

FINANCIAL INFORMATION FORUM

5 Hanover Square
New York, New York 10004

212-422-8568

Via Electronic Delivery

Thesys CAT
1740 Broadway
New York, N.Y.10010

March 29, 2018

Re: February 22, 2018 CAT Industry Member Reporting Specification, v0.2

Dear Thesys CAT,

The Financial Information Forum (FIF)¹ Consolidated Audit Trail Working Group (FIF CAT WG) has reviewed the referenced specification and has the following initial comments. FIF and its members will continue to review and provide further comment.

These new chapters (11 through 15) do not contain the expected level of detail and depth of information. Members require more detail both to assess the adequacy of the CAT Reporter interfaces and functions and to plan for the support structure that they will require to submit or repair reports. We were disappointed that the section on Connectivity (11.2) is empty.

A summary of FIF's major concerns with these new chapters are listed below. These concerns are discussed more fully within the Detailed Comments Appendix organized by the technical specification section where it was addressed.

- The SFTP-based submission process is cumbersome, exposes industry members to unnecessary complexity, and puts the burden of support on the CAT Reporter rather than imbedding more functionality into the Plan Processor.
- The process by which information is received/sent by the Plan Processor should be as automated as possible.
- All records in a submitted file should be scanned and identified with any errors so that any record in error can be corrected on the next repair file submission.
- All CAT Reporter test environments should have equivalent security to the production environment such that production data can be introduced into those environments.
- The repair process should not be tightly coupled with the originally submitted file. Repairs will not always be against the originally submitted file or can span multiple files.

FIF offers to discuss these comments with Thesys CAT at your earliest convenience.

¹ FIF (www.fif.com) was formed in 1996 to provide a centralized source of information on the implementation issues that impact the securities industry across the order lifecycle. Our participants include trading and back office service bureaus, broker-dealers, market data vendors and exchanges. Through topic-oriented working groups, FIF participants focus on critical issues and productive solutions to technology developments, regulatory initiatives, and other industry changes.

Sincerely,



Janet Early
Financial Information Forum

cc: Brett Redfearn, Director, Division of Trading and Markets, SEC
David Shillman, Associate Director, Division of Trading and Markets, SEC

Andre Frank, President, Thesys CAT
Michael Simon, Chair, CAT Operating Committee
Manisha Kimmel, Chair, CAT Advisory Committee
Alex Ravenel, Senior Manager, Deloitte
CAT Technical Team (Deloitte)

CAT NMS Plan Participants (BOX Options Exchange, CBOE BYX Exchange, CBOE BZX, CBOE EDGA Exchange, CBOE EDGX Exchange, CBOE C2 Exchange, CBOE Exchange, Chicago Stock Exchange, FINRA, Investors' Exchange, MIAX Options Exchange, MIAX PEARL, Nasdaq BX, Nasdaq GEMX, Nasdaq ISE, Nasdaq MRX, Nasdaq PHLX, New York Stock Exchange, NYSE Arca, NYSE American, NYSE National, The Nasdaq Stock Market

Appendix. Detailed Comments

11.1.1 File Names

Recommendations/Expected Actions

- FIF requests that the sequence number limit be expanded to at least a billion. Limiting the number of files that can be submitted to the CAT within any business day to 999,999 will be a problem for some firms, and seems to be an artificial limitation. Many broker-dealers are planning to send smaller files to the CAT throughout the day, in a “pseudo” real-time manner. In addition, this sequence number must also accommodate error correction files. We do not believe this is a sufficient size to accommodate this type of processing. We also do not think expanding the sequence number to a larger number should place any undue cost on the Plan Processor.
- FIF understands the need for a number in the file name to provide a unique qualifier within each day’s file submissions, but requests that the Thesys CAT dependency on the file sequence number determining the order of processing be eliminated. FIF believes this will introduce additional and unnecessary rejection of files during the ingestion process.
- There is a reference to Reporter IMID. Could the meaning for the term IMID (Industry Member ID) be defined in the specification?

11.1.2 Metadata Files

Recommendations/Expected Actions:

- FIF requests that metadata files also be supported in csv. CSV is a valid interface protocol to CAT and it seems unreasonable to support csv yet require that json be used for metadata files.
- FIF requests that the requirement to identify the last file transmitted for the day be eliminated. The field “IsKindDone” indicates the last file for this file kind for the day. Broker-dealers may not know the last file for the day. Or, even if they think they know the last file, some last minute change may arise requiring another file to be sent. We do not understand why this field is needed and do not believe it is practical for broker-dealers to support it.
- FIF requests that the requirement to transmit a metadata file with every file transmission be eliminated. Can Thesys CAT provide more information on why the metadata file is needed with every file transmission for integrity checking? Are there other methods that can be used, because it seems quite burdensome to require a metadata file for every report file submitted?

