InfoSphere Data Replication – CDC

Troubleshooting
Agenda

- How to approach a problem
- Collecting error information
- Troubleshooting Resources
- Questions
How to approach a problem
Understanding CDC Architecture

- Database Logs
- Source Engine
- Target Engine
- TCP/IP
- Monitoring and Configuration

- Database
- Web Services
- Message Queue
- Business Process
- Flat files
- Information Server

Oracle, DB2, SQL Server, etc
Identify Type of Problems

1) Communication
2) Security/Permissions
3) Database
4) System Problems
5) Application Problems
6) User Error
Collecting Diagnostic Information
Collecting information to Diagnose the problem

- Event logs
- System logs
- Product Tracing
- Communications Tracing
- Database level Tracing
- Support Assistant
- Management Console Debug Trace
- CDC Statistics
1) Check the Event Logs

- Source and Target side
- Subscription Level events
- Datastore level events

<table>
<thead>
<tr>
<th>Type</th>
<th>Message</th>
<th>Event ID</th>
<th>Object</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>Subscription SUB1 has started using the single scope staging store...</td>
<td>2922</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:21 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_1.1</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_1.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_3.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_4.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_5.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_6.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_7.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_8.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_9.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Mirroring has been initiated for table DFS0000Y.TABLE_10.2</td>
<td>44</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Communication to SUB1 on target host localhost is starting on the Data channel.</td>
<td>1221</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:22 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Subscription SUB1 is starting in Continuous Mirroring mode.</td>
<td>1463</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:17 AM</td>
</tr>
<tr>
<td>Information</td>
<td>The target reported time zone America/New York, IBM InfoSphere Change Data Capture will use source time zone America/New York, IBM InfoSphere Change Data Capture will use.</td>
<td>1591</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:19 AM</td>
</tr>
<tr>
<td>Information</td>
<td>Communication to SUB1 on target host localhost is starting on the Control channel.</td>
<td>1705</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:19 AM</td>
</tr>
<tr>
<td>Information</td>
<td>IBM InfoSphere Change Data Capture is starting.</td>
<td>52</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:19 AM</td>
</tr>
<tr>
<td>Information</td>
<td>IBM InfoSphere Change Data Capture is ending.</td>
<td>63</td>
<td>SUB1</td>
<td>Aug 10, 2010 9:59:14 AM</td>
</tr>
</tbody>
</table>
2) System Logs

- iSeries job log
- SQL Server application logs
- Netezza nzload, nzbad, pg.log and dbos.log
3) IIDR Engine Tracing (version 10.2 onwards)

- Significantly faster product tracing with less impact on system (< 70%)
- There are 2 types of trace points 1) ASCII readable format 2) Binary format
- Binary format is typically used by IBM support only and needs to be decoded into translated string data
- To decode your trace file:
  - `$CDC_HOME/bin/dmdecodetrace
    $CDC_HOME/instance/<instance_name>/log/on/trace* > decodedtrace.log`
- To enable tracing:
  - System parameter `global_trace_hours` (valid values 0 to 99) determine how many hours to trace data
  - The system parameter will adjust itself as time goes by
  - Minutes are optional (e.g. 1:22) and intended for smooth countdown
  - Longer use requires re-confirmation every 4 days
IIIDR Engine Tracing cont'd

- Unconditional tracing is produced regardless of the trace setting
- Conditional tracing is produced when tracing is enabled
- When trace is enabled, the number of trace points is big and normally the trace files wrap around very quickly. In these scenarios, we might be losing important information from the unconditional trace stream.
- Therefore, there are two streams of traces in different directories:
  - $CDC_HOME/instance/<instance_name>/log stores unconditional trace
  - $CDC_HOME/instance/<instance_name>/log/on stores both unconditional and conditional traces.

