

Subject: Sawmill Newsletter: Deleting Database Fields Using a Command Line Action
From: Greg Ferrar <greg@flowerfire.com>
Date: 2/21/12 11:02 AM
To: ferrar@flowerfire.com



Welcome to the Sawmill Newsletter!

You're receiving this newsletter because during the downloading or purchase of Sawmill, you checked the box to join our mailing list. If you wish to be removed from this list, please send an email, with the subject line of "UNSUBSCRIBE" to newsletter@sawmill.net (please include the entire message, as the identifying information is at the bottom).

News

Sawmill 8.5.5 is now available. This is the latest minor upgrade to Sawmill, improving reporting performance, adding some minor functionality, and fixing a few bugs. It is a free upgrade for Sawmill 8.5 users, and also a free upgrade for Sawmill 8.0/8.1 users (though your profiles and databases will need to be converted). You can download Sawmill 8.5.5 from <http://sawmill.net/download.html> . Please send your feedback about your experiences with Sawmill.

Sawmill 7 users can upgrade to Sawmill 8 for half of the license price; or if you have Premium Support, the upgrade is free. Major features of Sawmill 8 include support for Oracle and Microsoft SQL Server databases, a completely redesigned web interface, better multi-processor and multi-core support, and role-based authentication control.

This issue of the Sawmill Newsletter describes how to delete unneeded database fields, with all dependencies, to simplify the database and improve performance.

Get The Most Out Of Sawmill With Professional Services

Looking to get more out of your statistics from Sawmill? Running short on time, but need the information now to make critical business decisions? Our Professional Service Experts are available for just this situation and many others. We will assist in the initial installation of Sawmill using best practices; work with you to integrate and configure Sawmill to generate reports in the shortest possible time. We will tailor Sawmill to your environment, create a customized solution, be sensitive to your requirements and stay

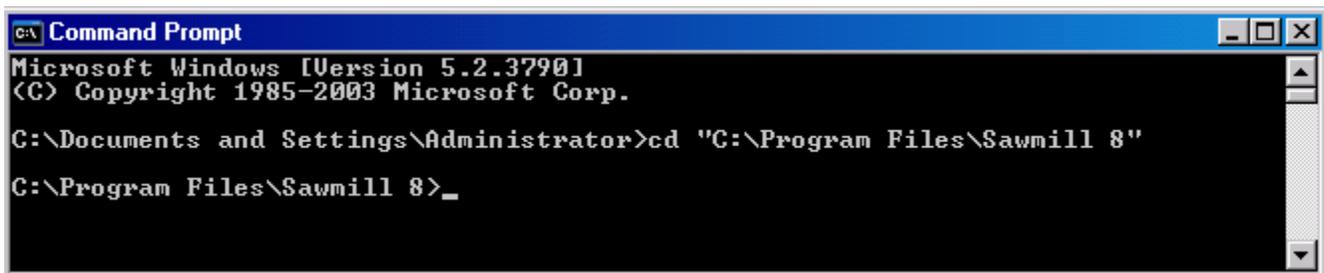
focused on what your business needs are. We will show you areas of Sawmill you may not even be aware of, demonstrating these methods will provide you with many streamlined methods to get you the information more quickly. Often you'll find that Sawmill's deep analysis can even provide you with information you've been after but never knew how to reach, or possibly never realized was readily available in reports. Sawmill is an extremely powerful tool for your business, and most users only exercise a fraction of this power. That's where our experts really can make the difference. Our Sawmill experts have many years of experience with Sawmill and with a large cross section of devices and business sectors. Our promise is to very quickly come up with a cost effective solution that fits your business, and greatly expand your ROI with only a few hours of fee based Sawmill Professional Services. For more information, a quote, or to speak directly with a Professional services expert contact consulting@flowerfire.com.

Tips & Techniques: Deleting Database Fields Using a Command Line Action

When Sawmill processes your log data, it creates a database of statistics which is queried when you click on reports in the user interface. This database persists and will grow larger if you do subsequent updates to add new log data. One way to manage the disk space used by a Sawmill database is to delete all unneeded database fields. These are the fields associated with reports that you never use. The simplest way to do this is to use the command line action `delete_database_field`, or `ddf` for short. When deleting a field with this action, not only will the data be removed from the main table, but the associated indices and cross-reference tables which are created to speed up reporting are also deleted, freeing up more disk space. In addition, all reports and report fields depending on the field are deleted.

