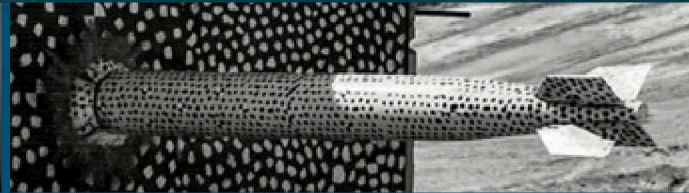


LDMS Genders



Ben Schwaller

LDMSCON 2019

- Genders configurations allows for a succinct way to configure LDMS in complex multi-cluster environments
- A single configuration file can be used for the admin, login, and computes nodes in a cluster
- Easy to separate different LDMS functionalities into their own LDMS instance / system controlled daemon
- Creates actual LDMS configuration files from more concise genders files

System Requirements / Build Configuration

- Need to have a genders rpm installed
- Need these enables in the configuration
 - --enable-libgenders
 - --enable-genderssystemd

Important man pages / file locations

man ldmsd-genders / ldms-attributes supplies all the information needed for genders configuration

- Ldmsd-genders gives an overview and describes environment variables, among other things
- Ldms-attributes gives gritty details on configuration for genders files

/etc/sysconfig/ldms.d/ (genderroot) contains all gender processing scripts

- genderroot/ClusterGenders/genders.* are the gender files
- genderroot/ClusterSecrets/ldmsauth.conf contains the ldms authentication for ovis auth
- genderroot/*.conf contain environment variable settings that reference the gender files

/var/run/ldmsd contains the actual ldmsd configuration files and environment variables derived from gender scripts

- Great for debugging genders

- The notion of load, configure, start used in normal ldms config files does not exist with genders config files
- Instead, hostnames are associated with an attribute
- Cluster1 ldmsd_port=411
- Attributes define hierarchical and local ldmsd behaviors
- Multiple hostnames can be associated with the same attribute which allows for the same file to be used on multiple hosts
 - Only attributes corresponding to the hostname that the daemon is running on are used in the ldmsd configuration
 - Different configurations are simply put on a different line
- Cluster-admin,Cluster[1-50] ldmsd_xprt=sock
- Cluster-login,Cluster[51-100] ldmsd_xprt=rdma

- `ldmsd` : enables sampling daemon
- `ldmsd_interval_default` : defines sampling interval time
- `ldmsd_offset_default` : defines offset from the interval
- `ldmsd_port` : defines port number for data server
- `ldmsd_xprt` : defines the protocol for the data server connections
- `ldmsd_metric_plugins` : list of sampler plugins to load
- `ldmsd_store_plugins` : list of store plugins to load
- `ldmsd_$(plugin)` : configuration line for a plugin
- `ldmsd_dbg` : defines the logging level
- `man ldms-attributes` has full list

Genders Config Notation

- Lists of items are separated by a colon
 - Cluster1 ldmsd_metric_plugins=vmstat:meminfo
 - Equivalent to
 - load name=vmstat; load name=meminfo
- Slashes are used as equal signs in a config notation
 - Only for the first slash after an equals sign or colon
 - Cluster1 ldmsd_filesingle=conf//etc/sysconfig/ldms.d/plugins-conf/filesingle.conf
 - Equivalent to
 - config name=filesingle conf=/etc/sysconfig/ldms.d/plugins-conf/filesingle.conf

Gender Config Notation (cont)

- %n can be used to insert the hostname into the attribute value
 - Cluster1 ldmsd_procstat=maxcpu/2:job_set/%n/jobinfo
 - Equivalent to
 - config name=procstat maxcpu=2 job_set=Cluster1/jobinfo
- Ampersand is used to separate list values rather than commas
 - Cluster1 ldmsd_procnetsdev=job_set/%n/jobinfo:ifaces/eno2&ib0
 - Equivalent to
 - config name=procnetdev job_set=Cluster1/jobinfo ifaces=eno2,ib0