11.1.3. Data Files

Recommendations/Expected Actions:

- FIF requests that the limitations that Thesys CAT is imposing on record sizes be expanded. There is a limit defined for a line to be 4095 bytes and that each CAT record must be contained on one line. Is it correct that a CAT report must be contained within a CAT record? If that assumption is correct, FIF believes this is not a reasonable limit for a CAT record, especially in the long term. Given complex order definitions, and aggregated orders and use of 255 byte free form text fields, and capture of order handling instructions, etc., that this limit could be exceeded for a CAT report.

11.1.3.2 CSV Conversion

Recommendations/Expected Actions:

- The text and example should match. The text discusses a New Order event example, yet the example is a Trade Break event.

11.2 Connectivity

Questions:

- This section is blank. When will information be provided on the connectivity requirements to CAT?

11.3 Transport Options

Recommendations/Expected Actions:

- FIF requests that all data interfaces between CAT and CAT Reporters support a protocol that allows automation of that interface for upload or download of the data, including report statistics and system status.
- Using the same standard for both upload and report/feedback/statistics services would greatly simplify automation for industry members and will lead to more reliable implementations.
- RESTful APIs over a secure communications channel (e.g., HTTPS), with clear success/error states and data models for inputs/outputs are an industry standard for interoperability.
- The API specifications should include clearly defined success/error codes and input/output data models.
- This will greatly simplify the automation of the interactions between CAT Reporters and CAT Processor, thus leading to more reliable implementations.

11.3.1 File Size, Encryption and Compression

Questions:

- The file size limit is expressed in terms of file size before compression and encryption. Shouldn't the limit be expressed in terms of final file size before send to CAT?
- When files are uploaded, what checks/validations are performed? E.g., are the following checks performed?
 - a. Virus scanning
 - b. Checking the file for malicious executable
 - c. Limiting the length of file name
 - d. Limiting the interval for sending files

Recommendations/Expected Actions:

- FIF requests that the 1GB file limitation for files submitted via the Web Portal be expanded, so that the Web Portal could be used as an alternate channel if a CAT Reporter was experiencing difficulties with SFTP transmissions.

11.3.2 SFTP Upload Process

Recommendations/Expected Actions"

- FIF requests redefinition of the file upload process. The upload process is clumsy, requiring the upload of two files into the "upload/transit" directory, and then the move to the "upload/complete" directory. Thesys CAT should be able to manage this process, once the transmittal of the files to the "upload/transit" directory is complete, without requiring the two step process by the CAT Reporter.
- In general, any transmittal or message transmission to Thesys should always be accompanied with an acknowledgement of receipt by Thesys.
- There should never be a state with CAT that the CAT is not aware of the file state, even if the state is an error state. It should never be "unknown". We did not understand the statement "If

files are removed by the client, the processing behavior is undefined. Files may or may not undergo processing and may or may not receive receipts”.

11.4.1 CAT Feedback

Questions:

- FIF understands that error rates and calculations for error rates are the responsibility of the Operating Committee; however, we would like to register our questions:
 - If a record is corrected 30 days after T+1, and if that record was not originally identified as “in error”, is that corrected record counted as an error against the firm?
 - For mismatched data errors, which firm does the error get reported to, and which firm does it get “counted” against? When the final corrections are successfully submitted, are the errors properly correlated with the firm who caused the error?
 - When will the Web Portal API be published?

Recommendations/Expected Actions:

- FIF recommends removing the limit of stopping a file scan when more than 10% of the records are in error. Thesys CAT should always complete the scan and validation of an entire file. This is another case of an artificial limit imposed by Thesys CAT on broker-dealer processing. OATS does not have that limit today. This limit may cause a broker-dealer to be late in CAT reporting because remaining records in a rejected file will not have been processed and by definition, will be considered late. And then when the file is resubmitted, those unprocessed records may have errors, which will require another cycle of repair and resubmission.
- FIF requests that an API be provided to download the status reports and report statistics automatically, either via SFTP or the Web Portal.
- FIF requests that industry standard RESTful APIs be provided to download status reports, report statistics and system status.
- The API specifications should include clearly defined success/error codes and input/output data models.
- This will greatly simplify the automation of the interactions between CAT Reporters and CAT Processor, thus leading to more reliable implementations.