The Command Line

This action will be performed from a command line. If Sawmill is installed on Windows, a command prompt window can be opened by using Start -> Run, and entering `cmd` in the dialog box. On Linux/Unix/Mac you can open a terminal window. In the following example, we will delete some fields from an Apache NCSA profile on Windows. Once a Command Prompt window is open, you will move to the Sawmill 8 directory. In this example, Sawmill is installed in the default location, `C:\Program Files`.



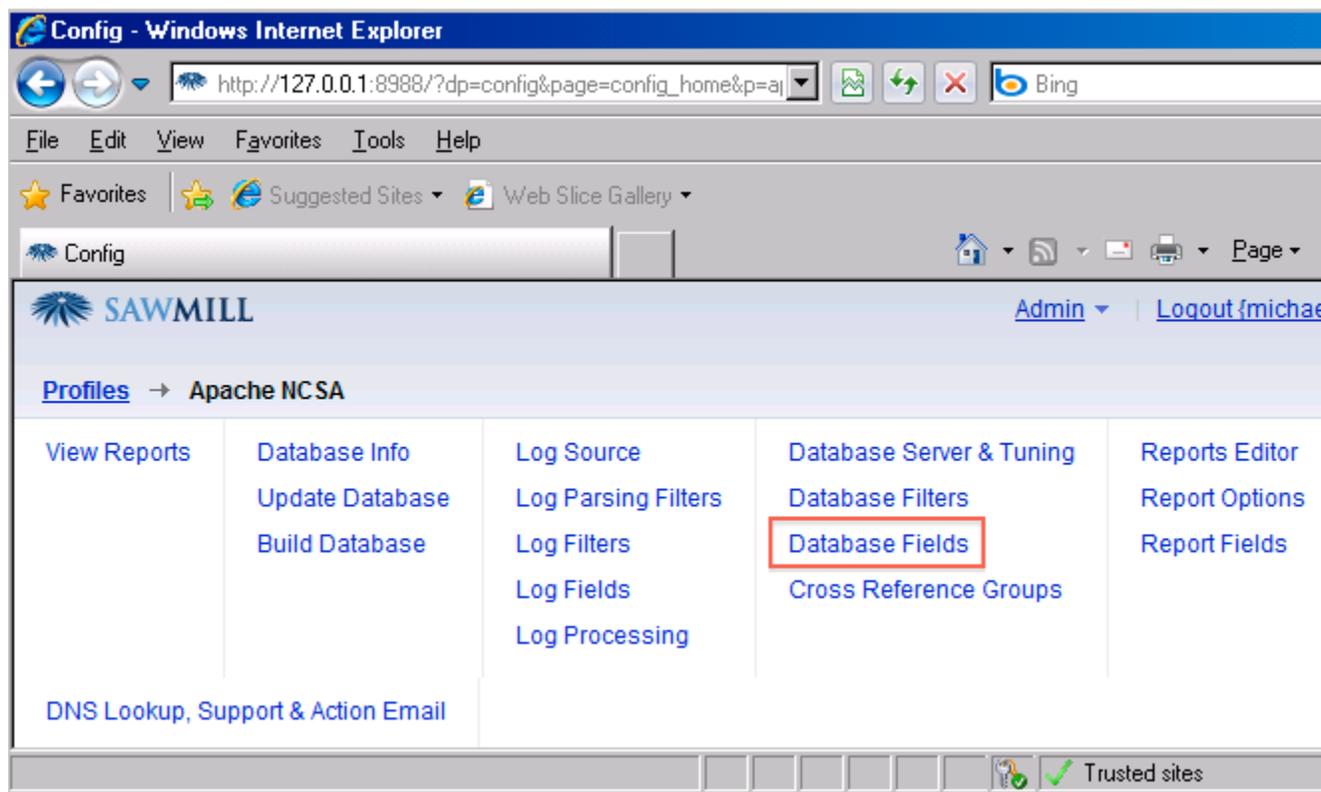
```
Microsoft Windows [Version 5.2.3790]
(C) Copyright 1985-2003 Microsoft Corp.

C:\Documents and Settings\Administrator>cd "C:\Program Files\Sawmill 8"
C:\Program Files\Sawmill 8>_
```

Now we are ready to enter a command. Deleting a database field requires three parameters to be specified: the profile to be used, the action, and the field name. In this example, the profile is named "Apache NCSA". The action is "delete_database_field", and the first field to be deleted is "Domain description".

