Abstract

This is the MySQL NDB Cluster Internals Manual, which contains information about the NDBCLUSTER storage engine that is not strictly necessary for running the NDB Cluster product, but can prove useful for development and debugging purposes. Topics covered in this Guide include, among others, communication protocols employed between nodes, file systems used by management nodes and data nodes, error messages, and debugging (DUMP) commands in the management client.

The information presented in this guide is current for recent releases of NDB Cluster up to and including NDB Cluster 8.0.18, now under development. Due to significant functional and other changes in NDB Cluster and its underlying APIs, you should not expect this information to apply to previous releases of the NDB Cluster software prior to NDB Cluster 7.2. Users of older NDB Cluster releases should upgrade to the latest available release of NDB Cluster 7.6, currently the most recent GA release series.

For more information about NDB 8.0, see What is New in NDB Cluster. For information regarding NDB 7.6, see What is New in NDB Cluster 7.6.

For legal information, see the Legal Notices.

For help with using MySQL, please visit the MySQL Forums, where you can discuss your issues with other MySQL users.

Third-party licensing information. This product may include third-party software, used under license. If you are using a Commercial release of NDB Cluster 7.5, see the MySQL NDB Cluster 7.5 Commercial Release License Information User Manual for licensing information relating to third-party software that may be included in this Commercial release. If you are using a Community release of NDB Cluster 7.5, see the MySQL NDB Cluster 7.5 Community Release License Information User Manual for licensing information relating to third-party software that may be included in this Community release.

Licensing information—MySQL NDB Cluster 7.6. If you are using a Developer Preview release of NDB Cluster 7.6, see the MySQL NDB Cluster 7.6 Community Release License Information User Manual for licensing information relating to third-party software that may be included in this Community release.

Document generated on: 2019-06-30 (revision: 62510)
# Table of Contents

Preface and Legal Notices ........................................................................................................... vii

1 NDB Cluster File Systems ........................................................................................................ 1
  1.1 NDB Cluster Data Node File System ................................................................................. 1
    1.1.1 NDB Cluster Data Node Data Directory Files ......................................................... 1
    1.1.2 NDB Cluster Data Node File System Directory Files ........................................... 2
    1.1.3 NDB Cluster Data Node Backup Data Directory Files ......................................... 3
    1.1.4 NDB Cluster Disk Data Files ................................................................................... 3
  1.2 NDB Cluster Management Node File System .................................................................. 4

2 NDB Cluster Management Client DUMP Commands ............................................................... 5
  2.1 DUMP 1 ........................................................................................................................... 8
  2.2 DUMP 13 .......................................................................................................................... 10
  2.3 DUMP 14 .......................................................................................................................... 11
  2.4 DUMP 15 .......................................................................................................................... 11
  2.5 DUMP 16 .......................................................................................................................... 11
  2.6 DUMP 17 ........................................................................................................................... 12
  2.7 DUMP 18 ........................................................................................................................... 12
  2.8 DUMP 20 ........................................................................................................................... 13
  2.9 DUMP 21 ........................................................................................................................... 13
  2.10 DUMP 22 ......................................................................................................................... 14
  2.11 DUMP 23 ......................................................................................................................... 14
  2.12 DUMP 24 ......................................................................................................................... 15
  2.13 DUMP 25 ......................................................................................................................... 16
  2.14 DUMP 70 ......................................................................................................................... 16
  2.15 DUMP 400 ......................................................................................................................... 16
  2.16 DUMP 401 ......................................................................................................................... 17
  2.17 DUMP 402 ......................................................................................................................... 18
  2.18 DUMP 403 ......................................................................................................................... 18
  2.19 DUMP 406 ......................................................................................................................... 19
  2.20 DUMP 908 ......................................................................................................................... 19
  2.21 DUMP 1000 ...................................................................................................................... 20
  2.22 DUMP 1223 ...................................................................................................................... 20
  2.23 DUMP 1224 ...................................................................................................................... 21
  2.24 DUMP 1225 ...................................................................................................................... 21
  2.25 DUMP 1226 ...................................................................................................................... 21
  2.26 DUMP 1228 ...................................................................................................................... 22
  2.27 DUMP 1229 ...................................................................................................................... 22
  2.28 DUMP 1332 ...................................................................................................................... 22
  2.29 DUMP 1333 ...................................................................................................................... 23
  2.30 DUMP 2300 ...................................................................................................................... 24
  2.31 DUMP 2301 ...................................................................................................................... 24
  2.32 DUMP 2302 ...................................................................................................................... 26
  2.33 DUMP 2303 ...................................................................................................................... 26
  2.34 DUMP 2304 ...................................................................................................................... 26
  2.35 DUMP 2305 ...................................................................................................................... 29
  2.36 DUMP 2308 ...................................................................................................................... 29
  2.37 DUMP 2315 ...................................................................................................................... 30
  2.38 DUMP 2350 ...................................................................................................................... 30
  2.39 DUMP 2352 ...................................................................................................................... 31
  2.40 DUMP 2354 ...................................................................................................................... 32
  2.41 DUMP 2398 ...................................................................................................................... 32
  2.42 DUMP 2399 ...................................................................................................................... 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.97 DUMP 7020</td>
<td>58</td>
</tr>
<tr>
<td>2.98 DUMP 7021</td>
<td>58</td>
</tr>
<tr>
<td>2.99 DUMP 7024</td>
<td>60</td>
</tr>
<tr>
<td>2.100 DUMP 7027</td>
<td>60</td>
</tr>
<tr>
<td>2.101 DUMP 7033</td>
<td>60</td>
</tr>
<tr>
<td>2.102 DUMP 7080</td>
<td>61</td>
</tr>
<tr>
<td>2.103 DUMP 7090</td>
<td>61</td>
</tr>
<tr>
<td>2.104 DUMP 7098</td>
<td>61</td>
</tr>
<tr>
<td>2.105 DUMP 7099</td>
<td>62</td>
</tr>
<tr>
<td>2.106 DUMP 7901</td>
<td>62</td>
</tr>
<tr>
<td>2.107 DUMP 8004</td>
<td>62</td>
</tr>
<tr>
<td>2.108 DUMP 8005</td>
<td>63</td>
</tr>
<tr>
<td>2.109 DUMP 8010</td>
<td>63</td>
</tr>
<tr>
<td>2.110 DUMP 8011</td>
<td>63</td>
</tr>
<tr>
<td>2.111 DUMP 8013</td>
<td>64</td>
</tr>
<tr>
<td>2.112 DUMP 9002</td>
<td>65</td>
</tr>
<tr>
<td>2.113 DUMP 9800</td>
<td>65</td>
</tr>
<tr>
<td>2.114 DUMP 9801</td>
<td>65</td>
</tr>
<tr>
<td>2.115 DUMP 9802</td>
<td>66</td>
</tr>
<tr>
<td>2.116 DUMP 9803</td>
<td>66</td>
</tr>
<tr>
<td>2.117 DUMP 10000</td>
<td>67</td>
</tr>
<tr>
<td>2.118 DUMP 10001</td>
<td>67</td>
</tr>
<tr>
<td>2.119 DUMP 10002</td>
<td>67</td>
</tr>
<tr>
<td>2.120 DUMP 10003</td>
<td>67</td>
</tr>
<tr>
<td>2.121 DUMP 11000</td>
<td>68</td>
</tr>
<tr>
<td>2.122 DUMP 12001</td>
<td>68</td>
</tr>
<tr>
<td>2.123 DUMP 12002</td>
<td>68</td>
</tr>
<tr>
<td>2.124 DUMP 12009</td>
<td>69</td>
</tr>
<tr>
<td>2.125 DUMP 103003</td>
<td>69</td>
</tr>
<tr>
<td>2.126 DUMP 103004</td>
<td>69</td>
</tr>
<tr>
<td>2.127 DUMP 103005</td>
<td>70</td>
</tr>
<tr>
<td>3 The NDB Communication Protocol</td>
<td>73</td>
</tr>
<tr>
<td>3.1 NDB Protocol Overview</td>
<td>73</td>
</tr>
<tr>
<td>3.2 NDB Protocol Messages</td>
<td>74</td>
</tr>
<tr>
<td>3.3 Operations and Signals</td>
<td>74</td>
</tr>
<tr>
<td>4 NDB Kernel Blocks</td>
<td>85</td>
</tr>
<tr>
<td>4.1 The BACKUP Block</td>
<td>86</td>
</tr>
<tr>
<td>4.2 The CMVMI Block</td>
<td>86</td>
</tr>
<tr>
<td>4.3 The DBACC Block</td>
<td>86</td>
</tr>
<tr>
<td>4.4 The DBDICT Block</td>
<td>87</td>
</tr>
<tr>
<td>4.5 The DBDIH Block</td>
<td>87</td>
</tr>
<tr>
<td>4.6 The DBINFO Block</td>
<td>88</td>
</tr>
<tr>
<td>4.7 The DBLQH Block</td>
<td>88</td>
</tr>
<tr>
<td>4.8 The DBSPJ Block</td>
<td>90</td>
</tr>
<tr>
<td>4.9 The DBTC Block</td>
<td>90</td>
</tr>
<tr>
<td>4.10 The DBTUP Block</td>
<td>92</td>
</tr>
<tr>
<td>4.11 The DBTXU Block</td>
<td>93</td>
</tr>
<tr>
<td>4.12 The DBUTIL Block</td>
<td>93</td>
</tr>
<tr>
<td>4.13 The LGMAN Block</td>
<td>94</td>
</tr>
<tr>
<td>4.14 The NDBCNTR Block</td>
<td>94</td>
</tr>
<tr>
<td>4.15 The NDBFS Block</td>
<td>94</td>
</tr>
<tr>
<td>4.16 The PGMAN Block</td>
<td>95</td>
</tr>
<tr>
<td>4.17 The QMGR Block</td>
<td>95</td>
</tr>
<tr>
<td>4.18 The RESTORE Block</td>
<td>96</td>
</tr>
</tbody>
</table>
Preface and Legal Notices

This is the MySQL NDB Cluster Internals Manual, which contains information about the NDBCLUSTER storage engine that is not strictly necessary for running the NDB Cluster product, but can prove useful for development and debugging purposes. Topics covered in this Guide include, among others, communication protocols employed between nodes, file systems used by management nodes and data nodes, error messages, and debugging (DUMP) commands in the management client.

NDB Cluster also provides support for the Memcache API; for more information, see ndbmemcache—Memcache API for NDB Cluster.

NDB Cluster 7.3 and later also provides support for applications written using Node.js. See MySQL NoSQL Connector for JavaScript, for more information.

The information presented in this guide is current for recent releases of NDB Cluster up to and including NDB Cluster 7.6.11. Due to significant functional and other changes in NDB Cluster and its underlying APIs, you should not expect this information to apply to previous releases of the NDB Cluster software prior to NDB Cluster 7.2. Users of older NDB Cluster releases should upgrade to the latest available GA release of NDB Cluster 7.5.

This guide also contains information relating to NDB Cluster 7.6, now under development, and available as a Developer Preview for testing and evaluation purposes. For more information, see What is New in NDB Cluster 7.6.

For legal information, see the Legal Notices.

Licensing information. This product may include third-party software, used under license. If you are using a Commercial release of NDB Cluster, see this document for licensing information, including licensing information relating to third-party software that may be included in this Commercial release. If you are using a Community release of NDB Cluster, see this document for licensing information, including licensing information relating to third-party software that may be included in this Community release.

Legal Notices

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info or visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs if you are hearing impaired.
Chapter 1 NDB Cluster File Systems

Table of Contents

1.1 NDB Cluster Data Node File System ................................................................. 1
   1.1.1 NDB Cluster Data Node Data Directory Files ............................................. 1
   1.1.2 NDB Cluster Data Node File System Directory Files .................................. 2
   1.1.3 NDB Cluster Data Node Backup Data Directory Files ................................. 3
   1.1.4 NDB Cluster Disk Data Files ...................................................................... 3
1.2 NDB Cluster Management Node File System ....................................................... 4

This chapter contains information about the file systems created and used by NDB Cluster data nodes and management nodes.

1.1 NDB Cluster Data Node File System

This section discusses the files and directories created by NDB Cluster nodes, their usual locations, and their purpose.

1.1.1 NDB Cluster Data Node Data Directory Files

An NDB Cluster data node's data directory (DataDir) contains at least 3 files. These are named as shown in the following list, where node_id is the node ID:

• ndb_node_id_out.log

Sample output:

WOPool::init(61, 9)
RWPool::init(82, 13)
RWPool::init(a2, 18)
RWPool::init(c2, 13)
RWPool::init(122, 17)
RWPool::init(142, 15)
WOPool::init(41, 8)
RWPool::init(e2, 12)
RWPool::init(102, 55)
WOPool::init(21, 8)
Dbdict:: name=sys/def/SYSTAB_0, id=0, obj_ptr_i=0
Dbdict:: name=sys/def/NDB$EVENTS_0, id=1, obj_ptr_i=1
m_active_buckets.set(0)

• ndb_node_id_signal.log

This file contains a log of all signals sent to or from the data node.

Note

This file is created only if the SendSignalId parameter is enabled, which is true only for --debug builds.
- `ndb_node_id.pid`

  This file contains the data node's process ID; it is created when the `ndbd` process is started.

  The location of these files is determined by the value of the `DataDir` configuration parameter.

### 1.1.2 NDB Cluster Data Node File System Directory Files

The location of this directory can be set using `FileSystemPath`; the directory itself is always named `ndb_nodeid_fs`, where `nodeid` is the data node's node ID. The file system directory contains the following files and directories:

**Files:**
- `data-nodeid.dat`
- `undo-nodeid.dat`

**Directories:**
- **LCP:** This directory holds 2 subdirectories, named 0 and 1, each of which contain local checkpoint data files, one per local checkpoint.

  These subdirectories each contain a number of files whose names follow the pattern `TNFM.Data`, where `N` is a table ID and `M` is a fragment number. Each data node typically has one primary fragment and one backup fragment. This means that, for an NDB Cluster having 2 data nodes, and with `NoOfReplicas` equal to 2, `M` is either 0 to 1. For a 4-node cluster with `NoOfReplicas` equal to 2, `M` is either 0 or 2 on node group 1, and either 1 or 3 on node group 2.

  When using `ndbmt` there may be more than one primary fragment per node. In this case, `M` is a number in the range of 0 to the number of LQH worker threads in the entire cluster, less 1. The number of fragments on each data node is equal to the number of LQH on that node times `NoOfReplicas`.

**Note**

Increasing `MaxNoOfExecutionThreads` does not change the number of fragments used by existing tables; only newly-created tables automatically use the new fragment count. To force the new fragment count to be used by an existing table after increasing `MaxNoOfExecutionThreads`, you must perform an `ALTER TABLE ... REORGANIZE PARTITION` statement (just as when adding new node groups).

- **Directories** named `D1` and `D2`, each of which contains 2 subdirectories:
  - **DBDICT:** Contains data dictionary information. This is stored in:
    - The file `P0.SchemaLog`
    - A set of directories `T0, T1, T2, ...,` each of which contains an `S0.TableList` file.
  - **Directories** named `D8, D9, D10,` and `D11`, each of which contains a directory named `DBLQH`. These contain the redo log, which is divided into four parts that are stored in these directories. with redo log part 0 being stored in `D8`, part 1 in `D9`, and so on.

    Within each directory can be found a `DBLQH` subdirectory containing the `N` redo log files; these are named `S0.Fraglog, S1.FragLog, S2.FragLog, ..., SN.FragLog`, where `N` is equal
to the value of the `NoOfFragmentLogFiles` configuration parameter. The default value for `NoOfFragmentLogFiles` is 16. The default size of each of these files is 16 MB, controlled by the `FragmentLogFileSize` configuration parameter.

The size of each of the four redo log parts is `NoOfFragmentLogFiles * FragmentLogFileSize`. You can find out how much space the redo log is using with `DUMP 2398` or `DUMP 2399`; see Section 2.41, “DUMP 2398”, and Section 2.42, “DUMP 2399”, for more information.

- **DBDIH**: This directory contains the file `PX.sysfile`, which records information such as the last GCI, restart status, and node group membership of each node; its structure is defined in `storage/ndb/src/kernel/blocks/dbdih/Sysfile.hpp` in the NDB Cluster source tree. In addition, the `SX.FragList` files keep records of the fragments belonging to each table.

### 1.1.3 NDB Cluster Data Node Backup Data Directory Files

NDB Cluster creates backup files in the directory specified by the `BackupDataDir` configuration parameter, as discussed in Using The NDB Cluster Management Client to Create a Backup.

See NDB Cluster Backup Concepts, for information about the files created when a backup is performed.

### 1.1.4 NDB Cluster Disk Data Files

[Note]

This section applies only to NDB Cluster in MySQL 5.1 and later. Previous versions of MySQL did not support Disk Data tables.

NDB Cluster Disk Data files are created (or dropped) by the user by means of SQL statements intended specifically for this purpose. Such files include the following:

- One or more `undo logfiles` associated with a `logfile group`
- One or more `datafiles` associated with a `tablespace` that uses the logfile group for undo logging

Both undo logfiles and datafiles are created in the data directory (`DataDir`) of each cluster data node. The relationship of these files with their logfile group and tablespace are shown in the following diagram:
Figure 1.1 NDB Cluster Disk Data Files (Tablespace, Datafiles; Logfile Group, Undo Files)

Disk Data files and the SQL commands used to create and drop them are discussed in depth in NDB Cluster Disk Data Tables.

1.2 NDB Cluster Management Node File System

The files used by an NDB Cluster management node are discussed in \texttt{ndb\_mgmd — The NDB Cluster Management Server Daemon}.
## Chapter 2 NDB Cluster Management Client DUMP Commands

### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 DUMP 1</td>
<td>8</td>
</tr>
<tr>
<td>2.2 DUMP 13</td>
<td>10</td>
</tr>
<tr>
<td>2.3 DUMP 14</td>
<td>11</td>
</tr>
<tr>
<td>2.4 DUMP 15</td>
<td>11</td>
</tr>
<tr>
<td>2.5 DUMP 16</td>
<td>11</td>
</tr>
<tr>
<td>2.6 DUMP 17</td>
<td>12</td>
</tr>
<tr>
<td>2.7 DUMP 18</td>
<td>12</td>
</tr>
<tr>
<td>2.8 DUMP 20</td>
<td>13</td>
</tr>
<tr>
<td>2.9 DUMP 21</td>
<td>13</td>
</tr>
<tr>
<td>2.10 DUMP 22</td>
<td>14</td>
</tr>
<tr>
<td>2.11 DUMP 23</td>
<td>14</td>
</tr>
<tr>
<td>2.12 DUMP 24</td>
<td>15</td>
</tr>
<tr>
<td>2.13 DUMP 25</td>
<td>16</td>
</tr>
<tr>
<td>2.14 DUMP 70</td>
<td>16</td>
</tr>
<tr>
<td>2.15 DUMP 400</td>
<td>16</td>
</tr>
<tr>
<td>2.16 DUMP 401</td>
<td>17</td>
</tr>
<tr>
<td>2.17 DUMP 402</td>
<td>18</td>
</tr>
<tr>
<td>2.18 DUMP 403</td>
<td>18</td>
</tr>
<tr>
<td>2.19 DUMP 406</td>
<td>19</td>
</tr>
<tr>
<td>2.20 DUMP 908</td>
<td>19</td>
</tr>
<tr>
<td>2.21 DUMP 1000</td>
<td>20</td>
</tr>
<tr>
<td>2.22 DUMP 1223</td>
<td>20</td>
</tr>
<tr>
<td>2.23 DUMP 1224</td>
<td>21</td>
</tr>
<tr>
<td>2.24 DUMP 1225</td>
<td>21</td>
</tr>
<tr>
<td>2.25 DUMP 1226</td>
<td>21</td>
</tr>
<tr>
<td>2.26 DUMP 1228</td>
<td>22</td>
</tr>
<tr>
<td>2.27 DUMP 1229</td>
<td>22</td>
</tr>
<tr>
<td>2.28 DUMP 1332</td>
<td>22</td>
</tr>
<tr>
<td>2.29 DUMP 1333</td>
<td>23</td>
</tr>
<tr>
<td>2.30 DUMP 2300</td>
<td>24</td>
</tr>
<tr>
<td>2.31 DUMP 2301</td>
<td>24</td>
</tr>
<tr>
<td>2.32 DUMP 2302</td>
<td>26</td>
</tr>
<tr>
<td>2.33 DUMP 2303</td>
<td>26</td>
</tr>
<tr>
<td>2.34 DUMP 2304</td>
<td>26</td>
</tr>
<tr>
<td>2.35 DUMP 2305</td>
<td>29</td>
</tr>
<tr>
<td>2.36 DUMP 2308</td>
<td>29</td>
</tr>
<tr>
<td>2.37 DUMP 2315</td>
<td>30</td>
</tr>
<tr>
<td>2.38 DUMP 2350</td>
<td>30</td>
</tr>
<tr>
<td>2.39 DUMP 2352</td>
<td>31</td>
</tr>
<tr>
<td>2.40 DUMP 2354</td>
<td>32</td>
</tr>
<tr>
<td>2.41 DUMP 2398</td>
<td>32</td>
</tr>
<tr>
<td>2.42 DUMP 2399</td>
<td>33</td>
</tr>
<tr>
<td>2.43 DUMP 2400</td>
<td>34</td>
</tr>
<tr>
<td>2.44 DUMP 2401</td>
<td>35</td>
</tr>
<tr>
<td>2.45 DUMP 2402</td>
<td>36</td>
</tr>
<tr>
<td>2.46 DUMP 2403</td>
<td>36</td>
</tr>
<tr>
<td>2.47 DUMP 2404</td>
<td>36</td>
</tr>
<tr>
<td>2.48 DUMP 2405</td>
<td>37</td>
</tr>
<tr>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>2.49</td>
<td>DUMP 2406</td>
</tr>
<tr>
<td>2.50</td>
<td>DUMP 2500</td>
</tr>
<tr>
<td>2.51</td>
<td>DUMP 2501</td>
</tr>
<tr>
<td>2.52</td>
<td>DUMP 2502</td>
</tr>
<tr>
<td>2.53</td>
<td>DUMP 2503 (OBSOLETE)</td>
</tr>
<tr>
<td>2.54</td>
<td>DUMP 2504</td>
</tr>
<tr>
<td>2.55</td>
<td>DUMP 2505</td>
</tr>
<tr>
<td>2.56</td>
<td>DUMP 2506 (OBSOLETE)</td>
</tr>
<tr>
<td>2.57</td>
<td>DUMP 2507</td>
</tr>
<tr>
<td>2.58</td>
<td>DUMP 2508</td>
</tr>
<tr>
<td>2.59</td>
<td>DUMP 2509</td>
</tr>
<tr>
<td>2.60</td>
<td>DUMP 2510</td>
</tr>
<tr>
<td>2.61</td>
<td>DUMP 2511</td>
</tr>
<tr>
<td>2.62</td>
<td>DUMP 2512</td>
</tr>
<tr>
<td>2.63</td>
<td>DUMP 2513</td>
</tr>
<tr>
<td>2.64</td>
<td>DUMP 2514</td>
</tr>
<tr>
<td>2.65</td>
<td>DUMP 2515</td>
</tr>
<tr>
<td>2.66</td>
<td>DUMP 2516</td>
</tr>
<tr>
<td>2.67</td>
<td>DUMP 2517</td>
</tr>
<tr>
<td>2.68</td>
<td>DUMP 2550</td>
</tr>
<tr>
<td>2.69</td>
<td>DUMP 2555</td>
</tr>
<tr>
<td>2.70</td>
<td>DUMP 2600</td>
</tr>
<tr>
<td>2.71</td>
<td>DUMP 2601</td>
</tr>
<tr>
<td>2.72</td>
<td>DUMP 2602</td>
</tr>
<tr>
<td>2.73</td>
<td>DUMP 2603</td>
</tr>
<tr>
<td>2.74</td>
<td>DUMP 2604</td>
</tr>
<tr>
<td>2.75</td>
<td>DUMP 2610</td>
</tr>
<tr>
<td>2.76</td>
<td>DUMP 5900</td>
</tr>
<tr>
<td>2.77</td>
<td>DUMP 7000</td>
</tr>
<tr>
<td>2.78</td>
<td>DUMP 7001</td>
</tr>
<tr>
<td>2.79</td>
<td>DUMP 7002</td>
</tr>
<tr>
<td>2.80</td>
<td>DUMP 7003</td>
</tr>
<tr>
<td>2.81</td>
<td>DUMP 7004</td>
</tr>
<tr>
<td>2.82</td>
<td>DUMP 7005</td>
</tr>
<tr>
<td>2.83</td>
<td>DUMP 7006</td>
</tr>
<tr>
<td>2.84</td>
<td>DUMP 7007</td>
</tr>
<tr>
<td>2.85</td>
<td>DUMP 7008</td>
</tr>
<tr>
<td>2.86</td>
<td>DUMP 7009</td>
</tr>
<tr>
<td>2.87</td>
<td>DUMP 7010</td>
</tr>
<tr>
<td>2.88</td>
<td>DUMP 7011</td>
</tr>
<tr>
<td>2.89</td>
<td>DUMP 7012</td>
</tr>
<tr>
<td>2.90</td>
<td>DUMP 7013</td>
</tr>
<tr>
<td>2.91</td>
<td>DUMP 7014</td>
</tr>
<tr>
<td>2.92</td>
<td>DUMP 7015</td>
</tr>
<tr>
<td>2.93</td>
<td>DUMP 7016</td>
</tr>
<tr>
<td>2.94</td>
<td>DUMP 7017</td>
</tr>
<tr>
<td>2.95</td>
<td>DUMP 7018</td>
</tr>
<tr>
<td>2.96</td>
<td>DUMP 7019</td>
</tr>
<tr>
<td>2.97</td>
<td>DUMP 7020</td>
</tr>
<tr>
<td>2.98</td>
<td>DUMP 7021</td>
</tr>
<tr>
<td>2.99</td>
<td>DUMP 7024</td>
</tr>
<tr>
<td>2.100</td>
<td>DUMP 7027</td>
</tr>
<tr>
<td>2.101</td>
<td>DUMP 7033</td>
</tr>
<tr>
<td>2.102</td>
<td>DUMP 7080</td>
</tr>
</tbody>
</table>
Warning

*Never use these commands on a production NDB Cluster except under the express direction of MySQL Technical Support. Oracle will not be held responsible for adverse results arising from their use under any other circumstances!*

**DUMP** commands can be used in the Cluster management client (**ndb_mgm**) to dump debugging information to the Cluster log. They are documented here, rather than in the *MySQL Manual*, for the following reasons:

- They are intended only for use in troubleshooting, debugging, and similar activities by MySQL developers, QA, and support personnel.

- Due to the way in which DUMP commands interact with memory, they can cause a running NDB Cluster to malfunction or even to fail completely when used.

- The formats, arguments, and even availability of these commands are not guaranteed to be stable. *All of this information is subject to change at any time without prior notice.*

- For the preceding reasons, **DUMP** commands are neither intended nor warranted for use in a production environment by end-users.

General syntax:

```
ndb_mgm> node_id DUMP code [arguments]
```

This causes the contents of one or more NDB registers on the node with ID *node_id* to be dumped to the Cluster log. The registers affected are determined by the value of *code*. Some (but not all) **DUMP** commands accept additional **arguments**; these are noted and described where applicable.
Individual **DUMP** commands are listed by their *code* values in the sections that follow.

