- Define a capacity planning process to be approved by the Operating Committee, with such process incorporating industry utility capacity metrics; and

- Develop a robust process to add capacity, including both the ability to scale the environment to meet the expected annual increases as well as to rapidly expand the environment should unexpected peaks in data volumes breach the defined capacity baseline. Capacity forecasts from systems, including OPRA, UTP, and CTA, must also be included for capacity planning purposes. This capacity planning process must be approved by the Operating Committee.

1.3.1 Monitoring Capacity Utilization and Performance Optimization

In order to manage the data volume, operational capacity planning must be conducted on a periodic basis. The Plan Processor must submit capacity-planning metrics to the Operating Committee for review to ensure that all parties are aware of the system processing capabilities and changes to assumptions. Changes to assumptions could lead to positive or negative adjustments in the costs charged to CAT Reporters. Reports that capture daily disk space, processing time, amount of data received and linkage completion times must be provided by the Plan Processor to the Operating Committee.

1.4 Data Retention Requirements

The Plan Processor must develop a formal record retention policy and program for the CAT, to be approved by the Operating Committee, which will, at a minimum:

- Contain requirements associated with data retention, maintenance, destruction, and holds;

- Comply with applicable SEC record-keeping requirements;

- Have a record hold program where specific CAT Data can be archived offline for as long as necessary;

- Store and retain both raw data submitted by CAT Reporters and processed data; and

- Make data directly available and searchable electronically without manual intervention for at least six years.

2. Data Management

The Plan Processor must develop data management policies and procedures to govern and manage CAT Data, reference data, and metadata contained in and used by the Central Repository.

The CAT must capture, store, and maintain current and historical reference data information. This master / reference database will include data elements such as, but not limited to, SRO-assigned market participant identifiers, product type, trading unit size, trade / quote minimum price variation, corporate actions, symbology changes, and changes in listings market center. The Plan Processor must support bi-temporal milestones (e.g., Effective Date and as-of-date) of the reference data.

CAT Reporters will submit data to the Central Repository with the listing exchange symbology format. The Central Repository must use the listing exchange symbology format for output of the linked data. Instrument validation must be included in the processing of data submitted by CAT Reporters.

The Central Repository must be able to link instrument data across any time period so that data can be properly displayed and linked regardless of changes to issue symbols or market class. The Plan Processor is required to create and maintain a symbol history and mapping table, as well as to provide a tool that will display a complete issue symbol history that will be accessible to CAT Reporters, Participants and the SEC. In addition, the Plan Processor will be required to create a start-of-day ("SOD") and end-of-day ("EOD") CAT reportable list of securities for use by CAT Reporters. This list must be available online and in a machine readable (e.g., .csv) format by 6 a.m. on each Trading Day.

Queries, reports, and searches for data that span dates where there are changes to reference data must automatically include data within the requested date range. For example, if a query is run for a symbol that had three issue symbol changes during the time window of the query parameters, the result set must automatically include data for all three symbols that were in use during the time window of the query.

The Plan Processor must also develop an end-to-end process and framework for technical, business and operational metadata.

2.1 Data Types and Sources

The Plan Processor will be responsible for developing detailed data and interface specifications for data to be submitted by CAT Reporters. These specifications will be contained in the Technical Specifications, the initial version of which will be presented to the Operating Committee for approval. The Technical Specifications must be designed to capture all of the data elements required by SEC Rule 613, as well as other information the Participants determine necessary to facilitate elimination of reporting systems that the CAT may cause to be redundant, such as EBS and OATS. In the future, new data sources such as public news may be added to the specifications.

CAT Reporters and Data Submitters will transmit data in an electronic data format(s) that will be defined by the Plan Processor. The Technical Specifications must include details for connectivity and electronic submission, transmission, retransmission and processing. It is possible that more than one format will be defined to support the various senders throughout the industry.

The Participants anticipate that some broker-dealers will not directly report to the CAT but will rely on other organizations to report on their behalf. However, the CAT will need to have the flexibility to adapt on a timely basis to changes in the number of entities that report CAT Data.

2.2 Data Feed Management

The Plan Processor must monitor and manage incoming and outgoing data feeds for, at a minimum, the following:

- Data files from each CAT Reporter and Data Submitter;
- Files that cover multiple trade dates (e.g., to account for clearing and changes);
- Full and partial file submissions that contain corrections from previously rejected files;
- Full and partial file submissions based on CAT Reporter; and
- Receipt and processing of market data feeds (SIP, OPRA, OCC).

The Plan Processor must also develop a process for detecting, managing, and mitigating duplicate file submissions. It must create and store operational logs of transmissions, success, and failure reasons in order to create reports for CAT Reporters, Participants, and the SEC. Outgoing data feeds must be logged and corresponding metadata elements must be monitored and captured.

2.2.1 Managing connectivity for data feeds (e.g., SIPs, broker-dealers and regulators)

The Plan Processor will be required to ensure that it provides all CAT Reporters with the ability to transmit CAT Data to the Central Repository as required to meet the reporting requirements. The Plan Processer is required to have a robust managed file transfer ("MFT") tool, including full monitoring, permissioning, auditing, security, high availability,[264] file integrity checks, identification of data transmission failures / errors, transmission performance metrics, multiple transmission protocols, Latency / network bottlenecks or delays, key management, etc. CAT Reporters must also have the ability to conduct manual data entry via a GUI interface or the uploading of a file, subject to a maximum record capacity, which will be defined by the Plan Processor in consultation with the Operating Committee.

3. Reporting and Linkage Requirements

All CAT Data reported to the Central Repository must be processed and assembled to create the complete lifecycle of each Reportable Event. Reportable Events must contain data elements sufficient to ensure the same regulatory coverage currently provided by existing regulatory reporting systems that have been identified as candidates for retirement.

[264] To be defined in the SLAs to be agreed to between the Participants and the Plan Processor, as detailed in Appendix D, Functionality of the CAT System.

Additionally, the Central Repository must be able to:

- Assign a unique CAT-Reporter-ID to all reports submitted to the system based on sub-identifiers, (e.g., MPIDs, ETPID, trading mnemonic) currently used by CAT Reporters in their order handling and trading processes.

- Handle duplicate sub-identifiers used by members of different Participants to be properly associated with each Participant.

- Generate and associate one or more Customer-IDs with all Reportable Events representing new orders received from a Customer(s) of a CAT Reporter. The Customer-ID(s) will be generated from a Firm Designated ID provided by the CAT Reporter for each such event, which will be included on all new order events.

- Accept time stamps on order events handled electronically to the finest level of granularity captured by CAT Reporters. Additionally, the CAT must be able to expand the time stamp field to accept time stamps to an even finer granularity as trading systems expand to capture time stamps in ever finer granularity. The Plan Processor must normalize all processed date/time CAT Data into a standard time zone/format.

In addition, the data required from CAT Reporters will include all events and data elements required by the Plan Processor in the Technical Specifications to build the:

- Life cycle of an order for defined events within a CAT Reporter;

- Life cycle of an order for defined events intra-CAT Reporter; and

- State of all orders across all CAT Reporters at any point in time.

The Plan Processor must use the "daisy chain approach" to link and create the order lifecycle. In the daisy chain approach, a series of unique order identifiers, assigned to all order events handled by CAT Reporters are linked together by the Central Repository and assigned a single CAT-generated CAT-Order-ID that is associated with each individual order event and used to create the complete lifecycle of an order.

By using the daisy chain approach the Plan Processor must be able to link all related order events from all CAT Reporters involved in the lifecycle of an order. At a minimum, the Central Repository must be able to create the lifecycle between:

- All order events handled within an individual CAT Reporter, including orders routed to internal desks or departments with different functions (e.g., an internal ATS);

- Customer orders to "representative" orders created in firm accounts for the purpose of facilitating a customer order (e.g., linking a customer order handled on a riskless principal basis to the street-side proprietary order);

- Orders routed between broker-dealers;

- Orders routed from broker-dealers to exchanges;

- Orders sent from an exchange to its routing broker-dealer;

- Executed orders and trade reports;

- Various legs of option/equity complex orders; and

- Order events for all equity and option order handling scenarios that are currently or may potentially be used by CAT Reporters, including:

    - Agency route to another broker-dealer or exchange;
    - Riskless principal route to another broker-dealer or exchange capturing within the lifecycle both the customer leg and street side principal leg;
    - Orders routed from one exchange through a routing broker-dealer to a second exchange;
    - Orders worked through an average price account capturing both the individual street side execution(s) and the average price fill to the Customer;
    - Orders aggregated with other orders for further routing and execution capturing both the street side executions for the aggregated order and the fills to each customer order;
    - Complex orders involving one or more options legs and an equity leg, with a linkage between the option and equity legs;
    - Complex orders containing more legs than an exchange's order management system can accept, causing the original order to be broken into multiple orders;
    - Orders negotiated over the telephone or via a negotiation system;
    - Orders routed on an agency basis to a foreign exchange;
    - Execution of customer order via allocation of shares from a pre-existing principal order;
    - Market maker quotes; and
    - Complex orders involving two or more options legs.