Command line syntax and field identifiers

Command line syntax requires all values to be lowercase alphanumeric characters with spaces and other characters changed to underscores. So when specifying the profile name "Apache NCSA" on a command line, the name changes to apache_ncsa. Likewise, each database field name has a special identifier that is used on the command line. These identifiers can be seen by going to the configuration options page of the profile and clicking "Database Fields".



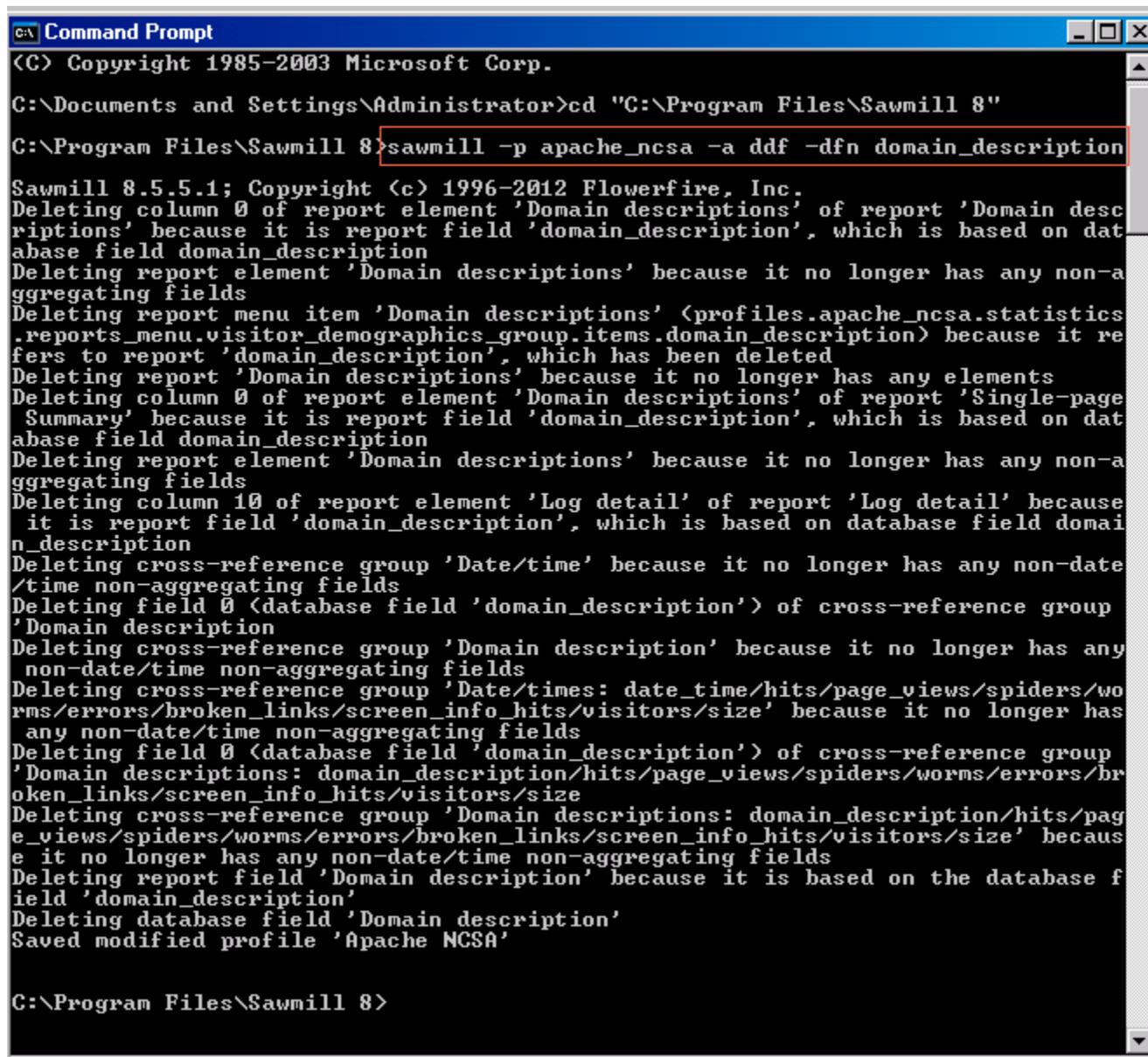
Now select the field and click "Show Identifier".