Each listing includes the following information:

- The *code* value
- The relevant **NDB** kernel block or blocks (see Chapter 4, **NDB Kernel Blocks**, for information about these)
- The **DUMP** code symbol where defined; if undefined, this is indicated using a triple dash: ***---***.
- Sample output; unless otherwise stated, it is assumed that each **DUMP** command is invoked as shown here:

  ```
  ndb_mgm> 2 DUMP code
  ```

  Generally, this is from the cluster log; in some cases, where the output may be generated in the node log instead, this is indicated. Where the **DUMP** command produces errors, the output is generally taken from the error log.

- Where applicable, additional information such as possible extra arguments, warnings, state or other values returned in the **DUMP** command's output, and so on. Otherwise its absence is indicated with “[N/A]”.

**Note**

**DUMP** command codes are not necessarily defined sequentially. For example, codes 2 through 12 are currently undefined, and so are not listed. However, individual **DUMP** code values are subject to change, and there is no guarantee that a given code value will continue to be defined for the same purpose (or defined at all, or undefined) over time.

There is also no guarantee that a given **DUMP** code—even if currently undefined—will not have serious consequences when used on a running **NDB** Cluster.

For information concerning other **ndb_mgm** client commands, see Commands in the **NDB Cluster Management Client**.

**Note**

**DUMP** codes in the following ranges are currently unused and thus unsupported:

- 3000 to 5000
- 6000 to 7000
- 13000 and higher

### 2.1 DUMP 1

<table>
<thead>
<tr>
<th>Code</th>
<th>Symbol</th>
<th>Kernel Block(s)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>---</td>
<td>QMGR</td>
<td>Dumps information about cluster start Phase 1 variables (see Section 5.4, “STTOR Phase 1”).</td>
</tr>
</tbody>
</table>
2.2 DUMP 13

Code 13

Symbol ---

Kernel Block(s) CMVMI, NDBCNTR

Description. Dump signal counter and start phase information.

Sample Output.

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Log Entry</th>
<th>Node Information</th>
<th>Description</th>
</tr>
</thead>
</table>
2.3 DUMP 14

Code 14
Symbol CommitAckMarkersSize
Kernel Block(s) DBLQH, DBTC

Description. Dumps free size in `commitAckMarkerPool`.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Source</th>
<th>Message</th>
</tr>
</thead>
</table>

Additional Information. [N/A]

2.4 DUMP 15

Code 15
Symbol CommitAckMarkersDump
Kernel Block(s) DBLQH, DBTC

Description. Dumps information in `commitAckMarkerPool`.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Source</th>
<th>Message</th>
</tr>
</thead>
</table>

Additional Information. [N/A]

2.5 DUMP 16

Code 16
Symbol DihDumpNodeRestartInfo
Kernel Block(s) DBDIH

Description. Provides node restart information.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Source</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-10-13</td>
<td>21:01:19</td>
<td>[MgmtSrvr]</td>
<td>-- Node 5: c_nodeStartMaster.blockGcp = 0, c_nodeStartMaster.wait = 0</td>
</tr>
<tr>
<td>2014-10-13</td>
<td>21:01:19</td>
<td>[MgmtSrvr]</td>
<td>-- Node 5: [ 0 : cfirstVerifyQueue = 0 clastVerifyQueue = 0 sz: 8193 ]</td>
</tr>
<tr>
<td>2014-10-13</td>
<td>21:01:19</td>
<td>[MgmtSrvr]</td>
<td>-- Node 6: c_nodeStartMaster.blockGcp = 0, c_nodeStartMaster.wait = 0</td>
</tr>
<tr>
<td>2014-10-13</td>
<td>21:01:19</td>
<td>[MgmtSrvr]</td>
<td>-- Node 6: [ 0 : cfirstVerifyQueue = 0 clastVerifyQueue = 0 sz: 8193 ]</td>
</tr>
</tbody>
</table>

Additional Information. [N/A]
2.6 DUMP 17

Code 17
Symbol DihDumpNodeStatusInfo
Kernel Block(s) DBDIH

Description. Dumps node status.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Module</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-10-13</td>
<td>21:02:28</td>
<td>[MgmtSrvr] INFO</td>
<td>-- Node 5: Node = 5 has status = 1</td>
</tr>
<tr>
<td>2014-10-13</td>
<td>21:02:28</td>
<td>[MgmtSrvr] INFO</td>
<td>-- Node 5: Node = 6 has status = 1</td>
</tr>
<tr>
<td>2014-10-13</td>
<td>21:02:28</td>
<td>[MgmtSrvr] INFO</td>
<td>-- Node 6: Node = 5 has status = 1</td>
</tr>
</tbody>
</table>

Additional Information. Possible node status values are shown in the following table:

Table 2.1 Node status values and names

<table>
<thead>
<tr>
<th>Value</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>NOT_IN_CLUSTER</td>
</tr>
<tr>
<td>1</td>
<td>ALIVE</td>
</tr>
<tr>
<td>2</td>
<td>STARTING</td>
</tr>
<tr>
<td>3</td>
<td>DIED_NOW</td>
</tr>
<tr>
<td>4</td>
<td>DYING</td>
</tr>
<tr>
<td>5</td>
<td>DEAD</td>
</tr>
</tbody>
</table>

2.7 DUMP 18

Code 18
Symbol DihPrintFragmentation
Kernel Block(s) DBDIH

Description. Prints one entry per table fragment; lists the table number, fragment number, log part ID, and the IDs of the nodes handling the primary and secondary replicas of this fragment.

Sample Output.

Node 5: Printing nodegroups --
Node 5: NG 0(0) ref: 4 [ cnt: 2 : 5 6 4294967040 4294967040 ]
Node 5: Printing fragmentation of all tables --
Node 5: Table 2 Fragment 0(1) LP: 0 — 5 6
Node 5: Table 2 Fragment 1(1) LP: 0 — 6 5
Node 5: Table 3 Fragment 0(2) LP: 1 — 5 6
Node 5: Table 3 Fragment 1(2) LP: 1 — 6 5
Node 6: Printing nodegroups --
Node 6: NG 0(0) ref: 4 [ cnt: 2 : 5 6 4294967040 4294967040 ]
Node 6: Printing fragmentation of all tables --
Node 6: Table 2 Fragment 0(1) LP: 0 — 5 6
Node 6: Table 2 Fragment 1(1) LP: 0 - 6 5
Node 6: Table 3 Fragment 0(2) LP: 1 - 5 6
Node 6: Table 3 Fragment 1(2) LP: 1 - 6 5

Additional Information. [N/A]

2.8 DUMP 20

<table>
<thead>
<tr>
<th>Code</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>--</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>BACKUP</td>
</tr>
</tbody>
</table>

**Description.** Prints the values of BackupDataBufferSize, BackupLogBufferSize, BackupWriteSize, and BackupMaxWriteSize

**Sample Output.**

```
```

**Additional Information.** This command can also be used to set these parameters, as in this example:

```
ndb_mgm> ALL DUMP 20 3 3 64 512
ALL DUMP 20 3 3 64 512
Sending dump signal with data:
0x00000014 0x00000003 0x00000003 0x00000040
0x00000200
Sending dump signal with data:
0x00000014 0x00000003 0x00000003 0x00000040
0x00000200
...
```

**Warning**

You must set each of these parameters to the same value on all nodes; otherwise, subsequent issuing of a `START BACKUP` command crashes the cluster.

2.9 DUMP 21

<table>
<thead>
<tr>
<th>Code</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>--</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>BACKUP</td>
</tr>
</tbody>
</table>

**Description.** Sends a `GSN_BACKUP_REQ` signal to the node, causing that node to initiate a backup.

**Sample Output.**

```
Node 2: Backup 1 started from node 2
```
Node 2: Backup 1 started from node 2 completed
StartGCP: 158515 StopGCP: 158518
#Records: 2061 #LogRecords: 0
Data: 35664 bytes Log: 0 bytes

Additional Information. [N/A]

2.10 DUMP 22

Code 22
Symbol backup_id
Kernel Block(s) BACKUP

Description. Sends a GSN_FREMOVEREQ signal to the node. This should remove the backup having backup ID backup_id from the backup directory; however, it actually causes the node to crash.

Sample Output.

Additional Information. It appears that any invocation of DUMP 22 causes the node or nodes to crash.

2.11 DUMP 23

Code 23
Symbol ---
Kernel Block(s) BACKUP

Description. Dumps all backup records and file entries belonging to those records.

Note
The example shows a single record with a single file only, but there may be multiple records and multiple file lines within each record.

Sample Output. With no backup in progress (BackupRecord shows as 0):

Node 2: BackupRecord 0: BackupId: 5 MasterRef: f70002 ClientRef: 0
Node 2: State: 2
Node 2: file 0: type: 3 flags: H'0'

While a backup is in progress (BackupRecord is 1):

Node 2: BackupRecord 1: BackupId: 8 MasterRef: f40002 ClientRef: 80010001
Node 2: State: 1
Node 2: file 3: type: 3 flags: H'1'
Node 2: file 2: type: 2 flags: H'1'
Node 2: file 0: type: 1 flags: H'9'
Node 2: BackupRecord 0: BackupId: 110 MasterRef: f70002 ClientRef: 0
Node 2: State: 2
Node 2: file 0: type: 3 flags: H'0'

Additional Information. Possible State values are shown in the following table:
Table 2.2 State values, the corresponding State, and a description of each State.

<table>
<thead>
<tr>
<th>Value</th>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>INITIAL</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>DEFINING</td>
<td>Defining backup content and parameters</td>
</tr>
<tr>
<td>2</td>
<td>DEFINED</td>
<td>DEFINE_BACKUP_CONF signal sent by slave, received on master</td>
</tr>
<tr>
<td>3</td>
<td>STARTED</td>
<td>Creating triggers</td>
</tr>
<tr>
<td>4</td>
<td>SCANNING</td>
<td>Scanning fragments</td>
</tr>
<tr>
<td>5</td>
<td>STOPPING</td>
<td>Closing files</td>
</tr>
<tr>
<td>6</td>
<td>CLEANING</td>
<td>Freeing resources</td>
</tr>
<tr>
<td>7</td>
<td>ABORTING</td>
<td>Aborting backup</td>
</tr>
</tbody>
</table>

Types are shown in the following table:

Table 2.3 File type values and names

<table>
<thead>
<tr>
<th>Value</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CTL_FILE</td>
</tr>
<tr>
<td>2</td>
<td>LOG_FILE</td>
</tr>
<tr>
<td>3</td>
<td>DATA_FILE</td>
</tr>
<tr>
<td>4</td>
<td>LCP_FILE</td>
</tr>
</tbody>
</table>

Flags are shown in the following table:

Table 2.4 Flag values and names

<table>
<thead>
<tr>
<th>Value</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x01</td>
<td>BF_OPEN</td>
</tr>
<tr>
<td>0x02</td>
<td>BF_OPENING</td>
</tr>
<tr>
<td>0x04</td>
<td>BF_CLOSING</td>
</tr>
<tr>
<td>0x08</td>
<td>BF_FILE_THREAD</td>
</tr>
<tr>
<td>0x10</td>
<td>BF_SCAN_THREAD</td>
</tr>
<tr>
<td>0x20</td>
<td>BF_LCP_META</td>
</tr>
</tbody>
</table>

2.12 DUMP 24

Code 24

Symbol ---

Kernel Block(s) BACKUP

Description. Prints backup record pool information.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Node</th>
<th>Message</th>
</tr>
</thead>
</table>
Additional Information. If 2424 is passed as an argument (for example, **2 DUMP 24 2424**), this causes an LCP.

### 2.13 DUMP 25

<table>
<thead>
<tr>
<th>Code</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>NdbcntrTestStopOnError</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>NDBCNTR</td>
</tr>
</tbody>
</table>

**Description.** Kills the data node or nodes.

**Sample Output.**

```
...```

**Additional Information.** [N/A]

### 2.14 DUMP 70

<table>
<thead>
<tr>
<th>Code</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>NdbcntrStopNodes</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>NDBCNTR</td>
</tr>
</tbody>
</table>

**Description.** Forces data node shutdown.

**Sample Output.**

```
...```

**Additional Information.** [N/A]

### 2.15 DUMP 400

<table>
<thead>
<tr>
<th>Code</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>NdbfsDumpFileStat-</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>NDBFS</td>
</tr>
</tbody>
</table>

**Description.** Provides NDB file system statistics.

**Sample Output.**

```
```
2.16 DUMP 401

**Code:** 401

**Symbol:** NdbfsDumpAllFiles

**Kernel Block(s):** NDBFS

**Description:** Prints NDB file system file handles and states (OPEN or CLOSED).

**Sample Output:**

```
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 0 (0x7f5ace0029f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 1 (0x7f5ace0100f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 2 (0x7f5ace01d780): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 3 (0x7f5ace02add0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 4 (0x7f5ace0387f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 5 (0x7f5ace045e40): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 7 (0x7f5ace060ae0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 8 (0x7f5ace06e130): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 9 (0x7f5ace07b780): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 10 (0x7f5ace088dd0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 11 (0x7f5ace0969f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 12 (0x7f5ace0a4040): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 13 (0x7f5ace0b1690): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 14 (0x7f5ace0bece0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 15 (0x7f5ace0c3c30): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 16 (0x7f5ace0d9980): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 17 (0x7f5ace0e6fd0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 18 (0x7f5ace0f4620): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 20 (0x7f5ace10f2c0): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 22 (0x7f5ace129f60): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 23 (0x7f5ace1375b0): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 24 (0x7f5ace144c00): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 26 (0x7f5ace15f8a0): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 27 (0x7f5ace16ce00): OPEN
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 28 (0x7f5ace030029f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 29 (0x7f5ace0300100f0): CLOSED
2014-10-13 21:17:37 [MgmtSrvr] INFO -- Node 5: 30 (0x7f5ace03000000): CLOSED
```

Additional Information.  [N/A]
2.17 DUMP 402

**Code**  402

**Symbol**  NdbfsDumpOpenFiles

**Kernel Block(s)**  NDBFS

**Description.**  Prints list of NDB file system open files.

**Sample Output.**

Node 2: NDBFS: Dump open files: 10
Node 2: 0 (0x8792f70): /usr/local/mysql/cluster/ndb_2_fs/D1/DBDIH/P0.sysfile
Node 2: 1 (0x8794590): /usr/local/mysql/cluster/ndb_2_fs/D2/DBDIH/P0.sysfile
Node 2: 2 (0x878ed10): /usr/local/mysql/cluster/ndb_2_fs/D8/DBLQH/S0.FragLog
Node 2: 3 (0x8790330): /usr/local/mysql/cluster/ndb_2_fs/D9/DBLQH/S0.FragLog
Node 2: 4 (0x8791950): /usr/local/mysql/cluster/ndb_2_fs/D10/DBLQH/S0.FragLog
Node 2: 5 (0x8795da0): /usr/local/mysql/cluster/ndb_2_fs/D11/DBLQH/S0.FragLog
Node 2: 6 (0x8797358): /usr/local/mysql/cluster/ndb_2_fs/D8/DBLQH/S1.FragLog
Node 2: 7 (0x8798978): /usr/local/mysql/cluster/ndb_2_fs/D9/DBLQH/S1.FragLog
Node 2: 8 (0x8799e98): /usr/local/mysql/cluster/ndb_2_fs/D10/DBLQH/S1.FragLog
Node 2: 9 (0x879b5b8): /usr/local/mysql/cluster/ndb_2_fs/D11/DBLQH/S1.FragLog

2.18 DUMP 403

**Code**  403

**Symbol**  NdbfsDumpIdleFiles

**Kernel Block(s)**  NDBFS

**Description.**  Prints list of NDB file system idle file handles.

**Sample Output.**

2014-10-13 21:18:48 [MgmtSrvr] INFO -- Node 5: 0 (0x7f5aec0029f0): CLOSED
2014-10-13 21:18:48 [MgmtSrvr] INFO -- Node 5: 1 (0x7f5aec0100f0): CLOSED
2014-10-13 21:18:48 [MgmtSrvr] INFO -- Node 5: 3 (0x7f5aec02add0): CLOSED
2.19 DUMP 406

**Code** 406

**Symbol** NdbfsDumpRequests

**Kernel Block(s)** NDBFS

**Description.** Includes NDBFS information in LCP and GCP stall reports written to data node logs.

**Sample Output.**

```
```

**Additional Information.** Added in NDB 7.5.14, 7.6.10, and 8.0.16. (Bug #28922609)

2.20 DUMP 908

**Code** 908

**Symbol** ---

**Kernel Block(s)** DBDIH, QMGR

**Description.** Causes heartbeat transmission information to be written to the data node logs. Useful in conjunction with setting the HeartbeatOrder parameter.
Sample Output.

| HB: pres:5 own:5 dyn:1-0 mxdyn:2 hb:6->5->6 node:dyih.cfg: 5:1-0,0 6:2-0,0 |

Additional Information. [N/A]

### 2.21 DUMP 1000

<table>
<thead>
<tr>
<th>Code</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>DumpPageMemory</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBACC, DBTUP</td>
</tr>
</tbody>
</table>

**Description.** Prints data node memory usage (ACC and TUP), as both a number of data pages, and the percentage of DataMemory and IndexMemory used.

**Sample Output.**

```
2014-10-15 12:06:29 [MgmtSrvr] INFO -- Node 5: Data usage is 0%(12 32K pages of total 32768)
2014-10-15 12:06:29 [MgmtSrvr] INFO -- Node 5: Index usage is 0%(24 8K pages of total 131104)
2014-10-15 12:06:29 [MgmtSrvr] INFO -- Node 6: Data usage is 0%(12 32K pages of total 32768)
```

**Note**

When invoked as ALL DUMP 1000, this command reports memory usage for each data node separately, in turn.

**Additional Information.** You can also use the ndb_mgm client command REPORT MEMORYUSAGE to obtain this information (see Commands in the NDB Cluster Management Client). You can also query the memoryusage table (in the ndbinfo database) for this information.

### 2.22 DUMP 1223

<table>
<thead>
<tr>
<th>Code</th>
<th>1223</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDICT</td>
</tr>
</tbody>
</table>

**Description.** Formerly, this killed the node. In NDB Cluster 7.4 and later, it has no effect.

**Sample Output.**
2.23 DUMP 1224

Code 1224
Symbol ---
Kernel Block(s) DBDICT

Description. Formerly, this killed the node. In NDB Cluster 7.4 and later, it has no effect.

Sample Output.

Additional Information. [N/A]

2.24 DUMP 1225

Code 1225
Symbol ---
 Kernel Block(s) DBDICT

Description. Formerly, this killed the node. In NDB Cluster 7.4, it has no effect.

Sample Output.

Additional Information. [N/A]

2.25 DUMP 1226

Code 1226
Symbol ---
 Kernel Block(s) DBDICT

Description. Prints pool objects to the cluster log.

Sample Output.


Additional Information. [N/A]
2.26 DUMP 1228

Code 1228
Symbol DictLockQueue
Kernel Block(s) DBDICT

Description. Dumps the contents of the NDB internal dictionary lock queue to the cluster log.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Node</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-10-15</td>
<td>12:14:08</td>
<td>5</td>
<td>DICT : c_sub_startstop _outstanding 0 _lock 0000000000000000</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>12:14:08</td>
<td>6</td>
<td>DICT : c_sub_startstop _outstanding 0 _lock 0000000000000000</td>
</tr>
</tbody>
</table>

Additional Information. [N/A]

2.27 DUMP 1229

Code 1229
Symbol DictDumpGetTabInfoQueue
Kernel Block(s) DBDICT

Description. Shows state of GETTABINFOREQ queue.

Sample Output.

```
ndb_mgm> ALL DUMP 1229
Sending dump signal with data:
0x000004cd
Sending dump signal with data:
0x000004cd
```

Additional Information. Full debugging output requires the relevant data nodes to be configured with DictTrace >= 2 and relevant API nodes with ApiVerbose >= 2. See the descriptions of these parameters for more information.

Added in NDB 7.4.12 and NDB 7.5.2. (Bug #20368450)

2.28 DUMP 1332

Code 1332
Symbol LqhDumpAllDefinedTabs
Kernel Block(s) DBACC

Description. Prints the states of all tables known by the local query handler (LQH) to the cluster log.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Node</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-10-15</td>
<td>12:15:07</td>
<td>5</td>
<td>Table 2 Status: 0 Usage: [ r: 0 w: 0 ]</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>12:15:07</td>
<td>5</td>
<td>fragment: 0 distKey: 0</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>12:15:07</td>
<td>5</td>
<td>fragment: 1 distKey: 0</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>12:15:07</td>
<td>5</td>
<td>Table 3 Status: 0 Usage: [ r: 0 w: 0 ]</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>12:15:07</td>
<td>5</td>
<td>fragment: 0 distKey: 0</td>
</tr>
</tbody>
</table>
2.29 DUMP 1333

Code

1333

Symbol

LqhDumpNoLogPages

Kernel Block(s)

DBACC
**Description.** Reports redo log buffer usage.

**Sample Output.**

```
```

**Additional Information.** The redo log buffer is measured in 32KB pages, so the sample output can be interpreted as follows:

- **Redo log buffer total.** \(1024 \times 32K = 32MB\)
- **Redo log buffer free.** \(960 \times 32KB = 31,457KB = \sim 30MB\)
- **Redo log buffer used.** \((1024 - 960) \times 32K = 2,097KB = \sim 2MB\)

# 2.30 DUMP 2300

**Code** 2300  
**Symbol** LqhDumpOneScanRec  
**Kernel Block(s)** DBLQH  

**Description.** Prints the given scan record. Syntax: `DUMP 2300 recordno`.  

**Sample Output.**

```
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 5: Dblqh::ScanRecord[1]: state=0, type=0, complStatus=0, scanNodeId=0
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 5:  apiBref=0x2f40005, scanAccPtr=0
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 5:  stpid=0, flag=2, lhold=0, lmode=0, num=134
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 5:  LcpScan=1  RowId(0:0)
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 6: Dblqh::ScanRecord[1]: state=0, type=0, complStatus=0, scanNodeId=0
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 6:  apiBref=0x2f40006, scanAccPtr=0
2014-10-15 12:33:35 [MgmtSrvr] INFO     -- Node 6:  LcpScan=1  RowId(0:0)
```

**Additional Information.** [N/A]

# 2.31 DUMP 2301

**Code** 2301  
**Symbol** LqhDumpAllScanRec  
**Kernel Block(s)** DBLQH  

**Description.** Dump all scan records to the cluster log.  

**Sample Output.** Only the first few scan records printed to the log for a single data node are shown here.
Additional Information. This DUMP code should be used sparingly if at all on an NDB Cluster in production, since hundreds or even thousands of scan records may be created on even a relatively small cluster that is not under load. For this reason, it is often preferable to print a single scan record using DUMP 2300.

The first line provides the total number of scan records dumped for this data node.
2.32 DUMP 2302

Code 2302

Symbol LqhDumpAllActiveScanRec

Kernel Block(s) DBLQH

Description. Dump only the active scan records from this node to the cluster log.

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Log Level</th>
<th>Message</th>
</tr>
</thead>
</table>

Additional Information. The first line in each block of output contains the total number of (active and inactive) scan records. If nothing else is written to the log, then no scan records are currently active.

2.33 DUMP 2303

Code 2303

Symbol LqhDumpLcpState

Kernel Block(s) DBLQH

Description. Dumps the status of a local checkpoint from the point of view of a DBLQH block instance.

Beginning with NDB 7.2.6, this command also dumps the status of the single fragment scan record reserved for this LCP. (Bug #13986128)

Sample Output.

<table>
<thead>
<tr>
<th>Date</th>
<th>Log Level</th>
<th>Message</th>
</tr>
</thead>
</table>
Sample Output. The following is taken from ndb_5_out.log in an NDB Cluster having 2 data nodes:

```
LP 0 blockInstance: 1 partNo: 0 state: 0 WW_Gci: 1 gcprec: -256 flq: 4294967040 4294967040 currfile: 0 tailFileNo: 0 logTailMbyte: 2 cnoOfLogPages: 1016 problems: 0x0
file 0(0) FileChangeState: 0 logFileStatus: 20 currentMbyte: 2 currentFilepage 75
file 1(1) FileChangeState: 0 logFileStatus: 20 currentMbyte: 0 currentFilepage 0
file 2(2) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 3(3) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 4(4) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 5(5) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 6(6) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 7(7) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 8(8) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 9(9) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 10(10) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 11(11) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 12(12) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 13(13) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 14(14) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 15(15) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0

LP 1 blockInstance: 1 partNo: 1 state: 0 WW_Gci: 1 gcprec: -256 flq: 4294967040 4294967040 currfile: 16 tailFileNo: 0 logTailMbyte: 2 cnoOfLogPages: 1016 problems: 0x0
file 0(0) FileChangeState: 0 logFileStatus: 20 currentMbyte: 2 currentFilepage 69
file 1(1) FileChangeState: 0 logFileStatus: 20 currentMbyte: 0 currentFilepage 0
file 2(2) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 3(3) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 4(4) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 5(5) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 6(6) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 7(7) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 8(8) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 9(9) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 10(10) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 11(11) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 12(12) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 13(13) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 14(14) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
file 15(15) FileChangeState: 0 logFileStatus: 1 currentMbyte: 0 currentFilepage 0
```

Description. This command causes all fragment log files and their states to be written to the data node's out file (in the case of the data node having the node ID 5, this would be ndb_5_out.log). The number of fragment log files is controlled by the NoOfFragmentLogFiles data node configuration parameter.
Additional Information. The next 2 tables provide information about file change state codes and log file status codes as shown in the previous example.