Additionally, the Central Repository must be able to:

- Link each order lifecycle back to the originating Customer;

- Integrate and appropriately link reports representing repairs of original submissions that are rejected by the CAT due to a failure to meet a particular data validation;

- Integrate into the CAT and appropriately link reports representing records that are corrected by a CAT Reporter for the purposes of correcting data errors not identified in the data validation process;

- Assign a single CAT-Order-ID to all events contained within the lifecycle of an order so that regulators can readily identify all events contained therein; and

- Process and link Manual Order Events with the remainder of the associated order lifecycle.

3.1 Timelines for Reporting

CAT Data for the previous Trading Day must be reported to the Central Repository by 8:00 a.m. Eastern Time on the Trading Day following the day the Industry Member receives such data; however, the Plan Processor must accept data prior to that deadline, including intra-day submissions.

3.2 Other Items

The Plan Processor must anticipate and manage order data processing over holidays, early market closures and both anticipated and unanticipated market closures. The Plan Processor must allow and enable entities that are not CAT Reporters (e.g., service bureaus) to report on behalf of CAT Reporters only upon being permissioned by the CAT Reporter, and must develop appropriate tools to facilitate this process.

3.3 Required Data Attributes for Order Records Submitted by CAT Reporters

At a minimum, the Plan Processor must be able to receive the data elements as detailed in the CAT NMS Plan.

4. Data Security

4.1 Overview

SEC Rule 613 requires that the Plan Processor ensure the security and confidentiality of all information reported to and maintained by the CAT in accordance with the policies, procedures and standards in the CAT NMS Plan.

The Plan Processor must have appropriate solutions and controls in place to ensure data confidentiality and security during all communication between CAT Reporters and Data Submitters and the Plan Processor, data extraction, manipulation and transformation, loading to and from the Central Repository and data maintenance by the CAT System. The Plan Processor must address security controls for data retrieval and query reports by Participant and the SEC. The solution must provide appropriate tools, logging, auditing and access controls for all components of the CAT System, such as but not limited to access to the Central Repository, access for CAT Reporters, access to rejected data, processing status and CAT Reporter performance and comparison statistics.

The Plan Processor must provide to the Operating Committee a comprehensive security plan that covers all components of the CAT System, including physical assets and personnel, and the training of all persons who have access to the Central Repository consistent with Article VI, Section 6.1(m). The security plan must be updated annually. The security plan must include an overview of the Plan Processor's network security controls, processes and procedures pertaining to the CAT Systems. Details of the security plan must document how the Plan Processor will protect, monitor and patch the environment; assess it for vulnerabilities as part of a managed process, as

well as the process for response to security incidents and reporting of such incidents. The security plan must address physical security controls for corporate, data center, and leased facilities where Central Repository data is transmitted or stored. The Plan Processor must have documented "hardening baselines" for systems that will store, process, or transmit CAT Data or PII data.

4.1.1 Connectivity and Data Transfer

The CAT System(s) must have encrypted internet connectivity. CAT Reporters must connect to the CAT infrastructure using secure methods such as private lines or (for smaller broker-dealers) Virtual Private Network connections over public lines. Remote access to the Central Repository must be limited to authorized Plan Processor staff and must use secure multi-factor authentication that meets or exceeds the Federal Financial Institutions Examination Council ("FFIEC") security guidelines surrounding authentication best practices.[265]

The CAT databases must be deployed within the network infrastructure so that they are not directly accessible from external end-user networks. If public cloud infrastructures are used, virtual private networking and firewalls/access control lists or equivalent controls such as private network segments or private tenant segmentation must be used to isolate CAT Data from unauthenticated public access.

4.1.2 Data Encryption

All CAT Data must be encrypted at rest and in flight using industry standard best practices (e.g., SSL/TLS) including archival data storage methods such as tape backup. Symmetric key encryption must use a minimum key size of 128 bits or greater (e.g., AES-128), larger keys are preferable. Asymmetric key encryption (e.g., PGP) for exchanging data between Data Submitters and the Central Repository is desirable.

[All PII data must be encrypted both at rest and in flight, including archival data storage methods such as tape backup.] Storage of unencrypted PII data is not permissible. PII encryption methodology must include a secure documented key management strategy such as the use of HSM(s). The Plan Processor must describe how PII encryption is performed and the key management strategy (e.g., AES-256, 3DES).

[CAT Data stored in a public cloud must be encrypted at rest. Non-PII CAT Data stored in a Plan Processor private environment is not required to be encrypted at rest.]

If public cloud managed services are used that would inherently have access to the data (e.g., BigQuery, S3, Redshift), then the key management surrounding the encryption of that data must be documented (particularly whether the cloud provider manages the keys, or if the Plan Processor maintains that control). Auditing and real-time monitoring of the service for when cloud provider personnel are able to access/decrypt CAT Data must be documented, as well as a response plan to address instances where unauthorized access to CAT Data is detected. Key

[265] Federal Financial Institutions Examination Council, Supplement to Authentication in an Internet Banking Environment (June 22, 2011), *available at* http://www.ffiec.gov/pdf/Auth-ITS-Final%206-22-11%20(FFIEC%20Formated).pdf.

management/rotation/revocation strategies and key chain of custody must also be documented in detail.

4.1.3 Data Storage and Environment

Data centers housing CAT Systems (whether public or private) must, at a minimum, be AICPA SOC 2 certified by [an independent third party auditor] a qualified third-party auditor that is not an affiliate of any of the Participants or the CAT Processor. The frequency of the audit must be at least once per year.

CAT compute infrastructure may not be commingled with other non-regulatory systems (or tenets, in the case of public cloud infrastructure). Systems hosting the CAT processing for any applications must be segmented from other systems as far as is feasible on a network level (firewalls, security groups, ACL's, VLAN's, authentication proxies/bastion hosts and similar). In the case of systems using inherently shared infrastructure/storage (e.g., public cloud storage services), an encryption/key management/access control strategy that effectively renders the data private must be documented.

The Plan Processor must include penetration testing and an application security code audit by a reputable (and named) third party prior to launch as well as periodically as defined in the SLA(s). Reports of the audit will be provided to the Operating Committee as well as remediation plan for identified issues. The penetration test reviews of the Central Repository's network, firewalls, and development, testing and production systems should help the CAT evaluate the system's security and resiliency in the face of attempted and successful systems intrusions.

4.1.4 Data Access

The Plan Processor must provide an overview of how access to PII and other CAT Data by Plan Processor employees and administrators is restricted. This overview must include items such as, but not limited to, how the Plan Processor will manage access to the systems, internal segmentation, multi-factor authentication, separation of duties, entitlement management, background checks, etc.

The Plan Processor must develop and maintain policies and procedures reasonably designed to prevent, detect, and mitigate the impact of unauthorized access or usage of data in the Central Repository. Such policies and procedures must be approved by the Operating Committee, and should include, at a minimum:

- Information barriers governing access to and usage of data in the Central Repository;
- Monitoring processes to detect unauthorized access to or usage of data in the Central Repository; and
- Escalation procedures in the event that unauthorized access to or usage of data is detected.

A Role Based Access Control ("RBAC") model must be used to permission user with access to different areas of the CAT System. The CAT System must support an arbitrary number

of roles with access to different types of CAT Data, down to the attribute level. The administration and management of roles must be documented. Periodic reports detailing the current list of authorized users and the date of their most recent access must be provided to Participants, the SEC and the Operating Committee. The reports of the Participants and the SEC will include only their respective list of users. The Participants [and the SEC] must provide a response to the report confirming that the list of users is accurate. The required frequency of this report will be defined by the Operating Committee. The Plan Processor must log every instance of access to Central Repository data by users.

Passwords stored in the CAT System must be stored according to industry best practices. Reasonable password complexity rules should be documented and enforced, such as, but not limited to, mandatory periodic password changes and prohibitions on the reuse of the recently used passwords.

Password recovery mechanisms must provide a secure channel for password reset, such as emailing a one-time, time-limited login token to a pre-determined email address associated with that user. Password recovery mechanisms that allow in-place changes or email the actual forgotten password are not permitted.

Any login to the system that is able to access PII data must follow non-PII password rules and must be further secured via multi-factor authentication ("MFA"). The implementation of MFA must be documented by the Plan Processor. MFA authentication capability for all logins [(including non-PII)] is required to be implemented by the Plan Processor.

4.1.5 Breach Management

The Plan Processor must develop policies and procedures governing its responses to systems or data breaches. Such policies and procedures will include a formal cyber incident response plan, and documentation of all information relevant to breaches.

The cyber incident response plan will provide guidance and direction during security incidents. The plan will be subject to approval by the Operating Committee. The plan may include items such as:

- Guidance on crisis communications;
- Security and forensic procedures;
- Customer notifications;
- "Playbook" or quick reference guides that allow responders quick access to key information;
- Insurance against security breaches;
- Retention of legal counsel with data privacy and protection expertise; and

- Retention of a Public Relations firm to manage media coverage.