- * Line numbers are shared in between unconditional and conditional trace points.
How to utilize Trace information

- Trace logs contain valuable information that is often overlooked when troubleshooting issues in the field
- Exceptions or error messages may be logged in the trace
- Useful to 'grep' for keywords such as 'exception', 'error', or specific subscription or table name
- Install level trace is found in
  - `<CDC install library>/log`
- Specific instance level trace is found in
  - `<CDC install library>/instance/<name of instance>/log`
4) Communications Tracing

- You need to ensure that there is sufficient disk space to contain a longer run.
- To enable, edit the comms.ini file by removing the ";" at the beginning of the ";name=commtrace.log" line to get comms traces. The comms.ini file can be found in /instance//conf/comms.ini. Note that for this change to take affect, you are required to restart the CDC instance.

- Comms file:
  - [LOG]
    ;name=commtrace.log
    stdout=no

  [SETTING]
  KEEP_ALIVE_TIMEOUT=20
  TCP_KEEP_ALIVE = 60
  LINK_CYCLE_TIMEOUT = 30
  SESSION_CLOSE_TIMEOUT = 30
Comms Issues

- To address additional communications issues you can adjust the following comms parameters in the comms.ini to increase timeout values.

  KEEP_ALIVE_TIMEOUT = 20
  TCP_KEEP_ALIVE = 60
  LINK_CYCLE_TIMEOUT = 30
  SESSION_CLOSE_TIMEOUT = 30
5) Database level Tracing

- Each database platform may have their own tools, settings used to generate database level tracing. Consult database documentation for more details.
- Running `tkprof` on the output generated in the Oracle `udump` directory will display the amount of time performing operations and waiting. Ideal to detect slow tables on the target.
- The product allows you to enable tracing for the Oracle platform for the following connections:
  - Target Apply
    - `mirror_apply_db_trace = true` (to enable, false by default)
  - Source Scrape
    - `mirror_scrape_db_trace = true` (to enable, false by default)
6) Support Assistant (Management Console)

- Management Console and the LUW engines provide support for collecting diagnostic information that assists support and development in identifying the cause of problems
  - Simplifies the process of turning on the right level of tracing
  - Enhances the Support Assistant dialog
  - Engine supports collecting trace information from the command-line

It is important to specify a range to collect detailed information including performance stats, traces, etc.
Support Assistant (Command line)

- **Command Line Usage:**
  - dmsupportinfo [-I <instance name>] [-L <locale>] [-t]
  - -I <Instance name>
  - -L locale if different from user locale (e.g. en_ca)
  - -t collect only in this time range "yyyy-mm-dd hh:mm:ss to yyyy-mm-dd hh:mm:ss". Note that the quotes are required, along with all date-time elements. Use of a time range is optional, but will reduce the size of the output file considerably. Also, a time range must be specified if you want performance statistics in order for IBM to perform performance analysis.
7) Management Console Debug Tracing

**IBM InfoSphere Change Data Capture Management Console**

- **Trace Options**
  - **Management Console Tracing**
    - Capture communication between Management Console, Access Server, and datastores.
    - Enable Management Console message capture
  - **Access Server Tracing**
    - Access Server diagnostic collection affects configuration and monitoring performance and should only be enabled for problem identification. The collection does not affect engine performance.
    - Enable Access Server message capture
    - Enable Access Server logging
    - Disable Access Server tracing and logging

© 2016 IBM Corporation
8) Performance Statistics (LUW engines)

- The CDC Java Engine provides real-time statistics enabled by the stats_collect system parameter.
  - Collected by default with a 30 second sampling interval. Use:
    • 10 seconds for small tests (under an hour)
    • 30 seconds (default) good for an all day test
    • 120 seconds to identify major issues (monthly)
  - Collected on a per subscription basis
  - Independent collection on both the source and target systems
  - Low impact on customer system to collect the data
  - Collected by dmsupportinfo when the –t option is specified
  - Upload dmsupportinfo output to a PMR for IBM support to provide analysis
Performance Statistics (z/OS)

- To enable SMF data:
  - Configure SMFINTERVAL=<interval in minutes>. For example, 1, may as well collect every minute which is the most frequent currently available. The overhead is negligible.
  - SMFTYPE needs to be in the range 128-255. This is a z/OS requirement for applications. Values less than 128 are reserved for z/OS.