**FileChangeState Codes**

**Table 2.5 FileChangeState codes states**

<table>
<thead>
<tr>
<th>Value</th>
<th>File Change State</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Content row 1, column 2</td>
</tr>
<tr>
<td>1</td>
<td>NOT_ONGOING</td>
</tr>
<tr>
<td>2</td>
<td>BOTH_WRITES_ONGOING</td>
</tr>
<tr>
<td>3</td>
<td>LAST_WRITE_ONGOING</td>
</tr>
<tr>
<td>4</td>
<td>WRITE_PAGE_ZERO_ONGOING</td>
</tr>
</tbody>
</table>

**LogFileStatus Codes**

**Table 2.6 LogFileStatus codes with log file status and descriptions**

<table>
<thead>
<tr>
<th>Value</th>
<th>Log File Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>LFS_IDLE</td>
<td>Log file record not in use</td>
</tr>
<tr>
<td>1</td>
<td>CLOSED</td>
<td>Log file closed.</td>
</tr>
<tr>
<td>2</td>
<td>OPENING_INIT</td>
<td>---</td>
</tr>
<tr>
<td>3</td>
<td>OPEN_SR_FRONTPAGE</td>
<td>Log file opened as front page of the log</td>
</tr>
<tr>
<td>4</td>
<td>OPEN_SR_LAST_FILE</td>
<td>Opening last log file</td>
</tr>
<tr>
<td>5</td>
<td>OPEN_SR_NEXT_FILE</td>
<td>Opening log file with information about GCPs</td>
</tr>
<tr>
<td>6</td>
<td>OPEN_EXEC_SR_START</td>
<td>Log file opened while executing log during system restart</td>
</tr>
<tr>
<td>7</td>
<td>OPEN_EXEC_SR_NEW_MBYTE</td>
<td>---</td>
</tr>
<tr>
<td>8</td>
<td>OPEN_SR_FOURTH_PHASE</td>
<td>---</td>
</tr>
<tr>
<td>9</td>
<td>OPEN_SR_FOURTH_NEXT</td>
<td>---</td>
</tr>
<tr>
<td>10</td>
<td>OPEN_SR_FOURTH ZERO</td>
<td>---</td>
</tr>
<tr>
<td>11</td>
<td>OPENING_WRITE_LOG</td>
<td>Log file opened while writing log</td>
</tr>
<tr>
<td>12</td>
<td>OPEN_EXEC_LOG</td>
<td>---</td>
</tr>
<tr>
<td>13</td>
<td>CLOSING_INIT</td>
<td>---</td>
</tr>
<tr>
<td>Value</td>
<td>Log File Status</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>14</td>
<td>CLOSING_SR</td>
<td>Log file closed as part of system restart. Currently trying to find where to start executing the log.</td>
</tr>
<tr>
<td>15</td>
<td>CLOSING_EXEC_SR</td>
<td>Log file closed as part of log execution during system restart.</td>
</tr>
<tr>
<td>16</td>
<td>CLOSING_EXEC_SR_COMPLETED</td>
<td>---</td>
</tr>
<tr>
<td>17</td>
<td>CLOSING_WRITE_LOG</td>
<td>Log file closed as part of writing log during normal operations.</td>
</tr>
<tr>
<td>18</td>
<td>CLOSING_EXEC_LOG</td>
<td>---</td>
</tr>
<tr>
<td>19</td>
<td>OPEN_INIT</td>
<td>---</td>
</tr>
<tr>
<td>20</td>
<td>OPEN</td>
<td>Log file open.</td>
</tr>
<tr>
<td>21</td>
<td>OPEN_SR_READ_INVALIDATE_PAGES</td>
<td>---</td>
</tr>
<tr>
<td>22</td>
<td>CLOSE_SR_READ_INVALIDATE_PAGES</td>
<td>---</td>
</tr>
<tr>
<td>23</td>
<td>OPEN_SR_WRITE_INVALIDATE_PAGES</td>
<td>---</td>
</tr>
<tr>
<td>24</td>
<td>CLOSE_SR_WRITE_INVALIDATE_PAGES</td>
<td>---</td>
</tr>
<tr>
<td>25</td>
<td>OPEN_SR_READ_INVALIDATE_SEARCH_FILES</td>
<td>---</td>
</tr>
<tr>
<td>26</td>
<td>CLOSE_SR_READ_INVALIDATE_SEARCH_FILES</td>
<td>---</td>
</tr>
<tr>
<td>27</td>
<td>CLOSE_SR_READ_INVALIDATE_SEARCH_LAST_FILE</td>
<td>---</td>
</tr>
<tr>
<td>28</td>
<td>OPEN_EXEC_LOG_CACHED</td>
<td>---</td>
</tr>
<tr>
<td>29</td>
<td>CLOSING_EXEC_LOG_CACHED</td>
<td>---</td>
</tr>
</tbody>
</table>

More information about how these codes are defined can be found in the source file `storage/ndb/src/kernel/blocks/dblqh/Dblqh.hpp`. See also Section 2.35, “DUMP 2305”.

### 2.35 DUMP 2305

**Code** 2305  
**Symbol** ---  
**Kernel Block(s)** DBLQH

**Description.** Show the states of all fragment log files (see Section 2.34, “DUMP 2304”), then kills the node.

**Sample Output.**

```plaintext
...
```

**Additional Information.** [N/A]

### 2.36 DUMP 2308

**Code** 2308  
**Symbol** ---  
**Kernel Block(s)** DBLQH

**Description.** Kills the node.
Sample Output.

... 

Additional Information. [N/A]

2.37 DUMP 2315

<table>
<thead>
<tr>
<th>Code</th>
<th>2315</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>LqhErrorInsert5042</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBLQH</td>
</tr>
</tbody>
</table>

**Description.** [Unknown]

**Sample Output.** [N/A]

**Additional Information.** [N/A]

2.38 DUMP 2350

**Code**

data_node_id

**Symbol**

---

**Kernel Block(s)**

---

**Description.** Dumps all operations on a given data node or data nodes, according to the type and other parameters defined by the operation filter or filters specified.

**Sample Output.** Dump all operations on data node 2, from API node 5:

```
ndb_mgm> 2 DUMP 2350 1 5
2011-11-01 13:16:49 [MgmSrvr] INFO     -- Node 2: OP[470]: Tab: 4 frag: 0 TC: 3 API: 5(0x8035)transid: 0x31c 0x3500500 op: SCAN state: InQueue
```

**Additional information.** Information about operation filter and operation state values follows.

**Operation filter values.** The operation filter (or filters) can take on the following values:

**Table 2.7 Filter values**

<table>
<thead>
<tr>
<th>Value</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>table ID</td>
</tr>
<tr>
<td>1</td>
<td>API node ID</td>
</tr>
<tr>
<td>2</td>
<td>2 transaction IDs, defining a range of transactions</td>
</tr>
<tr>
<td>3</td>
<td>transaction coordinator node ID</td>
</tr>
</tbody>
</table>

In each case, the ID of the object specified follows the specifier. See the sample output for examples.

**Operation states.** The “normal” states that may appear in the output from this command are listed here:

- Transactions:
• **Prepared**: The transaction coordinator is idle, waiting for the API to proceed
• **Running**: The transaction coordinator is currently preparing operations
• **Committing, Prepare to commit, Commit sent**: The transaction coordinator is committing
• **Completing**: The transaction coordinator is completing the commit (after commit, some cleanup is needed)
• **Aborting**: The transaction coordinator is aborting the transaction
• **Scanning**: The transaction coordinator is scanning

• **Scan operations**:
  • **WaitNextScan**: The scan is idle, waiting for API
  • **InQueue**: The scan has not yet started, but rather is waiting in queue for other scans to complete

• **Primary key operations**:
  • **In lock queue**: The operation is waiting on a lock
  • **Running**: The operation is being prepared
  • **Prepared**: The operation is prepared, holding an appropriate lock, and waiting for commit or rollback to complete

**Relation to NDB API.** It is possible to match the output of DUMP 2350 to specific threads or Ndb objects. First suppose that you dump all operations on data node 2 from API node 5, using table 4 only, like this:

```
ndb_mgm> 2 DUMP 2350 1 5 0 4
2011-11-01 13:16:49 [MgmSrvr] INFO -- Node 2: OP[470]: Tab: 4 frag: 0 TC: 3 API: 5(0x8035)transid: 0x31c 0x3500500 op: SCAN state: InQueue
```

Suppose you are working with an Ndb instance named MyNdb, to which this operation belongs. You can see that this is the case by calling the Ndb object’s `getReference()` method, like this:

```
printf("MyNdb.getReference(): 0x%x\n", MyNdb.getReference());
```

The output from the preceding line of code is:

```
MyNdb.getReference(): 0x80350005
```

The high 16 bits of the value shown corresponds to the number in parentheses from the OP line in the DUMP command’s output (8035). For more about this method, see `Ndb::getReference()`.

**2.39 DUMP 2352**

<table>
<thead>
<tr>
<th>Code</th>
<th>node_id 2352 operation_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>
DUMP 2354

Kernel Block(s)  ---

Description.  Gets information about an operation with a given operation ID.

Sample Output.  First, obtain a dump of operations. Here, we use DUMP 2350 to get a dump of all operations on data node 5 from API node 100:

```
2014-10-15 13:36:26 [MgmtSrvr] INFO -- Node 100: Event buffer status: used=1025KB(100%) alloc=1025KB(0%) mem
```

In this case, there is a single operation reported on node 2, whose operation ID is 3. To obtain the transaction ID and primary key for the operation having 3 as its ID, we use the node ID and operation ID with DUMP 2352 as shown here:

```
ndb_mgm> 5 DUMP 2352 3
```

The following is written to the cluster log:

```
```

Additional Information.  Use DUMP 2350 to obtain an operation ID. See Section 2.38, “DUMP 2350”, and the previous example.

2.40 DUMP 2354

Code 2354
Symbol LqhReportCopyInfo
Kernel Block(s) DBLQH

Description.  Prints a given scan fragment record, given the instance. The syntax is shown here:

```
DUMP 2354 recordno instanceno
```

Here, recordno is the scan fragment record number, and instanceno is the number of the instance.

Sample Output.

```
2014-10-13 16:30:57 [MgmtSrvr] INFO -- Node 5: LDM instance 1: CopyFrag complete. 0 frags, +0/-0 rows, 0 bytes/29362776 ms 0 bytes/s.
2014-10-13 16:30:57 [MgmtSrvr] INFO -- Node 6: LDM instance 1: CopyFrag complete. 0 frags, +0/-0 rows, 0 bytes/29362818 ms 0 bytes/s.
```

Additional Information.  This DUMP code was added in NDB 7.4.1.

2.41 DUMP 2398

Code node_id 2398
Symbol ---
Kernel Block(s) DBLQH

Description.  Dumps information about free space in log part files for the data node with the node ID node_id. The dump is written to the data node out log rather than to the cluster log.
Sample Output. Written to ndb_6_out.log:

| REDO part: 0 HEAD: file: 0 mbyte: 2 TAIL: file: 0 mbyte: 2 total: 256 free: 256 (mb) |
| REDO part: 1 HEAD: file: 0 mbyte: 2 TAIL: file: 0 mbyte: 2 total: 256 free: 256 (mb) |
| REDO part: 2 HEAD: file: 0 mbyte: 2 TAIL: file: 0 mbyte: 2 total: 256 free: 256 (mb) |
| REDO part: 3 HEAD: file: 0 mbyte: 2 TAIL: file: 0 mbyte: 2 total: 256 free: 256 (mb) |

Additional Information. Each line of the output has the following format (shown here split across two lines for legibility):

| REDO part: part_no HEAD: file: start_file_no mbyte: start_pos |
| TAIL: file: end_file_no mbyte: end_pos total: total_space free: free_space (mb) |

A data node's redo log is divided into four parts; thus, part_no is always a number between 0 and 3 inclusive. The parts are stored in the data node file system D8, D9, D10, and D11 directories with redo log part 0 being stored in D8, part 1 in D9, and so on (see Section 1.1.2, “NDB Cluster Data Node File System Directory Files”). Within each directory can be found a DBLQH subdirectory containing NoOfFragmentLogFiles files. The default value for NoOfFragmentLogFiles is 16. The default size of each of these files is 16 MB; this can be changed by setting the FragmentLogFileSize configuration parameter.

start_file_no indicates the number of the file and start_pos the point inside this file in which the redo log starts; for the example just shown, since part_no is 0, this means that the redo log starts at approximately 12 MB from the end of the file D8/DBLQH/S6.FragLog.

Similarly, end_file_no corresponds to the number of the file and end_pos to the point within that file where the redo log ends. Thus, in the previous example, the redo log's end point comes approximately 10 MB from the end of D8/DBLQH/S6.FragLog.

total_space shows the total amount of space reserved for part part_no of the redo log. This is equal to NoOfFragmentLogFiles * FragmentLogFileSize; by default this is 16 times 16 MB, or 256 MB. free_space shows the amount remaining. Thus, the amount used is equal to total_space - free_space; in this example, this is 256 - 254 = 2 MB.

Caution
It is not recommended to execute DUMP 2398 while a data node restart is in progress.

2.42 DUMP 2399

Code node_id 2399
Symbol ---
Kernel Block(s) DBLQH

Description. Similarly to DUMP 2398, this command dumps information about free space in log part files for the data node with the node ID node_id. Unlike the case with DUMP 2398, the dump is written to the cluster log, and includes a figure for the percentage of free space remaining in the redo log.

Sample Output.

ndb_mgm> 6 DUMP 2399
Sending dump signal with data:
DUMP 2400

Additional Information. Each line of the output uses the following format (shown here split across two lines for legibility):

```
```

`timestamp` shows when the command was executed by data node `node_id`. A data node's redo log is divided into four parts, which part is indicated by `part_no` (always a number between 0 and 3 inclusive). The parts are stored in the data node file system directories named `D8`, `D9`, `D10`, and `D11`; redo log part 0 is stored in `D8`, part 1 in `D9`, and so on. Within each of these four directories is a `DBLQH` subdirectory containing `NoOfFragmentLogFiles` fragment log files. The default value for `NoOfFragmentLogFiles` is 16. The default size of each of these files is 16 MB; this can be changed by setting the `FragmentLogFileSize` configuration parameter. (See Section 1.1.2, “NDB Cluster Data Node File System Directory Files”, for more information about the fragment log files.)

`start_file_no` indicates the number of the file and `start_pos` the point inside this file in which the redo log starts; for the example just shown, since `part_no` is 0, this means that the redo log starts at approximately 12 MB from the end of the file `D8/DBLQH/S6.FragLog`.

Similarly, `end_file_no` corresponds to the number of the file and `end_pos` to the point within that file where the redo log ends. Thus, in the previous example, the redo log's end point comes approximately 10 MB from the end of `D8/DBLQH/S6.FragLog`.

`total_space` shows the total amount of space reserved for part `part_no` of the redo log. This is equal to `NoOfFragmentLogFiles * FragmentLogFileSize`; by default this is 16 times 16 MB, or 256 MB. `free_space` shows the amount remaining. The amount used is equal to `total_space - free_space`; in this example, this is 256 - 254 = 2 MB. `free_pct` shows the ratio of `free_space` to `total_space`, expressed as whole-number percentage. In the example just shown, this is equal to 100 * (254 / 256), or approximately 99 percent.

2.43 DUMP 2400

Code 2400 `record_id`
Symbol `AccDumpOneScanRec`
Kernel Block(s) `DBACC`

Description. Dumps the scan record having record ID `record_id`.

Sample Output. From `ALL DUMP 2400 1` the following output is written to the cluster log:

```
2014-10-15 13:49:50 [MgmtSrvr] INFO     -- Node 5: Dbacc::ScanRec[1]: state=1, transid(0x0, 0x0)
```
2014-10-15 13:49:50 [MgmtSrvr] INFO -- Node 5:  scanUserP=0, startNoBuck=0, minBucketIndexToRescan=0, maxBucketIndexToRescan=0
2014-10-15 13:49:50 [MgmtSrvr] INFO -- Node 5:  scanBucketState=0, scanLockHeld=0, userBlockRef=0, scanMask=0 scanLockMode=0
2014-10-15 13:49:50 [MgmtSrvr] INFO -- Node 6: Dbacc::ScanRec[1]: state=1, transid(0x0, 0x0)
2014-10-15 13:49:50 [MgmtSrvr] INFO -- Node 6: scanUserP=0, startNoBuck=0, minBucketIndexToRescan=0, maxBucketIndexToRescan=0
2014-10-15 13:49:50 [MgmtSrvr] INFO -- Node 6: scanBucketState=0, scanLockHeld=0, userBlockRef=0, scanMask=0 scanLockMode=0

Additional Information. For dumping all scan records, see Section 2.44, “DUMP 2401”.

2.44 DUMP 2401

Code 2401
Symbol AccDumpAllScanRec
Kernel Block(s) DBACC

Description. Dumps all scan records for the node specified.

Sample Output.

2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: Dbacc::ScanRec[1]: state=1, transid(0x0, 0x0)
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: activeLocalFrag=0, nextBucketIndex=0
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: scanUserP=0, startNoBuck=0, minBucketIndexToRescan=0, maxBucketIndexToRescan=0
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: scanBucketState=0, scanLockHeld=0, userBlockRef=0, scanMask=0 scanLockMode=0
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: Dbacc::ScanRec[2]: state=1, transid(0x0, 0x0)
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: activeLocalFrag=0, nextBucketIndex=0
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: scanUserP=0, startNoBuck=0, minBucketIndexToRescan=0, maxBucketIndexToRescan=0
2014-10-15 13:52:06 [MgmtSrvr] INFO -- Node 5: scanBucketState=0, scanLockHeld=0, userBlockRef=0, scanMask=0 scanLockMode=0

Additional Information. Use this command with caution, as there may be a great many scans. If you want to dump a single scan record, given its record ID, see Section 2.43, “DUMP 2400”; for dumping all active scan records, see Section 2.45, “DUMP 2402”.

35
### 2.45 DUMP 2402

**Code**
2402

**Symbol**
AccDumpAllActiveScanRec

**Kernel Block(s)**
DBACC

**Description.** Dumps all active scan records.

**Sample Output.** Similar to that for DUMP 2400 and DUMP 2401. See Section 2.44, “DUMP 2401”.

**Additional Information.** To dump all scan records (active or not), see Section 2.44, “DUMP 2401”.

### 2.46 DUMP 2403

**Code**
2403

**Symbol**
AccDumpOneOperationRec

**Kernel Block(s)**
DBACC

**Description.** Dumps a given operation record, given its ID. No arguments other than this (and the node ID or ALL) are required.

**Sample Output.** (For ALL DUMP 2403 1:)

2014-10-15 13:56:26 [MgmtSrvr] INFO -- Node 5: Dbacc::operationrec[1]: transid(0x0, 0x306400)
2014-10-15 13:56:26 [MgmtSrvr] INFO -- Node 5: m_op_bits=0xffffffff, scanBits=0, reducedHashValue=ebe8
2014-10-15 13:56:26 [MgmtSrvr] INFO -- Node 6: Dbacc::operationrec[1]: transid(0xf, 0x806400)
2014-10-15 13:56:26 [MgmtSrvr] INFO -- Node 6: m_op_bits=0xffffffff, scanBits=0, reducedHashValue=a4f1

**Additional Information.** [N/A]

### 2.47 DUMP 2404

**Code**
2404

**Symbol**
AccDumpNumOpRecs

**Kernel Block(s)**
DBACC

**Description.** Prints the number of operation records (total number, and number free) to the cluster log.

**Sample Output.**
Additional Information. The total number of operation records is determined by the value set for the `MaxNoOfConcurrentOperations` configuration parameter.

2.48 DUMP 2405

<table>
<thead>
<tr>
<th>Code</th>
<th>2405</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>AccDumpFreeOpRecs</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>---</td>
</tr>
</tbody>
</table>

**Description.** Unknown: No output results if this command is called without additional arguments; if an extra argument is used, this command crashes the data node.

**Sample Output.**

(For 2 DUMP 2405 1:)

Time: Sunday 01 November 2015 - 18:33:54
Status: Temporary error, restart node
Message: Job buffer congestion (Internal error, programming error or missing error message, please report a bug)
Error: 2334
Error data: Job Buffer Full
Error object: APZJobBuffer.C
Program: ./libexec/ndbd
Pid: 27670
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]

2.49 DUMP 2406

<table>
<thead>
<tr>
<th>Code</th>
<th>2406</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>AccDumpNotFreeOpRecs</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBACC</td>
</tr>
</tbody>
</table>

**Description.** Unknown: No output results if this command is called without additional arguments; if an extra argument is used, this command crashes the data node.

**Sample Output.**

(For 2 DUMP 2406 1:)

Time: Sunday 01 November 2015 - 18:39:16
Status: Temporary error, restart node
Message: Job buffer congestion (Internal error, programming error or missing error message, please report a bug)
Error: 2334
Error data: Job Buffer Full
Error object: APZJobBuffer.C
Program: ./libexec/ndbd
Pid: 27956
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]
### 2.50 DUMP 2500

In NDB Cluster 7.4 and later, this DUMP code prints a set of scan fragment records to the cluster log.

<table>
<thead>
<tr>
<th>Code</th>
<th>2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcDumpSetOfScanFragRec</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
</tbody>
</table>

**Description.** This DUMP code uses the syntax shown here:

```
DUMP 2500 recordno numrecords dbtcinst [activeonly]
```

This prints `numrecords` records from DBTC instance `dbtcinst`, starting with the record having record number `recordno`. The last argument is optional; all of the others shown are required. `activeonly` is a boolean that determines whether or not to print only active records. If set to 1 (actually, any nonzero value), only active records are printed and ignore any free records not in use for the moment. 0 means all records are included. The default is 1.

**Sample Output.**

```
...
```

**Additional Information.** [N/A]

Prior to NDB Cluster 7.4, this DUMP code had a different symbol and function, as described in this table and the notes that follow.

<table>
<thead>
<tr>
<th>Code</th>
<th>2500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcDumpAllScanFragRec</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
</tbody>
</table>

**Description.** Kills the data node.

**Sample Output.**

```
Time: Sunday 01 November 2015 - 13:37:11
Status: Temporary error, restart node
Message: Assertion (Internal error, programming error or missing error message, please report a bug)
Error: 2301
Error data: ArrayPool<T>::getPtr
Error object: ../../../storage/ndb/src/kernel/vm/ArrayPool.hpp line: 345 (block: CMVMI)
Program: ./libexec/ndbd
Pid: 13237
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.21-ndb-7.3.7
```
Kernel Block(s)  DBTC

Description. No output if called without any additional arguments. With additional arguments, it kills the data node.

Sample Output. (For 2 DUMP 2501 1:)

Time: Sunday 01 November 2015 - 18:41:41
Status: Temporary error, restart node
Message: Assertion (Internal error, programming error or missing error message, please report a bug)
Error: 2301
Error data: ArrayPool<T>::getPtr
Error object: ../../../../../storage/ndb/src/kernel/vm/ArrayPool.hpp line: 345 (block: DBTC)
Program: ./libexec/ndbd
Pid: 28239
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]

2.52 DUMP 2502

In NDB Cluster 7.4 and later, this code can be used to print a set of scan records for a given DBTC block instance in the cluster log.

Code 2502
Symbol TcDumpAllScanRec
Kernel Block(s) DBTC

Description. This DUMP code uses the syntax shown here:

```
DUMP 2502 recordno numrecords dbtcinst [activeonly]
```

This prints `numrecords` scan records from DBTC instance number `dbtcinst`, starting with the record having record number `recno`. The last argument is optional; all of the others shown are required. `activeonly` is a boolean that determines whether or not to print only active records. If set to 1 (actually, any nonzero value), only active records are printed and ignore any free records not in use for the moment. 0 means all records are included. The default is 1.

NDB Cluster 7.3 and earlier:

Code 2502
Symbol TcDumpAllScanRec
Kernel Block(s) DBTC

Description. Dumps all scan records held by TC blocks.

Sample Output.

Node 2: TC: Dump all ScanRecord - size: 256
Node 2: Dbtc::ScanRecord[1]: state=0nextfrag=0, nofrag=0
Node 2: alen=0, para=0, receivedop=0, noOprePperFrag=0
Node 2: schv=0, tab=0, sproc=0
2.53 DUMP 2503 (OBsolete)

This DUMP code was removed in NDB 7.4.1.

- **Code**: 2503
- **Symbol**: TcDumpAllActiveScanRec
- **Kernel Block(s)**: DBTC

**Description.** Dumps all active scan records.

**Sample Output.**

Node 2: TC: Dump active ScanRecord - size: 256

2.54 DUMP 2504

- **Code**: 2504 record_id
- **Symbol**: TcDumpOneScanRec
- **Kernel Block(s)**: DBTC

**Description.** Dumps a single scan record having the record ID record_id. (For dumping all scan records, see Section 2.52, “DUMP 2502”.

**Sample Output.** (For 2 DUMP 2504 1)
Node 2: Dbtc::ScanRecord[1]: state=0, nextfrag=0, nofrag=0
Node 2: ailen=0, para=0, receivedop=0, noOprePperFrag=0
Node 2: schv=0, tab=0, sproc=0
Node 2: apiRec=-256, next=2

Additional Information. The attributes in the output of this command are described as follows:

• **ScanRecord.** The scan record slot number (same as `record_id`)

• **state.** One of the following values (found as `ScanState` in `Dbtc.hpp`):

<table>
<thead>
<tr>
<th>Value</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>IDLE</td>
</tr>
<tr>
<td>1</td>
<td>WAIT_SCAN_TAB_INFO</td>
</tr>
<tr>
<td>2</td>
<td>WAIT_AI</td>
</tr>
<tr>
<td>3</td>
<td>WAIT_FRAGMENT_COUNT</td>
</tr>
<tr>
<td>4</td>
<td>RUNNING</td>
</tr>
<tr>
<td>5</td>
<td>CLOSING_SCAN</td>
</tr>
</tbody>
</table>

• **nextfrag:** ID of the next fragment to be scanned. Used by a scan fragment process when it is ready for the next fragment.

• **nofrag:** Total number of fragments in the table being scanned.

• **ailen:** Length of the expected attribute information.

• **para:** Number of scan frag processes that belonging to this scan.

• **receivedop:** Number of operations received.

• **noOprePperFrag:** Maximum number of bytes per batch.

• **schv:** Schema version used by this scan.

• **tab:** The index or table that is scanned.

• **sproc:** Index of stored procedure belonging to this scan.

• **apiRec:** Reference to `ApiConnectRecord`

• **next:** Index of next `ScanRecord` in free list

### 2.55 DUMP 2505

<table>
<thead>
<tr>
<th>Code</th>
<th>2505</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcDumpOneApiConnectRec</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
</tbody>
</table>

**Description.** Prints the API connection record `recordno` from instance `instanceno`, using the syntax shown here:

DUMP 2505 `recordno instanceno`
Sample Output.

...  

Additional Information.  DUMP code 2505 was added in NDB 7.4.1.

2.56 DUMP 2506 (OBsolete)

Note
This DUMP code was removed in NDB 7.4.1.

Code 2506
Symbol TcDumpAllApiConnectRec
Kernel Block(s) DBTC
Description. [Unknown]

Sample Output.

Node 2: TC: Dump all ApiConnectRecord - size: 12288
Node 2: Dbtc::ApiConnectRecord[1]: state=0, abortState=0, apiFailState=0
Node 2: transid(0x0, 0x0), apiBref=0x10000002, scanRec--256
Node 2: ctcTimer=36057, apiTimer=0, counter=0, retcode=0, retsig=0
Node 2: lqhkeyconfrec=0, lqhkeyreqrec=0, tckeyrec=0
Node 2: next--256
Node 2: Dbtc::ApiConnectRecord[2]: state=0, abortState=0, apiFailState=0
Node 2: transid(0x0, 0x0), apiBref=0x10000002, scanRec--256
Node 2: ctcTimer=36057, apiTimer=0, counter=0, retcode=0, retsig=0
Node 2: lqhkeyconfrec=0, lqhkeyreqrec=0, tckeyrec=0
Node 2: next--256
Node 2: Dbtc::ApiConnectRecord[3]: state=0, abortState=0, apiFailState=0
Node 2: transid(0x0, 0x0), apiBref=0x10000002, scanRec--256
Node 2: ctcTimer=36057, apiTimer=0, counter=0, retcode=0, retsig=0
Node 2: lqhkeyconfrec=0, lqhkeyreqrec=0, tckeyrec=0
Node 2: next--256
Node 2: Dbtc::ApiConnectRecord[12287]: state=7, abortState=0, apiFailState=0
Node 2: transid(0x0, 0x0), apiBref=0xffffffff, scanRec--256
Node 2: ctcTimer=36308, apiTimer=0, counter=0, retcode=0, retsig=0
Node 2: lqhkeyconfrec=0, lqhkeyreqrec=0, tckeyrec=0
Node 2: next--256
Node 2: Dbtc::ApiConnectRecord[12287]: state=7, abortState=0, apiFailState=0
Node 2: transid(0x0, 0x0), apiBref=0xffffffff, scanRec--256
Node 2: ctcTimer=36308, apiTimer=0, counter=0, retcode=0, retsig=0
Node 2: lqhkeyconfrec=0, lqhkeyreqrec=0, tckeyrec=0
Node 2: next--256

Additional Information. If the default settings are used, the output from this command is likely to exceed the maximum log file size.