Documentation of information relevant to breaches should include:

- A chronological timeline of events from the breach throughout the duration of the investigation;
- Relevant information related to the breach (e.g., date discovered, who made the discovery, and details of the breach);
- Response efforts, involvement of third parties, summary of meetings/conference calls, and communication; and
- The impact of the breach, including an assessment of data accessed during the breach and impact on CAT Reporters.

4.1.6 PII Data Requirements

PII data must not be included in the result set(s) from online or direct query tools, reports or bulk data extraction. Instead, results will display existing non-PII unique identifiers (e.g., Customer-ID or Firm Designated ID). The PII corresponding to these identifiers can be gathered using the PII workflow described in Appendix D, Data Security, PII Data Requirements. By default, users entitled to query CAT Data are not authorized for PII access. The process by which someone becomes entitled for PII access, and how they then go about accessing PII data, must be documented by the Plan Processor. The chief regulatory officer, or other such designated officer or employee at each Participant [and the Commission] must, at least annually, review and certify that people with PII access have the appropriate level of access for their role.

Using the RBAC model described above, access to PII data shall be configured at the PII attribute level, following the "least privileged" practice of limiting access as much as possible.

PII data must be stored separately from other CAT Data. It cannot be stored with the transactional CAT Data, and it must not be accessible from public internet connectivity. A full audit trail of PII access (who accessed what data, and when) must be maintained. The Chief Compliance Officer and the Chief Information Security Officer shall have access to daily PII reports that list all users who are entitled for PII access, as well as the audit trail of all PII access that has occurred for the day being reported on.

4.2 <u>Industry Standards</u>

The following industry standards [, at a minimum,]<u>—which is not intended to be an exclusive list—</u>must be followed as such standards and requirements may be replaced by successor publications, or modified, amended, or supplemented and as approved by the Operating Committee (in the event of a conflict between standards, the more stringent standard shall apply, subject to the approval of the Operating Committee):

- <u>National Institute of Standards and Technology</u>:

- 800-23 – Guidelines to Federal Organizations on Security Assurance and Acquisition / Use of Test/Evaluated Products
- 800-53 – Security and Privacy Controls for Federal Information Systems and Organizations
- 800-115 – Technical Guide to Information Security Testing and Assessment
- 800-118 – Guide to Enterprise Password Management
- 800-133 – Recommendation for Cryptographic Key Generation
- 800-137 – Information Security Continuous Monitoring for Federal Information Systems and Organizations
- To the extent not specified above, all other provisions of the NIST Cyber Security Framework

- Federal Financial Institutions Examination Council:

  - Authentication Best Practices

- International Organization for Standardization:

  - ISO/IEC 27001 – Information Security Management

The Company shall endeavor to join the FS-ISAC and comparable bodies as the Operating Committee may determine. The FS-ISAC provides real time security updates, industry best practices, threat conference calls, xml data feeds and a member contact directory. The FS-ISAC provides the Company with the ability to work with the entire financial industry to collaborate for the purposes of staying up to date with the latest information security activities.

5. BCP / DR Process

5.1 Overview

The Plan Processor must develop and implement disaster recovery ("DR") and business continuity plans ("BCP") that are tailored to the specific requirements of the CAT environment, and which must be approved and regularly reviewed by the Operating Committee. The BCP must address the protection of data, service for the data submissions, processing, data access, support functions and operations. In the context of this document, BCP generally refers to how the business activities will continue in the event of a widespread disruption and the DR requirements refer to how the CAT infrastructure will be designed to support a full data center outage. In addition, the Plan Processor must have SLAs in place to govern redundancy (i.e., no single point of failure) of critical aspects of the CAT System (e.g., electrical feeds, network connectivity, redundant processors, storage units, etc.) and must have an architecture to support and meet the SLA requirements. Any SLAs between the Plan Processor and third parties must be approved by the Operating Committee.

5.2 Industry Standards

The following National Institute of Standards and Technology standards, at a minimum, must be followed in association with Disaster Recovery, in each case as such standards and requirements may be replaced by successor publications, or modified, amended, or supplemented and as approved by the Operating Committee:

- 800-34 – Contingency Planning for Federal Information Systems; and

- Specifically, the following sections as minimum requirements for designing and implementing BCP and DR plans:

    - Chapter 3: Information System Contingency Planning Process, which identifies seven steps to use when developing contingency plans;
    - Chapter 4: Information System Contingency Plan Development, which outlines the key elements of a contingency plan;
    - Chapter 5: Technical Contingency Planning Considerations (using the specific sections applicable to the Plan Processor's systems) which provides considerations specific to different types of technology; and
    - Other sections and the appendices should be taken into consideration as warranted.

In addition, the Plan Processor will need to develop a process to manage and report all breaches.

5.3 Business Continuity Planning

The Plan Processor will design a BCP that supports a continuation of the business activities required of the CAT in the event of a widespread disruption.

With respect to the team supporting CAT business operations, a secondary site must be selected that is capable of housing the critical staff necessary for CAT business operations. The site must be fully equipped to allow for immediate use. The selection of the site must take into account diversity in utility and telecommunications infrastructure as well as the ability for CAT staff to access the site in the event of transit shutdowns, closure of major roadways and other significant disruptions that may affect staff. Planning should consider operational disruption involving significant unavailability of staff.

A bi-annual test of CAT operations where CAT staff operates the facility from the secondary site is required. This will ensure that phone systems, operational tools and other help desk functions all work as expected and the Plan Processor still functions as usual even in the event of a disruption.

CAT operations staff must maintain, and annually test, remote access capabilities to ensure smooth operations during a site un-availability event. Certain critical staff may be required to report directly to the secondary office site. However, an effective telecommuting solution must be in place for all critical CAT operations staff. Furthermore, any telecommuting strategy must require a remote desktop style solution where CAT operations and data consoles remain at the

primary data center and must further ensure that CAT Data may not be downloaded to equipment that is not CAT-owned and compliant with CAT security requirements.

The BCP must identify critical third party dependencies. The Plan Processor will coordinate with critical suppliers regarding their arrangements and involve these parties in tests on an annual basis. Critical third party firms may be required to provide evidence of their BCP capabilities and testing.

The Plan Processor must conduct third party risk assessments at regular intervals to verify that security controls implemented are in accordance with NIST SP 800-53. These risk assessments must include assessment scheduling, questionnaire completion and reporting. The Plan Processor should provide assessment reports to the Operating Committee.

The Plan Processor will develop and annually test a detailed crisis management plan to be invoked following certain agreed disruptive circumstances.

The processing sites for business continuity must adhere to the "Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System."[266]

The Plan Processor will conduct an annual Business Continuity Audit using an Independent Auditor approved by the Operating Committee. The Independent Auditor will document all findings in a detailed report provided to the Operating Committee.

5.4 Disaster Recovery Requirements

The Plan Processor will implement a DR capability that will ensure no loss of data and will support the data availability requirements and anticipated volumes of the CAT.

A secondary processing site must be capable of recovery and restoration of services at the secondary site within a minimum of 48 hours, but with the goal of achieving next day recovery after a disaster event. The selection of the secondary site must consider sites with geographic diversity that do not rely on the same utility, telecom and other critical infrastructure services. The processing sites for disaster recovery and business continuity must adhere to the "Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System."

The secondary site must have the same level of availability / capacity / throughput and security (physical and logical) as the primary site. The requirement implies and expects that fully redundant connectivity between the primary and secondary processing sites be established and fully available. Further, given this recovery window, this connectivity must be used to replicate repositories between the primary and secondary sites. Finally, CAT Reporter and Data Submitter submissions must be replicated to the secondary site for possible replay if recent replications are incomplete. Replication must occur as deliveries complete to ensure that a widespread communications failure will have minimal impact to the state of the secondary site.

[266] *See* Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System (Apr. 8, 2003), *available at* http://www.sec.gov/news/studies/34-47638.htm.

On an annual basis, the Plan Processor must execute an industry DR test, which must include Plan Participants and a critical mass of non-Plan Participant CAT Reporters and Data Submitters. The tests must be structured such that all CAT Reporters and other Data Submitters can upload to the DR site and the data be ingested by the CAT Data loaders. All DR tests are required to realistically reflect the worst-case scenario.

Failover processes must be transparent to CAT Reporters, as well as failback. In the event of a site failover, CAT Reporters must be able to deliver their daily files without changing configuration. This avoids requiring all CAT Reporters to update configurations, which is an error-prone effort.

After a DR event, the primary processing site must be made available as quickly as possible. For short duration DR events, the primary site must be returned to primary within 48 hours after the DR event. Longer duration outages will have differing SLAs. The DR plan must include designs that allow the re-introduction of the primary site or the introduction of a new primary site as the event dictates and an indication of the time required for this re-introduction.