- You can confirm the actual settings by looking at CHCAUDIT in the product's spooled output.

  013/05/16 11:53:15 OSC CHC9105I SMF Record Type . . . . 200
  013/05/16 11:53:15 OSC CHC9105I SMF Record Interval . . 1 min.

- There are 4 types of CDC SMF records (SYSTEM, LOG CACHE, SOURCE, TARGET)
- CDC will always write SYSTEM records and LOG CACHE (if enabled)
- SOURCE and TARGET records are logged only when subscriptions are running
Additional Diagnostic information
Diagnosing a Communications Problem

1. Check the event logs on both source and target side
   1. Compare system times to correlate the events between both sides

2. Common Network Problems
   1. Firewall or VPN gateway closing idle connections – Enable TCP_KEEPALIVE_SECS parameter with short enough time
   2. Timeout delivering packets due to TCP re-transmissions (latency, congestion) – use a packet sniffer capture to investigate
   3. Firewall blocking connections – ensure ports are opened for CDC
   4. Missing DNS configuration while using host names – verify DNS configuration
Network Requirements

- A fully inclusive TCP/IP network path with adequate network bandwidth that connects the source and target installations of CDC
- Firewalls or network tools that do not interfere with or close CDC communication ports
- CDC user exits or notifications require reliable connections to the database or other applications such as e-mail servers
- Reliable network connections between the source or target deployments of CDC and Access Server.
- Sufficient network bandwidth when the product is configured to read remote database logs with a networked file system
How to verify if there is a network problem

- Check if target is reachable from source with ‘ping <target>’
- Check the next level of connection, from the source issue ‘telnet <target><CDC target-listener port>’
- Get output of ‘netstat –na’ from both source and target
- CDC comms trace can be enabled – remove the “;” at the beginning of “;name=commtrace.log”
- Run a test again with tcpdump or iptrace set up and send IBM support the tcp dump for investigation
What resource constraints to look for

- **Memory**
  - Ensure that an appropriate amount of resident memory is available to CDC. You can use ‘top’ (‘topas’ on AIX) command and look at the RES column.

- **CPU**
  - Can be an issue if the CPU on the system is maxed out and CDC is not getting reasonable time slices. On Unix, use the ‘nice’ command and set the priority higher. To modify the priority for a running process you would use ‘renice’.

- **Disk**
  - CDC stages data to disk if there is not enough memory to hold large transactions and is dependent on the disk speed that the source and target database are running on. Useful to have a periodic monitoring of system resources over a period of time. You can use tools such as vmstat, iostat, sar.
What to check for

- **Permissions**
  - Verify that the Oracle database user has sufficient privileges. A sample script of privileges is included in the product (createora-user.sql)
  - Verify that the user has permissions to the redo and archive logs

- **Verify listener is running**
  
  netstat –an | grep <portNumberUsedForListener>

  For example, if the port number is 52001 then *right after starting the listener* you would see the following:

  % netstat –an | grep 52001
  tcp 0 0 0.0.0.0:52001 0.0.0.0:* LISTEN

- **Verifying sufficient disk space**
  - Run the Unix df command to determine disk space availability
  - Recovery: Free up disk space if required and restart CDC instance
**Resources**

- Please refer to these resources FIRST before logging a PMR
- Use the resources provided by the *product* mentioned in this presentation
- Read documentation available on Knowledge Center
- Check if it's already been fixed. Fixes listed in the DeveloperWorks Blog
- Check Technology and DeveloperWorks
Backup
Old release details