2.57 DUMP 2507

Code 2507
<table>
<thead>
<tr>
<th>Symbol</th>
<th>TcSetTransactionTimeout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
<tr>
<td><strong>Description.</strong></td>
<td>Apparently requires an extra argument, but is not currently known with certainty.</td>
</tr>
<tr>
<td><strong>Sample Output.</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td><strong>Additional Information.</strong></td>
<td>[N/A]</td>
</tr>
</tbody>
</table>

**2.58 DUMP 2508**

<table>
<thead>
<tr>
<th>Code</th>
<th>2508</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcSetApplTransactionTimeout</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
<tr>
<td><strong>Description.</strong></td>
<td>Apparently requires an extra argument, but is not currently known with certainty.</td>
</tr>
<tr>
<td><strong>Sample Output.</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td><strong>Additional Information.</strong></td>
<td>[N/A]</td>
</tr>
</tbody>
</table>

**2.59 DUMP 2509**

<table>
<thead>
<tr>
<th>Code</th>
<th>2509</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>StartTcTimer</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
<tr>
<td><strong>Description.</strong></td>
<td>[Unknown]</td>
</tr>
<tr>
<td><strong>Sample Output.</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td><strong>Additional Information.</strong></td>
<td>[N/A]</td>
</tr>
</tbody>
</table>

**2.60 DUMP 2510**

<table>
<thead>
<tr>
<th>Code</th>
<th>2510</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>StopTcTimer</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTC</td>
</tr>
<tr>
<td><strong>Description.</strong></td>
<td>[Unknown]</td>
</tr>
<tr>
<td><strong>Sample Output.</strong></td>
<td></td>
</tr>
</tbody>
</table>
2.61 DUMP 2511

Code 2511
Symbol StartPeriodicTcTimer
Kernel Block(s) DBTC

Description. [Unknown]

Sample Output.

... Additional Information. [N/A]

2.62 DUMP 2512

Code 2512 delay
Symbol TcStartDumpIndexOpCount
Kernel Block(s) DBTC

Description. Dumps the value of MaxNoOfConcurrentIndexOperations, and the current resource usage, in a continuous loop. The delay time between reports can optionally be specified (in seconds), with the default being 1 and the maximum value being 25 (values greater than 25 are silently coerced to 25).

Sample Output. (Single report:)

Node 2: IndexOpCount: pool: 8192 free: 8192

Additional Information. There appears to be no way to disable the repeated checking of MaxNoOfConcurrentIndexOperations once started by this command, except by restarting the data node. It may be preferable for this reason to use DUMP 2513 instead (see Section 2.63, “DUMP 2513”).

2.63 DUMP 2513

Code 2513
Symbol TcDumpIndexOpCount
Kernel Block(s) ---

Description. Dumps the value of MaxNoOfConcurrentIndexOperations, and the current resource usage.

Sample Output.
2.64 DUMP 2514

Code 2514

Symbol TcDumpApiConnectRecSummary

Kernel Block(s) DBTC

Description. Provides information counts for allocated, seized, stateless, stateful, and scanning transaction objects for each API node. Available beginning with NDB 7.2.13. (Bug #15878085)

The syntax for this command is shown here:

DUMP 2514 [instanceno]

This command takes the DBTC instance number (instanceno) as an optional argument; if not specified, it defaults to 0. The instanceno is not needed if there is only one instance of DBTC.

Sample Output.

Start of ApiConnectRec summary (6144 total allocated)
  Api node 10 connect records seized : 0 stateless : 0 stateful : 0 scan : 0
  Api node 11 connect records seized : 2 stateless : 0 stateful : 0 scan : 0
  Api node 12 connect records seized : 1 stateless : 0 stateful : 0 scan : 0

The total number of records allocated depends on the number of transactions and a number of other factors, with the value of MaxNoOfConcurrentTransactions setting an upper limit. See the description of this parameter for more information.

Additional Information. There are two possible states for each record, listed here:

1. Available: In the per-data node pool, not yet seized by any API node
2. Seized: Reserved from the per-data node pool by a particular API

Seized nodes further be divided into a number of categories or sub-states, as shown in the following list:

- Ready: (Not counted here) Seized, ready for use; can be calculated for an API as # seized - (# stateless + # stateful + # scan)
- Stateless: Record was last used for a 'stateless' transaction, and is effectively ready
- Stateful: Record is in use by a transaction
- Scan: Record is in use for a scan (table or ordered index)

2.65 DUMP 2515

Code 2515

Symbol TcDumpSetOfApiConnectRec
### DUMP 2515

**Kernel Block(s)**: DBTC

**Description.** Prints a range of API connection records. The syntax is as shown here, where `recordno` is the number of the first record, `numrecords` is the number of records to be dumped, and `instanceno` is the block instance number:

```
DUMP 2515 recordno numrecords instanceno
```

- **Caution**
  - It is recommended not to print more than 10 records at a time using this DUMP code from a cluster under load.

**Sample Output.** ...

**Additional Information.** DUMP code 2515 was added in NDB 7.4.1.

## 2.66 DUMP 2516

<table>
<thead>
<tr>
<th>Code</th>
<th>2516</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcDumpOneTcConnectRec</td>
</tr>
</tbody>
</table>

**Kernel Block(s)**: DBTC

**Description.** Prints the TC connection record `recordno` from instance `instanceno`, using the syntax shown here:

```
DUMP 2516 recordno instanceno
```

To print a series of such records, use DUMP 2517.

**Sample Output.** ...

**Additional Information.** DUMP code 2516 was added in NDB 7.4.1.

## 2.67 DUMP 2517

<table>
<thead>
<tr>
<th>Code</th>
<th>2517</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TcDumpSetOfTcConnectRec</td>
</tr>
</tbody>
</table>

**Kernel Block(s)**: DBTC

**Description.** Prints a range of TC connection records. The syntax is as shown here, where `recordno` is the number of the first record, `numrecords` is the number of records to be dumped, and `instanceno` is the block instance number:

```
DUMP 2517 recordno numrecords instanceno
```
Caution
It is recommended not to print more than 10 records at a time using DUMP 2517 code from a cluster under load.

Sample Output.  ...  

Additional Information.  DUMP code 2517 was added in NDB 7.4.1.

2.68 DUMP 2550

Code  data_node_id 2550 transaction_filters
Symbol  ---
Kernel Block(s)  ---

Description.  Dumps all transaction from data node data_node_id meeting the conditions established by the transaction filter or filters specified.

Sample Output.  Dump all transactions on node 2 which have been inactive for 30 seconds or longer:

ndb_mgm> 2 DUMP 2550 4 30
2011-11-01 13:16:49 [MgmSrvr] INFO -- Node 2: TRX[123]: API: 5(0x8035) transid: 0x31c 0x3500500 inactive: 42s state:

Additional Information.  The following values may be used for transaction filters. The filter value must be followed by one or more node IDs or, in the case of the last entry in the table, by the time in seconds that transactions have been inactive:

Table 2.9 Data node transaction filter values and descriptions

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>API node ID</td>
<td>transaction filter</td>
</tr>
<tr>
<td>2</td>
<td>2 transaction IDs, defining a range of transactions</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>time transactions inactive (seconds)</td>
<td></td>
</tr>
</tbody>
</table>

2.69 DUMP 2555

Code  2555
Symbol  TcDumpPoolLevels
Kernel Block(s)  DBTC

Description.  Prints pool levels to the cluster log.

Sample Output.  ...

...
2.70 DUMP 2600

Code 2600

Symbol CmvmiDumpConnections

Kernel Block(s) CMVMI

Description. Shows status of connections between all cluster nodes. When the cluster is operating normally, every connection has the same status.

Sample Output.

Node 3: Connection to 1 (MGM) is connected
Node 3: Connection to 2 (MGM) is trying to connect
Node 3: Connection to 3 (DB) does nothing
Node 3: Connection to 4 (DB) is connected
Node 3: Connection to 7 (API) is connected
Node 3: Connection to 8 (API) is connected
Node 3: Connection to 9 (API) is trying to connect
Node 3: Connection to 10 (API) is trying to connect
Node 3: Connection to 11 (API) is trying to connect
Node 4: Connection to 1 (MGM) is connected
Node 4: Connection to 2 (MGM) is trying to connect
Node 4: Connection to 3 (DB) is connected
Node 4: Connection to 4 (DB) does nothing
Node 4: Connection to 7 (API) is connected
Node 4: Connection to 8 (API) is connected
Node 4: Connection to 9 (API) is trying to connect
Node 4: Connection to 10 (API) is trying to connect
Node 4: Connection to 11 (API) is trying to connect

Additional Information. The message *is trying to connect* actually means that the node in question was not started. This can also be seen when there are unused [api] or [mysql] sections in the config.ini file nodes configured—in other words when there are spare slots for API or SQL nodes.

2.71 DUMP 2601

Code 2601

Symbol CmvmiDumpLongSignalMemory

Kernel Block(s) CMVMI

Description. [Unknown]

Sample Output.

Node 2: Cmvmi: g_sectionSegmentPool size: 4096 free: 4096

Additional Information. [N/A]

2.72 DUMP 2602

Code 2602
### 2.73 DUMP 2603

<table>
<thead>
<tr>
<th>Code</th>
<th>2603</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>CmvmiTestLongSigWithDelay</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>CMVMI</td>
</tr>
</tbody>
</table>

**Description:** [Unknown]

**Sample Output.**

```
... 
```

**Additional Information.** [N/A]

### 2.74 DUMP 2604

<table>
<thead>
<tr>
<th>Code</th>
<th>2604</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>CmvmiDumpSubscriptions</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>CMVMI</td>
</tr>
</tbody>
</table>

**Description.** Dumps current event subscriptions.

**Note**
This output appears in the `ndb_node_id_out.log` file (local to each data node) and not in the management server (global) cluster log file.

**Sample Output.**

```
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- List subscriptions:
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Subscription: 0, nodeId: 1, ref: 0x80000001
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 0 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 1 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 2 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 3 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 4 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 5 Level 8
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 6 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 7 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 8 Level 15
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 9 Level 7
Sunday 01 November 2015 17:10:54 [ndbd] INFO -- Category 10 Level 7
```
Additional Information. The output lists all event subscriptions; for each subscription a header line and a list of categories with their current log levels is printed. The following information is included in the output:

- **Subscription**: The event subscription's internal ID
- **nodeID**: Node ID of the subscribing node
- **ref**: A block reference, consisting of a block ID from `storage/ndb/include/kernel/BlockNumbers.h` shifted to the left by 4 hexadecimal digits (16 bits) followed by a 4-digit hexadecimal node number. Block id `0x8000` appears to be a placeholder; it is defined as `MIN_API_BLOCK_NO`, with the node number part being 1 as expected
- **Category**: The cluster log category, as listed in *Event Reports Generated in NDB Cluster* (see also the file `storage/ndb/include/mgmapi/mgmapi_config_parameters.h`).
- **Level**: The event level setting (the range being 0 to 15).

### 2.75 DUMP 2610

<table>
<thead>
<tr>
<th>Code</th>
<th>2610</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>CmvmiSetKillerWatchdog</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>CMVMI</td>
</tr>
</tbody>
</table>

**Description.** Activate or deactivate the killer watchdog, which, on the next watchdog warning following activation, shuts down the data node where it occurred. This provides a trace log which includes a signal trace; if the node process was started with the ```--core-file``` option, core files are also generated when this occurs.

**Syntax:** `DUMP 2610 [value]`. Use 1 for the `value` or omit `value` altogether to activate; use 0 to deactivate.

**Sample Output.** Client:

```
ndb_mgm> ALL DUMP 2610 1
Sending dump signal with data:
0x0000000a32 0x00000001
Sending dump signal with data:
0x0000000a32 0x00000001
Sending dump signal with data:
0x0000000a32 0x00000001
Sending dump signal with data:
0x0000000a32 0x00000001

ndb_mgm> ALL DUMP 2610 0
Sending dump signal with data:
0x0000000a32 0x00000000
Sending dump signal with data:
0x0000000a32 0x00000000
Sending dump signal with data:
0x0000000a32 0x00000000
Sending dump signal with data:
0x0000000a32 0x00000000
```

Node log:
Additional Information.  Added in NDB 7.2.18 and NDB 7.3.7 (Bug #18703922).

2.76 DUMP 5900

Code 5900
Symbol LCPContinue
Kernel Block(s) DBLQH

Description.  [Unknown]

Sample Output.

...  

Additional Information.  [N/A]

2.77 DUMP 7000

Code 7000
Symbol ---
Kernel Block(s) DBDIH

Description.  Prints information on GCP state

Sample Output.

Node 2: ctimer = 299072, cgcpParticipantState = 0, cgcpStatus = 0
Node 2: coldGcpStatus = 0, coldGcpId = 436, cmasterState = 1
Node 2: cmasterTakeOverNode = 65535, ctcCounter = 299072

Additional Information.  [N/A]

2.78 DUMP 7001

Code 7001
Symbol ---
Kernel Block(s) DBDIH

Description.  Prints information on the current LCP state.

Sample Output.

Node 2: c_lcpState.keepGci = 1
Node 2: c_lcpState.lcpStatus = 0, clcpStopGcp = 1
Node 2: cgcpStartCounter = 7, cimmediateLcpStart = 0

Additional Information.  [N/A]
## 2.79 DUMP 7002

<table>
<thead>
<tr>
<th>Code</th>
<th>7002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>

### Kernel Block(s)

- **DBDIH**

**Description:** [Unknown]

**Sample Output.**

```
Node 2: cnoOfActiveTables = 4, cgcpDelay = 2000
Node 2: cdictblockref = 16384002, cfailurenr = 1
Node 2: con_lineNodes = 2, reference() = 16121858, creceivedfrag = 0
```

**Additional Information.** [N/A]

## 2.80 DUMP 7003

<table>
<thead>
<tr>
<th>Code</th>
<th>7003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>

### Kernel Block(s)

- **DBDIH**

**Description:** [Unknown]

**Sample Output.**

```
Node 2: cfirstAliveNode = 2, cgckptflag = 0
Node 2: clocalqhblockref = 16187394, clocaltcblockref = 16056322, cgcpOrderBlocked = 0
Node 2: cstarttype = 0, csystemnodes = 2, currentgcp = 438
```

**Additional Information.** [N/A]

## 2.81 DUMP 7004

<table>
<thead>
<tr>
<th>Code</th>
<th>7004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>

### Kernel Block(s)

- **DBDIH**

**Description:** [Unknown]

**Sample Output.**

```
Node 2: cmasterdihref = 16121858, cownNodeId = 2, cnewgcp = 438
Node 2: cnmdbStartReqBlockref = 16449538, cremainingfrags = 1268
Node 2: cntriblockref = 16449538, cgcpSameCounter = 16, coldgcp = 437
```

**Additional Information.** [N/A]

## 2.82 DUMP 7005

<table>
<thead>
<tr>
<th>Code</th>
<th>7005</th>
</tr>
</thead>
</table>
2.83 DUMP 7006

Code  7006
Symbol  ---
Kernel Block(s)  DBDIH
Description.  [Unknown]
Sample Output.

Node 2: crestartGci = 1

Additional Information.  [N/A]

2.84 DUMP 7007

Code  7007
Symbol  ---
Kernel Block(s)  DBDIH
Description.  [Unknown]
Sample Output.

Node 2: clcpDelay = 20, cgcpMasterTakeOverState = 0
Node 2: cmasterNodeId = 2
Node 2: cnoHotSpare = 0, c_nodeStartMaster.startNode = -256, c_nodeStartMaster.wait = 0

Additional Information.  [N/A]

2.85 DUMP 7008

Code  7008
Symbol  ---
Kernel Block(s)  DBDIH
### 2.86 DUMP 7009

**Code**  
7009

**Symbol**  
---

**Kernel Block(s)**  
DBDIH

**Description.**  
[Unknown]

**Sample Output.**

Node 2: cfirstDeadNode = -256, cstartPhase = 7, cnoReplicas = 2  
Node 2: cwaitLcpSr = 0

**Additional Information.**  
[N/A]

### 2.87 DUMP 7010

**Code**  
7010

**Symbol**  
---

**Kernel Block(s)**  
DBDIH

**Description.**  
[Unknown]

**Sample Output.**

Node 2: ccalcOldestRestorableGci = 1, cnoOfNodeGroups = 1  
Node 2: cstartGcpNow = 0  
Node 2: crestartGci = 1  
Node 2: c_blockCommit = 0, c_blockCommitNo = 0

**Additional Information.**  
[N/A]

### 2.88 DUMP 7011

**Code**  
7011

**Symbol**  
---

**Kernel Block(s)**  
DBDIH

**Description.**  
[Unknown]

**Sample Output.**
2.89 DUMP 7012

**Code** 7012  
**Symbol** c  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]  
**Sample Output.**  
Node 2: ParticipatingDIH = 0000000000000000  
Node 2: ParticipatingLQH = 0000000000000000  
Node 2: m_LCP_COMPLETE_REP.Counter_DIH = [SignalCounter: m_count=0 0000000000000000]  
Node 2: m_LCP_COMPLETE_REP.Counter_LQH = [SignalCounter: m_count=0 0000000000000000]  
Node 2: m_LAST_LCP_FRAG_ORD = [SignalCounter: m_count=0 0000000000000000]  
Node 2: m_LCP_COMPLETE_REP.From_Master_Received = 0  
**Additional Information.** [N/A]

2.90 DUMP 7013

**Code** 7013  
**Symbol** DihDumpLCPState  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]  
**Sample Output.**  
Node 2: lcpStatus = 0 (update place = 9843)  
Node 2: lcpStart = 1 lcpStopGcp = 1 keepGci = 1 oldestRestorable = 1  
Node 2: immediateLcpStart = 0 masterLcpNodeId = 2  
**Additional Information.** [N/A]
# DUMP 7014

## 2.91 DUMP 7014

**Code** 7014  
**Symbol** DihDumpLCPMasterTakeOver  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]  
**Sample Output.**

Node 2: c_lcpMasterTakeOverState.state = 0 updatePlace = 11756 failedNodeId = -202116109  
Node 2: c_lcpMasterTakeOverState.minTableId = 4092851187 minFragId = 4092851187

**Additional Information.** [N/A]

## 2.92 DUMP 7015

**Code** 7015  
**Symbol** ---  
**Kernel Block(s)** DBDIH

**Description.** Writes table fragment status output for NDB tables to the cluster log, in order of their table IDs. A starting table ID can optionally be specified, in which case tables having lower IDs than this are skipped; otherwise, status information for all NDB tables is included in the output.

**Sample Invocation/Output.** Invoking this command using the optional table ID argument gives the following output in the system shell:

```
shell> ndb_mgm -e 'ALL DUMP 2015 25'
Connected to Management Server at: localhost:1186
Sending dump signal with data: 0x00001b67 0x00000019
Sending dump signal with data: 0x00001b67 0x00000019
This causes Table 1 through Table 24 to be skipped in the output written into the cluster log, as shown here:
```

```bash
2016-07-21 13:08:29 [MgmtSrvr] INFO -- Node 5: Table 25: TabCopyStatus: 0 TabUpdateStatus: 0 TabLcpStatus: 3
2016-07-21 13:08:29 [MgmtSrvr] INFO -- Node 5: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
2016-07-21 13:08:29 [MgmtSrvr] INFO -- Node 5: Fragment 1: noLcpReplicas==0 0(on 6)=59(Idle) 1(on 5)=59(Idle)
```

## 2.93 DUMP 7016

**Code** 7016  
**Symbol** DihAllAllowNodeStart  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]

**Sample Output.**

```
...  
```

**Additional Information.** [N/A]

## 2.94 DUMP 7017

**Code** 7017  
**Symbol** DihMinTimeBetweenLCP  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]

**Sample Output.**

```
...  
```

**Additional Information.** [N/A]

## 2.95 DUMP 7018

**Code** 7018  
**Symbol** DihMaxTimeBetweenLCP  
**Kernel Block(s)** DBDIH  
**Description.** [Unknown]

**Sample Output.**

```
...  
```

**Additional Information.** [N/A]
2.96 DUMP 7019

<table>
<thead>
<tr>
<th>Code</th>
<th>7019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** Write the distributed data block's view of node failure handling for a failed node (given its node ID) into the cluster log. Execute as **ALL DUMP 7019 FailedNodeId**.

**Sample Output.**

...

**Additional Information.** Added in NDB 7.2.9. (Bug #14220269)

2.97 DUMP 7020

<table>
<thead>
<tr>
<th>Code</th>
<th>7020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** This command provides general signal injection functionality. Two additional arguments are always required:

1. The number of the signal to be sent
2. The number of the block to which the signal should be sent

In addition some signals permit or require extra data to be sent.

**Sample Output.**

...

**Additional Information.** [N/A]

2.98 DUMP 7021

<table>
<thead>
<tr>
<th>Code</th>
<th>7021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** Writes table fragment status information for a single NDB table to the cluster log. **DUMP 7015** is the same as this command, except that **DUMP 7015** logs the information for multiple (or all) NDB tables.

The table to obtain information for is specified by table ID. You can find the ID for a given table in the output of **ndb_show_tables**, as shown here:
<table>
<thead>
<tr>
<th>ID</th>
<th>Type</th>
<th>State</th>
<th>Logging Database</th>
<th>Schema</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>OrderedIndex</td>
<td>Online</td>
<td>No</td>
<td>sys</td>
<td>PRIMARY</td>
</tr>
<tr>
<td>1</td>
<td>IndexTrigger</td>
<td>Online</td>
<td>-</td>
<td>NDB$INDEX_11_CUSTOM</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>UserTable</td>
<td>Online</td>
<td>Yes</td>
<td>mysql</td>
<td>NDB$INDEX_15_CUSTOM</td>
</tr>
<tr>
<td>5</td>
<td>IndexTrigger</td>
<td>Online</td>
<td>-</td>
<td>NDB$INDEX_28_CUSTOM</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>OrderedIndex</td>
<td>Online</td>
<td>No</td>
<td>sys</td>
<td>PRIMARY</td>
</tr>
<tr>
<td>10</td>
<td>UserTable</td>
<td>Online</td>
<td>Yes</td>
<td>test</td>
<td>n1</td>
</tr>
<tr>
<td>27</td>
<td>UserTable</td>
<td>Online</td>
<td>Yes</td>
<td>c</td>
<td>t1</td>
</tr>
</tbody>
</table>

**Sample Invocation/Output.**  Using the table ID for table n1 found in the `ndb_show_tables` sample output shown previously (and highlighted therein), an invocation of this command might look like this when running `ndb_mgm` in the system shell:

```shell
dbshell> ndb_mgm -e 'ALL DUMP 7015 10'
Connected to Management Server at: localhost:1186
Sending dump signal with data:
0x00001b67 0x0000000a
Sending dump signal with data:
0x00001b67 0x00000000
```

This writes the following output to the cluster log:

```
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 5: Table 10: TabCopyStatus: 0 TabUpdateStatus: 0 TabLcpStatus: 3
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 5: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 5: Fragment 1: noLcpReplicas==0 0(on 6)=59(Idle) 1(on 5)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 6: Table 10: TabCopyStatus: 0 TabUpdateStatus: 0 TabLcpStatus: 3
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 6: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 6: Fragment 1: noLcpReplicas==0 0(on 6)=59(Idle) 1(on 5)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 7: Table 10: TabCopyStatus: 0 TabUpdateStatus: 0 TabLcpStatus: 3
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 7: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 7: Fragment 1: noLcpReplicas==0 0(on 6)=59(Idle) 1(on 5)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 8: Table 10: TabCopyStatus: 0 TabUpdateStatus: 0 TabLcpStatus: 3
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 8: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
2016-07-21 12:12:11 [MgmtSrvr] INFO -- Node 8: Fragment 1: noLcpReplicas==0 0(on 6)=59(Idle) 1(on 5)=59(Idle)
```

**Additional Information.**  More information about each of the fields written by DUMP 7021 into the cluster log is shown in the next few paragraphs. The enumerations are defined as properties of structure `TabRecord` in `storage/ndb/src/kernel/blocks/dbdih/Dbdih.hpp`.

- **TabCopyStatus** (table copy status) takes one of the following values: 0: CS_IDLE, 1: CS_SR_PHASE1_READ_PAGES, 2: CS_SR_PHASE2_READ_TABLE, 3: CS_SR_PHASE3_COPY_TABLE, 4: CS_REMOVE_NODE, 5: CS_LCP_READ_TABLE, 6: CS_COPY_TAB_REQ, 7: CS_COPY_NODE_STATE, 8: CS_ADD_TABLE_MASTER, 9: CS_ADD_TABLE_SLAVE, 10: CS_INVALIDATE_NODE_LCP, 11: CS_ALTER_TABLE, 12: CS_COPY_TO_SAVE, 13: CS_GET_TABINFO.


- **TabLcpStatus** (table local checkpoint status) takes one of the following values: 1: TLS_ACTIVE, 2: TLS_WRITING_TO_FILE, 3: TLS_COMPLETED.

Table fragment information is also provided for each node. This is similar to what is shown here:

```
Node 5: Fragment 0: noLcpReplicas==0 0(on 5)=59(Idle) 1(on 6)=59(Idle)
```
The node and fragment are identified by their IDs. \texttt{noLcpReplicas} represents the number of replicas remaining to be checkpointed by any ongoing LCP. The remainder of the line has the format shown here:

\begin{verbatim}
replica_id(on node_id)=lcp_id(status)
\end{verbatim}

\texttt{replica_id}, \texttt{node_id}, and \texttt{lcp_id} are the IDs of, respectively, the replica, node, and local checkpoint. \texttt{status} is always one of \texttt{Idle} or \texttt{Ongoing}.

### 2.99 DUMP 7024

<table>
<thead>
<tr>
<th>Code</th>
<th>7024</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** Determines whether tables are in their expected states.

**Sample Output.**

```
...
```

**Additional Information.** Added in NDB 7.2.17 and NDB 7.3.6. (Bug #18550318)

### 2.100 DUMP 7027

<table>
<thead>
<tr>
<th>Code</th>
<th>7027</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>DihStallLcpStart</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** Causes a local checkpoint to stall. Used for testing of LCP issues.

**Usage.** This command requires an additional argument \texttt{91919191} for activation. For example, to initiate an LCP stall on all nodes, execute the \texttt{DUMP} command shown here:

```
ALL DUMP 7027 91919191
```

To clear the stall and resume normal operation, invoke \texttt{DUMP 7027} with any argument \textit{other} than 91919191 (or even no additional argument at all).

**Additional Information.** Added in NDB 7.3.19, 7.4.17, 7.5.8, and 7.6.4. (Bug #26661468)

### 2.101 DUMP 7033

<table>
<thead>
<tr>
<th>Code</th>
<th>7033</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>DihFragmentsPerNode</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBDIH</td>
</tr>
</tbody>
</table>

**Description.** Prints the number of fragments on one or more data nodes. No arguments other than the node ID are used.
Sample Output.  Output from ALL DUMP 7033 on an NDB Cluster with two data nodes and NoOfReplicas=2:


Additional Information.  Added in NDB 7.4.1.

2.102 DUMP 7080

Code 7080
Symbol EnableUndoDelayDataWrite
Kernel Block(s) DBACC, DBDIH, DBTUP

Description.  [Unknown]

Sample Output.

...

Additional Information.  [N/A]

2.103 DUMP 7090

Code 7090
Symbol DihSetTimeBetweenGcp
Kernel Block(s) DBDIH

Description.  [Unknown]

Sample Output.

...

Additional Information.  [N/A]

2.104 DUMP 7098

Code 7098
Symbol ---
Kernel Block(s) DBDIH

Description.  [Unknown]

Sample Output.

Node 2: Invalid no of arguments to 7098 - startLcpRoundLoopLab - expected 2 (tableId, fragmentId)
2.105 DUMP 7099

Code 7099
Symbol DihStartLcpImmediately
Kernel Block(s) DBDIH

Description. Can be used to trigger an LCP manually.

Sample Output. In this example, node 2 is the master node and controls LCP/GCP synchronization for the cluster. Regardless of the node_id specified, only the master node responds:

Node 2: Local checkpoint 7 started. Keep GCI = 1003 oldest restorable GCI = 947
Node 2: Local checkpoint 7 completed

Additional Information. You may need to enable a higher logging level using the CLUSTERLOG ndb_mgm client command to have the checkpoint's completion reported, as shown here:

ndb_mgmgt; ALL CLUSTERLOG CHECKPOINT=8

2.106 DUMP 7901

Code 7901
Symbol ---
Kernel Block(s) DBDIH, DBLQH

Description. Provides timings of GCPs.