- There are several special attributes for managing groups of LDMS data sources
 - CLIENTOFLIST, AGGCLIENTOFLIST, BOOTNODELIST, LDMSALL
- CLIENTOFLIST gets the nodes associated with ldmsd_clientof list
 - Cluster-admin[1-2]: ldmsaggd=CLIENTOFLIST
 - Cluster[1-50] ldmsd_clientof=Cluster-admin1
 - Cluster[51-100] ldmsd_clientof=Cluster-admin2
- AGGCLIENTOFLIST can be used to define higher level aggregation relationships
 - Cluster-super-admin ldmsaggd=AGGCLIENTOFLIST
 - Cluster-admin[1-4] ldmsaggd_clientof=Cluster-super-admin
- If nodes cluster-admin[1-4] are themselves to be monitored then it should be
 - Cluster-super-admin ldmsaggd=AGGCLIENTOFLIST:CLIENTOFLIST
- BOOTNODELIST and LDMSALL are other less commonly used lists

Genders Environment Config Files

- Main file format is `ldmsd.Cluster.conf` and it contains environment variable overrides for other gender files for that Cluster daemon
 - `ldmsd.agg.conf` and `ldmsd.local.conf` are the given default files
- `man ldms-genders` has all environment variables possible
- `LDMS_GENDERS` defines a list of files to be used for genders queries
 - `LDMS_GENDERS_1|2` are used for gender queries about upstream `ldmsd` daemons not in the `LDMS_GENDERS` file list
 - `LDMS_AUTH_FILE="/etc/sysconfig/ldms.d/ClusterSecrets/ldmsauth.conf"`
 - `LDMS_GENDERS="/etc/sysconfig/ldms.d/ClusterGenders/genders.agg"`
 - `LDMS_GENDERS_1="/etc/sysconfig/ldms.d/ClusterGenders/genders.local"`
- `ldmsd.all_instances.conf` can be used as an environment variable generator for all configuration files
 - I.e. use a hostname case statement to set different `ldmsd` ports on the L2 aggregator and have the `ldmsd_port=$L2_PORT`

II Running with systemd

- Systemctl restart ldmsd runs ldmsd using the ldmsd.local.conf for configuration
 - References genders.local unless LDMS_GENDERS is specified in ldmsd.local.conf
 - Used for running local ldmsd sampling
- Systemctl restart ldmsd@Cluster
 - Run ldmsd using other configurations found in ldmsd.Cluster.conf
 - References genders.Cluster unless LDMS_GENDERS is specified in ldmsd.local.conf
 - Using ldmsd@agg is common practice for L1 aggregator
 - Personally I only use non agg/local gender files for multi-cluster L2 aggregation
- Detailed installation location information for each can be found in ldmsd-genders

Debugging LDMSD Genders

- First check is to see if genders file is interpretable
 - `nodeattr -f genders.Cluster -parse-check`
- `/var/run/ldmsd` contains many files helpful for debugging
 - Traditional ldms configuration files created from genders can be found in the above directory as `all-config.Cluster`
 - The `genderscratch` directory has copies of the used genders files
 - The `ldmsd.env.Cluster` files describe what environment variables were set for the associated daemon



Genders Demo



Example Config Files

Genders.local Example

#Ldmsd Config

blake[01-31],blake-rps ldmsd

blake[01-31],blake-rps
ldmsd_interval_default=1000000,ldmsd_offset_default=0

blake[01-31],blake-rps ldmsd_port=411,ldmsd_xprt=sock

blake[01-31],blake-rps ldmsd_host=blake-rps

blake[01-31],blake-rps ldmsd_mem_res=2M

#Sampler Config

blake[01-31],blake-rps
ldmsd_metric_plugins=jobinfo:meminfo:vmstat:procstat:procnetdev:opa2:procnfs:filesingle

blake-rps ldmsd_idbase=5210001

blake[01-31] ldmsd_idbase=5200001

blake[01-31],blake-rps ldmsd_meminfo=job_set/%n/jobinfo

blake[01-31],blake-rps ldmsd_vmstat=job_set/%n/jobinfo

blake[01-31],blake-rps ldmsd_opa2=job_set/%n/jobinfo

#blake[01-31],blake-rps ldmsd_procnfs=job_set/%n/jobinfo

blake[01-31],blake-rps
ldmsd_filesingle=job_set/%n/jobinfo:conf//etc/sysconfig/ldms.d/
plugins-conf/filesingle.conf

