dCache Beginners Course
Pool Management and Selection

Get to know the scientifics behind the dCache pools.

Karlsruhe Institute of Technology (KIT), Steinbuch Centre for Computing (SCC)

dCache's Namespace

- dCache keeps a *global namespace* for all stored files.
    - *Chimera* is the current provider for this.
    - Presents a single-rooted file system view, traditionally having `/pnfs` as the root directory.
    - Direct access to the locally mounted namespace is possible with OS-level tools (`ls`, `mkdir`, `mv`, etc.).
    - However, direct I/O-operations (`cp`, `cat` and the like) require the *NFSv4.1* service.
    - The namespace service also features *directory tags* – more about them later on.
- The inner dCache components talk to Chimera via a module called *PnfsManager*.
- From the user's perspective, it is not visible where actually each file gets stored.

Karlsruhe Institute of Technology (KIT)

Steinbuch Centre for Computing (SCC)

dCache's Pool Manager

- Management of the pools is done by the unique service *PoolManager*.
  - With every file transfer the PoolManager is involved. More precisely, the *Pool Selection Unit* (PSU) – a submodule of the PoolManager – is asked to choose a pool.
- It is the PoolManager that keeps the configuration about classifications of the pools.
  - Define pools that are writable/readable.
  - Define pools that may archive to/stage from tape backend.
  - Define pools that are (not) usable for particular users.
  - et cetera.
- PoolManager may be configured online, or via `PoolManager.conf`.
- In order to make accurate decisions, the PoolManager has a *cost calculation module* integrated.
  - But that is all you will get to know about it in this course. ☺

Pool Selection – Finding Transfer Candidates

- Pools are selected as a suitable transfer candidate, if their classification in PoolManager is matching the requirements of a transfer.
- Maintaining such a detailed table of match constraints and searching through it is resource intensive.
- Therefore, dCache takes another approach, by grouping pools and transfer criteria together.
- Virtually, the decision table still exists, but is much smaller.

| PSU		pool1	pool2	pool3	pool4
write	gsiftp	yes	no	no	no
read	gsiftp	no	yes	no	no
write	dcap	no	no	yes	no
read	dcap	no	no	no	yes

Pool Selection – PoolManager Links

- Pools are grouped together in *pool group*s (pgroup).
- Transfer criteria are called *unit*s and are grouped in *unit group*s (ugroup).
- *Link*s combine ugroups with pgroups.
- The PSU will test all links, whether their ugroup matches the transfer requirements.
- If so, all pools of the linked pgroup/-s possibly will get selected.

| PSU	pgroup1	pgroup2	pgroup3	pgroup4
ugroup1	link1			
ugroup2		link2		
ugroup3			link3	
ugroup4				link4

PoolManager: Units

- In order to group units, you must first create a ugroup.
  - `psu create ugroup` *`<name-of-unitgroup>`*
- *Network units* allow restriction by IP adress (ranges).
  - `psu create unit` *`-net`* *`<IP-address>`*`|`*`<net mask>`*
- *Protocol units* define transfer protocols and version numbers.
  - `psu create unit` *`-protocol`* *`<protocol-name>`*`|`*`<version-number>`*
- *Storage* and *Cache Class units* are used for file based selection via *storage* or *cache class*.
  - `psu create unit` *`-store`* *`<StoreName>`*`:`*`<StorageGroup>`*`@`*`<type-of-storage-system>`*
  - `psu create unit` *`-dcache`* *`<name-of-cache-class>`*
- Now you can add the units to the ugroup.
  - `psu addto ugroup` *`<name-of-unitgroup> <unit>`*

ugroup	unit -net	unit -protocol	unit -store	unit -dcache
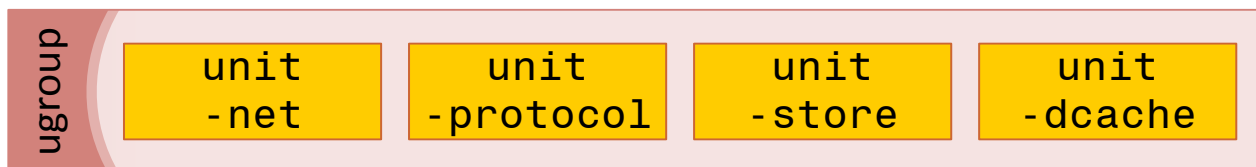
PoolManager: Pgroups