11.5 CAT Reporting Hours

Questions:

- If the Plan Processor does not meet its error reporting deadline, will the error correction deadline for CAT Reporters be extended to accommodate the missed deadline?

11.6 Security

Questions:

- Can additional information be provided on the multi-factor authentication that will be utilized?
- Will CAT Reporters be required to change on-line passwords, and if yes, what will be the frequency that will be enforced?
- Is this a classic rest API connection? Will this be between the CAT Reporter/Submitter and the Plan Processor?
 - If yes, please provide additional information on the authentication process for the API?
 - How is authorization performed?
 - How is input validation performed?

11.6.1 Data Security Standards

Questions:

- Can additional explanation be provided for the following statement “Upon retrieval by CAT, the CAT private key will be used to decrypt the file? Should the reporter desire to be able to decrypt the data at another time, the reporter should encrypt with both the reporter’s public key, as well as the CAT public key”?
- Can Thesys specifically define the “limited access roles” which will be supported for automated access? What functions can/cannot be done via automated roles?
- Can Thesys explain how the key exchange will work?
- What version of TLS is being used?

12. Feedback and Corrections

Questions:

- When will the details on the programmable API be published?

12.1 Feedback Files

Questions:

- When will the Reporter Web Portal interface and screens be published?
- What is the expected availability timing at each stage of processing? Can each processing stage be outlined in a table?
- What does this statement mean “other stages may take significant time, meaning that feedback could be delayed”? Does this mean that the error correction cycle for the CAT Reporter may be reduced? If there is a delay in providing error correction feedback, will the deadline for CAT Reporter submission for error corrections be extended in kind?

12.2.1 File Acknowledgement Feedback

Questions:

- What is the timeout period when waiting on a metafile?

12.3.2 Basic File Integrity Feedback

Questions:

- Will Thesys CAT provide all of the checks and edits it will be performing, so that Industry Members can duplicate those checks and do thorough testing before submission of the CAT report file? Could Thesys provide a test package that could be run by CAT Reporters/Submitters that would replicate the tests that the Plan Processor would perform on submission of a CAT report file?

12.4 Order Events Files

Questions:

- Much more detail is needed to understand the validations to be performed. E.g., what specific linkage checks will be performed, against what keys? What matching tests will be performed, if any? If mismatches (other than linkages) will not be performed by Thesys CAT, are these tests being dropped by regulators as not being required? Without matching criteria tests, won’t this degrade the quality of the data being collected?
- Can a scenario be provided which demonstrates the expected procedure for correcting linkage errors, especially when spanning two firms? Exactly what information will be provided in identifying the linkage error? Will both sides of the mismatched linkage be provided to the firm

identified “in error”? Will the other firm of the paired linkage be notified in any way of the linkage error? How will Thesys CAT decide which firm in a linkage error is at fault? Are both firms in a linkage error notified that the linkage was determined to be “in error”, with one of the two firms to be “in fault”? If that firm determines that the other firm is at fault, does it need to notify Thesys CAT that the error was mistakenly assigned to them? Are their error statistics corrected to remove that error?

- What errors on records will be classified as warnings, but processing continues and which errors will cause processing to stop?
- As per the document: *“There may be cases where the system has a preferred method for reporting some order events but will accept other methods (especially if there is a change or transition to something new). In such cases where the preferred method is detectable, but not detrimental, a warning may be generated to inform the reporter what is happening, but the record will still be accepted and processed by the system.”* This is unclear - it would be beneficial to have an example or additional details on what this refers to.

Recommendations/Expected Actions:

- As stated earlier, FIF objects to the rejection of a file after 10% of the records within the file have been rejected. FIF requests that this restriction be eliminated. OATS today performs error checking on the whole file, allowing the best opportunity for records to be included within in audit trail on time, and allowing any linkages to be completed. Stopping the processing of ingesting a file after 10% of the records are in error can mean that a good percentage of the records in the file will not have been scanned, so when the file is resubmitted after error correction, some of those records which were not scanned in the first submission will now be scanned for the first time and errors could be found. This will result in yet a second submission, which could have been prevented if the first file had been completely scanned. More late submissions, poorer linkage statistics on T+1, and higher error rates for CAT Reporters will all result due to this limitation. The goal should be to scan the entire file, ingesting as many records as possible, even if the file is marked as “rejected” because it had a greater than 10% error rate for records in the file.

