
HighWaterMark DBPUT flag

Ken Paul

World-wide Support Manager

Adager Corporation

Sun Valley, Idaho 83353-3000 U.S.A.

Tel. (208) 726-9100 Fax (208) 726-8191

ken@adager.com http://www.adager.com

This article will deal with a new flag known as the HighWaterMark DBPUT or HWMPUT flag for short. This is a database level flag which only applies to Detail datasets and is disabled by default. The SHOW <dbname> FLAGS command within DBUTIL will show you the current setting of this flag. The ENABLE/DISABLE <dbname> HWMPUT commands will allow you to turn it on or off.

DBPUT on detail datasets

Before I talk about what effect this flag will have on your DBPUTs, I want to first review the structure of a Detail dataset and how DBPUT currently works.

Every Detail dataset has a UserLabel associated with it which contains three pieces of information: the HighWaterMark, the FreeEntry Count and the Head of the Delete Chain. When a Detail dataset is first built (let's say with a capacity of 100) the HighWaterMark is initialized to zero the FreeEntry Count is set to 100 and the Head of the Delete Chain is also initialized to zero.

When a DBPUT is issued for a Detail dataset, IMAGE first looks at the FreeEntry Count to make sure it is not zero. IMAGE then looks to see if the Head of the Delete Chain is non-zero. If the Head of the Delete Chain is non-zero, IMAGE places the new entry at that location and decrements the FreeEntry Count by one. If that location was pointing to the next deleted entry then that address is made the Head of the Delete Chain. If the Head of the Delete Chain is zero, IMAGE adds one to the HighWaterMark and places the new entry at that location and decrements the FreeEntry Count by one.

When a DBDELETE is issued for a Detail dataset, IMAGE zeroes out the deleted entry and places the Head of the Delete Chain in the first two words of the deleted entry and places the address of the newly deleted entry as the Head of the Delete Chain. The FreeEntry Count is then incremented. The Delete Chain can then be thought of as a singly-linked list where the entry which was most recently deleted will be where the next new entry will be added or First Out First In (FOFI).

Motivation for the HWMPUT flag

Some IMAGE users over the years were not happy that IMAGE checked for existence of a Delete Chain before going to the HighWaterMark. They felt that IMAGE should add new entries to a Detail dataset at the HighWaterMark first and only when the HighWaterMark reached the capacity should DBPUT then start using the Delete Chain.

This is what the HWMPUT flag does. It changes the precedence which IMAGE uses when adding a new entry to a Detail dataset. When the HWMPUT flag is disabled (default), IMAGE will first use the Delete Chain and only after exhausting it will IMAGE use the HighWaterMark. When the HWMPUT flag is enabled, IMAGE will first use the HighWaterMark until it reaches the capacity and then use the Delete Chain.

The inside story

There are a few things that I want to stress about the HWM-PUT flag. First, it is a DATABASE flag and not a DATASET flag. You enable the WHOLE database for this functionality, not a given Detail dataset. Second, if entries are never deleted from a Detail dataset then enabling this option will have no effect because the dataset will never have a Delete Chain and IMAGE will always use the HighWaterMark.

The question that some of you are probably asking is, "Why would this feature make any difference with my database?" The best example that I can think of is a LINE-ITEM Detail dataset where there may be an average of 20 line-items for every order. If this dataset has a lot of activity with single line-items being deleted and whole new orders being added then you may want to enable this flag. If a lot of single line-items are being deleted throughout the day then you may have a very long Delete Chain with deleted entries located all over the dataset. When you add a new order with 20 line-items, these 20 entries may end up scattered all over the dataset. With the HWMPUT flag enabled you have a better chance that all 20 line-items stay physically next to each other so retrieval of this order will be

faster. In other words, the HWMPUT flag can increase data locality within the dataset.

I highly recommend though that the HWMPUT flag be used in conjunction with scheduled repacks of the detail datasets using Adager's Repack Dataset function (DetPack). This will eliminate all of the holes caused by the deleted entries and ensure that the active entries are all packed at the beginning which will cause serial reads to take the least amount of time.

Don't forget to do periodic repacking of detail datasets

Another time to use the HWMPUT flag is when you have to do a large amount of archiving off of a Detail dataset and you can not follow it up with a Repack for another week or another month. This might be a good time to enable the HWMPUT flag so that your new entries are not being scattered all over the Detail dataset.

Archives

The HWMPUT flag is not for everyone and should not be thought of as some magical panacea. It should be used with careful thought and knowledge of how your particular database works. Fred White of Adager thinks that the HWMPUT flag is a Pandora's Box waiting to be abused. We would like to know what you think about this new feature. If you have any examples (pro or con) for using HWMPUT we would love to hear them.

Pandora's box?