www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Prolog, Epilog and X11 forwarding

Carles Fenoy
carles.fenoy@bsc.es

Barcelona, 05 February 2015

INTRODUCTION

INTRODUCTION

- Prolog: Script or application run at the beginning of an event
- Epilog: Script or application run at the end of an event

Parameter (slurm.conf)	Location	Invoked by	User	When executed
Prolog	Compute node or front end node	Slurmd daemon	SlurmdUser (usually root)	First job or job step initiation on that node
PrologSlurmctld	Head node, where slurmctld runs	Slurmctld daemon	SlurmctldUser	At job allocation
Epilog	Compute node or front end node	Slurmd daemon	SlurmdUser	At job termination
EpilogSlurmctld	Head node	Slurmctld daemon	Slurmctld User	At job termination

INTRODUCTION

Parameter (slurm.conf)	Location	Invoked by	User	When executed
SrunProlog or srun --prolog	srun invocation node	srun command	User invoking srun command	Prior to launching job step
TaskProlog	Compute node	slurmstepd daemon		
srun --task-prolog				
TaskEpilog				Completion job step
srun --task-epilog				
SrunEpilog or srun --epilog	srun invocation node	srun command		

Introduction

- Prolog Flags
  - Alloc: Execute prolog script at job allocation instead of job start. This forces the execution just when allocation is created, as for salloc instead of waiting for the first job step to run there.
  - NoHold: Prevents salloc to wait for the prolog to wait before returning control to the user.
- If Prolog is too slow it may cause a failure in job execution. The parameter BatchStartTimeout prevents a job from being killed before the prolog ends.
- ResvProlog and ResvEpilog for reservations start and end events.

Order of execution

1. **pre_launch_priv()** Function in TaskPlugin
2. **pre_launch()** Function in TaskPlugin
3. **TaskProlog**
4. **user prolog** Job step specific task program defined using **srun**'s **--task-prolog** option or **SLURM_TASK_PROLOG**
5. Execute the job step's task
6. **user epilog** Job step specific task program defined using **srun**'s **--task-epilog** option or **SLURM_TASK_EPILOG**
7. **TaskEpilog**
8. **post_term()** Function in TaskPlugin

Environment Variables

- Environment Variables available in Prolog and Epilog scripts

BASIL_RESERVATION_ID
MPIRUN_PARTITION
SLURM_ARRAY_JOB_ID
SLURM_ARRAY_TASK_ID
SLURM_JOB_ACCOUNT
SLURM_JOB_CLUSTER_NAME
SLURM_JOB_CONSTRAINTS
SLURM_JOB_DERIVED_EC
SLURM_JOB_EXIT_CODE
SLURM_JOB_EXIT_CODE2
SLURM_JOB_GID
SLURM_JOB_GROUP
SLURM_JOB_ID
SLURM_JOB_NAME
SLURM_JOB_NODELIST
SLURM_JOB_PARTITION
SLURM_JOB_UID
SLURM_JOB_USER

X11 FORWARDING

X11 FORWARDING

- X11 Forwading allows the usage of GUI though ssh connections.
- Some applications require the usage of a GUI in order to run
- Uses authentication based on magic cookie stored on a file in users home directory

X11 FORWARDING

- Our solution:
  - Our submit wrapper sets an environmental variable if the x11 directive is set " **#@ x11 = 1**"

```
SPANK_X11=login1/unix:10
```

  - A SPANK plugin on job startup adds the host to a Xauth authorization file and creates the ssh tunnel from the master node to the login

```
ssh -o UserKnownHostsFile=no \
    -L 127.0.0.1:6010:127.0.0.1:6010 -N login1
```

X11 FORWARDING

- Other solution:
  - SPANK plugin adds a new parameter to sbatch and srun (--x11=[batch|first|last|all])
  - It is also possible to use it with the environment variable SLURM_SPANK_X11


BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

USE CASES

Use Case 1

Nodes Health Check

- Run a script to check that everything is working properly
- If there is a problem, requeue the job and DRAIN the node to avoid other jobs trying to start on this node.
- Set the drain reason to indicate the issue found

