

Banking



Brian Christiansen
SchedMD

SLUG 2019

Copyright 2019 SchedMD LLC
<https://www.schedmd.com>

- 
- Overview
 - Credits
 - Charging/Refunding
 - Usage Resetting
 - Decayed time/credits
 - Charge Rates
 - Time-based credits

Overview



- Customers switching from Gold/MAM
- Discuss how it's done in Slurm
- Wrappers haven been written to mimic banking nomenclature
 - SLURM Bank
 - https://slurm.schedmd.com/slurm_ug_2012/SUG2012-TCHPC-sbank.pdf

Credits

- Credits == Limits
 - `GrpTRESMins=<TRES>=#` (# of credits)
- Add credits
 - `sacctmgr modify <entity> set grptresmins=<TRES>=(current + new)`
- Usage/Balance
 - `scontrol show assoc`
 - `GrpTRESMins=TRES=LIMIT(USAGE),...`
 - N == No Limit
 - `GrpTRESMins=cpu=10(2),mem=N(1200)`

scontrol show assoc

```
$ scontrol show assoc users=brian accounts=stuff flags=assoc
```

```
...
ClusterName=lappy Account=stuff UserName= Partition= Priority=0 ID=3
SharesRaw/Norm/Level/Factor=2147483647/0.00/5/0.00
UsageRaw/Norm/Efctv=20612.12/1.00/1.00
ParentAccount=root(1) Lft=86 DefAssoc=No
GrpJobs=N(40) GrpJobsAccrue=N(29)
GrpSubmitJobs=N(69) GrpWall=N(69.59)
GrpTRES=cpu=500(80),mem=N(16000),energy=N(0),node=N(10),billing=N(160),...
GrpTRESMins=cpu=1000(171),mem=N(34350),energy=N(0),node=N(69),billing=N(343),...
GrpTRESRunMins=cpu=N(80),mem=N(16000),energy=N(0),node=N(40),billing=N(160),...
```

```
...
ClusterName=lappy Account=stuff UserName=brian(1003) Partition= Priority=0 ID=4
SharesRaw/Norm/Level/Factor=1/0.20/5/0.20
UsageRaw/Norm/Efctv=20612.12/1.00/1.00
ParentAccount= Lft=91 DefAssoc=Yes
GrpJobs=N(40) GrpJobsAccrue=N(29)
GrpSubmitJobs=N(69) GrpWall=N(69.59)
GrpTRES=cpu=N(80),mem=N(16000),energy=N(0),node=N(10),billing=N(160),...
GrpTRESMins=cpu=N(171),mem=N(34350),energy=N(0),node=N(69),billing=N(343),...
GrpTRESRunMins=cpu=N(80),mem=N(16000),energy=N(0),node=N(40),billing=N(160),...
MaxJobs= MaxJobsAccrue= MaxSubmitJobs= MaxWallP=
```

Copyright 2019 SchedMD
www.schedmd.com

Charging/Refunding

- Charging

- Jobs are charged for use
- AccountingStorageEnforce=associations,limits,qos,**safe**
 - Only run job if the job can run to completion with available GrpTRESMins
 - Otherwise, the job could be killed after the job is launched and exceeds the limit

- Refunding

- Find how much time the job used
 - `(sacct -j <jobid> -o CPUTimeRaw) / 60`
- Increase limit by time used

Usage Reset



- **PriorityResetUsagePeriod**
 - **NONE** - Never clear historic usage. The default value.
 - **NOW** - Clear the historic usage now. Executed at startup and reconfiguration time.
 - **DAILY** - Cleared every day at midnight.
 - **WEEKLY** - Cleared every week on Sunday at time 00:00.
 - **MONTHLY** - Cleared on the first day of each month at time 00:00.
 - **QUARTERLY** - Cleared on the first day of each quarter at time 00:00.
 - **YEARLY** - Cleared on the first day of each year at time 00:00.
- **Manaully with sacctmgr**
 - `sacctmgr mod account acctA set rawusage=0`

Decayed time/credits

- By default GrpTRESMins is decayed
- A couple of ways to handle hard limits
 - PriorityDecayHalfLife=0
 - No decay at all -- including fairshare
 - Use QOS with NoDecay QOS flag
 - QOS usage not decayed
 - Account/User usage still decayed

Billing TRES

- Use Billing TRES to define what a "credit" is.
- Limit against Billing TRES
 - `sacctmgr mod account acctA set grptresmins=billing=#`
- TRESBillingWeights == Charge Rates
 - Per partition
 - `PartitionName=gpu ... TRESBillingWeights=cpu=1,gres/gpu=2 # gpus charged at 2x`
 - **Note** that a QOS's UsageFactor can also affect the rate

Time-based credits



- Slurm doesn't natively do this
- Have to write an external program/db to manage this
 - and any other policies related to credits (e.g. roll over)

Questions?



Copyright 2019 SchedMD LLC
<https://www.schedmd.com>