6. Data Availability

6.1 Data Processing

CAT order events must be processed within established timeframes to ensure data can be made available to Participants' regulatory staff and the SEC in a timely manner. The processing timelines start on the day the order event is received by the Central Repository for processing. Most events must be reported to the CAT by 8:00 a.m. Eastern Time the Trading Day after the order event occurred (referred to as transaction date). The processing timeframes below are presented in this context. All events submitted after T+1 (either reported late or submitted later because not all of the information was available) must be processed within these timeframes based on the date they were received.

The Participants require the following timeframes (Figure A) for the identification, communication and correction of errors from the time an order event is received by the processor:

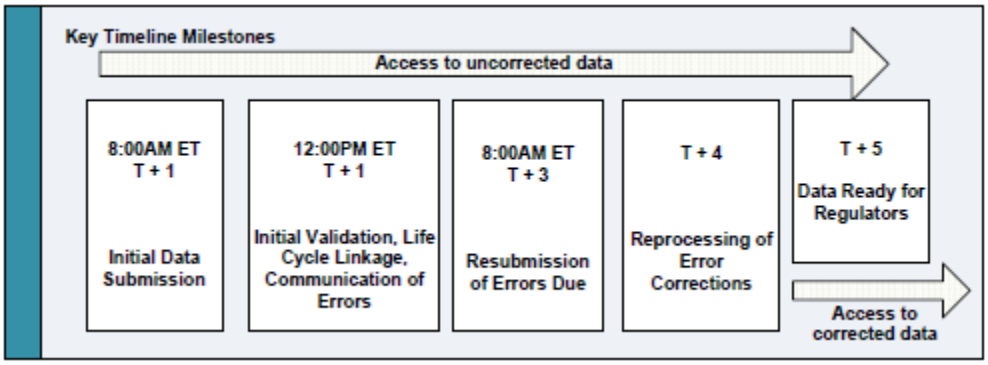
- Noon Eastern Time T+1 (transaction date + one day) – Initial data validation, lifecycle linkages and communication of errors to CAT Reporters;

- 8:00 a.m. Eastern Time T+3 (transaction date + three days) – Resubmission of corrected data; and

- 8:00 a.m. Eastern Time T+5 (transaction date + five days) – Corrected data available to Participant regulatory staff and the SEC.

Late submissions or re-submissions (after 8:00 a.m.) may be considered to be processed that day if it falls within a given time period after the cutoff. This threshold will be determined by the Plan Processor and approved by the Operating Committee. In the event that a significant portion of the data has not been received as monitored by the Plan Processor, the Plan Processor may decide to halt processing pending submission of that data.

Figure A: CAT Central Repository Data Processing Timelines



6.2 Data Availability Requirements

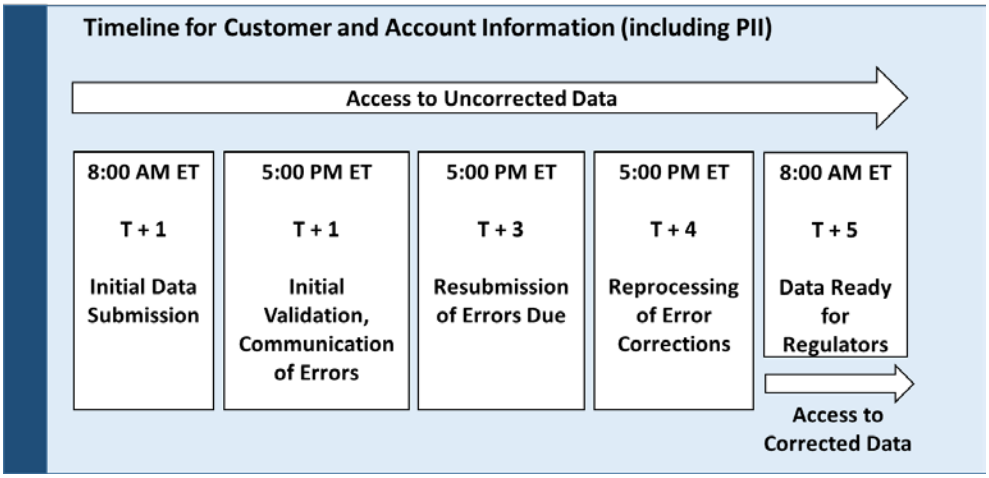
Prior to 12:00 p.m. Eastern Time on T+1, raw unprocessed data that has been ingested by the Plan Processor must be available to Participants' regulatory staff and the SEC.

Between 12:00 p.m. Eastern Time on T+1 and T+5, access to all iterations of processed data must be available to Participants' regulatory staff and the SEC.

The Plan Processor must provide reports and notifications to Participant regulatory staff and the SEC regularly during the five-day process, indicating the completeness of the data and errors. Notice of major errors or missing data must be reported as early in the process as possible. If any data remains un-linked after T+5, it must be available and included with all linked data with an indication that the data was not linked.

If corrections are received after T+5, Participants' regulatory staff and the SEC must be notified and informed as to how re-processing will be completed. The Operating Committee will be involved with decisions on how to re-process the data; however, this does not relieve the Plan Processor of notifying the Participants' regulatory staff and the SEC.

Figure B: Customer and Account Information (Including PII)



CAT PII data must be processed within established timeframes to ensure data can be made available to Participants' regulatory staff and the SEC in a timely manner. Industry Members submitting new or modified Customer information must provide it to the Central Repository no later than 8:00 a.m. Eastern Time on T+1. The Central Repository must validate the data and generate error reports no later than 5:00 p.m. Eastern Time on T+[3]1. The Central Repository must process the resubmitted data no later than 5:00 p.m. Eastern Time on T+4. Corrected data must be resubmitted no later than 5:00 p.m. Eastern Time on T+3. The Central Repository must process the resubmitted data no later than 5:00 p.m. Eastern Time on T+4. Corrected data must be available to regulators no later than 8:00 a.m. Eastern Time on T+5.

Customer information that includes PII data must be available to regulators immediately upon receipt of initial data and corrected data, pursuant to security policies for retrieving PII.

7. Receipt of Data from Reporters

7.1 Receipt of Data Transmission

Following receipt of data files submitted by the CAT Reporter or Data Submitter, the Plan Processor must send an acknowledgement of data received to the CAT Reporter and Data Submitter, if applicable. Such acknowledgment will enable CAT Reporters to create an audit trail of their submissions and allow for tracing of data breakdowns when data is not received. At a minimum, the receipt acknowledgement will include:

- SRO-Assigned Market Participant Identifier;
- Date of Receipt;
- Time of Receipt;
- File Identifier; and
- Value signifying the acknowledgement of receipt, but not processing, of the file.

7.2 Data Validation

The Plan Processor will implement data validations at the file and individual record level for data received by the Plan Processor including customer data. If a record does not pass basic validations, such as syntax rejections, then it must be rejected and sent back to the CAT Reporter as soon as possible, so it can repair and resubmit.[267] The required data validations may be amended based on input from the Operating Committee and the Advisory Committee. All identified exceptions will be reported back to the CAT Reporter submitting the data and/or the CAT Reporter on whose behalf the data was submitted.

[267] If needed – data validation may be a process with an initial validation phase for data errors and a subsequent validation phase later in processing where more time is needed to assess the context of the record in relation to data that may be submitted to the CAT later in the submission window. The Plan Processor must have an additional "matching" process for the purposes of linking together order data passed between CAT Reporters.

The data validations must include the following categories and must be explained in the Technical Specifications document:

- File Validations – Confirmation of file transmission and receipt are in the correct formats. This includes validation of header and trailers on the submitted report, confirmation of a valid SRO-Assigned Market Participant Identifier, and verification of the number of records in the file.

- Validation of CAT Data – Syntax and context checks, including:

    - Format checks:
        - Check that the data is entered in the specified format
    - Data Type checks:
        - Check that the data type of each attribute is as per specification
    - Consistency checks:
        - Check that all attributes for a record of a specified type are consistent
    - Range/logic checks:
        - Range check – Validate that each attribute for every record has a value within specified limits
        - Logic check – Validate that the values provided against each attribute are associated with the event type they represent
    - Data validity checks:
        - Validate that each attribute for every record has an acceptable value
    - Completeness checks:
        - Verify that each mandatory attribute for every record is not null
    - Timeliness checks:
        - Verify that records were submitted within the submission timelines

- Linkage Validation[268] – Process by which related CAT Reportable Events are in a linked daisy chain method

CAT Reporters must have the ability to correct, replace or delete records that have passed initial validations within the CAT.

After the Central Repository has processed the data, the Plan Processor must provide daily statistics, including at a minimum, the following information:

- SRO-Assigned Market Participant Identifier;

- Date of Submission;

- Number of files received;

- Number of files accepted;

[268] A linkage validation error should only populate for the CAT Reporter that the Plan Processor determines to have broken the link.