12.4.1 Order Event Feedback

Questions:

- “..any order that is not fully linked or terminated (executed, canceled, or flagged as routed to a foreign destination) will be flagged as an error”. Can additional information, and examples, be provided on this termination process?

12.5 Corrections

Questions:

- Please define the difference between a correction and a repair.
- What exactly is the process when a correction is received beyond the error correction timeframe? Is that change stored in the database but not imbedded in the daisy chain of events for trade date? What happens if that correction affects the linkages relating to another firm?
- Do corrections count against the error rate for the day?

12.5.1 Repair Process

Recommendations/Expected Actions:

- FIF requests that the tight coupling in the repair process with the originally submitted file be removed. Repairs will not always be against the originally submitted file or can span multiple files.
- FIF recommends that each record in the originally submitted file be identified with a record identifier, and that the record identifier be the key used for repairs. This is done in OATS today. The identification of the repair record via an index in a file is very clumsy for processing. While submitting repairs (Corrections/Deletions) the file name of the correction file must match the file name in which the record was originally submitted. The specification further specifies that the index (line number) of the record in the original file must be referenced while submitting the correction. In order to fulfill this requirement a mapping of the record and the file in which record was submitted and index of the record in the file must be maintained. This seems unnecessary.
- FIF recommends that a file can contain records from multiple business days. For example, in today's OATS world we frequently have a scenario where a trail was missed in a file. When we receive an Order Trail from the previous day we simply club the records with the current day processing and it is submitted to OATS. With the requirement that Thesys CAT is proposing, all records for each business days must be separately processed so that they can be written to businessDate specific file or the timestamp of each order that comes in would need to be check and a separate file created whenever a new date is encountered. Thesys CAT could easily do this processing by removing the restriction that the business date on the records within the file must match the business date in the file name.
- FIF recommends that if an invalid record type is found in a repair file, that the invalid record be rejected but that the file is not rejected and all other repair records in that file should be processed.
- The example specified for the repair file name does not appear to be correct (see section 12.5.1 on page 28). Based on the details in document the base name of an original file has to be **<CAT Reporter ID>_<Date>_[<Group>_]<FileKind>_<File Number>** e.g. **MYID_20170101_OrderEvents_000123** and base name for the correction for same file should be **MYID_20170101_OrderEventsCorrections_000123**. However, the example specifies: **MYID_20170101_OrderEvents_000123** (original file) and **MYID_20170101_OrderEventsCorrections_000001** (correction file name).

12.5.1.2 Correction Record

Questions:

- Will there be any accommodation for shorthand in the interface? E.g., in json, for capacity, could A be specified for Agency and P for Principal? If yes, can the long and short versions of each parameter be included in the specification?

13. Testing

Questions:

- Very little information has been provided about the test environment, test tools or testing services that will be provided by Thesys. When will additional details be provided about the test environment?
- Will Thesys CAT be supporting coordinated industry testing, including both CAT Reporters and exchanges in a planned and coordinated industry-wide testing procedure?

Recommendations/Expected Actions:

- . It is imperative that both function/system test environments for CAT Reporters, as well as UAT environments have the same security procedures and infrastructure as production. What exactly is meant by “Please note that reporters should not submit highly sensitive data, such as PII data, into the testing environment”? Does this mean that CAT Reporters cannot submit production data into the test environment? This is a very normal practice; for example, to verify repairs, a production repair file will be tested in a test environment before submission to CAT – the very point being to use exactly the data to be submitted in production to verify that the repairs are correct

14. CAT Help Desk and Support

Recommendations/Expected Actions:

- As mentioned earlier, FIF requests that Thesys CAT provide an automated mechanism for the retrieval and processing of reporting statistics and system status.
- FIF requests that MPID or Group (from the file name) be added as a filter for reporting statistics. Today, FINRA statistics can be viewed by MPID, This will allow a better level of granularity for understanding their CAT performance making adjustments/corrections for a firm.

15.1 Failure Codes

Questions:

- When will a complete list of warning and failure codes be provided?
- If a failure for the Metadata file (Failure code META) is reported, can only the metadata file be resubmitted or must the order event file associated with this metadata file also be resubmitted?