LDMS User’s Group Meeting

July 21, 2019

Agenda

- SOS Storage Plugin Indexes
- Configuring the SOS storage plugin
- Data management with SOS
- OVIS 4.3 Updates

SOS Storage Plugin Indexes

- job_comp_time
  - Orders data by job_id by component_id, by timestamp
  - Useful for analysis by job
  - Logically identical to SQL: ORDER BY job_id, component_id, timestamp;
- comp_time
  - Orders data by component_id by timestamp
  - Useful for analysis by component
  - Logically identical to SQL: ORDER BY component_id, timestamp;
- Keys that include multiple values are called *composite keys*
- There are other indexes, but they will be deprecated in 4.3, e.g.
  - timestamp, component_id, job_id
  - rationale:
    - orderings by timestamp, but random by component and job are not useful
    - the same holds for simple orderings by component and job
    - indexes are compute intensive and consume storage

Timestamp

- The *time* part of the *job_comp_time* and *comp_time* composite keys
- An ldms_timestamp
  - tv_secs - seconds
  - tv_usecs - microseconds
- Value is the time the sampler finished updating the metric set with the latest values
  - Updated in the metric set by the ldms_transaction_end() function
- Does not require any plugin configuration

Component ID

- Identifies the system “component” or “device” on which the measurement was taken
- REQUIRES sampler plugin configuration
- The *comp* part of the *job_comp_time* and *comp_time* composite key
- Should never be unset (defaults to zero)
  - Changing to -1 in 4.3 to support direct NID Number → component_id mapping
- Currently, GUI uses zero as a wild-card
  - Wild card changing to -1 in 4.3
- Originally conceived to encode not only the node, but also internal components within a node, i.e. CPU, DIMM, etc…
- OVIS 4.3 includes support for *component mapping*

Job ID

- The job scheduler's unique job identifier
- REQUIRES sampler plugin configuration
- The *job* part of the *job_comp_time* composite key
- Used to query a time series for a job
- Has the default value of zero (0)
  - A value of zero means that there was no job scheduled on the node when the sample was taken
- Taken from another metric set that is updated by one of three samplers:
  - One job per node
    - jobinfo – Available since version 3
    - job_info – Available since version 3
  - Multiple jobs per node
    - slurm – Available in version 4.3

OVIS 4.3 Update

- Release Plan
  - Beta early September
  - Public github release early October
- Usability/Bug Fixes
  - Zero length configuration option values result in error/log message
  - -1 replaces 0 for *default/unset* component_id
  - If the job_set is initially missing, it will be searched for again at each sample interval
    - Avoids configuration issue where the job sampler is configured after a dependent metric set
  - Metric sets will be deleted as opposed to just unregistered
- Features
  - Multi-tenant (multiple jobs on one compute) capable Spank plugin
  - Multi-tenant job sampler
  - Mutli-tenant PAPI sampler
  - Storage plugins for new job and PAPI sampler that separate arrays into individual storage records for ease of analysis
  - Metric set info and meta-data is provided by ldms_xprt_dir() removing the need to perform ldms_xprt_lookup() to obtain this information

Sampler Plugin Configuration

- The component_id value and job set name from which to obtain the Job ID are specified when the sampler plugin is configured
- Example:

```
config name=slurm_sampler instance=${HOSTNAME}/slurm …
config name=meminfo component=${COMPONENT_ID} job_set=${HOSTNAME}/slurm …
```

- The ${<name>} syntax will obtain the value from the LDMS Daemon's environment

Common Sampler Configuration Problems

- The *component_id* configuration attribute is required, or a configuration error will be generated, however:

  config name=meminfo component_id= instance=${HOSTNAME}/meminfo …

  - will not generate an error but ‘’ is interpreted as zero by the parser
    - In OVIS-4.3 an error will be generated for configuration attribute values with a length of zero
- The *job_set* is not present at the time the sampler was configured
  - This results in a warning in the log and no job data associated with the metric set
  - In OVIS-4.3 the specified job set name will be searched for at the top of each sample interval
    - This will result in repeated log warnings if the job_set name is mis-configured

Sampler Configuration - continued

- A systemctl environment file is convenient for specifying the value of node specific environment variables
- /var/ldmsd.env

```
COMPONENT_ID=10001
…
```

- /lib/systemd/system/ldmsd.service

```
# shared environment variables
EnvironmentFile=/opt/ovis/etc/ldms/ldmsd-sampler.env
# node specific environment variables
EnvironmentFile=/var/ldmsd.env
…
```

SOS Storage Plugin Configuration

- One configuration parameter, *path*
- Sets the path in which the *container* will be created/used
- The *container* name comes from storage policy *container* configuration parameter
- Example:

```
load name=store_sos
config name=store_sos path=/local/store/DATA/Monday
strgp_add name=meminfo plugin=store_sos container=ldms_data
schema=meminfo
strgp_start name=meminfo
```

- Data will be stored in the SOS Container:

```
/local/store/DATA/Monday/ldms_data
```

SOS Data Management

- Local/fast storage needed to handle metric set insertion rate
- Data must be moved to slower/remote storage due to local storage limitations
- Solution is to periodically reconfigure the store_sos plugin to use a new path for the container, ex. at midnight:

```
config name=store_sos path=/local/store/DATA/Tuesday
```

- Subsequent data is stored in `/local/store/DATA/Tuesday/ldms_data`

SOS Data Management

- Example script in ovis/ldms/scripts/ovis-roll-over.py

```
usage: ovis-roll-over.py [-h] --src_path DB-PATH --dst_path DB-PATH
                         [--prev_link DB-PATH] [--next_link DB-PATH]
                        --next_name DB-PATH [--cfg_host PORT-NO]
                         [--cfg_port PORT-NO] [--cfg_auth_file PATH] [--save]

Rotate OVIS SOS Storage Containers.

optional arguments:
  -h, --help            show this help message and exit
  --src_path DB-PATH    The directory for the current container.
  --dst_path DB-PATH    The path where the current container will be archived.
  --prev_link DB-PATH   The link pointing to the previous LDMSD storage
                        subdirectory.
  --next_link DB-PATH   The link pointing to the next LDMS storage
                        subdirectory.
  --next_name DB-PATH   The name of the subdirectory in src_path where new
                        LDMSD data will be placed.
  --cfg_host PORT-NO    The hostname/ip-address where the ldmsd is listening.
  --cfg_port PORT-NO    The port on where the ldmsd is listening.
  --cfg_auth_file PATH  The path to the file containing the shared secret.
  --save                Specify if old data is to be preserved.
```