
Run harvest from the command line in Digitool

Overview

Harvest can be run from the command line under the following scenarios:

1. To force the harvest of new items so that they will be immediately available in resource discovery,
2. To Force harvest between defined dates and times.
3. To force a from-scratch re-build of the resource discovery indexes.
(Note that this deletes the existing resource discovery indexes and rebuilds them from scratch, during which time resource discovery shall be unavailable to the end users)

Harvest commands

The following commands should be run in the terminal of your Digitool server as dtl user:

Immediate harvest of new items

This will result in the harvest of all new or changed items since the last harvest. the time of the previous harvest can be determined using the util:

Harvest command:

```
>csh -f $dtl_proc/p_harvest_02 GEN01,A,,,,,2 > & $dtle_scratch/exl_harvest_2.log &
```

Progress can be viewed in the log \$dtle_scratch/exl_harvest_from_scratch_1.log

Harvest all objects modified between specified dates & times

The command format is thus:

```
>csh -f $dtl_proc/p_harvest_02 GEN01,A,YYYYMMDD,YYYYMMDD,HHMM,HHMM,2 > & $dtle_scratch/exl_harvest_from_scratch_1.log &
```

The order of dates & times is Start Date, End Date, Start Time, End time

In this example we shall harvest all items with a modification date between a minute past midnight on 14th June 2016 & 10:00am on 20th June 2016:

```
>csh -f $dtl_proc/p_harvest_02 GEN01,A,20160614,20160620,0001,1000,2 > &
$dtle_scratch/exl_harvest_from_scratch_1.log &
```

Progress can be viewed in the log \$dtle_scratch/exl_harvest_from_scratch_1.log

Harvest from scratch

WARNING: This deletes the existing resource discovery indexes and rebuilds them from scratch, during which time resource discovery shall be unavailable to the end users. This could potentially take several hours depending on the size of your repository and whether or not you index full text.

```
>csh -f $dtl_proc/p_harvest_02 GEN01,C,,,,,2 > & $dtle_scratch/exl_harvest_from_scratch_1.log &
```

Progress can be viewed in the log \$dtle_scratch/exl_harvest_from_scratch_1.log

Appendix:

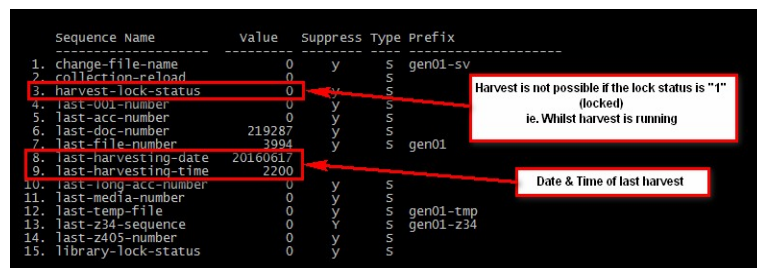
Pre-Check before harvesting

```
>dlib gen01
```

```
>util
```

From the Menu select G > 2

Example output of util > G > 2:



Sequence Name	value	Suppress	Type	Prefix
1. change-file-name	0	y	s	gen01-sv
2. collection-reload	0	y	s	s
3. harvest-lock-status	0	y	s	s
4. last-001-number	0	y	s	s
5. last-acc-number	0	y	s	s
6. last-doc-number	219287	y	s	s
7. last-file-number	3994	y	s	gen01
8. last-harvesting-date	20160617	y	s	s
9. last-harvesting-time	2200	y	s	s
10. last-long-acc-number	0	y	s	s
11. last-media-number	0	y	s	s
12. last-temp-file	0	y	s	gen01-tmp
13. last-z34-sequence	0	y	s	gen01-z34
14. last-z405-number	0	y	s	s
15. library-lock-status	0	y	s	s

If the harvest lock status is 1 "Locked" then it will not be possible to run a new harvest. When the status is locked either a harvest is currently running, or a previous process terminated unexpectedly. If you try to run a harvest with a lock in place

you will see entries like these in the harvest log:

```
Library is locked for harvesting. I'll sleep 10 seconds and try again.
```

See this article about how to clear this situation:

[Harvest hangs](#)

Notes

Please be aware that there may be large time intervals where nothing is written to the harvest logs when harvesting a large set of data.

You can monitor active progress by updates to the files in this path which will be constantly updated during the harvest process.

```
>dlib gen01
```

```
>ds (Switches to temporary harvest file directory)
```

```
>ls -lrth | tail -20
```