Sample Output.

...

2.107 DUMP 8004

Code 8004
Symbol ---
Kernel Block(s) SUMA

Description. Dumps information about subscription resources.

Sample Output.

Node 2: Suma: c_subscriberPool size: 260 free: 258
Node 2: Suma: c_tablePool size: 130 free: 128
Node 2: Suma: c_subscriptionPool size: 130 free: 128
Node 2: Suma: c_syncPool size: 2 free: 2
Node 2: Suma: c_dataBufferPool size: 1009 free: 1005
Node 2: Suma: c_metaSubscribers count: 0
Node 2: Suma: c_removeDataSubscribers count: 0
Additional Information. When subscriberPool ... free becomes and stays very low relative to subscriberPool ... size, it is often a good idea to increase the value of the MaxNoOfTables configuration parameter (\( \text{subscriberPool} = 2^\*\text{MaxNoOfTables} \)). However, there could also be a problem with API nodes not releasing resources correctly when they are shut down. DUMP 8004 provides a way to monitor these values.

2.108 DUMP 8005

<table>
<thead>
<tr>
<th>Code</th>
<th>8005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>SUMA</td>
</tr>
</tbody>
</table>

Description. [Unknown]

Sample Output.

| Node 2: Bucket 0 10-0 switch gci: 0 max_acked_gci: 2961 max_gci: 0 tail: -256 head: -256 |
| Node 2: Bucket 1 00-0 switch gci: 0 max_acked_gci: 2961 max_gci: 0 tail: -256 head: -256 |

Additional Information. [N/A]

2.109 DUMP 8010

<table>
<thead>
<tr>
<th>Code</th>
<th>8010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>SUMA</td>
</tr>
</tbody>
</table>

Description. Writes information about all subscribers and connected nodes to the cluster log.

Sample Output. In this example, node 1 is a management node, nodes 2 and 3 are data nodes, and nodes 4 and 5 are SQL nodes (which both act as replication masters).


For each data node, this DUMP command prints two hexadecimal numbers. These are representations of bitfields having one bit per node ID, starting with node ID 0 for the rightmost bit (0x01).

The subscriber nodes bitmask (\(c\_subscriber\_nodes\)) has the significant hexadecimal digits 30 (decimal 48), or binary 110000, which equates to nodes 4 and 5. The connected nodes bitmask (\(c\_connected\_nodes\)) has the significant hexadecimal digits 32 (decimal 50). The binary representation of this number is 1100010, which has 1 as the second, fifth, and sixth digits (counting from the right), and so works out to nodes 1, 4, and 5 as the connected nodes.

2.110 DUMP 8011

<table>
<thead>
<tr>
<th>Code</th>
<th>8011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>
Kernel Block(s)  SUMA

**Description.** Writes information to the cluster log about all subscribers in the cluster. When using this information, you should keep in mind that a table may have many subscriptions, and a subscription may have more than one subscriber. The output from DUMP 8011 includes the following information:

- **For each table:** The table ID, version number, and total number of subscribers
- **For each subscription to a given table:** The subscription ID
- **For each subscriber belonging to a given subscription:** The subscriber ID, sender reference, sender data, and subscription ID

**Sample Output.** (From cluster log:)

```
```

---

### 2.111 DUMP 8013

**Code**  8013  
**Symbol**  ---  
**Kernel Block(s)**  SUMA

**Description.** Writes a dump of all lagging subscribers to the cluster log.

**Sample Output.** This example shows what is written to the cluster log after ALL DUMP 8013 is executed on a 4-node cluster:

```
```

**Additional Information.** Added in NDB 7.2.13. (Bug #16203623)
2.112 DUMP 9002

Code 9002
Symbol DumpTsman
Kernel Block(s) TSMAN

Description. [Unknown]

Sample Output.

Additional Information. [N/A]

2.113 DUMP 9800

Code 9800
Symbol DumpTsman
Kernel Block(s) TSMAN

Description. Kills data node.

Sample Output.

Time: Sunday 01 November 2015 - 18:32:53
Status: Temporary error, restart node
Message: Internal program error (failed ndbrequire) (Internal error, programming error or missing error message, please report a bug)
Error: 2341
Error data: tsman.cpp
Error object: TSMAN (Line: 1413) 0x0000000a
Program: ./libexec/ndbd
Pid: 29658
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]

2.114 DUMP 9801

Code 9801
Symbol ---
Kernel Block(s) TSMAN

Description. Kills data node.

Sample Output.

Time: Sunday 01 November 2015 - 18:35:48
Status: Temporary error, restart node
Message: Internal program error (failed ndbrequire) (Internal error,
programming error or missing error message, please report a bug)
Error: 2341
Error data: tsman.cpp
Error object: TSMAN (Line: 1844) 0x0000000a
Program: ./libexec/ndbd
Pid: 30251
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]

2.115 DUMP 9802

<table>
<thead>
<tr>
<th>Code</th>
<th>9802</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>

Kernel Block(s) TSMAN

Description. Kills data node.

Sample Output.

Time: Sunday 01 November 2015 - 18:39:30
Status: Temporary error, restart node
Message: Internal program error (failed ndbrequire) (Internal error, programming error or missing error message, please report a bug)
Error: 2341
Error data: tsman.cpp
Error object: TSMAN (Line: 1413) 0x0000000a
Program: ./libexec/ndbd
Pid: 30482
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8

Additional Information. [N/A]

2.116 DUMP 9803

<table>
<thead>
<tr>
<th>Code</th>
<th>9803</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>---</td>
</tr>
</tbody>
</table>

Kernel Block(s) TSMAN

Description. Kills data node.

Sample Output.

Time: Sunday 01 November 2015 - 18:41:32
Status: Temporary error, restart node
Message: Internal program error (failed ndbrequire) (Internal error, programming error or missing error message, please report a bug)
Error: 2341
Error data: tsman.cpp
Error object: TSMAN (Line: 2144) 0x0000000a
Program: ./libexec/ndbd
Pid: 30712
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1
Version: Version 5.6.27-ndb-7.4.8
### 2.117 DUMP 10000

<table>
<thead>
<tr>
<th>Code</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>DumpLgman</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>LGMAN</td>
</tr>
</tbody>
</table>

**Description.** [Unknown]

**Sample Output.**

...  

**Additional Information.** [N/A]

### 2.118 DUMP 10001

<table>
<thead>
<tr>
<th>Code</th>
<th>10001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>LgmanDumpUndoStateClusterLog</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>LGMAN</td>
</tr>
</tbody>
</table>

**Description.** [Unknown]

**Sample Output.**

...  

**Additional Information.** Added in NDB 7.3.19, 7.4.17, 7.5.8, and 7.6.4. (Bug #26365433)

### 2.119 DUMP 10002

<table>
<thead>
<tr>
<th>Code</th>
<th>10002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>LgmanDumpUndoStateLocalLog</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>LGMAN</td>
</tr>
</tbody>
</table>

**Description.** [Unknown]

**Sample Output.**

...  

**Additional Information.** Added in NDB 7.3.19, 7.4.17, 7.5.8, and 7.6.4. (Bug #26365433)

### 2.120 DUMP 10003

<table>
<thead>
<tr>
<th>Code</th>
<th>10003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>LgmanCheckCallbacksClear</td>
</tr>
</tbody>
</table>
### 2.121 DUMP 11000

<table>
<thead>
<tr>
<th>Code</th>
<th>11000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>DumpPgman</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>LGMAN</td>
</tr>
</tbody>
</table>

#### Description
[Unknown]

#### Sample Output
```
...
```

#### Additional Information
Added in NDB 7.3.19, 7.4.17, 7.5.8, and 7.6.4. (Bug #26365433)

### 2.122 DUMP 12001

<table>
<thead>
<tr>
<th>Code</th>
<th>12001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TuxLogToFile</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTUX</td>
</tr>
</tbody>
</table>

#### Description
[Unknown]

#### Sample Output
```
...
```

#### Additional Information
[N/A]

### 2.123 DUMP 12002

<table>
<thead>
<tr>
<th>Code</th>
<th>12002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
<td>TuxSetLogFlags</td>
</tr>
<tr>
<td>Kernel Block(s)</td>
<td>DBTUX</td>
</tr>
</tbody>
</table>

#### Description
[Unknown]

#### Sample Output
```
...
```
2.124 DUMP 12009

**Code** 12009  
**Symbol** `TuxMetaDataJunk`  
**Kernel Block(s)** `DBTUX`

**Description.** Kills data node.

**Sample Output.**

Time: Sunday 01 November 2015 - 19:49:59  
Status: Temporary error, restart node  
Message: Error OS signal received (Internal error, programming error or missing error message, please report a bug)  
Error: 6000  
Error data: Signal 6 received; Aborted  
Error object: main.cpp  
Program: ./libexec/ndbd  
Pid: 13784  
Trace: /usr/local/mysql/cluster/ndb_2_trace.log.1  
Version: Version 5.6.27-ndb-7.4.8

2.125 DUMP 103003

**Code** 103003  
**Symbol** `CmvmiRelayDumpStateOrd`  
**Kernel Block(s)** `CMVMI`

**Description.** Sends a `DUMP` command using another node. The syntax is shown here, and explained in the paragraph following:

```
node_id DUMP 103003 other_node_id dump_cmd_no [args]
```

*node_id* is the ID of the node on which the command is issued (as usual). *other_node_id* is the ID of the node where the command is actually executed. *dump_cmd_no* is the number of the `DUMP` command to be executed on the other node; *args* represents any additional arguments required by that `DUMP` command.

**Sample Output.** (Output is dependent on the command that is sent.)

**Additional Information.** This command is particularly useful because it can be employed to send a `DUMP` command to an API node, since API nodes are connected to data nodes only, and not to the management server.

This command was added in NDB 8.0.18.

2.126 DUMP 103004

**Code** 103004
DUMP 103005

Symbol: CmvmiDummy
Kernel Block(s): CMVMI

Description. Logs a CmvmiSendDummySignal signal sent using DUMP 103005 in the event logger. Includes information about the size of the signal, number and sizes of signal sections, and the node which sent it. Syntax is shown here:

```
node_id DUMP 103004 to_id from_id padding frag_size sections length1 [length2 ..]
```

Arguments are described in the following list:
- `node_id`: ID if the node where the command is executed
- `to_id`: node ID of the destination of the signal
- `from_id`: node ID of the signal's origin
- `padding`: padding size
- `frag_size`: fragment size
- `sections`: number of sections
- `length1`, `length2`, `...`: lengths of each of the sections

Sample Output.

```
...
```

Additional Information. Added in NDB 8.0.18.

2.127 DUMP 103005

Code: 103005
Symbol: CmvmiSendDummy
Kernel Block(s): CMVMI

Description. Sends a CmvmiSendDummySignal having a given size and sections to the specified node. This is used to verify that messages with certain signal sizes and sections can be sent and received —this is also recorded by the event logger. The resulting log entry should be matched with that generated when receiving DUMP 103004 (see same). Syntax is shown here:

```
node_id DUMP 103005 to_id from_id padding frag_size sections length1 [length2 ..]
```

Arguments are described in the following list:
- `node_id`: ID if the node where the command is executed
- `to_id`: node ID of the destination of the signal
- `from_id`: node ID of the signal's origin
- `padding`: padding size
- **frag_size**: fragment size
- **sections**: number of sections
- \( length_1[, length_2][,...] \): lengths of each of the sections

**Sample Output.**

... 

**Additional Information.**  Added in NDB 8.0.18.
Chapter 3 The NDB Communication Protocol

Table of Contents

3.1 NDB Protocol Overview ................................................................. 73
3.2 NDB Protocol Messages ............................................................... 74
3.3 Operations and Signals ................................................................. 74

This chapter provides information about the protocol used for communication between data nodes and API nodes in an NDB Cluster to perform various operations such as data reads and writes, committing and rolling back transactions, and handling of transaction records.

3.1 NDB Protocol Overview

NDB Cluster data and API nodes communicate with one another by passing messages to one another. The sending of a message from one node and its reception by another node is referred to as a signal; the NDB Protocol is the set of rules governing the format of these messages and the manner in which they are passed.

An NDB message is typically either a request or a response. A request indicates that an API node wants to perform an operation involving cluster data (such as retrieval, insertion, updating, or deletion) or transactions (commit, roll back, or to fetch or release a transaction record). A request is, when necessary, accompanied by key or index information. The response sent by a data node to this request indicates whether or not the request succeeded and, where appropriate, is accompanied by one or more data messages.

Request types. A request is represented as a REQ message. Requests can be divided into those handling data and those handling transactions:

• Data requests. Data request operations are of three principal types:
  1. Primary key lookup operations are performed through the exchange of TCKEY messages.
  2. Unique key lookup operations are performed through the exchange of TCINDX messages.
  3. Table or index scan operations are performed through the exchange of SCANTAB messages.

Data request messages are often accompanied by KEYINFO messages, ATTRINFO messages, or both sorts of messages.

• Transactional requests. These may be divided into two categories:
  1. Commits and rollbacks, which are represented by TC_COMMIT and TCROLLBACK request messages, respectively.
  2. Transaction record requests, consisting of transaction record acquisition and release, are handled through the use of, respectively, TCSEIZE and TCRELEASE request messages.

Response types. A response indicates either the success or the failure of the request to which it is sent in reply:

• A response indicating success is represented as a CONF (confirmation) message, and is often accompanied by data, which is packaged as one or more TRANSID_AI messages.

• A response indicating failure is represented as a REF (refusal) message.
For more information about these message types and their relationship to one another, see Section 3.2, “NDB Protocol Messages”.

3.2 NDB Protocol Messages

This section describes the NDB Protocol message types, their function, and their structure.

**Naming Conventions.** Message names are constructed according to a simple pattern which should be readily apparent from the discussion of request and response types in the previous section. These are shown in the following matrix:

<table>
<thead>
<tr>
<th>Operation Type</th>
<th>Request (REQ)</th>
<th>Response: Success (CONF)</th>
<th>Response: Failure (REF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Key Lookup (TCKEY)</td>
<td>TCKEYREQ</td>
<td>TCKEYCONF</td>
<td>TCKEYREF</td>
</tr>
<tr>
<td>Unique Key Lookup (TCINDEX)</td>
<td>TCINDEXREQ</td>
<td>TCINDEXCONF</td>
<td>TCINDEXREF</td>
</tr>
<tr>
<td>Table or Index Scan (SCANTAB)</td>
<td>SCANTABREQ</td>
<td>SCANTABCONF</td>
<td>SCANTABREF</td>
</tr>
<tr>
<td>Result Retrieval (SCAN_NEXT)</td>
<td>SCAN_NEXTREQ</td>
<td>SCANTABCONF</td>
<td>SCANTABREF</td>
</tr>
<tr>
<td>Transaction Record Acquisition (TCSEIZE)</td>
<td>TCSEIZEREQ</td>
<td>TCSEIZECONF</td>
<td>TCSEIZEREF</td>
</tr>
<tr>
<td>Transaction Record Release (TCRELEASE)</td>
<td>TCRELEASEREQ</td>
<td>TCRELEASECONF</td>
<td>TCRELEASEREF</td>
</tr>
</tbody>
</table>

CONF and REF are shorthand for “confirmed” and “refused”, respectively.

Three additional types of messages are used in some instances of inter-node communication. These message types are listed here:

1. A **KEYINFO** message contains information about the key used in a TCKEYREQ or TCINDEXREQ message. It is employed when the key data does not fit within the request message. KEYINFO messages are also sent for index scan operations in which bounds are employed.

2. An **ATTRINFO** message contains nonkey attribute values which does not fit within a TCKEYREQ, TCINDEXREQ, or SCANTABREQ message. It is used for:
   - Supplying attribute values for inserts and updates
   - Designating which attributes are to be read for read operations
   - Specifying optional values to read for delete operations

3. A **TRANSID_AI** message contains data returned from a read operation; in other words, it is a result set (or part of one).

3.3 Operations and Signals

In this section we discuss the sequence of message-passing that takes place between a data node and an API node for each of the following operations:
• Primary key lookup
• Unique key lookup
• Table scan or index scan
• Explicit commit of a transaction
• Rollback of a transaction
• Transaction record handling (acquisition and release)

**Primary key lookup.** An operation using a primary key lookup is performed as shown in the following diagram:

**Figure 3.1 Messages Exchanged In A Primary Key Lookup**

**Note**

* and + are used here with the meanings “zero or more” and “one or more”, respectively.
The steps making up this process are listed and explained in greater detail here:

1. The API node sends a TCKEYREQ message to the data node. In the event that the necessary information about the key to be used is too large to be contained in the TCKEYREQ, the message may be accompanied by any number of KEYINFO messages carrying the remaining key information. If additional attributes are used for the operation and exceed the space available in the TCKEYREQ, or if data is to be sent to the data node as part of a write operation, then these are sent with the TCKEYREQ as any number of ATTRINFO messages.

2. The data node sends a message in response to the request, according to whether the operation succeeded or failed:

   • If the operation was successful, the data node sends a TCKEYCONF message to the API node. If the request was for a read operation, then TCKEYCONF is accompanied by a TRANSID_AI message, which contains actual result data. If there is more data than can be contained in a single TRANSID_AI can carry, more than one of these messages may be sent.

   • If the operation failed, then the data node sends a TCKEYREF message back to the API node, and no more signalling takes place until the API node makes a new request.

Unique key lookup. This is performed in a manner similar to that performed in a primary key lookup:

1. A request is made by the API node using a TCINDXREQ message which may be accompanied by zero or more KEYINFO messages, zero or more ATTRINFO messages, or both.

2. The data node returns a response, depending on whether or not the operation succeeded:

   • If the operation was a success, the message is TCINDXCONF. For a successful read operation, this message may be accompanied by one or more TRANSID_AI messages carrying the result data.

   • If the operation failed, the data node returns a TCINDXREF message.

The exchange of messages involved in a unique key lookup is illustrated in the following diagram:
Figure 3.2 Messages Exchanged In A Unique Key Lookup

**UNIQUE KEY LOOKUP**

- **REQUEST**
  - API Node → TCINDEXREQ
  - KEYINFO* → ATTRINFO* → Data Node

- **RESPONSE (SUCCESS)**
  - API Node → TCINDEXCONF
  - TRANSID_AI+ → Data Node

- **RESPONSE (FAILURE)**
  - API Node → TCINDEXREF → Data Node

**Table scans and index scans.** These are similar in many respects to primary key and unique key lookups, as shown here:
1. A request is made by the API node using a `SCAN_TABREQ` message, along with zero or more `ATTRINFO` messages. `KEYINFO` messages are also used with index scans in the event that bounds are used.

2. The data node returns a response, according to whether or not the operation succeeded:
   - If the operation was a success, the message is `SCAN_TABCONF`. For a successful read operation, this message may be accompanied by one or more `TRANSID_AI` messages carrying the result data. However—unlike the case with lookups based on a primary or unique key—it is often necessary to fetch multiple results from the data node. Requests following the first are signalled by the API node using a `SCAN_NEXTREQ`, which tells the data node to send the next set of results (if there are more results). This is shown here:
Figure 3.4 Fetching Multiple Result Data Sets Following A Table Or Index Scan Read Operation

FETCHING RESULTS

REQUEST

API Node  Messages  Data Node

SCANNEXTREQ

RESPONSE
(RETURN RESULT)

API Node  Messages  Data Node

SCAN_TABCONF
TRANSID_AT+

RESPONSE
(NO MORE RESULTS)

API Node  Messages  Data Node

SCAN_TABREF

- If the operation failed, the data node returns a SCAN_TABREF message.

SCAN_TABREF is also used to signal to the API node that all data resulting from a read has been sent.

Committing and rolling back transactions. The process of performing an explicit commit follows the same general pattern as shown previously. The API node sends a TC_COMMITREQ message to the data node, which responds with either a TC_COMMITCONF (on success) or a TC_COMMITREF (if the commit failed). This is shown in the following diagram:
Some operations perform a `COMMIT` automatically, so this is not required for every transaction.

Rolling back a transaction also follows this pattern. In this case, however, the API node sends a `TCROLLBACKTREQ` message to the data node. Either a `TCROLLBACKCONF` or a `TCROLLBACKREF` is sent in response, as shown here:
Handling of transaction records. Acquiring a transaction record is accomplished when an API node transmits a TCSEIZEREQ message to a data node and receives a TCSEIZECONF or TCSEIZEREF in return, depending on whether or not the request was successful. This is depicted here:
The release of a transaction record is also handled using the request-response pattern. In this case, the API node's request contains a **TCRELEASEREQ** message, and the data node's response uses either a **TCRELEASECONF** (indicating that the record was released) or a **TCRELEASEREF** (indicating that the attempt at release did not succeed). This series of events is illustrated in the next diagram:
TRANSACTION RECORD RELEASE

REQUEST

API Node  Messages  Data Node

RESPONSE (SUCCESS)

API Node  Messages  Data Node

RESPONSE (FAILURE)

API Node  Messages  Data Node
Chapter 4 NDB Kernel Blocks

Table of Contents

4.1 The BACKUP Block ................................................................. 86
4.2 The CMVMI Block ................................................................. 86
4.3 The DBACC Block ................................................................. 86
4.4 The DBDICT Block ............................................................... 87
4.5 The DBDIH Block ................................................................. 87
4.6 The DBINFO Block ............................................................... 87
4.7 The DBLQH Block ................................................................. 88
4.8 The DBSPJ Block ................................................................. 90
4.9 The DBTC Block ................................................................. 90
4.10 The DBTUP Block ............................................................... 92
4.11 The DBTUX Block ............................................................... 93
4.12 The DBUTIL Block ............................................................ 93
4.13 The LGMAN Block ............................................................. 94
4.14 The NDBCNTR Block .......................................................... 94
4.15 The NDBFS Block ............................................................... 94
4.16 The PGMAN Block ............................................................. 95
4.17 The QMGR Block ............................................................... 95
4.18 The RESTORE Block .......................................................... 96
4.19 The SUMA Block ............................................................... 96
4.20 The THRMAN Block .......................................................... 96
4.21 The TRPMAN Block .......................................................... 96
4.22 The TSMAN Block ............................................................. 97
4.23 The TRIX Block ................................................................. 97

This chapter provides information about the major software modules making up the NDB kernel. The files containing the implementations of these blocks can be found in several directories under `storage/ndb/src/kernel/blocks/` in the NDB Cluster source tree.

As described elsewhere, the NDB kernel makes use of a number of different threads to perform various tasks. Kernel blocks are associated with these threads as shown in the following table:

Table 4.1 NDB kernel blocks and NDB kernel threads

<table>
<thead>
<tr>
<th>Thread (ThreadConfig Name)</th>
<th>Kernel Blocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main (main)</td>
<td>CMVMI (primary), DBINFO, DBDICT, DBDIH, NDBCNTR, QMGR, DBUTIL</td>
</tr>
<tr>
<td>LDM (ldm)</td>
<td>DICT, DBACC, DBLQH (primary), DBCP, BACKUP, TSMA, LGMA, PGMAN, RESTORE</td>
</tr>
<tr>
<td>TC (tc)</td>
<td>DBTC (primary), TRIX</td>
</tr>
<tr>
<td>Replication (rep)</td>
<td>SUMA (primary), DBSPJ</td>
</tr>
<tr>
<td>Receiver (recv)</td>
<td>CMVMI</td>
</tr>
<tr>
<td>Sender (send)</td>
<td>CMVMI</td>
</tr>
<tr>
<td>I/O (io)</td>
<td>NDBFS</td>
</tr>
</tbody>
</table>

You can obtain more information about these threads from the documentation for the ThreadConfig data node configuration parameter.
The BACKUP Block

4.1 The BACKUP Block

This block is responsible for handling online backups and checkpoints. It is found in storage/ndb/src/kernel/blocks/backup/, and contains the following files:

- `Backup.cpp`: Defines methods for node signal handling; also provides output methods for backup status messages to user.
- `BackupFormat.hpp`: Defines the formats used for backup data, .CTL, and log files.
- `Backup.hpp`: Defines the `Backup` class.
- `BackupInit.cpp`: Actual `Backup` class constructor is found here.
- `Backup.txt`: Contains a backup signal diagram (text format). Somewhat dated (from 2003), but still potentially useful for understanding the sequence of events that is followed during backups.
- `FsBuffer.hpp`: Defines the `FsBuffer` class, which implements the circular data buffer that is used (together with the NDB file system) for reading and writing backup data and logs.
- `read.cpp`: Contains some utility functions for reading log and checkpoint files to STDOUT.

4.2 The CMVMI Block

This block is responsible for configuration management between the kernel blocks and the NDB virtual machine, as well as the cluster job que and cluster transporters. It is found in storage/ndb/src/kernel/blocks/cmvmi, and contains these files:

- `Cmvmi.cpp`: Implements communication and reporting methods for the `Cmvmi` class.
- `Cmvmi.hpp`: Defines the `Cmvmi` class.

During startup, this block allocates and touches most of the memory needed for buffers used by the NDB kernel, such as those defined by `IndexMemory`, `DataMemory`, and `DiskPageBufferMemory`. At that time, CMVMI also gets the starting order of the nodes, and performs a number of functions whereby software modules can affect the runtime environment.

4.3 The DBACC Block

Also referred to as the ACC block, this is the access control and lock management module. It is also responsible for storing primary key and unique key hash indexes are stored. This block is found in storage/ndb/src/kernel/blocks/dbacc, which contains the following files:

- `Dbacc.hpp`: Defines the `Dbacc` class, along with structures for operation, scan, table, and other records.
- `DbaccInit.cpp`: `Dbacc` class constructor and destructor; methods for initialising data and records.
- `DbaccMain.cpp`: Implements `Dbacc` class methods.

The ACC block handles database index structures, which are stored in 8K pages. Database locks are also handled in the ACC block.

When a new tuple is inserted, the TUP block stores the tuple in a suitable space and returns an index (a reference to the address of the tuple in memory). ACC stores both the primary key and this tuple index of the tuple in a hash table.
The DBDICT Block

Like the TUP block, the ACC block implements part of the checkpoint protocol. It also performs undo logging. It is implemented by the Dbacc class, which is defined in storage/ndb/src/kernel/blocks/dbacc/DbaccMain.hpp.

See also Section 4.10, "The DBTUP Block".

4.4 The DBDICT Block

This block, the data dictionary block, is found in storage/ndb/src/kernel/blocks/dbdict. Data dictionary information is replicated to all DICT blocks in the cluster. This is the only block other than DBTC to which applications can send direct requests.

DBDICT is responsible for managing metadata (using the master data node) including the definitions for tables, columns, indexes, tablespaces, log files, log file groups, and other data objects.

This block is implemented in the following files:

- CreateIndex.txt: Contains notes about processes for creating, altering, and dropping indexes and triggers.

- Dbdict.cpp: Implements structure for event metadata records (for NDB$EVENTS_0), as well as methods for system start and restart, table and schema file handling, and packing table data into pages. Functionality for determining node status and handling node failures is also found here. In addition, this file implements data and other initialisation routines for Dbdict.

- DictLock.txt: Implementation notes: Describes locking of the master node’s DICT against schema operations.

- printSchemaFile.cpp: Contains the source for the ndb_print_schema_file utility.

- Slave_AddTable.sfl: A signal log trace of a table creation operation for DBDICT on a nonmaster node.

- CreateTable.txt: Notes outlining the table creation process (dated).

- CreateTable.new.txt: Notes outlining the table creation process (updated version of CreateTable.txt).

- Dbdict.hpp: Defines the Dbdict class; also creates the NDB$EVENTS_0 table. Also defines a number of structures such as table and index records, as well as for table records.