The screenshot shows the Sawmill configuration interface in a Windows Internet Explorer browser. The browser address bar displays the URL `http://127.0.0.1:8988/?dp=config&page=database_fields&p`. The page title is "Config - Windows Internet Explorer". The interface includes a navigation menu with "Profiles" and "Apache NCSA Database Fields". A toolbar contains "Save Changes", "New Database Field", "Duplicate", "Delete", and "Undo All Changes". A list of "Database Fields" is shown on the left, with "Domain description" selected. The right panel displays configuration options for "Domain description", including "Field name", "Derivation method", "Log field", "Type", "Aggregation method", "Category", "Field length", "Integer bits", "Index", "Suppress levels", and "Itemnums hash function".

Now the identifier syntax that is to be used on the command line is displayed.

A list of all database field identifiers can also be displayed by running "sawmill -p *profile* -a ldf" from the command line.

We are now ready to enter the command.



```

C:\> Copyright 1985-2003 Microsoft Corp.

C:\Documents and Settings\Administrator>cd "C:\Program Files\Sawmill 8"

C:\Program Files\Sawmill 8>sawmill -p apache_ncsa -a ddf -dfn domain_description

Sawmill 8.5.5.1; Copyright (c) 1996-2012 Flowerfire, Inc.
Deleting column 0 of report element 'Domain descriptions' of report 'Domain descriptions' because it is report field 'domain_description', which is based on database field domain_description
Deleting report element 'Domain descriptions' because it no longer has any non-aggregating fields
Deleting report menu item 'Domain descriptions' (profiles.apache_ncsa.statistics.reports_menu.visitor_demographics_group.items.domain_description) because it refers to report 'domain_description', which has been deleted
Deleting report 'Domain descriptions' because it no longer has any elements
Deleting column 0 of report element 'Domain descriptions' of report 'Single-page Summary' because it is report field 'domain_description', which is based on database field domain_description
Deleting report element 'Domain descriptions' because it no longer has any non-aggregating fields
Deleting column 10 of report element 'Log detail' of report 'Log detail' because it is report field 'domain_description', which is based on database field domain_description
Deleting cross-reference group 'Date/time' because it no longer has any non-date/time non-aggregating fields
Deleting field 0 (database field 'domain_description') of cross-reference group 'Domain description'
Deleting cross-reference group 'Domain description' because it no longer has any non-date/time non-aggregating fields
Deleting cross-reference group 'Date/times: date_time/hits/page_views/spiders/worms/errors/broken_links/screen_info_hits/visitors/size' because it no longer has any non-date/time non-aggregating fields
Deleting field 0 (database field 'domain_description') of cross-reference group 'Domain descriptions: domain_description/hits/page_views/spiders/worms/errors/broken_links/screen_info_hits/visitors/size'
Deleting cross-reference group 'Domain descriptions: domain_description/hits/page_views/spiders/worms/errors/broken_links/screen_info_hits/visitors/size' because it no longer has any non-date/time non-aggregating fields
Deleting report field 'Domain description' because it is based on the database field 'domain_description'
Deleting database field 'Domain description'
Saved modified profile 'Apache NCSA'

C:\Program Files\Sawmill 8>

```

This command starts a Sawmill process (on Linux/Unix/Mac, the command would typically start with `./sawmill`), defines the profile with `-p` (`apache_ncsa` in this example), defines the action (`delete_database_field`, or `ddf`) with `-a`, and defines the database field name with `-dfn` (`domain_description`, in this example). The text that follows the executed command shows Sawmill removing the associated report columns, report fields, cross-reference group, etc. If you want to remove more fields, you can simply hit the up arrow on your keyboard which will reproduce the last command. Then you only need to back delete the field name following `-dfn`, type another field name and hit enter. Repeat for each field you want to remove.

It's important to know that this command line action only changes the profile configuration, and does not affect a database that might have already been built before deleting the fields. So in order to actually reduce the size of an existing database, you will need to rebuild it after deleting the fields. During a rebuild, and subsequent updates, only the remaining fields in the profile will be created in the database. If you're creating a new profile, the deletions should be done before the initial build.

Get The Most Out Of Sawmill With Professional Services

This newsletter describes the use of the `delete_database_field` action to simplify the database. This is only one of a wide range of advance database tuning available, including simplifying field values, tuning cross-references and indices, adding log filters to simplify individual field values or discard rows, and more. Our Sawmill Experts can help you get the most out of Sawmill, by tuning your database for maximum performance and minimum size. Contact consulting@sawmill.net for more information.