Mardy Matt: Beyond The DOME

Matt, for the Technical Meeting.
26/7/19

DOME - what's the difference?

- Essentially turns DPM into an xroot plugin - so all the DPM communication is via xroot. If your xroot service on your headnode isn't running, DOME (i.e. your DPM) isn't running.
- A lot of the old, much loved admin tools don't really work any more (or if they do it's an accident of still running in "legacy mode").
- Instead you have the dmlite-shell.
- Your DPM can now do clever things, with more support to do Third-Party Copies and Checksums with all your protocols.
- But most of all, your DPM will be only be supported beyond September this year.

Upgrading to DOME now - the Pros

- DOME is now on version 1.13 - this has fixed most of the bugs (many of them performance related).
- Greater support for TPC and Checksumming allows greater flexibility in choosing what protocols to use - important with SRM on the way out.
- (Anecdotally) DOME does many things faster and smoother then the old DPM - things like draining and metadata operations.
- Quotatokens are inherently a better way of going about enforcing user limits then Space Tokens ever were.

The Cons

- DOME is a very different beast to admin - a lot of the stuff you know about running a DPM will be out the window (especially once we have to switch "Legacy Mode" off).
- Related to this, the combined DOME knowledge base - whilst growing all the time - is still relatively small compared to veteran status of the DPM Old Guard.
- AIUI There is no way to move to DOME without a downtime (due to DB operations).
    - Although this might not be so bad if you want to move your headnode to C7 at the same time.

Upgrading

With reference to: https://twiki.cern.ch/twiki/bin/view/DPM/DpmSetup1100

- Like any other big update - have a cuppa and make a plan, set aside some time for it and set a downtime (I'd go for for a day - you can always end them early).
- DOME allows you to set things up on the headnode behind the scenes before making the switch - referred to as running dormently. Until you switch to using the dome_adapter.conf your DPM just runs as normal. You can cut down your downtime by really leaning into this.
- The action that requires a downtime is the script that aligns your space tokens and directory sizes to your quotatokens. Your DPM has to be off for this, and it can take a few hours.

Upgrade Tips

- Life is a lot easier if you use the DPM puppet modules (we run them locally at Lancaster, essentially as a YAIM replacement).
  - I ran things a mix of puppet and manual updates here at Lancaster - if I had my time again it would be puppet all the way.
  - The DPM docs do assume you have more then a novice level of puppet experience though.
- I would use this opportunity to enable gridftp redirection and xroot checksumming/delegation at the same time:

  configure_dpm_xrootd_checksum =>   true,
  configure_dpm_xrootd_delegation =>   true,
  gridftp_redirect => true,

Token Effort

IMO Quotatokens >> Spacetokens.

- But they are a very different way of thinking about things.
- Spacetokens are weird, abstract buckets a user (with the right permissions) can "assign" their files to.
- Quotatokens are aligned to directories in the DPM namespace - if you write into a directory you "fill up" the appropriate quotatoken (with who can fill a quotatoken decided by the ACLs on the directory.
  - This means that every writable directory in DPM must have a quotatoken (with sufficent space) assigned to it or users will not be able to write there.
  - Typically this means giving each VO directory its own VODISK quotatoken.

Token Effort 2

- During the move to DOME there are tools to align Space and Quota Tokens (it's a big step in the upgrade process).
- During this process you essentially assign namespace paths to existing space tokens during the conversion process.
  - Luckily atlas made this easy as they have a separate directories for each space token.
    - A top tip though - the ATLASDATADISK token should set to the "root" atlas directory: /dpm/domain.ac.uk/home/atlas
  - All other VOs tend to have one space token, so during the conversion this should be set to use the appropriate VO root directory.
  - And to reiterate - VOs without a spacetoken need to have a quotatoken assigned to their area in the namespace - ideally one with a bit more space then they're currently using (unless you don't want them to write until they clean things up!).

Token Effort 3

- Quotatokens can be "nested" - atlas are the best example for this.

  .../home/atlas   ← ATLASDATADISK

  .../home/atlas/atlaslocalgroupdisk ← ATLASLOCALGROUPDISK

  .../home/atlas/atlasscratchdisk ← ATLASSCRATCHDISK

- If I was to write into atlas/ or atlas/atlasdatadisk/ I would be writing to the ATLASDATADISK token, but writing into the other two directories would write to their assigned quotatokens instead.

Dmlite in the Shell

The main way of interacting with your DPM post-DOMEing is the dmlite-shell

- A bit of a misnomer as it's not got much to do with dmlite, and it's not much of a shell.
- Invoke with `dmlite-shell` to enter it. Typing help gets your a list of commands, help <command name> to get a bit of syntax detail. Ctrl-D to get out.
  - Or you can invoke it a single command mode - dmlite-shell -e 'command', or a simple script mode, dmlite-shell -s script.sh, where it runs each line in the script file in a non-intelligent fashion.
- This replaces your dpns-* commands, your do-something-to-an-fs commands, your drains and your getting file metadata.
- It's also how you create/modify quotatokens.

Some useful commands (and missing ones).

- >info <path/to/file> gets you the file metadata - including checksum information and location on disk
- >qryconf is the new dpm-qryconf - the output is much the same as before.
- >drainfs/drainserver are the drain commands - note that you have to explicitly tell it to not do a dry run.

And what's missing?

- No simple dpm-dpns-to-disk (info gives gives a lot of noise)
- No dpm-disk-to-dpns (don't lose a disk server!).

Useful files to know about.

- /etc/domehead.conf (/etc/domedisk.conf on disk nodes)
- /etc/dmlite.conf.d/domeadapter.conf (should be the only non-empty file in that directory).
- /etc/gridftp.conf
- /etc/xrootd/xrootd-dpmredir.cfg (or /etc/xrootd/xrootd-dpmdisk.cfg on disk nodes).
- /etc/xrootd/dpmxrd-sharedkey.dat
- /etc/httpd/conf.d/zlcgdm-dav.conf

Other features and gotchas.

- I advise switching gridftp redirection on when enabling DOME.
  - Needed for when we switch off SRM.
  - This can lead to some odd behaviour, as metadata requests are redirected to disk servers at random - all nodes need to know about all possible VOs in their vomsdirs.
- Likewise I would advise enabling at xrootd checksumming.
  - Much easier to do in puppet, as the configs are fiddly.
  - Allows jobs to upload using xrootd.
- Finally I would also enable the Storage Resource Reporting (to be ready for a life without SRM).
  - Just a cron job that runs a python script to produce a nice json.
  - https://fal-pygrid-30.lancs.ac.uk/dpm/lancs.ac.uk/home/atlas/storagesummary.json

Next steps after you've DOME'd.

So all our experience is with running DOME in "legacy mode".

- SRM, dpnsdaemon, dpm, rfiod all still running.
  - But of course come the end of September security support for these go away.
- You're mainly running all those gubbins to keep srm going.
  - Although actually there are some trepidation about turning off all the legacy stuff.
- Step one though will be to switch off SRM.
  - Hence why you want gridftp redirection - to still be able to do WAN TPCs (xroot and http not ready yet).
  - Atlas sites will want to move away from using SRM in their AGIS settings.
  - Other user groups will need to be notified.
  - Then your srm will need to be removed from your DPM bdii, and the entry for it removed from the gocdb.

Any questions?

Some links and useful emails:

https://twiki.cern.ch/twiki/bin/view/DPM/DpmSetup1100

https://twiki.cern.ch/twiki/bin/view/LCG/DPMupgrade

https://twiki.cern.ch/twiki/bin/view/LCG/ThirdPartyCopy

dpm-users-forum@cern.ch

dpm-devel@cern.ch