

Maintenance and monitoring during large-scale data changes and updates

- **Product:** Voyager, Alma
 - **Relevant for Installation Type:** Multi-Tenant Direct, Dedicated-Direct, Local, TotalCare
-

Question

- We are preparing our Voyager data for Alma and will be making large-scale updates to our records.
- We are making a large amount of updates to our bibs before sending them to a vendor for processing.
- We are planning an authority control work project.

What maintenance and monitoring should we do in Voyager while we're making massive changes to our data?

Answer

There are a few considerations when doing heavy data change work on your Voyager database.

Consider opening a Case with Support to let us know what you are doing so that Support can assist in keeping your server working properly.

Before embarking on such a project, check your available disk space (the `df -h` command will work) and cleanup any unnecessary files.

Oracle archive logs

Considerations

Every update to the database writes to an Oracle log. When making large quantities of changes, these logs can become quite large and fill up disk space.

Best practices

- Before engaging in a major data change/update process, test a subset of records. For example,
 1. Identify a set of records that represents 5-10% of the total number of records to be updated.
 2. Check the current available disk space. [Follow these instructions for clearing disk space.](#)
 3. Process the record change as you intend to for the full set.
 4. Check the disk space available when the update/change is complete. This will give a sense of how much space is required for the full set/for every change of the size tested.
- Some sites with a large number of bibs or a large number of Voyager xxxdb instances may need Oracle tuning. Contact Support if your institution falls into this category.

Keyword indexing

Considerations

Every update to a bibliographic or holdings record writes to the dynamic.dc file in /m1/voyager/xxxxdb/data (for bibs) or /m1/voyager/xxxxdb/mfhd.data (for holdings). The more updates are made, the larger the files grow. There are two impacts of this behavior:

- As the file grows, the speed of keyword indexing slows, making jobs/processes that update the file take longer and longer.
- If left unmonitored, the file can reach its max size of 2.0 GB, and can no longer be written to at that point, meaning records can no longer be saved/updated until a regen is run.

Best practices

- Disable keyword indexing when this option is available, i.e., for Bulk Import and Global Data Change jobs. Records will be keyword indexed when a regen is run later. See: [Does -KADDKEY only apply to bib records?](#) and [Should I keyword index when I run a data change job?](#)
- Check keyword file sizes daily during data work, multiple times daily during intense periods of activity. See: [How do I determine if keyword regen is needed?](#)
- Plan daily or weekly regens to stay ahead of keyword indexing issues. See: [UTIL Menu: Running a keyword regen](#)
- After authority control work or other massive data change projects, you may require a *Full Regen* of all your indexes (not just keyword). Contact Support to schedule this.

-
- **Article last edited:** 19-Sep-2020