GWME-7.1.0-1 - Removing duplicate host entries in gwcollagedb database prior to upgrade from 7.0.2.2

Issue

Previous versions of GroundWork Monitor allowed the addition of hosts in the gwcollage database which differed only in case but not in spelling. For example it was permitted to have both localhost and LOCALHOST as separate identities. In 7.0.2 SP03 and 7.1 this is not valid. These duplicate hostnames must be deleted before you can succeed in upgrading to a GroundWork version > 7.0.2 SP2. Failing to do the steps here will produce a broken system.

Name	Size	Creator	Creation Date	Comment
de-dup-scripts.tar.gz	10 kB	NotSupportContact-Mark Carey	Dec 06, 2016 21:12	MD5: 0c50e2291a3c06a9ba7d34fcb1c7bb6b

Classes of Duplicate Host

Duplicate Hosts Which Do Not Have the Same Services

Example host "LOCALHOST" with service "icmp_ping" and host "localhost" with service "http_alive":
The 7.1 installer binary and 7.0.2 SP03 install script each include "uniqueify-host-hostnames.sql" (see GWMON-12039). This script can successfully merge this class of duplicate host with no user action.

Duplicate Hosts Which Each Have One or More of the Same Services

Example host "LOCALHOST" with service "icmp_ping" and host "localhost" with service "icmp_ping"
This occurs when a hostname is renamed with different case and the services remain unchanged. In this use case, "uniqueify-host-hostnames.sql" will not merge the duplicate hosts, leaving them in the database. Subsequent steps in the upgrade will then fail because the table includes violation of the constraint that requires case insensitive uniqueness.

What Results

Where the duplicate hostnames have different services we just detach the service information from the older host and reattach those services and associated records to the newer host; then we delete the older host. This way you lose no state or detail message.

If the same service names are seen on both hosts the problem is not that easy to fix since the service name is also duplicated, and the association is already present on the newer host. We have to delete all the service records from the older host, as there is no way to merge them, one must go.

In both cases, the resolution will delete some records but updating status, events, and performance metrics from the various feeders and agents will resume to the surviving name and its associated services.

Action to Take

To resolve, first identify whether you have either duplicate host scenario using the de-duper_indentify.sql script:

```
source /usr/local/groundwork/scripts/setenv.sh
psql -d gwcollagedb -f show-similar-hosts.sql
```

If the output of this shows no names, return to the upgrade process. But if the result shows a list of duplicate host names you MUST do the next step. Remove the duplicates by running the de-duper_merge.sql script:

```
source /usr/local/groundwork/scripts/setenv.sh
psql -d gwcollagedb -f merge-similar-hosts.sql
```

Verify that a successful deduplication was accomplished by re-running the show-similar-hosts.sql script. It should return an empty list indicating no duplicate names. You may then return to the upgrade process.