- Pools and pgroups are created very similar.
    - `psu create pgroup <poolgroup>`
    - `psu create pool <pool>`
    - `psu addto pgroup <poolgroup> <pool>`
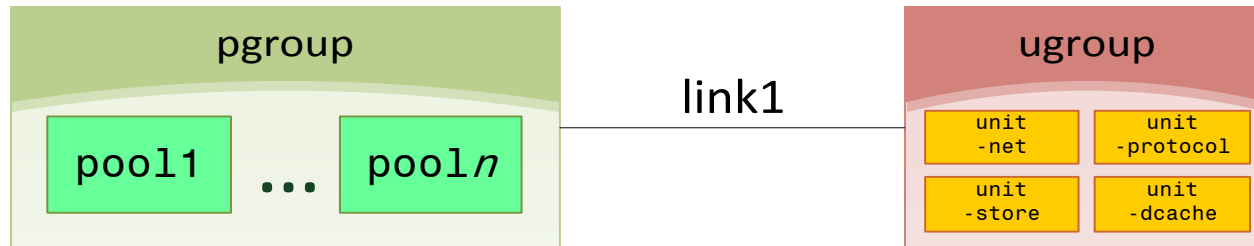
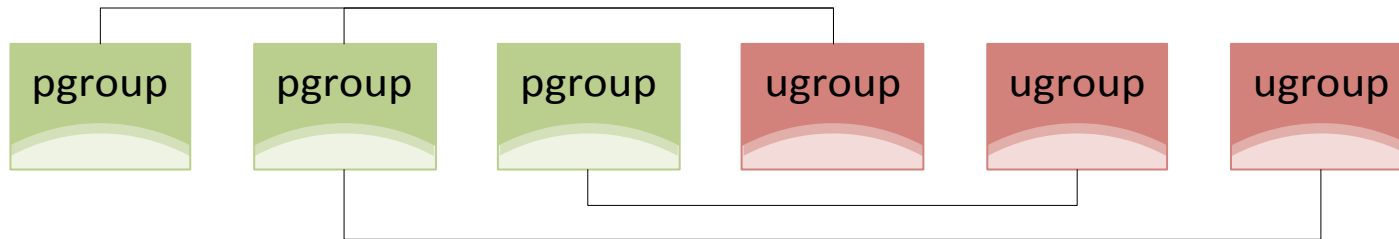- Finally, create links like this.
    - `psu create link <link> <name-of-unitgroup>`
    - `psu set link <link> -readpref=<10> -writepref=<0> -cachepref=<10> -p2ppref=<-1>`
    - `psu add link <link> <pgroup>`

PoolManager: Links

- ugroups as well as pgroups may be reused several times in different links.

- A link may contain only one ugroup, but unlimited pgroups.

- dCache uses preference values instead of plain yes/no statements.
  - A preference of 0 disables pools for the respective task.
  - p2ppref of -1 means "use the same value as for readpref".
  - All other values are meaningless, but dCache will check all links in descending preference order, if there is at least one pool online and useable.

Karlsruhe Institute of Technology (KIT)

Steinbuch Centre for Computing (SCC)

PoolManager: Store And Cache Units

- Store and cache units are also regarded to as *storage* and *cache tags*.

- If files are tagged with the same string, they will match the store/cache unit.

- In order to tag a file, the sysadmin needs to create special, hidden files in the /pnfs filesystem.

```
[root]# cd /pnfs/domain/experiment-a/
[root]# cat ".(tag)(OSMTemplate)"
StoreName myStore
[root]# cat ".(tag)(sGroup)"
STRING
```

- Look up the online documentation of this feature and you will find several examples illustrating the possibilities with the tools you learned so far.

Example

Look at what `PoolManager.conf` there is for your newly installed dCache.

- However, right now there is no such file and you will have to generate it.
- Log in to the Administration Interface and then the PoolManager cell.
- Execute 'save' once, which will generate the `PoolManager.conf` file with the default configuration.