- DropTable.txt: Implementation notes for the process of dropping a table.

- Dbdict.txt: Implementation notes for creating and dropping events and NdbEventOperation objects.

- Event.txt: A copy of Dbdict.txt.

- Master_AddTable.sfl: A signal log trace of a table creation operation for DBDICT on the master node.


4.5 The DBDIH Block

This block provides data distribution management services for distribution information about each table, table partition, and replica of each partition. It is also responsible for handling of local and global checkpoints. DBDIH also manages node and system restarts. This block is implemented in the following files, all found in the directory storage/ndb/src/kernel/blocks/dbdih:
The DBINFO Block

- **Dbdih.hpp**: This file contains the definition of the Dbdih class, as well as the FileRecordPtr type, which is used to keep storage information about a fragment and its replicas. If a fragment has more than one backup replica, then a list of the additional ones is attached to this record. This record also stores the status of the fragment, and is 64-byte aligned.

- **DbdihMain.cpp**: Contains definitions of Dbdih class methods.

- **printSysfile/printSysfile.cpp**: Older version of the printSysfile.cpp in the main dbdih directory.

- **DbdihInit.cpp**: Initializes Dbdih data and records; also contains the class destructor.

- **LCP.txt**: Contains developer notes about the exchange of messages between DIH and LQH that takes place during a local checkpoint.

- **printSysfile.cpp**: This file contains the source code for ndb_print_sys_file. For information about using this utility, see ndb_print_sys_file — Print NDB System File Contents.

- **Sysfile.hpp**: Contains the definition of the Sysfile structure; in other words, the format of an NDB system file. See Chapter 1, NDB Cluster File Systems, for more information about NDB system files.

This block often makes use of BACKUP blocks on the data nodes to accomplish distributed tasks, such as global checkpoints and system restarts.

This block is implemented as the Dbdih class, whose definition may be found in the file storage/ndb/src/kernel/blocks/dbdih/Dbdih.hpp.

### 4.6 The DBINFO Block

The DBINFO block provides support for the ndbinfo information database used to obtain information about data node internals.

An API node communicates with this block to retrieve ndbinfo data using DBINFO_SCANREQ and DBINFO_SCANCONF signals. The API node communicates with DBINFO on the master data node, which communicates with DBINFO on the remaining data nodes. The DBINFO block on each data node fetches information from the other kernel blocks on that node, including DBACC, DBTUP, BACKUP, DBTC, SUMA, DBUTIL, TRIX, DBTUX, DBDICT, CMVMI, DBLQH, LGMAN, PGMAN, DBSPJ, THRMAN, TRPMAN, and QMGR. The local DBINFO then sends the information back to DBINFO on the master node, which in turn passes it back to the API node.

This block is implemented in the file storage/ndb/src/kernel/blocks/dbinfo/Dbinfo.hpp as the Dbinfo class. The file Dbinfo.cpp in the same directory defines the methods of this class (mostly signal handlers). Also in the dbinfo directory is a text file DbinfoScan.txt which provides information about DBINFO messaging.

### 4.7 The DBLQH Block

This is the local, low-level query handler block, which manages data and transactions local to the cluster's data nodes, and acts as a coordinator of 2-phase commits. It is responsible (when called on by the transaction coordinator) for performing operations on tuples, accomplishing this task with help of DBACC block (which manages the index structures) and DBTUP (which manages the tuples). It is made up of the following files, found in storage/ndb/src/kernel/blocks/dblqh:

- **Dblqh.hpp**: Contains the Dblqh class definition. The code itself includes the following modules:
  - **Start/Restart Module**. This module handles the following start phases:
    - **Start phase 1.** Load block reference and processor ID
The DBLQH Block

- **Start phase 2.** Initiate all records within the block; connect \texttt{LQH} with \texttt{ACC} and \texttt{TUP}.

- **Start phase 4.** Connect each \texttt{LQH} with every other \texttt{LQH} in the database system. For an initial start, create the fragment log files. For a system restart or node restart, open the fragment log files and find the end of the log files.

- **Fragment addition and deletion module.** Used by the data dictionary to create new fragments and delete old fragments.

- **Execution module.** This module handles the reception of \texttt{LQHKEYREQ} messages and all processing of operations on behalf of this request. This also involves reception of various types of \texttt{ATTRINFO} and \texttt{KEYINFO} messages, as well as communications with \texttt{ACC} and \texttt{TUP}.

- **Log module.** The log module handles the reading and writing of the log. It is also responsible for handling system restarts, and controls system restart in \texttt{TUP} and \texttt{ACC} as well.

- **Transaction module.** This module handles the commit and completion phases.

- **TC failure module.** Handles failures in the transaction coordinator.

- **Scan module.** This module contains the code that handles a scan of a particular fragment. It operates under the control of the transaction coordinator and orders \texttt{ACC} to perform a scan of all tuples in the fragment. \texttt{TUP} performs the necessary search conditions to insure that only valid tuples are returned to the application.

- **Node recovery module.** This is used when a node has failed, copying the effected fragment to a new fragment replica. It also shuts down all connections to the failed node.

- **LCP module.** This module handles execution and control of local checkpoints in \texttt{TUP} and \texttt{ACC}. It also interacts with \texttt{DIH} to determine which global checkpoints are recoverable.

- **Global checkpoint module.** Assists \texttt{DIH} in discovering when GCPs are recoverable, and handles the \texttt{GCP_SAVEREQ} message requesting that \texttt{LQH} save a given GCP to disk and provide a notification of when this has been done.

- **File handling module.** This includes a number of sub-modules:
  - Signal reception
  - Normal operation
  - File change
  - Initial start
  - System restart, Phase 1
  - System restart, Phase 2
  - System restart, Phase 3
  - System restart, Phase 4
  - Error
4.8 The DBSPJ Block

This block implements multiple cursors in the NDB kernel, providing handling for joins pushed down from SQL nodes. It contains the following files, which can be found in the directory `storage/ndb/src/kernel/blocks/dbspj`:

- **Dbspj.hpp**: Defines the `Dbspj` class.
- **DbspjInit.cpp**: `Dbspj` initialization.
- **DbspjMain.cpp**: Handles conditions pushed down from API and signal passing between `DBSPJ` and the `DBLQH` and `DBTC` kernel blocks.
- **DbspjProxy.hpp**
- **DbspjProxy.cpp**

4.9 The DBTC Block

This is the transaction coordinator block, which handles distributed transactions and other data operations on a global level (as opposed to `DBLQH` which deals with such issues on individual data nodes). In the source code, it is located in the directory `storage/ndb/src/kernel/blocks/dbtc`, which contains these files:

- **Dbtc.hpp**: Defines the `Dbtc` class and associated constructs, including the following:
  - **Trigger and index data** (`TcDefinedTriggerData`). A record forming a list of active triggers for each table. These records are managed by a trigger pool, in which a trigger record is seized whenever a trigger is activated, and released when the trigger is deactivated.
  - **Fired trigger data** (`TcFiredTriggerData`). A record forming a list of fired triggers for a given transaction.
  - **Index data** (`TcIndexData`). This record forms lists of active indexes for each table. Such records are managed by an index pool, in which each index record is seized whenever an index is created, and released when the index is dropped.
  - **API connection record** (`ApiConnectRecord`). An API connect record contains the connection record to which the application connects. The application can send one operation at a time. It can send a new operation immediately after sending the previous operation. This means that several
operations can be active in a single transaction within the transaction coordinator, which is achieved
by using the API connect record. Each active operation is handled by the TC connect record; as soon
as the TC connect record has sent the request to the local query handler, it is ready to receive new
operations. The LQH connect record takes care of waiting for an operation to complete; when an
operation has completed on the LQH connect record, a new operation can be started on the current
LQH connect record. ApiConnectRecord is always 256-byte aligned.

- **Transaction coordinator connection record (TcConnectRecord).** A TcConnectRecord keeps all information required for carrying out a transaction; the transaction controller establishes
  connections to the different blocks needed to carry out the transaction. There can be multiple records
  for each active transaction. The TC connection record cooperates with the API connection record for
  communication with the API node, and with the LQH connection record for communication with any
  local query handlers involved in the transaction. TcConnectRecord is permanently connected to a
  record in DBDICT and another in DIH, and contains a list of active LQH connection records and a list
  of started (but not currently active) LQH connection records. It also contains a list of all operations that
  are being executed with the current TC connection record. TcConnectRecord is always 128-byte
  aligned.

- **Cache record (CacheRecord).** This record is used between reception of a TCKEYREQ and sending
  of LQHKEYREQ (see Section 3.3, “Operations and Signals”). This is a separate record, so as to improve
  the cache hit rate and as well as to minimize memory storage requirements.

- **Host record (HostRecord).** This record contains the “alive” status of each node in the system, and
  is 128-byte aligned.

- **Table record (TableRecord).** This record contains the current schema versions of all tables in the
  system.

- **Scan record (ScanRecord).** Each scan allocates a ScanRecord to store information about the
  current scan.

- **Data buffer (DatabufRecord).** This is a buffer used for general data storage.

- **Attribute information record (AttrbufRecord).** This record can contain one (1) ATTRINFO signal,
  which contains a set of 32 attribute information words.

- **Global checkpoint information record (GcpRecord).** This record is used to store the
  globalcheckpoint number, as well as a counter, during the completion phase of the transaction. A
  GcpRecord is 32-byte aligned.

- **TC failure record (TC_FAIL_RECORD).** This is used when handling takeover of TC duties from a
  failed transaction coordinator.

- **DbtcInit.cpp:** Handles allocation and deallocation of Dbtc indexes and data (includes class
destructor).

- **DbtcMain.cpp:** Implements Dbtc methods.

---

**Note**

Any data node may act as the transaction coordinator.

The **DBTC** block is implemented as the **Dbtc** class.

The transaction coordinator is the kernel interface to which applications send their requests. It establishes
connections to different blocks in the system to carry out the transaction and decides which node
The DBTUP Block

will handle each transaction, sending a confirmation signal on the result to the application so that the application can verify that the result received from the TUP block is correct.

This block also handles unique indexes, which must be co-ordinated across all data nodes simultaneously.

4.10 The DBTUP Block

This is the tuple manager, which manages the physical storage of cluster data. It consists of the following files found in the directory `storage/ndb/src/kernel/blocks/dbtup`:

- **AttributeOffset.hpp**: Defines the `AttributeOffset` class, which models the structure of an attribute, permitting up to 4096 attributes, all of which are nullable.
- **DbtupDiskAlloc.cpp**: Handles allocation and deallocation of extents for disk space.
- **DbtupIndex.cpp**: Implements methods for reading and writing tuples using ordered indexes.
- **DbtupScan.cpp**: Implements methods for tuple scans.
- **tuppage.cpp**: Handles allocating pages for writing tuples.
- **tuppage.hpp**: Defines structures for fixed and variable size data pages for tuples.
- **DbtupAbort.cpp**: Contains routines for terminating failed tuple operations.
- **DbtupExecQuery.cpp**: Handles execution of queries for tuples and reading from them.
- **DbtupMeta.cpp**: Handle table operations for the `Dbtup` class.
- **DbtupStoredProcDef.cpp**: Module for adding and dropping procedures.
- **DbtupBuffer.cpp**: Handles read/write buffers for tuple operations.
- **DbtupFixAlloc.cpp**: Allocates and frees fixed-size tuples from the set of pages attached to a fragment. The fixed size is set per fragment; there can be only one such value per fragment.
- **DbtupPageMap.cpp**: Routines used by `Dbtup` to map logical page IDs to physical page IDs. The mapping needs the fragment ID and the logical page ID to provide the physical ID. This part of `Dbtup` is the exclusive user of a certain set of variables on the fragment record; it is also the exclusive user of the struct for page ranges (the `PageRange` struct defined in `Dbtup.hpp`).
- **DbtupTabDesMan.cpp**: This file contains the routines making up the table descriptor memory manager. Each table has a descriptor, which is a contiguous array of data words, and which is allocated from a global array using a “buddy” algorithm, with free lists existing for each $2^N$ words.
- **Notes.txt**: Contains some developers’ implementation notes on tuples, tuple operations, and tuple versioning.
- **Undo_buffer.hpp**: Defines the `Undo_buffer` class, used for storage of operations that may need to be rolled back.
- **Undo_buffer.cpp**: Implements some necessary `Undo_buffer` methods.
- **DbtupCommit.cpp**: Contains routines used to commit operations on tuples to disk.
- **DbtupGen.cpp**: This file contains `Dbtup` initialization routines.
- **DbtupPagMan.cpp**: This file implements the page memory manager’s “buddy” algorithm. `PageMan` is invoked when fragments lack sufficient internal page space to accommodate all the data they are requested to store. It is also invoked when fragments deallocate page space back to the free area.
The DBTUX Block

- **DbtupTrigger.cpp**: The routines contained in this file perform handling of NDB internal triggers.
- **DbtupDebug.cpp**: Used for debugging purposes only.
- **Dbtup.hpp**: Contains the Dbtup class definition. Also defines a number of essential structures such as tuple scans, disk allocation units, fragment records, and so on.
- **DbtupRoutines.cpp**: Implements Dbtup routines for reading attributes.
- **DbtupVarAlloc.cpp**
- **test_varpage.cpp**: Simple test program for verifying variable-size page operations.

This block also monitors changes in tuples.

### 4.11 The DBTUX Block

This kernel block provides local management of ordered indexes. It consists of the following files found in the `storage/ndb/src/kernel/blocks/dbtux` directory:

- **DbtuxCmp.cpp**: Implements routines to search by key versus node prefix or entry. The comparison starts at a given attribute position, which is updated by the number of equal initial attributes found. The entry data may be partial, in which case CmpUnknown may be returned. The attributes are normalized and have a variable size, given in words.
- **DbtuxGen.cpp**: Implements initialization routines used in node starts and restarts.
- **DbtuxMaint.cpp**: Contains routines used to maintain indexes.
- **DbtuxNode.cpp**: Implements routines for node creation, allocation, and deletion operations. Also assigns lists of scans to nodes.
- **DbtuxSearch.cpp**: Provides routines for handling node scan request messages.
- **DbtuxTree.cpp**: Routines for performing node tree operations.
- **Times.txt**: Contains some (old) performance figures from tests runs on operations using ordered indexes. Of historical interest only.
- **DbtuxDebug.cpp**: Debugging code for dumping node states.
- **Dbtux.hpp**: Contains Dbtux class definition.
- **DbtuxMeta.cpp**: Routines for creating, setting, and dropping indexes. Also provides means of aborting these operations in the event of failure.
- **DbtuxScan.cpp**: Routines for performing index scans.
- **DbtuxStat.cpp**: Implements methods for obtaining node statistics.
- **tuxstatus.html**: 2004-01-30 status report on ordered index implementation. Of historical interest only.

### 4.12 The DBUTIL Block

This block provides internal interfaces to transaction and data operations, performing essential operations on signals passed between nodes. This block implements transactional services which can then be used
by other blocks. It is also used in building online indexes, and is found in `storage/ndb/src/kernel/blocks/dbutil`, which includes these files:

- `DbUtil.cpp`: Implements `Dbutil` class methods
- `DbUtil.hpp`: Defines the `Dbutil` class, used to provide transactional services.
- `DbUtil.txt`: Implementation notes on utility protocols implemented by `DBUTIL`.

Among the duties performed by this block is the maintenance of sequences for backup IDs and other distributed identifiers.

### 4.13 The LGMAN Block

This block, the log group manager, is responsible for handling the undo logs for Disk Data tables. It is implemented in these files in the `storage/ndb/src/kernel/blocks` directory:

- `lgman.cpp`: Implements `Lgman` for adding, dropping, and working with log files and file groups.
- `lgman.hpp`: Contains the definition for the `Lgman` class, used to handle undo log files. Handles allocation of log buffer space.

### 4.14 The NDBCNTR Block

This is a cluster management block that handles block initialisation and configuration. During the data node startup process, it takes over from the QMGR block and continues the process. It also assists with graceful (planned) shutdowns of data nodes. This block is implemented in `storage/ndb/src/kernel/blocks/ndbcntr`, which contains these files:

- `Ndbcntr.hpp`: Defines the `Ndbcntr` class used to implement cluster management functions.
- `NdbcntrInit.cpp`: Initializers for `Ndbcntr` data and records.
- `NdbcntrMain.cpp`: Implements methods used for starts, restarts, and reading of configuration data.
- `NdbcntrSysTable.cpp`: NDBCNTR creates and initializes system tables on initial system start. The tables are defined in static structs in this file.

### 4.15 The NDBFS Block

This block provides the NDB file system abstraction layer, and is located in the directory `storage/ndb/src/kernel/blocks/ndbfs`, which contains the following files:

- `AsyncFile.hpp`: Defines the `AsyncFile` class, which represents an asynchronous file. All actions are executed concurrently with the other activities carried out by the process. Because all actions are performed in a separate thread, the result of an action is sent back through a memory channel. For the asynchronous notification of a finished request, each call includes a request as a parameter. This class is used for writing or reading data to and from disk concurrently with other activities.
- `AsyncFile.cpp`: Defines the actions possible for an asynchronous file, and implements them.
- `Filename.hpp`: Defines the `Filename` class. Takes a 128-bit value (as a array of four longs) and makes a file name out of it. This file name encodes information about the file, such as whether it is a file or a directory, and if the former, the type of file. Possible types include data file, fragment log, fragment list, table list, schema log, and system file, among others.
- `Filename.cpp`: Implements `set()` methods for the `Filename` class.
The PGMAN Block

- **MemoryChannelTest/MemoryChannelTest.cpp**: Basic program for testing reads from and writes to a memory channel (that is, reading from and writing to a circular buffer).

- **OpenFiles.hpp**: Implements an `OpenFiles` class, which provides some convenience methods for determining whether or not a given file is already open.

- **VoidFs.cpp**: Used for diskless operation. Generates a “dummy” acknowledgment to write operations.

- **CircularIndex.hpp**: The `CircularIndex` class, defined in this file, serves as the building block for implementing circular buffers. It increments as a normal index until it reaches maximum size, then resets to zero.

- **CircularIndex.cpp**: Contains only a single `#define`, not actually used at this time.

- **MemoryChannel.hpp**: Defines the `MemoryChannel` and `MemoryChannelMultipleWriter` classes, which provide a pointer-based channel for communication between two threads. It does not copy any data into or out of the channel, so the item that is put in can not be used untill the other thread has given it back. There is no support for detecting the return of an item.

- **MemoryChannel.cpp**: “Dummy” file, not used at this time.

- **Ndbfs.hpp**: Because an `NDB` signal request can result in multiple requests to `AsyncFile`, one class (defined in this file) is responsible for keeping track of all outstanding requests, and when all are finished, reporting the outcome back to the sending block.

- **Ndbfs.cpp**: Implements initialization and signal-handling methods for the `Ndbfs` class.

- **Pool.hpp**: Creates and manages a pool of objects for use by `Ndbfs` and other classes in this block.

- **AsyncFileTest/AsyncFileTest.cpp**: Test program, used to test and benchmark functionality of `AsyncFile`.

### 4.16 The PGMAN Block

This block provides page and buffer management services for Disk Data tables. It includes these files:

- **diskpage.hpp**: Defines the `File_formats`, `Datafile`, and `Undofile` structures.

- **diskpage.cpp**: Initializes sero page headers; includes some output reoutines for reporting and debugging.

- **pgman.hpp**: Defines the `Pgman` class implementing a number of page and buffer services, including page entries and requests, page replacement, page lists, page cleanup, and other page processing.

- **pgman.cpp**: Implements `Pgman` methods for initialization and various page management tasks.

- **PgmanProxy.hpp**

- **PgmanProxy.cpp**

### 4.17 The QMGR Block

This is the logical cluster management block, and handles node membership in the cluster using a heartbeat mechanism. `QMGR` is responsible for polling the data nodes when a data node failure occurs and determining that the node has actually failed and should be dropped from the cluster. This block contains the following files, found in storage/ndb/src/kernel/blocks/qmgr:
• **Qmgr.hpp**: Defines the Qmgr class and associated structures, including those used in detection of node failure and cluster partitioning.

• **QmgrInit.cpp**: Implements data and record initialization methods for Qmgr, as well as its destructor.

• **QmgrMain.cpp**: Contains routines for monitoring of heartbeats, detection and handling of “split-brain” problems, and management of some startup phases.

• **timer.hpp**: Defines the Timer class, used by NDB to keep strict timekeeping independent of the system clock.

This block also assists in the early phases of data node startup.

The QMGR block is implemented by the Qmgr class, whose definition is found in the file `storage/ndb/src/kernel/blocks/qmgr/Qmgr.hpp`.

### 4.18 The RESTORE Block

This block is implemented in the files `restore.hpp`, `restore.cpp`, `RestoreProxy.hpp`, and `RestoreProxy.cpp` in the `storage/ndb/src/kernel/blocks` directory. It handles restoration of the cluster from online backups. It is also used to restore local checkpoints as part of the process of starting a data node.

### 4.19 The SUMA Block

The cluster subscription manager, which handles event logging and reporting functions. It also figures prominently in NDB Cluster Replication. SUMA consists of the following files, found in the directory `storage/ndb/src/kernel/blocks/suma/`:

• **Suma.hpp**: Defines the Suma class and interfaces for managing subscriptions and performing necessary communications with other SUMA (and other) blocks.

• **SumaInit.cpp**: Performs initialization of DICT, DIH, and other interfaces

• **Suma.cpp**: Implements subscription-handling routines.

• **Suma.txt**: Contains a text-based diagram illustrating SUMA protocols.

### 4.20 The THRMAN Block

This is the thread management block, and executes in every NDB kernel thread. This block is also used measure thread CPU usage and to write this and other information into the threadblocks and threadstat tables in the ndbinfo information database.

The THRMAN block is implemented as the Thrman class, in the file `storage/ndb/src/kernel/blocks/thrman.hpp.thrman.cpp`, found in the same directory, defines a `measure_cpu_usage()` method of this class for measuring the CPU usage of a given thread. It also defines a `execDBINFO_SCANREQ()` method, which writes this and other information about the thread such as its thread ID and block number to the threadblocks and threadstat tables.

### 4.21 The TRPMAN Block

This is the signal transport management block of the NDB kernel, implemented in `storage/ndb/src/kernel/blocks/trpman.hpp` as the Trpman class, whose methods are defined in `trpman.cpp`, also in the `blocks` directory.
TRPMAN is also responsible for writing rows to the ndbinfo.transporters table.

### 4.22 The TSMAN Block

This is the tablespace manager block for Disk Data tables, implemented in the following files from storage/ndb/src/kernel/blocks:

- **tsman.hpp**: Defines the Tsman class, as well as structures representing data files and tablespaces.
- **tsman.cpp**: Implements Tsman methods.

### 4.23 The TRIX Block

This kernel block is responsible for the handling of internal triggers and unique indexes. TRIX, like DBUTIL, is a utility block containing many helper functions for building indexes and handling signals between nodes. It is implemented in the following files, all found in the directory storage/ndb/src/kernel/blocks/trix:

- **Trix.hpp**: Defines the Trix class, along with structures representing subscription data and records (for communicating with SUMA) and node data and ists (needed when communicating with remote TRIX blocks).
- **Trix.cpp**: Implements Trix class methods, including those necessary for taking appropriate action in the event of node failures.
Chapter 5 NDB Cluster Start Phases

Table of Contents

5.1 Initialization Phase (Phase -1) ................................................................. 99
5.2 Configuration Read Phase (STTOR Phase -1) ........................................ 100
5.3 STTOR Phase 0 ...................................................................................... 101
5.4 STTOR Phase 1 ...................................................................................... 102
5.5 STTOR Phase 2 ...................................................................................... 105
5.6 NDB_STTOR Phase 1 .............................................................................. 105
5.7 STTOR Phase 3 ...................................................................................... 105
5.8 NDB_STTOR Phase 2 .............................................................................. 105
5.9 STTOR Phase 4 ...................................................................................... 106
5.10 NDB_STTOR Phase 3 ............................................................................. 106
5.11 STTOR Phase 5 ...................................................................................... 107
5.12 NDB_STTOR Phase 4 ............................................................................. 107
5.13 NDB_STTOR Phase 5 ............................................................................. 107
5.14 NDB_STTOR Phase 6 ............................................................................. 108
5.15 STTOR Phase 6 ...................................................................................... 108
5.16 STTOR Phase 7 ...................................................................................... 109
5.17 STTOR Phase 8 ...................................................................................... 109
5.18 NDB_STTOR Phase 7 ............................................................................. 109
5.19 STTOR Phase 9 ...................................................................................... 109
5.20 STTOR Phase 101 .................................................................................. 109
5.21 System Restart Handling in Phase 4 ..................................................... 109
5.22 START_MEREQ Handling ................................................................. 110

The start of an NDB Cluster data node is processed in series of phases which is synchronised with other nodes that are starting up in parallel with this node as well as with nodes already started. The next several sections of this chapter describe each of these phases in detail.

5.1 Initialization Phase (Phase -1)

Before the data node actually starts, a number of other setup and initialization tasks must be done for the block objects, transporters, and watchdog checks, among others.

This initialization process begins in storage/ndb/src/kernel/main.cpp with a series of calls to globalEmulatorData.theThreadConfig->doStart(). When starting ndbd with the -n or --nostart option there is only one call to this method; otherwise, there are two, with the second call actually starting the data node. The first invocation of doStart() sends the START_ORD signal to the CMVMI block (see Section 4.2, "The CMVMI Block"); the second call to this method sends a START_ORD signal to NDBCNTR (see Section 4.14, "The NDBCNTR Block").

When START_ORD is received by the NDBCNTR block, the signal is immediately transferred to NDBCNTR's MISSRA sub-block, which handles the start process by sending a READ_CONFIG_REQ signals to all blocks in order as given in the array readConfigOrder:

1. NDBFS
2. DBTUP
3. DBACC
4. DBTC
5. DBLQH
6. DBTUX
7. DBDICT
8. DBDIH
9. NDBCNTR
10. QMGR
11. TRIX
12. BACKUP
13. DBUTIL
14. SUMA
15. TSMAN
16. LGMAN
17. PGMAN
18. RESTORE

NDBFS is permitted to run before any of the remaining blocks are contacted, in order to make sure that it can start the CMVMI block’s threads.

5.2 Configuration Read Phase (STTOR Phase -1)

The READ_CONFIG_REQ signal provides all kernel blocks an opportunity to read the configuration data, which is stored in a global object accessible to all blocks. All memory allocation in the data nodes takes place during this phase.

Note

Connections between the kernel blocks and the NDB file system are also set up during Phase 0. This is necessary to enable the blocks to communicate easily which parts of a table structure are to be written to disk.

NDB performs memory allocations in two different ways. The first of these is by using the allocRecord() method (defined in storage/ndb/src/kernel/vm/SimulatedBlock.hpp). This is the traditional method whereby records are accessed using the ptrCheckGuard macros (defined in storage/ndb/src/kernel/vm/pc.hpp). The other method is to allocate memory using the setSize() method defined with the help of the template found in storage/ndb/src/kernel/vm/CArray.hpp.

These methods sometimes also initialize the memory, ensuring that both memory allocation and initialization are done with watchdog protection.

Many blocks also perform block-specific initialization, which often entails building linked lists or doubly-linked lists (and in some cases hash tables).
Many of the sizes used in allocation are calculated in the `Configuration::calcSizeAlt()` method, found in `storage/ndb/src/kernel/vm/Configuration.cpp`.

Some preparations for more intelligent pooling of memory resources have been made. `DataMemory` and disk records already belong to this global memory pool.