- Number of files rejected;
- Number of total order events received;
- Number of order events accepted;
- Number of order events rejected;
- Number of each type of report received;
- Number of each type of report accepted;
- Number of each type of report rejected;
- Number of customer records received;
- Number of total customer records accepted;
- Number of total customer records rejected;
- Number of unknown accounts;
- Number of late submissions;
- Order-IDs rejected;
- Reason(s) for rejection;
- Number of records attempted to be matched;
- Number of records matched; and
- Percentage of records matched.

Individual records contained in files that do not pass the file validation process must not be included for further processing. Once a file passes the initial validation, individual records contained therein may then be processed for further validation. Individual records that do not pass the data validation processes will not be included in the final audit trail but must be retained. Additionally, records not passing the validations will not be included for matching processes.

7.3 Exception Management

The Plan Processor must capture rejected records for each CAT Reporter and make them available to the CAT Reporter. The "rejects" file must be accessible via an electronic file format and the rejections and daily statistics must be available via a web interface. The Plan Processor must provide functionality for CAT Reporters to amend any exceptions.

The Plan Processor must support bulk error correction. Rejected records can be resubmitted as a new file with appropriate indicators to identify the rejection record, which is being repaired. The Plan Processor will then reprocess repaired records.

A GUI must be available for CAT Reporters to make updates to individual records or attributes and must include, at a minimum, the:

- Count of each type of rejection;
- Reason for each rejection;
- Ability to download the rejections;
- Firm assigned order ID of each rejection;
- Details of each rejection;
- Type of report rejected; and
- Repair status.

The Plan Processor must support bulk replacement of records, and reprocess such replaced records. The Plan Processor must provide CAT Reporters with documentation that detail the process how to amend and upload records that fail the validations that are outlined as part of Section 7.4. The Plan Processor must maintain a detailed audit trail capturing corrections to and replacements of records.

The Plan Processor will provide CAT Reporters with their error reports as they become available, and daily statistics will be provided after data has been uploaded and validated by the Plan Processor. The Plan Processor must support a continuous validation and feedback model so that CAT Reporters can identify and correct rejections on an ongoing basis. The rejected reports will include descriptive details, or codes related to descriptive details, as to why each data record was rejected by the Plan Processor.

On a monthly basis, the Plan Processor must produce and publish reports detailing performance and comparison statistics for CAT Reporters,[269] similar to the Report Cards published for OATS presently. This will enable CAT Reporters to assess their performance in relation to their industry peers and help them assess the risk related to their reporting of transmitted data.

Breaks in intermittent lifecycle linkages must not cause the entire lifecycle to break nor cause a reject to the CAT Reporter that correctly reported.

[269] *See* Appendix C, Error Communication, Correction, and Processing.

7.4 Error Corrections

Error corrections must be able to be submitted and processed at any time, including timeframes after the standard repair window. Additionally, in order to make corrections, CAT Reporters must have access to the Central Repository over weekends.

CAT Reporters must be able to submit error corrections for data errors identified by CAT Reporters that passed format validations.

Additionally, the Plan Processor must:

- Provide feedback as to the reason(s) for errors;
- Prevent a linkage break between reports from resulting in additional events being rejected;
- Allow broken linkages to be repaired without having to submit or resubmit additional reports;
- Allow error corrections to be submitted both via online and bulk uploads or via file submission;
- Support auto-correction of identified errors and notify reporters of any auto-corrections;
- Support group repairs (i.e., the wrong issue symbol affecting multiple reports).

7.5 Data Ingestion

Data submitted to the Central Repository, including rejections and corrections, must be stored in repositories designed to hold information based on the classification of the CAT Reporter (i.e., whether the CAT Reporter is a Participant, a broker-dealer, or a third party Data Submitter). After ingestion by the Central Repository, the Raw Data must be transformed into a format appropriate for data querying and regulatory output.

8. Functionality of the CAT System

8.1 Regulator Access

The Plan Processor must provide Participants' regulatory staff and the SEC with access to all CAT Data for regulatory purposes only. Participants' regulatory staff and the SEC will access CAT Data to perform functions, including economic analyses, market structure analyses, market surveillance, investigations, and examinations.

The CAT must be able to support, at a minimum, 3,000 regulatory users within the system. It is estimated that approximately 20% of all users will use the system on a daily or weekly basis while approximately 10% of all users will require advanced regulator-user access, as described below. Furthermore, it is estimated that there may be approximately 600 concurrent users

accessing the CAT at any given point in time. These users must be able to access and use the system without an unacceptable decline in system performance.[270]

As stated in Appendix D, Data Security, the Plan Processor must be able to support an arbitrary number of user roles. Defined roles must include, at a minimum:

- Basic regulator users – Individuals with approved access who plan to use the Central Repository to run basic queries (e.g., pulling all trades in a single stock by a specific party).

- Advanced regulator users – Individuals with approved access who plan to use the Central Repository to construct and run their own complex queries.

Regulators will have access to processed CAT Data through two different methods, an online-targeted query tool and user-defined direct queries and bulk extracts.

8.1.1 Online Targeted Query Tool

The online targeted query tool will provide authorized users with the ability to retrieve processed and/or validated (unlinked) data via an online query screen that includes the ability to choose from a variety of pre-defined selection criteria. Targeted queries must include date(s) and/or time range(s), as well as one or more of a variety of fields, including the following:

- Instrument(s);

- Related instruments (e.g., single stock and all options with for the stock);

- Data type (executions, orders, cancelations, quotes, etc.);

- Product type (equity, option, etc.);

- Processed data, unlinked data or both;

- Listing market;

- Exchange;

- CAT-Reporter-ID(s) – CAT assigned and Participant assigned;

- Customer-ID(s) – CAT assigned and CAT Reporter assigned;

- CAT-Order-ID(s) – CAT assigned and CAT Reporter assigned;

- ISO flag;

[270] Specific performance requirements will be included in the SLA.

- Put/call;
- Strike price (include ability to select range);
- Size;
- Price;
- Side;
- Short-sale identifier;
- Time-in-force (IOC, GTC, etc.);
- Orders, quotes, BBOs or trades above or below a certain size;
- Orders, quotes, BBOs or trades within a range of prices;
- Canceled orders and/or trades;
- CAT Reporters exceeding specified volume or percentage of volume thresholds in a single instrument or market-wide during a specified period of time;
- CAT Reporter correction rate over time;
- Audit trail of order linkages;
- Corporate action events;
- Instrument history; and
- Others to be defined.

The tool must provide a record count of the result set, the date and time the query request is submitted, and the date and time the result set is provided to the users. In addition, the tool must indicate in the search results whether the retrieved data was linked or unlinked (e.g., using a flag). In addition, the online targeted query tool must not display any PII data. Instead, it will display existing non-PII unique identifiers (e.g., Customer-ID or Firm Designated ID). The PII corresponding to these identifiers can be gathered using the PII workflow described in Appendix D, Data Security, PII Data Requirements. The Plan Processor must define the maximum number of records that can be viewed in the online tool as well as the maximum number of records that can be downloaded. Users must have the ability to download the results to .csv, .txt, and other formats, as applicable. These files will also need to be available in a compressed format (e.g., .zip, .gz). Result sets that exceed the maximum viewable or download limits must return to users a message informing them of the size of the result set and the option to choose to have the result set returned via an alternate method.

The Plan Processor must define a maximum number of records that the online targeted query tool is able to process. The minimum number of records that the online targeted query tool is able to process is 5,000 (if viewed within the online query tool) or 10,000 (if viewed via a downloadable file).

Once query results are available for download, users are to be given the total file size of the result set and an option to download the results in a single or multiple file(s). Users that select the multiple file option will be required to define the maximum file size of the downloadable files. The application will then provide users with the ability to download the files. This functionality is provided to address limitations of end-user network environment that may occur when downloading large files.

The tool must log submitted queries and parameters used in the query, the user ID of the submitter, the date and time of the submission, as well as the delivery of results. The Plan Processor will use this logged information to provide monthly reports to each Participant and the SEC of its respective metrics on query performance and data usage of the online query tool. The Operating Committee must receive all monthly reports in order to review items, including user usage and system processing performance.

8.1.2 Online Targeted Query Tool Performance Requirements

For targeted search criteria, the minimum acceptable response times will be increments of less than one minute. For the complex queries that either scan large volumes of data (e.g., multiple trade dates) or return large result sets (>1M records), the response time must generally be available within 24 hours of the submission of the request. Regardless of the complexity of the criteria used within the online query tool, any query request for data within one business date of a 12-month period must return results within 3 hours.

Performance requirements listed below apply to data:

- Online targeted query tool searches that include equities and options trade data only in the search criteria must meet minimum requirements, including:
    - Returning results within 1 minute for all trades and related lifecycle events for a specific Customer or CAT Reporter with the ability to filter by security and time range for a specified time window up to and including an entire day;
    - Returning results within 30 minutes for all trades and related lifecycle events for a specific Customer or CAT Reporter in a specified date range (maximum 1 month);
    - Returning results within 6 hours for all trades and related lifecycle events for a specific Customer or CAT Reporter in a specified date range (maximum 12-month duration from the most recent 24 months); and
    - Returning results for the full 6 years of data for all trades and lifecycle events across daily, weekly, and multi-year periods.