#Plugin settings

blake[01-31],blake-rps
ldmsd_procstat=maxcpu/2:job_set/%n/jobinfo:schema/procstat

blake[01-31],blake-rps
ldmsd_procnetsdev=job_set/%n/jobinfo:ifaces/eno2&ib0

Genders.agg Example

#Ldmsd config

```
blake-rps ldmsd_port=412,ldmsd_xprt=sock  
blake[01-31],blake-rps ldmsd_clientof=blake-rps  
blake[01-31],blake-rps ldmsaggd=CLIENTOFLIST  
blake-rps  
ldmsd_schemas_store_csv=vmstat:procstat:procnetdev:  
meminfo:jobinfo:dstat:procnfs:filesingle:opa2  
blake-rps ldmsaggd_mem_res=1G  
blake-rps ldmsd_store_plugins=store_csv  
blake-rps ldmsd_log=/var/log  
blake-rps ldmsd_dbg=ERROR
```

#Storage config

```
blake-rps ldmsd_store_csv=path//mnt/netapp1/lun26-  
ldms/blake:altheader/0:rolltype/1:rollover/86400:buffer/0
```

#Collect dstat info about the admin aggregator daemon as well

```
blake-rps ldmsd_metric_plugins=dstat  
blake-rps ldmsd_dstat=stat/1:statm/1:mmalloc/1:io/1
```

Genders.L2.blake example

#Ldmsd config

```
mon ldmsd_port=${L2_PORT},ldmsd_xprt=sock
mon ldmsaggd_interval_default=${L2_INTERVAL}
mon ldmsaggd_offset_default=${L2_OFFSET}

blake-rps ldmsd_clientof=mon

mon ldmsaggd=CLIENTOFLIST

mon ldmsaggd_mem_res=1G

mon ldmsd_log=/var/log

mon ldmsd_dbg=ERROR
```

#Storage config

```
mon ldmsd_store_plugins=store_csv:store_sos

mon
ldmsd_schemas_store_csv=vmstat:procstat:procnetdev:meminfo
:jobinfo:dstat:opa2:dstat:procnfs:filesingle

mon
ldmsd_store_csv=path//scratch_A/blake/:altheader/0:rolltype/1:
rollover/86400:buffer/0

mon
ldmsd_schemas_store_sos=vmstat:procstat:procnetdev:meminfo
:jobinfo:dstat:opa2:dstat:procnfs:filesingle

mon ldmsd_store_sos=path//scratch_A/blake/${SOS_DATE}
```

Ldmsd.blake.conf

#this file is symlinked to ldmsd.agg.conf so individual cluster files are easy to create

LDMS_AUTH_FILE="/etc/sysconfig/ldms.d/ClusterSecrets/ldmsauth.conf"

LDMS_AUTH_TYPE=ovis

LDMS_GENDERS="/etc/sysconfig/ldms.d/ClusterGenders/genders.L2.\${LDMSCLUSTER}"

LDMS_GENDERS_1="/etc/sysconfig/ldms.d/ClusterGenders/genders.L1.\${LDMSCLUSTER}"

#genders.agg on blake-rps is symlinked as genders.L1.blake on mon

Ldmsd.all_instances.conf

```
export SOS_DATE=`date '+%Y-%m-%d'`;
clusterlist="blake caraway"
case $LDMSCLUSTER in
blake)
    export L2_PORT=413
    ;;
caraway)
    export L2_PORT=414
    ;;
*)
esac
```

```
base_interval=60000000
case $LDMSCLUSTER in
blake)
    export L2_INTERVAL=$base_interval
    export L2_OFFSET=1200000
    ;;
caraway)
    export L2_INTERVAL=$base_interval
    export L2_OFFSET=2200000
    ;;
*)
esac
```