5.3 STTOR Phase 0

Most NDB kernel blocks begin their start phases at STTOR Phase 1, with the exception of NDBFS and NDBCNTR, which begin with Phase 0, as can be seen by inspecting the first value for each element in the `ALL_BLOCKS` array (defined in `src/kernel/blocks/ndbcntr/NdbcntrMain.cpp`). In addition, when the STTOR signal is sent to a block, the return signal `STTORRY` always contains a list of the start phases in which the block has an interest. Only in those start phases does the block actually receive a STTOR signal.

STTOR signals are sent out in the order in which the kernel blocks are listed in the `ALL_BLOCKS` array. While NDBCNTR goes through start phases 0 to 255, most of these are empty.

Both activities in Phase 0 have to do with initialization of the NDB file system. First, if necessary, NDBFS creates the file system directory for the data node. In the case of an initial start, NDBCNTR clears any existing files from the directory of the data node to ensure that the DBDIH block does not subsequently discover any system files (if DBDIH were to find any system files, it would not interpret the start correctly as an initial start). (See also Section 4.5, “The DBDIH Block”.)

Each time that NDBCNTR completes the sending of one start phase to all kernel blocks, it sends a `NODE_STATE_REP` signal to all blocks, which effectively updates the NodeState in all blocks.

Each time that NDBCNTR completes a nonempty start phase, it reports this to the management server; in most cases this is recorded in the cluster log.

Finally, after completing all start phases, NDBCNTR updates the node state in all blocks using a `NODE_STATE_REP` signal; it also sends an event report advising that all start phases are complete. In addition, all other cluster data nodes are notified that this node has completed all its start phases to ensure all nodes are aware of one another's state. Each data node sends a `NODE_START_REP` to all blocks; however, this is significant only for DBDIH, so that it knows when it can unlock the lock for schema changes on DBDICT.

Note

In the following table, and throughout this text, we sometimes refer to STTOR start phases simply as “start phases” or “Phase N” (where N is some number). NDB_STTOR start phases are always qualified as such, and so referred to as “NDB_STTOR start phases” or “NDB_STTOR phases”.

<table>
<thead>
<tr>
<th>Kernel Block</th>
<th>Receptive Start Phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBFS</td>
<td>0</td>
</tr>
<tr>
<td>DBTC</td>
<td>1</td>
</tr>
<tr>
<td>DBDIH</td>
<td>1</td>
</tr>
<tr>
<td>DBLQH</td>
<td>1, 4</td>
</tr>
<tr>
<td>DBACC</td>
<td>1</td>
</tr>
<tr>
<td>DBTUP</td>
<td>1</td>
</tr>
</tbody>
</table>
5.4 STTOR Phase 1

This is one of the phases in which most kernel blocks participate (see the table in Section 5.3, "STTOR Phase 0"). Otherwise, most blocks are involved primarily in the initialization of data—for example, this is all that DBTC does.

Many blocks initialize references to other blocks in Phase 1. DBLQH initializes block references to DBTUP, and DBACC initializes block references to DBTUP and DBLQH. DBTUP initializes references to the blocks DBLQH, TSMAN, and LGMAN.

NDBCNTR initializes some variables and sets up block references to DBTUP, DBLQH, DBACC, DBTC, DBDIH, and DBDICT; these are needed in the special start phase handling of these blocks using NDB_STTOR signals, where the bulk of the node startup process actually takes place.

If the cluster is configured to lock pages (that is, if the LockPagesInMainMemory configuration parameter has been set), CMVMI handles this locking.

The QMGR block calls the initData() method (defined in storage/ndb/src/kernel/blocks/qmgr/QmgrMain.cpp) whose output is handled by all other blocks in the READ_CONFIG_REQ phase (see Section 5.1, "Initialization Phase (Phase -1)"). Following these initializations, QMGR sends the DIH_RESTARTREQ signal to DBDIH, which determines whether a proper system file exists; if it does, an initial start is not being performed. After the reception of this signal comes the process of integrating the node among the other data nodes in the cluster, where data nodes enter the cluster one at a time. The first one to enter becomes the master; whenever the master dies the new master is always the node that has been running for the longest time from those remaining.

QMGR sets up timers to ensure that inclusion in the cluster does not take longer than what the cluster's configuration is set to permit (see Controlling Timeouts, Intervals, and Disk Paging for the relevant configuration parameters), after which communication to all other data nodes is established. At this point,
a **CM_REGREQ** signal is sent to all data nodes. Only the president of the cluster responds to this signal; the president permits one node at a time to enter the cluster. If no node responds within 3 seconds then the president becomes the master. If several nodes start up simultaneously, then the node with the lowest node ID becomes president. The president sends **CM_REGCONF** in response to this signal, but also sends a **CM_ADD** signal to all nodes that are currently alive.

Next, the starting node sends a **CM_NODEINFOREQ** signal to all current “live” data nodes. When these nodes receive that signal they send a **NODE_VERSION_REP** signal to all API nodes that have connected to them. Each data node also sends a **CM_ACKADD** to the president to inform the president that it has heard the **CM_NODEINFOREQ** signal from the new node. Finally, each of the current data nodes sends the **CM_NODEINFOCONF** signal in response to the starting node. When the starting node has received all these signals, it also sends the **CM_ACKADD** signal to the president.

When the president has received all of the expected **CM_ACKADD** signals, it knows that all data nodes (including the newest one to start) have replied to the **CM_NODEINFOREQ** signal. When the president receives the final **CM_ACKADD**, it sends a **CM_ADD** signal to all current data nodes (that is, except for the node that just started). Upon receiving this signal, the existing data nodes enable communication with the new node; they begin sending heartbeats to it and including in the list of neighbors used by the heartbeat protocol.

The **start** struct is reset, so that it can handle new starting nodes, and then each data node sends a **CM_ACKADD** to the president, which then sends a **CM_ADD** to the starting node after all such **CM_ACKADD** signals have been received. The new node then opens all of its communication channels to the data nodes that were already connected to the cluster; it also sets up its own heartbeat structures and starts sending heartbeats. It also sends a **CM_ACKADD** message in response to the president.

The signalling between the starting data node, the already “live” data nodes, the president, and any API nodes attached to the cluster during this phase is shown in the following diagram:
As a final step, `QMGR` also starts the timer handling for which it is responsible. This means that it generates a signal to blocks that have requested it. This signal is sent 100 times per second even if any one instance of the signal is delayed.

The `BACKUP` kernel block also begins sending a signal periodically. This is to ensure that excessive amounts of data are not written to disk, and that data writes are kept within the limits of what has been specified in the cluster configuration file during and after restarts. The `DBUTIL` block initializes the transaction identity, and `DBTUX` creates a reference to the `DBTUP` block, while `PGMAN` initializes pointers to the `LGMAN` and `DBTUP` blocks. The `RESTORE` kernel block creates references to the `DBLQH` and `DBTUP` blocks to enable quick access to those blocks when needed.
5.5 STTOR Phase 2

The only kernel block that participates in this phase to any real effect is NDBCNTR.

In this phase NDBCNTR obtains the current state of each configured cluster data node. Messages are sent to NDBCNTR from QMGR reporting the changes in status of any the nodes. NDBCNTR also sets timers corresponding to the StartPartialTimeout, StartPartitionTimeout, and StartFailureTimeout configuration parameters.

The next step is for a CNTR_START_REQ signal to be sent to the proposed master node. Normally the president is also chosen as master. However, during a system restart where the starting node has a newer global checkpoint than that which has survived on the president, then this node will take over as master node, even though it is not recognized as the president by QMGR. If the starting node is chosen as the new master, then the other nodes are informed of this using a CNTR_START_REF signal.

The master withholds the CNTR_START_REQ signal until it is ready to start a new node, or to start the cluster for an initial restart or system restart.

When the starting node receives CNTR_START_CONF, it starts the NDB_STTOR phases, in the following order:

1. DBLQH
2. DBDICT
3. DBTUP
4. DBACC
5. DBTC
6. DBDIH

5.6 NDB_STTOR Phase 1

DBDICT, if necessary, initializes the schema file. DBDIH, DBTC, DBTUP, and DBLQH initialize variables. DBLQH also initializes the sending of statistics on database operations.

5.7 STTOR Phase 3

DBDICT initializes a variable that keeps track of the type of restart being performed.

NDBCNTR executes the second of the NDB_STTOR start phases, with no other NDBCNTR activity taking place during this STTOR phase.

5.8 NDB_STTOR Phase 2

The DBLQH block enables its exchange of internal records with DBTUP and DBACC, while DBTC permits its internal records to be exchanged with DBDIH. The DBDIH kernel block creates the mutexes used by the NDB kernel and reads nodes using the READ_NODESREQ signal. With the data from the response to this signal, DBDIH can create node lists, node groups, and so forth. For node restarts and initial node restarts, DBDIH also asks the master for permission to perform the restart. The master will ask all “live” nodes if they are prepared to permit the new node to join the cluster. If an initial node restart is to be performed, then all LCPs are invalidated as part of this phase.
LCPs from nodes that are not part of the cluster at the time of the initial node restart are not invalidated. The reason for this is that there is never any chance for a node to become master of a system restart using any of the LCPs that have been invalidated, since this node must complete a node restart—including a local checkpoint—before it can join the cluster and possibly become a master node.

The CMVMI kernel block activates the sending of packed signals, which occurs only as part of database operations. Packing must be enabled prior to beginning any such operations during the execution of the redo log or node recovery phases.

The DBTUX block sets the type of start currently taking place, while the BACKUP block sets the type of restart to be performed, if any (in each case, the block actually sets a variable whose value reflects the type of start or restart). The SUMA block remains inactive during this phase.

The PGMAN kernel block starts the generation of two repeated signals, the first handling cleanup. This signal is sent every 200 milliseconds. The other signal handles statistics, and is sent once per second.

5.9 STTOR Phase 4

Only the DBLQH and NDBCNTR kernel blocks are directly involved in this phase. DBLQH allocates a record in the BACKUP block, used in the execution of local checkpoints using the DEFINE_BACKUP_REQ signal. NDBCNTR causes NDB_STTOR to execute NDB_STTOR phase 3; there is otherwise no other NDBCNTR activity during this STTOR phase.

5.10 NDB_STTOR Phase 3

The DBLQH block initiates checking of the log files here. Then it obtains the states of the data nodes using the READ_NODESREQ signal. Unless an initial start or an initial node restart is being performed, the checking of log files is handled in parallel with a number of other start phases. For initial starts, the log files must be initialized; this can be a lengthy process and should have some progress status attached to it.

Note

From this point, there are two parallel paths, one continuing to restart and another reading and determining the state of the redo log files.

The DBDIH block requests information about the cluster data nodes using the READ_NODESREQ signal. DBACC resets the system restart flag if this is not a system restart; this is used only to verify that no requests are received from DBTUX during system restart. DBTC requests information about all nodes by means of the READ_NODESREQ signal.

DBDIH sets an internal master state and makes other preparations exclusive to initial starts. In the case of an initial start, the nonmaster nodes perform some initial tasks, the master node doing once all nonmaster nodes have reported that their tasks are completed. (This delay is actually unnecessary since there is no reason to wait while initializing the master node.)

For node restarts and initial node restarts no more work is done in this phase. For initial starts the work is done when all nodes have created the initial restart information and initialized the system file.

For system restarts this is where most of the work is performed, initiated by sending the NDB_STARTREQ signal from NDBCNTR to DBDIH in the master. This signal is sent when all nodes in the system restart have reached this point in the restart. This we can mark as our first synchronization point for system restarts, designated WAITPOINT_4_1.

For a description of the system restart version of Phase 4, see Section 5.21, “System Restart Handling in Phase 4”.
After completing execution of the NDB_STARTREQ signal, the master sends a CNTR_WAITREP signal with WAITPOINT_4_2 to all nodes. This ends NDB_STTOR phase 3 as well as (STTOR) Phase 4.

5.11 STTOR Phase 5

All that takes place in Phase 5 is the delivery by NDBCNTR of NDB_STTOR phase 4; the only block that acts on this signal is DBDIH that controls most of the part of a data node start that is database-related.

5.12 NDB_STTOR Phase 4

Some initialization of local checkpoint variables takes place in this phase, and for initial restarts, this is all that happens in this phase.

For system restarts, all required takeovers are also performed. Currently, this means that all nodes whose states could not be recovered using the redo log are restarted by copying to them all the necessary data from the "live" data nodes.

For node restarts and initial node restarts, the master node performs a number of services, requested to do so by sending the START_MEREQ signal to it. This phase is complete when the master responds with a START_MECONF message, and is described in Section 5.22, "START_MEREQ Handling".

After ensuring that the tasks assigned to DBDIH tasks in the NDB_STTOR phase 4 are complete, NDBCNTR performs some work on its own. For initial starts, it creates the system table that keeps track of unique identifiers such as those used for AUTO_INCREMENT. Following the WAITPOINT_4_1 synchronization point, all system restarts proceed immediately to NDB_STTOR phase 5, which is handled by the DBDIH block. See Section 5.13, "NDB_STTOR Phase 5", for more information.

5.13 NDB_STTOR Phase 5

For initial starts and system restarts this phase means executing a local checkpoint. This is handled by the master so that the other nodes will return immediately from this phase. Node restarts and initial node restarts perform the copying of the records from the primary replica to the starting replicas in this phase. Local checkpoints are enabled before the copying process is begun.

Copying the data to a starting node is part of the node takeover protocol. As part of this protocol, the node status of the starting node is updated; this is communicated using the global checkpoint protocol. Waiting for these events to take place ensures that the new node status is communicated to all nodes and their system files.

After the node's status has been communicated, all nodes are signaled that we are about to start the takeover protocol for this node. Part of this protocol consists of Steps 3 - 9 during the system restart phase as described below. This means that restoration of all the fragments, preparation for execution of the redo log, execution of the redo log, and finally reporting back to DBDIH when the execution of the redo log is completed, are all part of this process.

After preparations are complete, copy phase for each fragment in the node must be performed. The process of copying a fragment involves the following steps:

1. The DBLQH kernel block in the starting node is informed that the copy process is about to begin by sending it a PREPARE_COPY_FRAGREQ signal.

2. When DBLQH acknowledges this request a CREATE_FRAGREQ signal is sent to all nodes notify them of the preparation being made to copy data to this replica for this table fragment.
3. After all nodes have acknowledged this, a \texttt{COPY_FRAGREQ} signal is sent to the node from which the data is to be copied to the new node. This is always the primary replica of the fragment. The node indicated copies all the data over to the starting node in response to this message.

4. After copying has been completed, and a \texttt{COPY_FRAGCONF} message is sent, all nodes are notified of the completion through an \texttt{UPDATE_TOREQ} signal.

5. After all nodes have updated to reflect the new state of the fragment, the \texttt{DBLQH} kernel block of the starting node is informed of the fact that the copy has been completed, and that the replica is now up-to-date and any failures should now be treated as real failures.

6. The new replica is transformed into a primary replica if this is the role it had when the table was created.

7. After completing this change another round of \texttt{CREATE_FRAGREQ} messages is sent to all nodes informing them that the takeover of the fragment is now committed.

8. After this, process is repeated with the next fragment if another one exists.

9. When there are no more fragments for takeover by the node, all nodes are informed of this by sending an \texttt{UPDATE_TOREQ} signal sent to all of them.

10. Wait for the next complete local checkpoint to occur, running from start to finish.

11. The node states are updated, using a complete global checkpoint. As with the local checkpoint in the previous step, the global checkpoint must be permitted to start and then to finish.

12. When the global checkpoint has completed, it will communicate the successful local checkpoint of this node restart by sending an \texttt{END_TOREQ} signal to all nodes.

13. A \texttt{START_COPYCONF} is sent back to the starting node informing it that the node restart has been completed.

14. Receiving the \texttt{START_COPYCONF} signal ends NDB_STTOR phase 5. This provides another synchronization point for system restarts, designated as \texttt{WAITPOINT_5_2}.

\textbf{Note}

The copy process in this phase can in theory be performed in parallel by several nodes. However, all messages from the master to all nodes are currently sent to single node at a time, but can be made completely parallel. This is likely to be done in the not too distant future.

In an initial and an initial node restart, the \texttt{SUMA} block requests the subscriptions from the \texttt{SUMA} master node. \texttt{NDBCNTR} executes NDB_STTOR phase 6. No other \texttt{NDBCNTR} activity takes place.

### 5.14 NDB\_STTOR Phase 6

In this NDB\_STTOR phase, both \texttt{DBLQH} and \texttt{DBDICT} clear their internal representing the current restart type. The \texttt{DBACC} block resets the system restart flag; \texttt{DBACC} and \texttt{DBTUP} start a periodic signal for checking memory usage once per second. \texttt{DBTC} sets an internal variable indicating that the system restart has been completed.

### 5.15 STTOR Phase 6

The \texttt{NDBCNTR} block defines the cluster's node groups, and the \texttt{DBUTIL} block initializes a number of data structures to facilitate the sending keyed operations can be to the system tables. \texttt{DBUTIL} also sets up a single connection to the \texttt{DBTC} kernel block.
5.16 STTOR Phase 7

In QMGR the president starts an arbitrator (unless this feature has been disabled by setting the value of the ArbitrationRank configuration parameter to 0 for all nodes—see Defining an NDB Cluster Management Server, and Defining SQL and Other API Nodes in an NDB Cluster, for more information; note that this currently can be done only when using NDB Cluster Carrier Grade Edition). In addition, checking of API nodes through heartbeats is activated.

Also during this phase, the BACKUP block sets the disk write speed to the value used following the completion of the restart. The master node during initial start also inserts the record keeping track of which backup ID is to be used next. The SUMA and DBTUX blocks set variables indicating start phase 7 has been completed, and that requests to DBTUX that occurs when running the redo log should no longer be ignored.

5.17 STTOR Phase 8

NDB_STTOR executes NDB_STTOR phase 7; no other NDBCNT activity takes place.

5.18 NDB_STTOR Phase 7

If this is a system restart, the master node initiates a rebuild of all indexes from DBDICT during this phase.

The CMVMI kernel block opens communication channels to the API nodes (including MySQL servers acting as SQL nodes). Indicate in globalData that the node is started.

5.19 STTOR Phase 9

NDBCNT resets some start variables.

5.20 STTOR Phase 101

This is the SUMA handover phase, during which a GCP is negotiated and used as a point of reference for changing the source of event and replication subscriptions from existing nodes only to include a newly started node.

5.21 System Restart Handling in Phase 4

This consists of the following steps:

1. The master sets the latest GCI as the restart GCI, and then synchronizes its system file to all other nodes involved in the system restart.

2. The next step is to synchronize the schema of all the nodes in the system restart. This is performed in 15 passes. The problem we are trying to solve here occurs when a schema object has been created while the node was up but was dropped while the node was down, and possibly a new object was even created with the same schema ID while that node was unavailable. In order to handle this situation, it is necessary first to re-create all objects that are supposed to exist from the viewpoint of the starting node. After this, any objects that were dropped by other nodes in the cluster while this node was "dead" are dropped; this also applies to any tables that were dropped during the outage. Finally, any tables that have been created by other nodes while the starting node was unavailable are re-created on the starting node. All these operations are local to the starting node. As part of this process, it is also necessary to ensure that all tables that need to be re-created have been created locally and that the proper data structures have been set up for them in all kernel blocks.

After performing the procedure described previously for the master node the new schema file is sent to all other participants in the system restart, and they perform the same synchronization.
3. All fragments involved in the restart must have proper parameters as derived from DBDIH. This causes a number of START_FRAGREQ signals to be sent from DBDIH to DBLQH. This also starts the restoration of the fragments, which are restored one by one and one record at a time in the course of reading the restore data from disk and applying in parallel the restore data read from disk into main memory. This restores only the main memory parts of the tables.

4. Once all fragments have been restored, a START_RECREQ message is sent to all nodes in the starting cluster, and then all undo logs for any Disk Data parts of the tables are applied.

5. After applying the undo logs in LGMAN, it is necessary to perform some restore work in TSMAN that requires scanning the extent headers of the tablespaces.

6. Next, it is necessary to prepare for execution of the redo log, which log can be performed in up to four phases. For each fragment, execution of redo logs from several different nodes may be required. This is handled by executing the redo logs in different phases for a specific fragment, as decided in DBDIH when sending the START_FRAGREQ signal. An EXEC_FRAGREQ signal is sent for each phase and fragment that requires execution in this phase. After these signals are sent, an EXEC_SRREQ signal is sent to all nodes to tell them that they can start executing the redo log.

7. Prior to executing the redo log, it is necessary to calculate where to start reading and where the end of the REDO log should have been reached. The end of the REDO log should be found when the last GCI to restore has been reached.

8. After completing the execution of the redo logs, all redo log pages that have been written beyond the last GCI to be restore are invalidated. Given the cyclic nature of the redo logs, this could carry the invalidation into new redo log files past the last one executed.

9. After the completion of the previous step, DBLQH report this back to DBDIH using a START_RECONF message.

10. When the master has received this message back from all starting nodes, it sends a NDB_STARTCONF signal back to NDBCNTR.

11. The NDB_STARTCONF message signals the end of STTOR phase 4 to NDBCNTR, which is the only block involved to any significant degree in this phase.

### 5.22 START_MEREQ Handling

The first step in handling START_MEREQ is to ensure that no local checkpoint is currently taking place; otherwise, it is necessary to wait until it is completed. The next step is to copy all distribution information from the master DBDIH to the starting DBDIH. After this, all metadata is synchronized in DBDICT (see Section 5.21, “System Restart Handling in Phase 4”).

After blocking local checkpoints, and then synchronizing distribution information and metadata information, global checkpoints are blocked.

The next step is to integrate the starting node in the global checkpoint protocol, local checkpoint protocol, and all other distributed protocols. As part of this the node status is also updated.

After completing this step the global checkpoint protocol is permitted to start again, the START_MECONF signal is sent to indicate to the starting node that the next phase may proceed.
Chapter 6 NDB Schema Object Versions

NDB supports online schema changes. A schema object such as a Table or Index has a 4-byte schema object version identifier, which can be observed in the output of the ndb_desc utility (see ndb_desc — Describe NDB Tables), as shown here (emphasized text):

```
shell> ndb_desc -c 127.0.0.1 -d test t1
-- t1 --
Version: 33554434
Fragment type: HashMapPartition
K Value: 6
Min load factor: 78
Max load factor: 80
Temporary table: no
Number of attributes: 3
Number of primary keys: 1
Length of frm data: 269
Row Checksum: 1
Row GCI: 1
SingleUserMode: 0
ForceVarPart: 1
FragmentCount: 4
ExtraRowGciBits: 0
ExtraRowAuthorBits: 0
TableStatus: Retrieved
HashMap: DEFAULT-HASHMAP-240-4
-- Attributes --
c1 Int PRIMARY KEY DISTRIBUTION KEY AT=FIXED ST=MEMORY AUTO_INCR
c2 Int NULL AT=FIXED ST=MEMORY
c6 Varchar(50;latin1_swedish_ci) NOT NULL AT=SHORT_VAR ST=MEMORY
-- Indexes --
PRIMARY KEY(c1) - UniqueHashIndex
PRIMARY(c1) - OrderedIndex
NDBT_ProgramExit: 0 - OK
```

The schema object version identifier (or simply “schema version”) is made up of a major version and a minor version; the major version occupies the (single) least significant byte of the schema version, and the minor version the remaining (3 most significant) bytes. You can see these two components more easily when viewing the schema version in hexadecimal notation. In the example output just shown, the schema version is shown as 33554434, which in hexadecimal (filling in leading zeroes as necessary) is 0x02000002; this is equivalent to major version 2, minor version 2. Adding an index to table t1 causes the schema version as reported by ndb_desc to advance to 50331650, or 0x03000002 hexadecimal, which is equivalent to major version 2 (3 least significant bytes 00 00 02), minor version 3 (most significant byte 03). Minor schema versions start with 0 for a newly created table.

In addition, each NDB API database object class has its own getObjectVersion() method that, like Object::getObjectVersion(), returns the object's schema object version. This includes instances, not only of Object, but of Table, Index, Column, LogfileGroup, Tablespace, Datafile, and UndoFile, as well as Event. (However, NdbBlob::getVersion() has a purpose and function that is completely unrelated to that of the methods just listed.)

Schema changes which are considered backward compatible—such as adding a DEFAULT or NULL column at the end of a table—cause the table object's minor version to be incremented. Schema changes which are not considered backward compatible—such as removing a column from a table—cause the major version to be incremented.
Note

While the implementation of an operation causing a schema major version change may actually involve 2 copies of the affected table (dropping and recreating the table), the final outcome can be observed as an increase in the table's major version.

Queries and DML operations which arrive from NDB clients also have an associated schema version, which is checked at the start of processing in the data nodes. If the schema version of the request differs from the affected database object's latest schema version only in its minor version component, the operation is considered compatible and is allowed to proceed. If the schema version differs in the major schema version then it will be rejected.

This mechanism allows the schema to be changed in the data nodes in various ways, without requiring a synchronized schema change in clients. Clients need not move on to the new schema version until they are ready to do so. Queries and DML operations can thus continue uninterrupted.

The NDB API and schema object versions. An NDB API application normally uses an NdbDictionary object associated with an Ndb object to retrieve schema objects. Schema objects are retrieved on demand from the data nodes; signalling is used to obtain the table or index definition; then, a local memory object is constructed which the application can use. NDB internally caches schema objects, so that each successive request for the same table or index by name does not require signalling.

Global schema cache. To avoid the need to signal to the data nodes for every schema object lookup, a schema cache is used for each Ndb_cluster_connection. This is referred to as the global schema cache. It is global in terms of spanning multiple Ndb objects. Instantiated table and index objects are automatically put into this cache to save on future signalling and instantiation costs. The cache maintains a reference count for each object; this count is used to determine when a given schema object can be deleted. Schema objects can have their reference counts modified by explicit API method calls or local schema cache operations.

Local schema cache. In addition to the per-connection global schema cache, each Ndb object's NdbDictionary object has a local schema cache. This cache contains pointers to objects held in the global schema cache. Each local schema cache holding a reference to a schema object in the global schema cache increments the global schema cache reference count by 1. Having a schema cache that is local to each Ndb object allows schema objects to be looked up without imposing any locks. The local schema cache is normally emptied (reducing global cache reference counts in the process) when its associated Ndb object is deleted.

Operation without schema changes. Normal operation proceeds as follows in the cases listed below:

A. A table is requested by some client (Ndb object) for the first time. The local cache is checked; the attempt results in a miss. The global cache is then also checked (using a lock), and the result is another miss.

Since there were no cache hits, the data node is sent a signal; the node's response is used to instantiate the table object. A pointer to the instantiated data object is added to the global cache; another such pointer is added to the local cache, and the reference count is set to 1. A pointer to the table is returned to the client.

B. A second client (a different Ndb object) requests access to the same table, also by name. A check of the local cache results in a miss, but a check of the global cache yields a hit.

As a result, an object pointer is added to the local cache, the global reference count is incremented—so that its value is now 2—and an object pointer is returned to the client. No new pointer is added to the global cache.
C. **For a second time, the second client requests access to same table by name.** The local cache is checked, producing a hit. An object pointer is immediately returned to the client. No pointers are added to the local or global caches, and the object's reference count is not incremented (and so the reference count remains constant at 2).

D. **Second client deletes Ndb object.** Objects in this client's local schema cache have their reference counts decremented in global cache.

This sets the global cache reference count to 1. Since it is not yet 0, no action is yet taken to remove the parent Ndb object.

**Schema changes.** Assuming that an object's schema never changes, the schema version first retrieved is used for the lifetime of the application process, and the in-memory object is deleted only when all local cache references (that is, all references to Ndb objects) have been deleted. This is unlikely to occur other than during a shutdown or cluster connection reset.