- Online targeted query tool searches that include equities and options order and National Best Bid and National Best Offer data in search criteria must meet minimum requirements, including:

    - Returning results within 5 minutes for all orders and their complete lifecycles for a single security from a specific Participant across all markets (note: a Participant could have multiple participant identifiers) in a specified time window not to exceed 10 minutes for a single date;
    - Returning results within 5 minutes for all orders, cancelations, and the National Best Bid and National Best Offer (or the protected best bid and offer) at the time the order is created for a single security in a specified time window not to exceed 10 minutes for a single date;
    - Returning results within 5 minutes for all equity and options orders, cancelations, and executions from a specific market participant in a single underlying instrument in a specified time window not to exceed 10 minutes for a single date;
    - Returning results within 5 minutes for all orders, quotes, routes, cancelations and trades (complete life-cycle) for related instruments (e.g., single stock and all options series for the same stock) in a specified time window not to exceed 10 minutes for a single date;
    - Returning results within 5 minutes for all orders and quotes entered during a specific time period by a list of specific CAT Reporters, with the ability to drill down to show the complete life-cycle must return results in a specified time window not to exceed 10 minutes for a single date; and
    - Returning results within 5 minutes for all orders and quotes entered during a specific time period for a specified list of instruments must return results in a specified time window not to exceed 10 minutes for a single date.

The online targeted query tool architecture must include an automated application-level resource management component. This feature must manage query requests to balance the workload to ensure the response times for targeted and complex queries meet the defined response times. The resource management function will categorize and prioritize query requests based on the input parameters, complexity of the query, and the volume of data to be parsed in the query. Additionally, the source of the query may also be used to prioritize the processing. The Plan Processor must provide details on the prioritization plan of the defined solution for online query requests.

The online targeted query tool must support parallel processing of queries. At a minimum, the online targeted query tool must be able to process up to 300 simultaneous query requests with no performance degradation.

8.1.3 Online Targeted Query Tool Access and Administration

Access to CAT Data is limited to authorized regulatory users from the Participants and the SEC. Authorized regulators from the Participants and the SEC may access all CAT Data, with the exception of PII data. A subset of the authorized regulators from the Participants and the SEC will have permission to access and view PII data. The Plan Processor must work with the Participants

and SEC to implement an administrative and authorization process to provide regulator access. The Plan Processor must have procedures and a process in place to verify the list of active users on a regular basis.

A two-factor authentication is required for access to CAT Data. PII data must not be available via the online targeted query tool or the user-defined direct query interface.

8.2 User-Defined Direct Queries and Bulk Extraction of Data

The Central Repository must provide for direct queries, bulk extraction, and download of data for all regulatory users. Both the user-defined direct queries and bulk extracts will be used by regulators to deliver large sets of data that can then be used in internal surveillance or market analysis applications. The data extracts must use common industry formats.

Direct queries must not return or display PII data. Instead, they will return existing non-PII unique identifiers (e.g., Customer-ID or Firm Designated ID). The PII corresponding to these identifiers can be gathered using the PII workflow described in Appendix D, Data Security, PII Data Requirements.

Participants and regulators must have the ability to create, save, and schedule dynamic queries that will run directly against processed and/or unlinked CAT Data. The examples below demonstrate robust usage of the CAT Data to perform a variety of complex query, surveillance, and market analysis use cases. User-defined direct queries will be used to perform tasks such as market reconstruction, behavioral analysis, and cross-market surveillance.

The method(s) for providing this capability is dependent upon the architecture of the CAT and will be defined by the final solution. The CAT cannot be web-based due to the volumes of data that could be extracted.

The Participants are agnostic as to how user-defined direct queries or bulk extracts are implemented as long as the solution provides an open API that allows regulators to use analytic tools (e.g., R, SAS, Python, Tableau) and can use ODBC/JDBC drivers to access the CAT Data. Queries invoked through the open API must be auditable. The CAT System must contain the same level of control, monitoring, logging and reporting as the online targeted query tool. The Plan Processor may define a limited set of basic required fields (e.g., date and at least one other field such as symbol, CAT-Reporter ID, or CAT-Customer-ID) that regulators must use in direct dynamic queries.

The Plan Processor must provide procedures and training to regulators that will use the direct query feature. The Plan Processor may choose to require that user-defined direct query users participate in mandatory training sessions.

The bulk extract feature will replace the current Intermarket Surveillance Group (ISG) ECAT and COATS compliance data files that are currently processed and provided to Participants for use in surveillance applications. These files are used extensively across all Participants in a variety of surveillance applications and are a critical data input to many surveillance algorithms. With the initial implementation of the CAT, opportunities exist to improve the content and depth of information available in these data files. The Plan Processor will need to work with ISG to

define new layouts that will include additional data elements that will be available in the CAT Data.

The Plan Processor is responsible for providing data models and data dictionaries for all processed and unlinked CAT Data.

8.2.1 User-Defined Direct Query Performance Requirements

The user-defined direct query tool is a controlled component of the production environment made available to allow the Participants' regulatory staff and the SEC to conduct queries. The user-defined direct query tool must:

- Provide industry standard programmatic interface(s) that allows Participants' regulatory staff and the SEC with the ability to create, save, and run a query;
- Provide query results that are extractable / downloadable and can be used to refine subsequent queries;
- Support complex, multistage queries;
- Run at a minimum 3,000 queries on a daily basis. Of these, it is anticipated that roughly 60% would be simple queries (e.g., pulling of all trades in a given symbol traded during a certain time period) and 40% would be complex (e.g., looking for quotes or orders more than 5% away from the National Best Bid and National Best Offer);
- Process and run approximately 1,800 queries concurrently;
- Support SQL 92 as well as recursive queries with common table expressions (recursive CTEs), bulk load utility, interface for dimension management, windowing functions, JBDC and ODBC, or provide another API with equal or greater query capabilities, so long as ODBC and JDBC are supported. Support for stored procedures and user-defined functions are optional;
- Include data presentation tools / query tools that support query results that produce data sets ranging from less than 1 gigabyte to at least 10 terabytes or more of uncompressed data;
- Provide query owners with the ability to schedule queries;
- Provide query owners with the ability to cancel a query during execution or prior to the scheduled running of a query;
- Provide Participants with a means to view all saved queries owned by the Participants as well as the scheduling of query executions (for queries that have been scheduled);

- Provide an automated delivery method of scheduled query results to the appropriate Participant. Delivery methods must comply with all information security guidelines (encryption, etc.);

- Provide technical expertise to assist regulators with questions and/or functionality about the content and structure of the CAT query capability;

- Include workload balancer to allow prioritization and processing of queries and delivery of results; and

- Support parallel processing of queries. At a minimum, the user-defined direct query tool must be able to process up to 300 simultaneous query requests with no performance degradation.

8.2.2 Bulk Extract Performance Requirements

For bulk extracts of an entire day of data, the minimum acceptable transfer time of equity and options data is four hours. This requirement assumes that there are no limitations within the regulator's own network environment that will prevent the Plan Processor from meeting this requirement.

A consideration was made to require an online Report Center that would include pre-canned reports that could be delivered to regulators or pulled upon request. The reports would be predefined based on requirements developed by Participants and the SEC. Due to the added complexity and the lack of quantifiable use cases, the Participants determined that this was something that may be useful in the future but not at the initial implementation and launch of the CAT. This will be reassessed when broker-dealers begin submitting data to the CAT.

It is envisioned that non-Participant CAT Reporters will be unable to access their data submissions through bulk data exports with the initial implementation of CAT. Only Participants and the SEC will have access to full lifecycle corrected bulk data exports.

Extraction of data must be consistently in line with all permissioning rights granted by the Plan Processor. Data returned must be encrypted, password protected and sent via secure methods of transmission. In addition, PII data must be masked unless users have permission to view the data that has been requested.

The Plan Processor must have an automated mechanism in place to monitor user-defined direct query usage. This monitoring must include automated alerts to notify the Plan Processor of potential issues with bottlenecks or excessively long queues for queries or data extractions. The Plan Processor must provide details as to how the monitoring will be accomplished and the metrics that will be used to trigger alerts.

The user-defined direct query and bulk extraction tool must log submitted queries and parameters used in the query, the user ID of the submitter, the date and time of the submission and the date and time of the delivery of results. The Plan Processor will use this logged information to

provide monthly reports to the Operating Committee, Participants and the SEC of their respective usage of the online query tool.

The bulk extract tool must support parallel processing of queries. At a minimum, the bulk extract tool must be able to process up to 300 simultaneous query requests with no performance degradation.

