

DSB Production Root Cause Analysis

For Production Outage on 2nd & 3rd Jan 2018

Prepared by: Technical.Support@anna-dsb.com
Date: 29th January 2018

TABLE OF CONTENTS

Impact assessment & Categorization 2
 Introduction..... 2
 Executive Summary - Findings and Root Cause 2
 Corrective Actions Taken 3
 Event Description..... 3
 Chronology of Events / Timeline..... 4

Revision History

Version	Date	Reason
1.0	29 th January 2018	Document reviewed and finalized

IMPACT ASSESSMENT & CATEGORIZATION

Significant Impact (Severity Two - S2)

Critical loss of application functionality or performance resulting in high number of users unable to perform their normal functions. Major feature/product failure; inconvenient workaround or no workaround exists. The program is usable but limited.

Condition 1: A key component of the solution, an application across all users, a set of users or intermittent network degradation or instability leading to performance and degradation of service.

Condition 2: An incident which is not yet S1, but might lead to a potential S1 incident.

Condition 3: Partial users at a particular location are affected but not all the users in all locations

INTRODUCTION

The purpose of this Root Cause Analysis (RCA) is to determine the causes that contributed to the recent “Insufficient Privileges” and “Something went wrong” error messages encountered by some clients in the DSB production environment on 2nd January 2018 at 5:49 am UTC and 3rd January 2018 at 5:00 am UTC. This RCA determines what happened during the event, how it happened, and why it happened. To accomplish this, an investigation took place internally between the DSB support, Development teams and senior analysts to ascertain the primary root cause or a list of root causes that contributed to this issue.

EXECUTIVE SUMMERY - FINDINGS AND ROOT CAUSE

Tuesday 2nd January 2018

This was due to capacity issues on SOLR (<http://lucene.apache.org/solr/>) as DSB participants connected to the system at the same time. The fix for this was scheduled for the upcoming weekend but due to the outage on the 3rd it was brought forward as an emergency break/fix.

“Insufficient Privileges” occurred because the high CPU on SOLR machines caused a significant slowdown in SOLR responses to other DSB services, resulting in internal timeouts and leaving some ISIN creation requests in an intermediate state.

Wednesday 3rd January 2018

It was noted at 5am UTC that the SOLR services experienced high memory usage and the OS had restarted the SOLR processes on two servers. The service recovered and there was no outage. The decision was made to perform a rolling upgrade of the SOLR servers to higher capacity servers.

After the upgrade of the SOLR servers it was noted during the validation that the Cordra (<https://cordra.org/>) services used for workflow management were producing intermittent errors with some of the connections to the SOLR services. A rolling restart of the Cordra servers remediated this issue, implemented in such a way as to ensure the service remained available during the restart period.

An after effect of the timeouts experienced during the high-load on SOLR services was that some ISIN creation requests were left in an intermediate state. The result was that multiple ISIN creation requests sent at around the same time for the same instrument opened up a race condition that allowed some duplicate ISINs to be created.

CORRECTIVE ACTIONS TAKEN

- 2nd Jan SOLR upgrade planned if system goes down.
- 3rd Jan Upgrade the capacity of the SOLR server.
- A mitigation plan is being put in place to deal with the duplicate ISIN issue should it occur during any future periods of SOLR capacity constraints.
- Alerting criteria for monitoring changed to give preemptive warning on SOLR capacity.

EVENT DESCRIPTION

On 2nd January 2018 at 5:49 am UTC, the production environment during start up experienced a very high load from new connections that caused extremely high CPU on machines running the Apache SOLR service. This caused internal timeouts between the SOLR service and other components of the service, which resulted in intermittent “Insufficient Privileges” error messages for some clients submitting FIX or ReST messages for ISIN’s previously requested between 2nd Jan 5:49 UTC and 2nd Jan 5:51 UTC. The load on the system subsided by 7:53 UTC, after which normal service resumed.

“Something went wrong” error messages were intermittently experienced by clients when searching or creating ISIN’s due to the Cordra service not connecting to the Solr services correctly. This was due to the Solr service being unresponsive during this period.

During this period, GUI users experienced intermittent disconnections and the ISIN creation and search service was available on an intermittent basis only.

As a result of the high load experienced that day, the DSB technology team planned the implementation of an upgrade procedure for the hardware infrastructure of the SOLR services, scheduled for the following weekend of 6th Jan.

On 3rd January 2018 at 5:00 am UTC, the production environment experienced a similarly high load on the SOLR service as a result of an unprecedented number of connections and request messages. The decision was made to immediately implement the hardware upgrade of the SOLR services prior to its scheduled execution on the weekend of 6th Jan. A permission update was also scheduled to address the “Insufficient Privileges” issue. The hardware upgrade and permission update were both implemented and validated to be working by 8:00am UTC.

During this period, GUI users experienced intermittent disconnections and the ISIN creation and search service was available on an intermittent basis only.

CHRONOLOGY OF EVENTS / TIMELINE

Tuesday 2nd January 2018

5:49 AM UTC – Tuesday 2nd January 2018

DSB Technical team start investigation. SOLR applications were identified as high memory usage >80% and high CPU usage which were not normalizing as expected

5:51 AM UTC – Tuesday 2nd January 2018

SOLR services restarted on an individual rolling basis so as to provide continuous availability of the service to users (albeit on an intermittent basis)

7:53 AM UTC -Tuesday 2nd January 2018

Production system fully validated as working normally.

Wednesday 3rd January 2018

5:00 AM UTC – Wednesday 3rd January 2018

SOLR services identified as running at high capacity.
Decision was made to proceed rolling upgrade capacity of servers.

5:10 AM UTC – Wednesday 3rd January 2018

Rolling capacity upgrade starts so as to provide continuous availability to users (albeit on an intermittent service basis)

6:05 AM UTC – Wednesday 3rd January 2018

Upgrade of SOLR servers complete

6:22 AM UTC – Wednesday 3rd January 2018

Permission change to fix “insufficient privileges” starts

6:41 AM UTC – Wednesday 3rd January 2018

Issue identified with the CNRI Cordra services from interacting with SOLR causing intermittent failures for some user requests
Investigation started

7:01 AM UTC – Wednesday 3rd January 2018

Decision to rolling restart x9 Cordra servers enacted

8:00 AM UTC – Wednesday 3rd January 2018

Cordra rolling restart completed and system fully validated.