If an object's schema changes in a backward-compatible way while an application is running, this has the following affects:

- The minor version at the data nodes is incremented. (Ongoing DML operations using the old schema version still succeed.)
- NDB API clients subsequently retrieving the latest version of the schema object then fetch the new schema version.
- NDB API clients with cached older versions do not use the new schema version unless and until their local and global caches are invalidated.
- NDB API clients subscribing to events can observe a **TE_ALTER** event for the table in question, and can use this to trigger schema object cache invalidations.
- Each local cache entry can be removed by calling `removeCachedTable()` or `removeCachedIndex()`. This removes the entry from the local cache, and decrements the reference count in the global cache. When (and if) the global cache reference count reaches zero, the old cached object can be deleted.
- Alternatively, local cache entries can be removed, and the global cache entry invalidated, by calling `invalidateTable()` or `invalidateIndex()`. Subsequent calls to `getTable()` or `getIndex()` for this and other clients return the new schema object version by signalling the data nodes and instantiating a new object.
- New Ndb objects fill their local table caches on demand from the global table cache as normal. This means that, once an old schema object has been invalidated in the global cache, such objects retrieve the latest table objects known at the time that the table objects are first cached.

When an incompatible schema change is made (that is, a schema major version change), NDB API requests using the old version fail as soon as the new version is committed. This can also be used as a trigger to retrieve a new schema object version.

The rules governing the handling of schema version changes are summarized in the following list:

- An online schema change (minor version change) does not affect existing clients (Ndb objects); clients can continue to use the old schema object version
- If and only if a client voluntarily removes cached objects by making API calls can it then observe the new schema object version.
• As Ndb objects remove cached objects and are deleted, the reference count on the old schema object version decreases.

• When this reference count reaches 0, the object can be deleted.

Implications of the schema object lifecycle. The lifespan of a schema object (such as a Table or Index) is limited by the lifetime of the Ndb object from which it is obtained. When the parent Ndb object of a schema object is deleted, the reference count which keeps the Ndb object alive is decremented. If this Ndb object holds the last remaining reference to a given schema object version, the deletion of the Ndb object can also result in the deletion of the schema object. For this reason, no other threads can be using the object at this time.

Care must be exercised when pointers to schema objects are held in the application and used between multiple Ndb objects. A schema object should not be used beyond the lifespan of the Ndb object which created it.

Applications can respond, asynchronously and independently of each other, to backward-compatible schema changes, moving to the new schema only when necessary. Different threads can operate on different schema object versions concurrently.

It is thus very important to ensure that schema objects do not outlive the Ndb objects used to create them. To help prevent this from happening, you can take any of the following actions to invalidate old schema objects:

• To trigger invalidation when and as needed, use NDB API TE_ALTER events (see Event::TableEvent).

• Use an external trigger to initiate invalidation.

• Perform a periodic invalidation explicitly.

Invalidating the caches in any of these ways allows applications to obtain new versions of schema objects as required.

It is also worth noting that not all NDB API Table getter methods return pointers; many of them (in addition to Table::getName()) return table names. Such methods include Index::getTable(), NdbOperation::getTableName(), Event::getTableName(), and NdbDictionary::getRecordTableName().
# Chapter 7 NDB Cluster API Errors

## Table of Contents

7.1 Data Node Error Messages ................................................................. 115  
7.1.1 ndbd Error Codes ................................................................. 115  
7.1.2 ndbd Error Classifications ...................................................... 119  
7.2 NDB Transporter Errors ................................................................. 120  

This section provides a listing of exit codes and messages returned by a failed data node (ndbd or ndbmtld) process, as well as NDB transporter error log messages.

For information about error handling and error codes for the NDB API, see NDB API Errors and Error Handling. For information about error handling and error codes for the MGM API, see MGM API Errors, as well as The ndb_mgm_error Type.

## 7.1 Data Node Error Messages

This section contains exit codes and error messages given when a data node process stops prematurely.

### 7.1.1 ndbd Error Codes

This section lists all the error messages that can be returned when a data node process halts due to an error, arranged in most cases according to the affected NDB kernel block.

For more information about kernel blocks, see Chapter 4, NDB Kernel Blocks.

The meanings of the values given in the Classification column of each of the following tables is given in Section 7.1.2, “ndbd Error Classifications”.

#### 7.1.1.1 General Errors

This section contains ndbd error codes that are either generic in nature or otherwise not associated with a specific NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_GENERIC</td>
<td>XRE</td>
<td>Generic error</td>
</tr>
<tr>
<td>NDBD_EXIT_PRGERR</td>
<td>XIE</td>
<td>Assertion</td>
</tr>
<tr>
<td>NDBD_EXIT_NODE_NOTIN_CONFIG</td>
<td>XCE</td>
<td>Node ID in the configuration has the wrong type (that is, it is not a data node)</td>
</tr>
<tr>
<td>NDBD_EXIT_SYSTEM</td>
<td>XIE</td>
<td>System error, node killed during node restart by other node</td>
</tr>
<tr>
<td>NDBD_EXIT_INDEX</td>
<td>XIE</td>
<td>Array index out of range</td>
</tr>
<tr>
<td>NDBD_EXIT_ARBIT_SHUTDOWN</td>
<td>XAE</td>
<td>Node lost connection to other nodes and cannot form a unpartitioned cluster, please investigate if there are error(s) on other node(s)</td>
</tr>
<tr>
<td>NDBD_EXIT_PARTITIONED_SHUTDOWN</td>
<td>XAE</td>
<td>Partitioned cluster detected. Please check if cluster is already running</td>
</tr>
<tr>
<td>Error Code</td>
<td>Error Classification</td>
<td>Error Text</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>NDBD_EXIT_NODE_DECLARED_DEAD</td>
<td>XAE</td>
<td>Node declared dead. See error log for details</td>
</tr>
<tr>
<td>NDBD_EXIT_POINTER_TOO_LARGE</td>
<td>XIE</td>
<td>Pointer too large</td>
</tr>
<tr>
<td>NDBD_EXIT_SR_OTHER_NODE_FAILED</td>
<td>XRE</td>
<td>Another node failed during system restart, please investigate error(s) on other node(s)</td>
</tr>
<tr>
<td>NDBD_EXIT_NODE_NOT_DEAD</td>
<td>XRE</td>
<td>Internal node state conflict, most probably resolved by restarting node again</td>
</tr>
<tr>
<td>NDBD_EXIT_SR_REDOLOG</td>
<td>XFI</td>
<td>Error while reading the REDO log</td>
</tr>
<tr>
<td>NDBD_EXIT_SR_SCHEMAFILE</td>
<td>XFI</td>
<td>Error while reading the schema file</td>
</tr>
<tr>
<td>2311</td>
<td>XIE</td>
<td>Conflict when selecting restart type</td>
</tr>
<tr>
<td>NDBD_EXIT_NO_MORE_UNDOLOG</td>
<td>XCE</td>
<td>No more free UNDO log, increase UndoIndexBuffer</td>
</tr>
<tr>
<td>NDBD_EXIT_SR_UNDOLOG</td>
<td>XFI</td>
<td>Error while reading the data pages and UNDO log</td>
</tr>
<tr>
<td>NDBD_EXIT_SINGLEUSER_MODE</td>
<td>XRE</td>
<td>Data node is not allowed to get added to the cluster while it is in single user mode</td>
</tr>
<tr>
<td>NDBD_EXIT_MEMALLOC</td>
<td>XRE</td>
<td>Memory allocation failure, please decrease some configuration parameters</td>
</tr>
<tr>
<td>NDBD_EXIT_BLOCK_BNR_ZERO</td>
<td>XIE</td>
<td>Send signal error</td>
</tr>
<tr>
<td>NDBD_EXIT_TIME_QUEUE_SHORT</td>
<td>XIE</td>
<td>Error in short time queue</td>
</tr>
<tr>
<td>NDBD_EXIT_TIME_QUEUE_LONG</td>
<td>XIE</td>
<td>Error in long time queue</td>
</tr>
<tr>
<td>NDBD_EXIT_TIME_QUEUE_DELAY</td>
<td>XIE</td>
<td>Error in time queue, too long delay</td>
</tr>
<tr>
<td>NDBD_EXIT_TIME_QUEUE_INDEX</td>
<td>XIE</td>
<td>Time queue index out of range</td>
</tr>
<tr>
<td>NDBD_EXIT_BLOCK_BNZERO</td>
<td>XRE</td>
<td>Send signal error</td>
</tr>
<tr>
<td>NDBD_EXIT_WRONG_PRIOR_LEVEL</td>
<td>XRE</td>
<td>Wrong priority level when sending signal</td>
</tr>
<tr>
<td>NDBD_EXIT_NDBREQUIRE</td>
<td>XRE</td>
<td>Internal program error (failed ndbrequire)</td>
</tr>
<tr>
<td>NDBD_EXIT_NDBASSERT</td>
<td>XRE</td>
<td>Internal program error (failed ndbassert)</td>
</tr>
<tr>
<td>NDBD_EXIT_ERROR_INSERT</td>
<td>XRE</td>
<td>Error insert executed</td>
</tr>
<tr>
<td>NDBD_EXIT_INVALID_CONFIG</td>
<td>XIE</td>
<td>Invalid configuration received from Management Server</td>
</tr>
<tr>
<td>NDBD_EXIT_RESOURCE_ALLOC_ERROR</td>
<td>XRE</td>
<td>Resource allocation error, please review the configuration</td>
</tr>
<tr>
<td>NDBD_EXIT_NO_MORE_REDOLOG</td>
<td>XRE</td>
<td>Fatal error due to end of REDO log. Increase NoOfFragmentLogFiles or FragmentLogFileSize</td>
</tr>
<tr>
<td>NDBD_EXIT_OS_SIGNAL_RECEIVED</td>
<td>XIE</td>
<td>Error OS signal received</td>
</tr>
<tr>
<td>NDBD_EXIT_SR_RESTARTCONFLICT</td>
<td>XRE</td>
<td>Partial system restart causing conflicting file systems</td>
</tr>
</tbody>
</table>

### 7.1.1.2 VM Errors

This section contains ndbd error codes that are associated with problems in the VM (virtual machine) NDB kernel block.
Table 7.2 ndbd errors associated with the VM (virtual machine) NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_OUT_OF_LONG_SIGNAL_MEMORY</td>
<td>XCR</td>
<td>Signal lost, out of long signal memory, please increase LongMessageBuffer</td>
</tr>
<tr>
<td>NDBD_EXIT_WATCHDOG_TERMINATE</td>
<td>XIE</td>
<td>WatchDog terminate, internal error or massive overload on the machine running this node</td>
</tr>
<tr>
<td>NDBD_EXIT_SIGNAL_LOST_SEND_BUFFER_FULL</td>
<td>XCR</td>
<td>Signal lost, out of send buffer memory, please increase SendBufferMemory or lower the load</td>
</tr>
<tr>
<td>NDBD_EXIT_SIGNAL_LOST</td>
<td>XIE</td>
<td>Signal lost (unknown reason)</td>
</tr>
<tr>
<td>NDBD_EXIT_ILLEGAL_SIGNAL</td>
<td>XIE</td>
<td>Illegal signal (version mismatch a possibility)</td>
</tr>
</tbody>
</table>

7.1.1.3 NDBCNTR Errors

This section contains ndbd error codes that are associated with problems in the NDBCNTR (initialization and configuration) NDB kernel block.

Table 7.3 ndbd errors associated with the NDBCNTR NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_RESTART_TIMEOUT</td>
<td>XRE</td>
<td>Total restart time too long, consider increasing StartFailureTimeout or investigate error(s) on other node(s)</td>
</tr>
<tr>
<td>NDBD_EXIT_RESTART_DURING_SHUTDOWN</td>
<td>XRE</td>
<td>Node started while node shutdown in progress. Please wait until shutdown complete before starting node</td>
</tr>
</tbody>
</table>

7.1.1.4 DIH Errors

This section contains ndbd error codes that are associated with problems in the DIH (distribution handler) NDB kernel block.

Table 7.4 ndbd errors associated the DIH (distribution handler) NDB kernel block

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_MAX_CRASHED_REPLICAS</td>
<td>XFL</td>
<td>Too many crashed replicas (8 consecutive node restart failures)</td>
</tr>
<tr>
<td>NDBD_EXIT_MASTER_FAILURE_DURING_RESTART</td>
<td>XRE</td>
<td>Unhandled master failure during node restart</td>
</tr>
<tr>
<td>NDBD_EXIT_LOST_NODE_GROUP</td>
<td>XAE</td>
<td>All nodes in a node group are unavailable</td>
</tr>
<tr>
<td>NDBD_EXIT_NO_RESTOREABLE_REPLICA</td>
<td>XFR</td>
<td>Unable to find a restorable replica</td>
</tr>
</tbody>
</table>

7.1.1.5 ACC Errors

This section contains ndbd error codes that are associated with problems in the ACC (access control and lock management) NDB kernel block.
Table 7.5 ndbd errors associated with the ACC (access control and lock management) NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_SR_OUT_OF_INDEXMEMORY</td>
<td>XCR_INDEXMEMORY</td>
<td>Out of index memory during system restart, please increase IndexMemory</td>
</tr>
</tbody>
</table>

7.1.1.6 TUP Errors

This section contains ndbd error codes that are associated with problems in the TUP (tuple management) NDB kernel block.

Table 7.6 ndbd errors associated with the TUP (tuple management) NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_SR_OUT_OF_DATAMEMORY</td>
<td>XCR_DATAMEMORY</td>
<td>Out of data memory during system restart, please increase DataMemory</td>
</tr>
</tbody>
</table>

7.1.1.7 LQH Errors

There is currently one ndbd error code associated with the LQH kernel block. This error code was added in NDB 7.2.6, and is shown in the following table:

Table 7.7 ndbd errors associated with the LQH kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_LCP_SCAN_WATCHDOG_FAIL</td>
<td>XIE_WATCHDOG_FAIL</td>
<td>LCP fragment scan watchdog detected a problem. Please report a bug.</td>
</tr>
</tbody>
</table>

At the lowest level, an LCP comprises a series of fragment scans. Scans are requested by the DBDIH Master using an LCP_FRAG_ORD signal to the DBLQH kernel block. DBLQH then asks the BACKUP block to perform a scan of the fragment, recording the resulting data to disk. This scan is run through the DBLQH block. See also Section 4.7, “The DBLQH Block”.

7.1.1.8 NDBFS Errors

This section contains ndbd error codes that are associated with problems in the NDBFS (filesystem) NDB kernel block.

Most of these errors will provide additional information, such as operating system error codes, when they are generated.

Table 7.8 ndbd errors associated with the NDBFS (filesystem) NDB kernel block.

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_AFS_NOPATH</td>
<td>XIE</td>
<td>No file system path</td>
</tr>
<tr>
<td>2802</td>
<td>XIE</td>
<td>Channel is full</td>
</tr>
<tr>
<td>2803</td>
<td>XIE</td>
<td>No more threads</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_PARAMETER</td>
<td>XIE_PARAMETER</td>
<td>Bad parameter</td>
</tr>
</tbody>
</table>
### ndbd Error Classifications

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDBD_EXIT_AFS_INVALIDPATH</td>
<td>XCE</td>
<td>Illegal file system path</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_MAXOPEN</td>
<td>XCR</td>
<td>Max number of open files exceeded, please increase MaxNoOfOpenFiles</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_ALREADY_OPEN</td>
<td>XIE</td>
<td>File has already been opened</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_ENVIRONMENT</td>
<td>XIE</td>
<td>Environment error using file</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_TEMPNO_ACCESS</td>
<td>XIE</td>
<td>Temporary on access to file</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_FILEFULL</td>
<td>XFF</td>
<td>The file system is full</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_PERMISSION_DENIED</td>
<td>XCE</td>
<td>Received permission denied for file</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_INVALID_PARAM</td>
<td>XCE</td>
<td>Invalid parameter for file</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_UNKNOWN</td>
<td>XIE</td>
<td>Unknown file system error</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_NORESOURCES</td>
<td>XIE</td>
<td>System reports no more file system resources</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_NOSUCH_FILE</td>
<td>XIE</td>
<td>File not found</td>
</tr>
<tr>
<td>NDBD_EXIT_AFS_READUNDERFLOW</td>
<td>XIE</td>
<td>Read underflow</td>
</tr>
<tr>
<td>NDBD_EXIT_INVALIDLCP_FILE</td>
<td>XIE</td>
<td>Invalid LCP</td>
</tr>
<tr>
<td>NDBD_EXIT_INSUFFICIENT_NODES</td>
<td>XRE</td>
<td>Insufficient nodes for system restart</td>
</tr>
<tr>
<td>NDBD_EXIT_UNSUPPORTED_VERSION</td>
<td>XRE</td>
<td>Unsupported version</td>
</tr>
<tr>
<td>NDBD_EXIT_RESTORE_SCHEMA</td>
<td>XCR</td>
<td>Failure to restore schema</td>
</tr>
<tr>
<td>NDBD_EXIT_GRACEFUL_SHUTDOWN_ERROR</td>
<td>XNE</td>
<td>Graceful shutdown not 100% possible due to mixed ndbd versions</td>
</tr>
</tbody>
</table>

#### 7.1.1.9 Sentinel Errors

A special case, to handle unknown or previously unclassified errors. You should always report a bug using [http://bugs.mysql.com/](http://bugs.mysql.com/) if you can repeat a problem giving rise to this error consistently.

**Table 7.9 ndbd sentinel errors**

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>XUE</td>
<td>No message slogan found (please report a bug if you get this error code)</td>
</tr>
</tbody>
</table>

#### 7.1.2 ndbd Error Classifications

This section lists the classifications for the error messages described in Section 7.1.1, “ndbd Error Codes”.

**Table 7.10 ndbd node errors**

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>XNE</td>
<td>Success</td>
<td>No error</td>
</tr>
<tr>
<td>XUE</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>XIE</td>
<td>XST_R</td>
<td>Internal error, programming error or missing error message, please report a bug</td>
</tr>
</tbody>
</table>
## 7.2 NDB Transporter Errors

This section lists error codes, names, and messages that are written to the cluster log in the event of transporter errors.

### Table 7.11 Transporter errors

<table>
<thead>
<tr>
<th>Error Code</th>
<th>Error Classification</th>
<th>Error Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00</td>
<td>TE_NO_ERROR</td>
<td>No error</td>
</tr>
<tr>
<td>0x01</td>
<td>TE_ERROR_CLOSING_SOCKET</td>
<td>Error found during closing of socket</td>
</tr>
<tr>
<td>0x02</td>
<td>TE_ERROR_IN_SELECT_BEFORE_ACCEPT</td>
<td>Error found while accept. The transporter will retry</td>
</tr>
<tr>
<td>0x03</td>
<td>TE_INVALID_MESSAGE_LENGTH</td>
<td>Error found in message (invalid message length)</td>
</tr>
<tr>
<td>0x04</td>
<td>TE_INVALID_CHECKSUM</td>
<td>Error found in message (checksum)</td>
</tr>
<tr>
<td>0x05</td>
<td>TE_COULD_NOT_CREATE_SOCKET</td>
<td>Error found while creating socket(can't create socket)</td>
</tr>
<tr>
<td>0x06</td>
<td>TE_COULD_NOT_BIND_SOCKET</td>
<td>Error found while binding server socket</td>
</tr>
<tr>
<td>0x07</td>
<td>TE_LISTEN_FAILED</td>
<td>Error found while listening to server socket</td>
</tr>
<tr>
<td>0x08</td>
<td>TE_ACCEPT_RETURN_ERROR</td>
<td>Error found during accept(accept return error)</td>
</tr>
<tr>
<td>Error Code</td>
<td>Error Classification</td>
<td>Error Text</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>------------</td>
</tr>
<tr>
<td>0x0b</td>
<td>TE_SHM_DISCONNECT</td>
<td>The remote node has disconnected</td>
</tr>
<tr>
<td>0x0c</td>
<td>TE_SHM_IPC_STAT</td>
<td>Unable to check shm segment</td>
</tr>
<tr>
<td>0x0d</td>
<td>TE_SHM_UNABLE_TO_CREATE_SEGMENT</td>
<td>Unable to create shm segment</td>
</tr>
<tr>
<td>0x0e</td>
<td>TE_SHM_UNABLE_TO_ATTACH_SEGMENT</td>
<td>Unable to attach shm segment</td>
</tr>
<tr>
<td>0x0f</td>
<td>TE_SHM_UNABLE_TO_REMOVE_SEGMENT</td>
<td>Unable to remove shm segment</td>
</tr>
<tr>
<td>0x10</td>
<td>TE_TOO_SMALL_SIGID</td>
<td>ID too small</td>
</tr>
<tr>
<td>0x11</td>
<td>TE_TOO_LARGE_SIGID</td>
<td>ID too large</td>
</tr>
<tr>
<td>0x12</td>
<td>TE_WAIT_STACK_FULL</td>
<td>Wait stack was full</td>
</tr>
<tr>
<td>0x13</td>
<td>TE_RECEIVE_BUFFER_FULL</td>
<td>Receive buffer was full</td>
</tr>
<tr>
<td>0x14</td>
<td>TE_SIGNAL_LOST_SEND_BUFFER_FULL</td>
<td>Send buffer was full, but sleeping for a while solved</td>
</tr>
<tr>
<td>0x15</td>
<td>TE_SIGNAL_LOST</td>
<td>Send failed for unknown reason (signal lost)</td>
</tr>
<tr>
<td>0x16</td>
<td>TE_SEND_BUFFER_FULL</td>
<td>Send buffer was full, but trying to force send fails</td>
</tr>
<tr>
<td>0x0017</td>
<td>TE_SCI_LINK_ERROR</td>
<td>There is no link from this node to the switch</td>
</tr>
<tr>
<td>0x18</td>
<td>TE_SCI_UNABLE_TO_START_SEQUENCE</td>
<td>Could not start a sequence, because system resources are exumed or no sequence has been created</td>
</tr>
<tr>
<td>0x19</td>
<td>TE_SCI_UNABLE_TO_REMOVE_SEQUENCE</td>
<td>Could not remove a sequence</td>
</tr>
<tr>
<td>0x1a</td>
<td>TE_SCI_UNABLE_TO_CREATE_SEQUENCE</td>
<td>Could not create a sequence, because system resources are exumed. Must reboot</td>
</tr>
<tr>
<td>0x1b</td>
<td>TE_SCI_UNRECOVERABLE_DATA_TFX_ERROR</td>
<td>Tried to send data on redundant link but failed</td>
</tr>
<tr>
<td>0x1c</td>
<td>TE_SCI_CANNOT_INIT_LOCALSEGMENT</td>
<td>Cannot initialize local segment</td>
</tr>
<tr>
<td>0x1d</td>
<td>TE_SCI_CANNOT_MAP_REMOTESEGMENT</td>
<td>Cannot map remote segment</td>
</tr>
<tr>
<td>0x1e</td>
<td>TE_SCI_UNABLE_TO_UNMAP_SEGMENT</td>
<td>Cannot free the resources used by this segment (step 1)</td>
</tr>
<tr>
<td>0x1f</td>
<td>TE_SCI_UNABLE_TO_REMOVE_SEGMENT</td>
<td>Cannot free the resources used by this segment (step 2)</td>
</tr>
<tr>
<td>0x20</td>
<td>TE_SCI_UNABLE_TO_DISCONNECT_SEGMENT</td>
<td>Cannot disconnect from a remote segment</td>
</tr>
<tr>
<td>0x21</td>
<td>TE_SHM_IPC_PERMANENT</td>
<td>Shm ipc Permanent error</td>
</tr>
<tr>
<td>0x22</td>
<td>TE_SCI_UNABLE_TO_CLOSE_CHANNEL</td>
<td>Cannot close the sci channel and the resources allocated</td>
</tr>
</tbody>
</table>
Appendix A NDB Internals Glossary

This appendix contains terms and abbreviations that are found in or useful to understanding the NDB source code.

**ACC.** ACCELERATOR or ACCess manager. Handles hash indexes of primary keys, providing fast access to records. See Section 4.3, “The DBACC Block”.

**API node.** In NDB terms, this is any application that accesses cluster data using the NDB API, including mysqld when functioning as an API node. (MySQL servers acting in this capacity are also referred to as “SQL nodes”.) Sometimes abbreviated informally as “API”. See NDB Cluster Nodes, Node Groups, Replicas, and Partitions.

**BACKUP.** In the NDB kernel, the block having this name performs online backups and checkpoints. For more information, see Section 4.1, “The BACKUP Block”.

**CMVMI.** Stands for Cluster Manager Virtual Machine Interface. An NDB kernel handling nonsignal requests to the operating system, as well as configuration management, interaction with the cluster management server, and interaction between various kernel blocks and the NDB virtual machine. See Section 4.2, “The CMVMI Block”, for more information.

**CNTR.** Stands for restart CoordInaToR. See Section 4.14, “The NDBCNTR Block”, for more information.

**DBINFO.** The Database Information block provides support for the ndbinfo information database used to obtain information about data node internals. See Section 4.6, “The DBINFO Block”.

**DBTC.** The transaction coordinator (also sometimes written simply as TC). See Section 4.9, “The DBTC Block”, for more information.

**DICT.** The NDB data DICTIONary kernel block. Also DBDICT. See Section 4.4, “The DBDICT Block”.

**DIH.** DIstribution Handler. An NDB kernel block. See Section 4.5, “The DBDIH Block”.

**LDM.** Local Data Manager. This set of NDB kernel blocks executes the code that manages the data handled on a given data node. It includes the DBTUP, DBACC, DBLQH, DBTUX, BACKUP, TSMAN, LGMAN, PGMAN, and RESTORE blocks.

Each such set of modules is referred to as an LDM instance, and is responsible for tuple storage, hash and T-tree indexes, page buffer and tablespace management, writing and restoring local checkpoints, and Disk Data log management. A data node can have multiple LDM instances, each of which can be distributed among a set of threads. Each LDM instance works with its own partition of the data.

**LGMAN.** The Log Group MANager NDB kernel block, used for NDB Cluster Disk Data tables. See Section 4.13, “The LGMAN Block”.

**LQH.** Local Query Handler. NDB kernel block, discussed in Section 4.7, “The DBLQH Block”.

**MGM.** ManaGeMent node (or management server). Implemented as the ndb_mgmd server daemon. Responsible for passing cluster configuration information to data nodes and performing functions such as starting and stopping nodes. Accessed by the user by means of the cluster management client (ndb_mgm). A discussion of management nodes can be found in ndb_mgmd — The NDB Cluster Management Server Daemon.

**NDB_STTOR.** NDB STarT Or Restart
QMGR. The cluster management block in the NDB kernel. It responsibilities include monitoring heartbeats from data and API nodes. See Section 4.17, “The QMGR Block”, for more information.

RBR. Row-Based Replication. NDB Cluster Replication is row-based replication. See NDB Cluster Replication.

STTOR. STarT Or Restart

SUMA. The cluster SUbscription MAner. See Section 4.19, “The SUMA Block”.

TC. Transaction Coordinator. See Section 4.9, “The DBTC Block”.

TRIX. Stands for TRansactions and IndeXes, which are managed by the NDB kernel block having this name. See Section 4.23, “The TRIX Block”.

TSMAN. Table space manager. Handles tablespaces for NDB Cluster Disk Data. See Section 4.22, “The TSMAN Block”, for more information.

TUP. TUPle. Unit of data storage. Also used (along with DBTUP) to refer to the NDB kernel’s tuple management block, which is discussed in Section 4.10, “The DBTUP Block”.

Index

D
DUMP commands
   NDB Cluster, 7

E
error messages
   NDB API, 115
errors
   MGM API, 115
   NDB API, 115

M
MGM API
   errors, 115

N
NDB API
   error messages, 115
   errors, 115
NDB Cluster
   DUMP commands, 7
ndb_mgm
   DUMP commands, 7