Article

[Jeff Morgan](#) · Oct 17, 2016 2m read

Why Keep Integrity Is Important When Purging HealthShare/Ensemble Data

When running the built in Ensemble Purge task (Ens.Util.Tasks.Purge) there are three parameters DaysToKeep, BodiesToo, and KeepIntegrity. This article focuses on the KeepIntegrity boolean parameter, but more information about running this task can be found here:

http://docs.intersystems.com/latest/csp/docbook/DocBook.UI.Page.cls?KEY=...

KeepIntegrity determines which Ensemble Messages are marked for deletion within the time specified by DaysToKeep.

With KeepIntegrity marked as **true**, the Purge will only mark "completed" messages for deletion. A completed message has a status of Complete, Error, Aborted, or Discarded. Any incomplete messages will not be marked for deletion.

When setting KeepIntegrity to **true** we must be aware of the consequences. Specifically, the purge will take longer. This is due to the overhead of checking each messages status value. Speeding up the purge has been a known reason for people to mark KeepIntegrity false.

Additionally, marking KeepIntegrity **false** can also have suboptimal effects. The most important consequence is that in flight messages will be deleted. In current released versions of HealthShare/Ensemble, this includes systems processes that rely on messages, such as the Scheduler. This will cause the Scheduler to fail and require a restart of the Production to resolve.

In light of these possible downsides, a strategy to manage the deletion of data is to utilize a double Purge Task configuration.

The first Purge Task has a low number for DaysToKeep (such as 7) while KeepIntegrity is **true**. This will keep all messages that are incomplete while still deleting the majority of messages so that the database remains at a manageable size.

While the other task takes care of most of the message, the second Purge task will have KeepIntegrity **false** and a large DaysToKeep (maybe 90). This will purge all messages older than the days to keep. Frequently, messages older than this value are irrelevant as they have either been resent, or handled outside of HealthShare/Ensemble.

As for most settings in HealthShare/Ensemble, the suggested value for a specific setting depends on your utilization of the software.

#Ensemble #System Administration