8.3 Identifying Latency and Communicating Latency Warnings to CAT Reporters

The Plan Processor will measure and monitor Latency within the CAT network. Thresholds for acceptable levels of Latency will be identified and presented to the Operating Committee for approval. The Plan Processor will also define policies and procedures for handling and the communication of data feed delays to CAT Reporters, the SEC, and Participants' regulatory staff that occur in the CAT. Any delays will be posted for public consumption, so that CAT Reporters may choose to adjust the submission of their data appropriately, and the Plan Processor will provide approximate timelines for when system processing will be restored to normal operations.

8.4 Technical Operations

The Plan Processor will develop policies, procedures, and tools to monitor and manage the performance of the Central Repository, to be approved by the Operating Committee. Such policies, procedures, and tools will include, at a minimum:

- Monitoring and management of system availability and performance, to include both Online Targeted Query Tool and User-Defined Direct Queries;
- Monitoring and management of query tool usage (e.g., to identify long-running or "stuck" queries); and
- Segregation of query queues by regulator or Participant (i.e., one regulator or Participant's queries should not prevent another regulator or Participant's queries from running).

8.5 System SLAs

Service Level Agreements for system and operational performance will be established for areas, including the following:

- Linkage and order event processing performance;
- Query performance and response times;
- System availability;
- User support/help desk performance;

- Application, network, and data security performance; and
- Development, change management, and implementation processes and timelines.

The actual terms of the SLAs will be negotiated between the Plan Participants and the eventual Plan Processor.

9. CAT Customer and Customer Account Information

9.1 Customer and Customer Account Information Storage

The CAT must capture and store Customer and Customer Account Information in a secure database physically separated from the transactional database. The Plan Processor will maintain information of sufficient detail to uniquely and consistently identify each Customer across all CAT Reporters, and associated accounts from each CAT Reporter. The following attributes, at a minimum, must be captured:

- Social security number (SSN) or Individual Taxpayer Identification Number (ITIN);
- Date of birth;
- Current name;
- Current address;
- Previous name; and
- Previous address.

For legal entities, the CAT must capture the following attributes:

- Legal Entity Identifier (LEI) (if available);
- Tax identifier;
- Full legal name; and
- Address.

The Plan Processor must maintain valid Customer and Customer Account Information for each trading day and provide a method for Participants' regulatory staff and the SEC to easily obtain historical changes to that information (e.g., name changes, address changes, etc.).

The Plan Processor will design and implement a robust data validation process for submitted Firm Designated ID, Customer Account Information and Customer Identifying Information, and must continue to process orders while investigating Customer information mismatches. Validations should:

- Confirm the number of digits on a SSN,
- Confirm date of birth, and
- Accommodate the situation where a single SSN is used by more than one individual.

The Plan Processor will use the Customer information submitted by all broker-dealer CAT Reporters to assign a unique Customer-ID for each Customer. The Customer-ID must be consistent across all broker-dealers that have an account associated with that Customer. This unique CAT-Customer-ID will not be returned to CAT Reporters and will only be used internally by the CAT.

Broker-Dealers will initially submit full account lists for all active accounts to the Plan Processor and subsequently submit updates and changes on a daily basis. In addition, the Plan Processor must have a process to periodically receive full account lists to ensure the completeness and accuracy of the account database. The Central Repository must support account structures that have multiple account owners and associated Customer information (joint accounts, managed accounts, etc.), and must be able to link accounts that move from one CAT Reporter to another (e.g., due to mergers and acquisitions, divestitures, etc.).

9.2 Required Data Attributes for Customer Information Data Submitted by Industry Members

At a minimum, the following Customer information data attributes must be accepted by the Central Repository:

- Account Owner Name;
- Account Owner Mailing Address;
- Account Tax Identifier (SSN, TIN, ITIN);
- Market Identifiers (Larger Trader ID, LEI);
- Type of Account;
- Firm Identifier Number;
    - The number that the CAT Reporter will supply on all orders generated for the Account;
- Prime Broker ID;
- Bank Depository ID; and
- Clearing Broker.

9.3 Customer-ID Tracking

The Plan Processor will assign a CAT-Customer-ID for each unique Customer. The Plan Processor will determine a unique Customer using information such as SSN and DOB for natural persons or entity identifiers for Customers that are not natural persons and will resolve discrepancies. Once a CAT-Customer-ID is assigned, it will be added to each linked (or unlinked) order record for that Customer.

Participants and the SEC must be able to use the unique CAT-Customer-ID to track orders from any Customer or group of Customers, regardless of what brokerage account was used to enter the order.

9.4 Error Resolution for Customer Data

The Plan Processor must design and implement procedures and mechanisms to handle both minor and material inconsistencies in Customer information. The Central Repository needs to be able to accommodate minor data discrepancies such as variations in road name abbreviations in searches. Material inconsistencies such as two different people with the same SSN must be communicated to the submitting CAT Reporters and resolved within the established error correction timeframe as detailed in Section 8.

The Central Repository must have an audit trail showing the resolution of all errors. The audit trail must, at a minimum, include the:

- CAT Reporter submitting the data;
- Initial submission date and time;
- Data in question or the ID of the record in question;
- Reason identified as the source of the issue, such as:
    - duplicate SSN, significantly different Name;
    - duplicate SSN, different DOB;
    - discrepancies in LTID; or
    - others as determined by the Plan Processor;
- Date and time the issue was transmitted to the CAT Reporter, included each time the issue was re-transmitted, if more than once;
- Corrected submission date and time, including each corrected submission if more than one, or the record ID(s) of the corrected data or a flag indicating that the issue was resolved and corrected data was not required; and
- Corrected data, the record ID, or a link to the corrected data.

10. User Support

10.1 CAT Reporter Support

The Plan Processor will provide technical, operational and business support to CAT Reporters for all aspects of reporting. Such support will include, at a minimum:

- Self-help through a web portal;
- Direct support through email and phone;
- Support contact information available through the internet; and
- Direct interface with Industry Members and Data Submitters via industry events and calls, industry group meetings and informational and training sessions.

The Plan Processor must develop tools to allow each CAT Reporter to:

- Monitor its submissions;
- View submitted transactions in a non-bulk format (i.e., non-downloadable) to facilitate error corrections;
- Identify and correct errors;
- Manage Customer and Customer Account Information;
- Monitor its compliance with CAT reporting requirements; and
- Monitor system status.

The Plan Processor will develop and maintain communication protocols (including email messaging) and a secure website to keep CAT Reporters informed as to their current reporting status, as well as issues with the CAT that may impact CAT Reporters' ability to submit or correct data. The website will use user authentication to prevent users for seeing information about firms other than their own, and will contain:

- Daily reporting statistics for each CAT Reporter,[271] including items such as:
    - SRO-Assigned Market Participant Identifier;
    - Date of submission;
    - Number of files received;
    - Number of files accepted;
    - Number of files rejected;
    - Number of total order events received;
    - Number of order events accepted;

[271] Each CAT Reporter or Data Submitter must only be able to view its own data and data it submits on behalf of others.

- Number of order events rejected;
- Number of each type of report received;
- Number of each type of report accepted;
- Number of each type of report rejected;
- Number of total customer records accepted;
- Number of total customer records rejected;
- Order-IDs rejected;
- Reason for rejection;
- Number of records attempted to be matched;
- Number of records matched;
- Percentage of records matched;
- Number of customer records received;
- Number of unknown accounts;
- Latest view of statistics inclusive of re-submissions to get a trade-date view of exceptions and correction statistics available for CAT Reporters to know when everything for a given trade date has been completed; and
- Most recent CAT Reporter Compliance Report Card, as defined in section 12.4;

- CAT System status, system notifications, system maintenance, and system outages; and

- A mechanism for submitting event data and correcting and resubmitting rejections or inaccurate data.

The Plan Processor will develop and maintain a public website containing comprehensive CAT reporting information, including:

- Technical Specifications;

- Reporting guidance (e.g., FAQs);

- Pending rule changes affecting CAT reporting;

- CAT contact information;

- Availability of test systems;

- Testing plans;

- Proposed changes to the CAT; and

- Fee schedule.

The Plan Processor will develop and maintain a mechanism for assigning CAT Reporter-IDs. A mechanism will also be developed and maintained to change CAT Reporter-IDs should this be necessary (e.g., due to a merger), with the expectations that such changes should be infrequent. Changes to CAT-Reporter-IDs must be reviewed and approved by the Plan Processor.

Initially, non-Participant CAT Reporters will not have access to their data submissions through bulk data exports with the initial implementation of the Central Repository. Only Participants and the SEC will have access to full lifecycle corrected bulk data exports. Non-Participant CAT Reporters will be able to view their submissions online in a read-only, non-exportable format to facilitate error identification and correction. Data Submitters will be able to export bulk file rejections for repair and error correction purposes.

The Plan Processor will define methods by which it will consult with and inform CAT Reporters and industry groups on updates and changes to user support.