- **CDC Engine Tracing (version 6.x)**
  - Enabling tracing in V6 has impact on performance. Do not enable tracing when troubleshooting performance issues.
  - Remember to disable tracing after diagnosing issue
  - **Enabling Tracing**
    - `global_trace_until= <mm/dd/yy>, <mm/dd/yy hh:mm AM/PM>`
    - e.g: 1/14/08, or 1/14/08 1:34 PM (pay attention to spaces). This action will turn the trace on until January 14, 2008 (12AM in the first case, and 1:34pm in the second case). A restart of the instance is not required.
    - Specify parameter in Management Console’s system parameter window OR
    - Via Command Line utility
      - `dmset -I <INSTANCE_NAME> [<parameter_name]=[<parameter_value>]] [-L <locale>]`
  - **Enabling via vmargs files**
    - If you want to enable tracing before the datastore has been configured via Management Console or Command line utility, you can specify the tracing parameters in the *.vmargs files found in the /<install folder>/conf directory
    - For example, in the configuration tool,
      - `Add -Dcom.datamirror.trace.until= <mm/dd/yy>, to the beginning of dmconfigurets.vmargs in the conf directory`
Old release details

- **CDC Engine Tracing (version 6.x) (cont’d)**
  - Disabling Tracing
    - Using either MC or dmset command-line tool delete `global_trace_until` property or set it to an expired date. No need to restart ICDC
  - Location of Trace files
    - Trace files for ICDC can be found in `/log` and `/instance//log` directories. The files are text files and should be compressed before sending over networks.
  - Increasing Trace File sizes
    - By default trace files are limited to 1Mb each and 5 Mb of disk space in total. When individual file limit is reached, a new file is created. When the total limit is reached, older files are deleted. To change these numbers, modify the following properties: `global_trace_files_each_mb` and/or `global_trace_files_total_mb`. No need to restart ICDC.
Legal Disclaimer

• © IBM Corporation 2016. All Rights Reserved.
• The information contained in this publication is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this publication, it is provided AS IS without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this publication or any other materials. Nothing contained in this publication is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.
• References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. Product release dates and/or capabilities referenced in this presentation may change at any time at IBM’s sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.
• If the text contains performance statistics or references to benchmarks, insert the following language; otherwise delete:
  Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user’s job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.
• If the text includes any customer examples, please confirm we have prior written approval from such customer and insert the following language; otherwise delete:
  All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.
• Please review text for proper trademark attribution of IBM products. At first use, each product name must be the full name and include appropriate trademark symbols (e.g., IBM Lotus® Sametime® Unyte™). Subsequent references can drop “IBM” but should include the proper branding (e.g., Lotus Sametime Gateway, or WebSphere Application Server). Please refer to http://www.ibm.com/legal/copytrade.shtml for guidance on which trademarks require the ® or ™ symbol. Do not use abbreviations for IBM product names in your presentation. All product names must be used as adjectives rather than nouns. Please list all of the trademarks that you use in your presentation as follows; delete any not included in your presentation. IBM, the IBM logo, Lotus, Lotus Notes, Notes, Domino, Quickr, Sametime, WebSphere, UC2, PartnerWorld and Lotusphere are trademarks of International Business Machines Corporation in the United States, other countries, or both. Unyte is a trademark of WebDialogs, Inc., in the United States, other countries, or both.
• If you reference Adobe® in the text, please mark the first use and include the following; otherwise delete:
  Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
• If you reference Java™ in the text, please mark the first use and include the following; otherwise delete:
  Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
• If you reference Microsoft® and/or Windows® in the text, please mark the first use and include the following, as applicable; otherwise delete:
  Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.
• If you reference Intel® and/or any of the following Intel products in the text, please mark the first use and include those that you use as follows; otherwise delete:
  Intel, Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
• If you reference UNIX® in the text, please mark the first use and include the following; otherwise delete:
  UNIX is a registered trademark of The Open Group in the United States and other countries.
• If you reference Linux® in your presentation, please mark the first use and include the following; otherwise delete:
  Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both. Other company, product, or service names may be trademarks or service marks of others.
• If the text/graphics include screenshots, no actual IBM employee names may be used (even your own), if your screenshots include fictitious company names (e.g., Renovations, Zeta Bank, Acme) please update and insert the following; otherwise delete: All references to [insert fictitious company name] refer to a fictitious company and are used for illustration purposes only.