The Plan Processor will define pre- and post-production support programs to minimize the Error Rate and help CAT Reporters to meet their compliance thresholds. Such pre-production support program shall include, but are not limited to, the following activities:

- Educational programs – Includes the following:
    - Publication and industry-wide communication (including FAQs) of the Technical Specifications, including:
        - Appropriate definitions / expected usages for each value in field format
        - All available attribute values for each field
    - Establishment of a dedicated help desk for Reporters to contact;
    - Industry participation in order linkage methodologies;
        - Include information on new order / trade types;
    - Hosting of industry educational calls; and
    - Hosting of industry-wide training.
- Registration – Requires all firms to:
    - Register and be certified as CAT Reporters;
    - Attend industry-wide training;
    - Establish internal controls to capture potential misreporting scenarios; and
    - Work with the Plan Processor to understand scenario-based reporting and expected outputs.
- Communications Plan – A strong communications plan of the timeline to reporting go-live shall:
    - Include communication on how Error Rates and Compliance Thresholds are calculated; and
    - Describe how errors will be communicated back to CAT Reporters.
- Industry-wide testing – Industry-wide test results must be available for all CAT Reporters.

- As mentioned in Appendix C, Objective Milestones to Assess Progress, appropriate time must be provided between Technical Specification publication and production go-live.
- Ample testing time must be provided.
- Appropriate scenario-based testing, including all three validation processes, shall be established.
- A separate test environment for CAT Reporters that mirrors the production environment shall be provided.

Post-production support program activities shall include, but are not limited to the following:

- Issuing a monthly Report Card on reporting statistics, with information on how reporters stand against similar entities;

- Publishing daily reporting statistics;

- Maintaining Technical Specifications with defined intervals for new releases/updates;

- Posting FAQs and other informational notices to be updated as necessary;

- Hosting of industry educational calls;

- Hosting of industry-wide training;

- Emailing outliers, meaning firms significantly reporting outside of industry standards;

- Conducting annual assessments of dedicated help desk to determine appropriate staffing levels;

- Using the test environment prior to releasing new code to production; and

- Imposing CAT Reporter requirements:

    - Attendance/participation of industry testing sessions;
    - Attendance in industry educational calls; and
    - Attendance in industry-wide training.

10.2 <u>CAT User Support</u>

The Plan Processor will develop a program to provide technical, operational and business support to CAT users, including Participants' regulatory staff and the SEC. The CAT help desk will provide technical expertise to assist regulators with questions and/or functionality about the content and structure of the CAT query capability.

The Plan Processor will develop tools, including an interface, to allow users to monitor the status of their queries and/or reports. Such website will show all in-progress queries/reports, as well as the current status and estimated completion time of each query/report.

The Plan Processor will develop communication protocols to notify regulators of CAT System status, outages and other issues that would affect Participants' regulatory staff and the SEC's ability to access, extract, and use CAT Data. At a minimum, Participants' regulatory staff and the SEC must each have access to a secure website where they can monitor CAT System status, receive and track system notifications, and submit and monitor data requests.

The Plan Processor will develop and maintain documentation and other materials as necessary to train regulators in the use of the Central Repository, including documentation on how to build and run reporting queries.

10.3 CAT Help Desk

The Plan Processor will implement and maintain a help desk to support broker-dealers, third party CAT Reporters, and Participant CAT Reporters (the "CAT Help Desk"). The CAT Help Desk will address business questions and issues, as well as technical and operational questions and issues. The CAT Help Desk will also assist Participants' regulatory staff and the SEC with questions and issues regarding obtaining and using CAT Data for regulatory purposes.

The CAT Help Desk must go live within a mutually agreed upon reasonable timeframe after the Plan Processor is selected, and must be available on a 24x7 basis, support both email and phone communication, and be staffed to handle at minimum 2,500 calls per month. Additionally, the CAT Help Desk must be prepared to support an increased call volume at least for the first few years. The Plan Processor must create and maintain a robust electronic tracking system for the CAT Help Desk that must include call logs, incident tracking, issue resolution escalation.

CAT Help Desk support functions must include:

- Setting up new CAT Reporters, including the assignment of CAT-Reporter-IDs and support prior to submitting data to CAT;
- Managing CAT Reporter authentication and entitlements;
- Managing CAT Reporter and third party Data Submitters testing and certification;
- Managing Participants and SEC authentication and entitlements;
- Supporting CAT Reporters with data submissions and data corrections, including submission of Customer and Customer Account Information;
- Coordinating and supporting system testing for CAT Reporters;
- Responding to questions from CAT Reporters about all aspects of CAT reporting, including reporting requirements, technical data transmission questions, potential changes to SEC Rule 613 that may affect the CAT, software/hardware updates and

upgrades, entitlements, reporting relationships, and questions about the secure and public websites;

- Responding to questions from Participants' regulatory staff and the SEC about obtaining and using CAT Data for regulatory purposes, including the building and running of queries; and

- Responding to administrative issues from CAT Reporters, such as billing.

10.4 CAT Reporter Compliance

The Plan Processor must include a comprehensive compliance program to monitor CAT Reporters' adherence to SEC Rule 613. The Chief Compliance Officer will oversee this compliance program, and will have responsibility for reporting on compliance by CAT Reporters to the Participants. The compliance program covers all CAT Reporters, including broker-dealers and Participants.

As a fundamental component of this program, the Plan Processor will identify on a daily basis all CAT Reporters exceeding the maximum allowable Error Rate established by the Participants. The Error Rate will initially be set by the CAT NMS Plan, and will be reviewed and adjusted on an ongoing basis by the Operating Committee. Error Rates will be based on timeliness, correctness, and linkages.

The Plan Processor will, on an ongoing basis, analyze reporting statistics and Error Rates and recommend to Participants proposed changes to the maximum allowable Error Rates established by the Participants. All CAT Reporters exceeding this threshold will be notified that they have exceeded the maximum allowable Error Rate and will be informed of the specific reporting requirements that they did not fully meet (e.g., timeliness or rejections).

The Plan Processor will develop and publish CAT Reporter compliance report cards on a periodic basis to assist CAT Reporters in monitoring overall compliance with CAT reporting requirements. The Plan Processor will also recommend criteria and processes by which CAT Reporters will be fined for inaccurate, incomplete, or late submissions. The compliance report cards will include the following information:

- Number of inaccurate transactions submitted;

- Number of incomplete transactions submitted; and

- Number of transactions submitted later than reporting deadlines.

The CAT Reporter compliance program will include reviews to identify CAT Reporters that may have failed to submit order events to the CAT, as well as to ensure CAT Reporters correct all identified errors even if such errors do not exceed the maximum allowable Compliance Threshold.

The Plan Processor will, on a monthly basis, produce and provide reports containing performance and comparison statistics as needed to each Participant on its members' CAT

reporting compliance thresholds so that Participants can monitor their members' compliance with CAT reporting requirements and initiate disciplinary action when appropriate. The Plan Processor will also produce and provide, upon request from the Participants and the SEC, reports containing performance and comparison statistics as needed on each CAT Reporter's compliance thresholds so that the Participants or the SEC may take appropriate action if a Participant fails to comply with its CAT reporting obligations.

The Plan Processor will produce and make available on a monthly basis reports for all CAT Reporters, benchmarking their performance and comparison statistics against similar peers. The reports will be anonymized such that it will not be possible to determine the members of the peer group to which the CAT Reporter was compared.

The Plan Processor will produce and make available to regulators on a monthly basis a report detailing Error Rates, transaction volumes, and other metrics as needed to allow regulators to oversee the quality and integrity of CAT Reporter reporting to the Central Repository.

11. Upgrade Process and Development of New Functionality

11.1 CAT Functional Changes

The Plan Processor must propose a process governing the determination to develop new functionality, which process must be reviewed and approved by the Operating Committee. The process must, at a minimum:

- Contain a mechanism by which changes can be suggested to the Operating Committee by Advisory Committee members, the Participants, or the SEC;
- Contain a defined process for developing impact assessments, including implementation timelines, for proposed changes; and
- Contain a mechanism by which functional changes which the Plan Processor wishes to undertake can be reviewed and approved by the Operating Committee.

The Plan Processor shall not unreasonably withhold, condition, or delay implementation of any changes or modifications reasonably requested by the Operating Committee.

11.2 CAT Infrastructure Changes

The Plan Processor must implement a process to govern changes to CAT. This process must contain provisions for:

- Business-as-usual changes (e.g., replacing failed hardware, adding capacity to deal with expected increases in transaction volumes) that would require the Plan Processor to provide the Operating Committee with a summary report (e.g., infrastructure changes, acquired costs, etc.); and
- Isolated infrastructure changes (e.g., moving components of the system from a self-hosted to an Infrastructure-as-a-Service provider) that would require the Plan

Processor to provide a request to the Operating Committee for review and approval before commencing any actions.

11.3 Testing of New Changes

The Plan Processor must implement a process governing user testing of changes to CAT functionality and infrastructure, which process must be reviewed and approved by the Operating Committee. The process must:

- Define the process by which changes are to be tested by CAT Reporters[and regulators];
- Define the criteria by which changes will be approved prior to their deployment into the production environment(s); and
- Define the environment(s) to be used for user testing.