Use Case 2

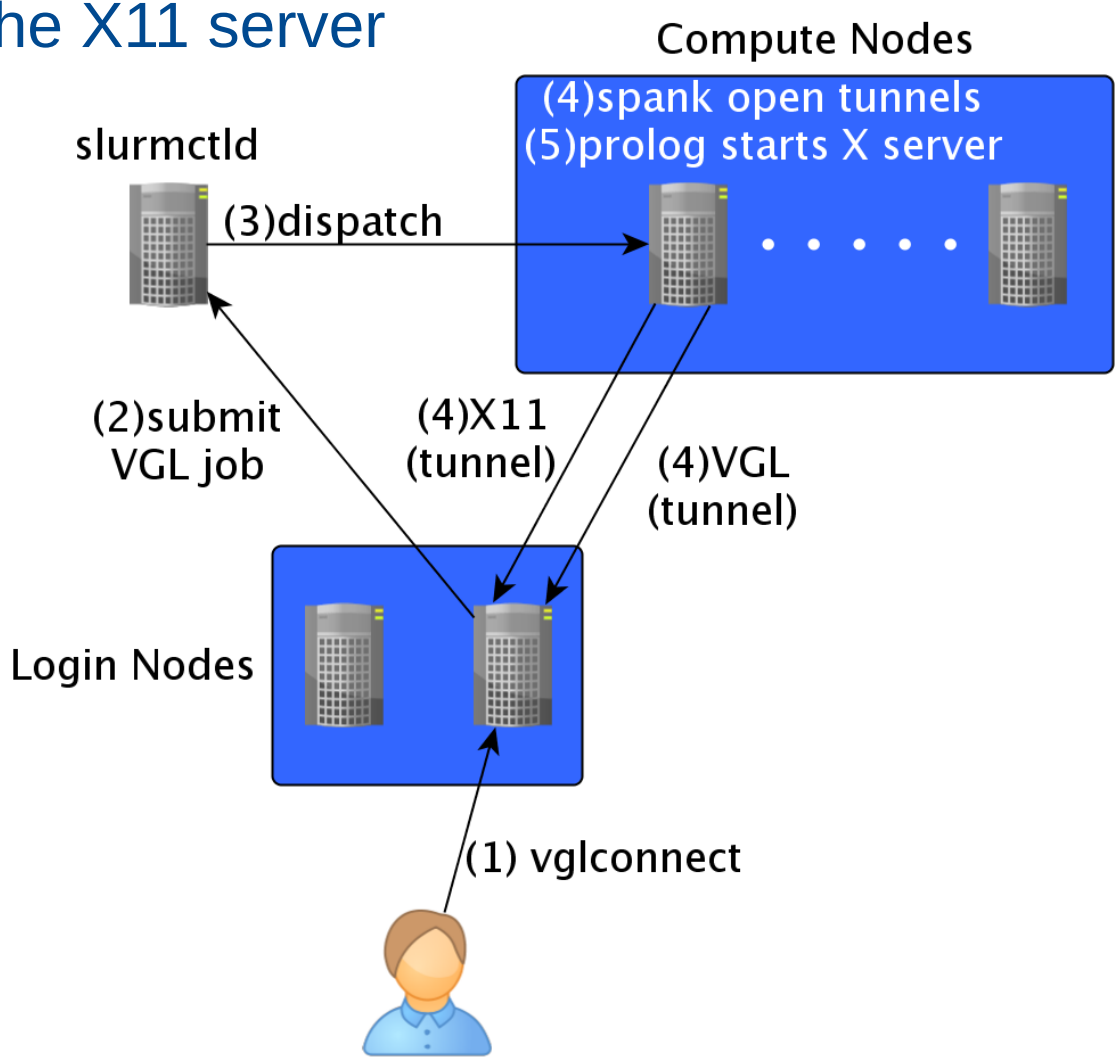
- Prepare the user environment to guarantee the proper operation of the compute nodes
  - Prolog:
    - Create a tmp directory for the job
    - Set the TMPDIR environment variable to this directory
    - Create a shm directory for the job
    - Set the RAMDISK environment variable to this directory
  - Epilog:
    - Remove TMPDIR and RAMDISK directories

Use Case 3

- Change some libraries for using OpenCL on Xeon processors instead of the Nvidia cards
  - Prolog
    - Check if an environment variable is set
    - Add the softlink to the Intel libraries in the OpenCL vendors directory
    - If variable not set, remove (if exists) the intel library from OpenCL vendors directory

Use Case 4

- VirtualGL used access the X11 server of the nodes
- Spank plugin to redirect X11 and virtualgl connection to login node



www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thank you!

For further information please contact
carles.fenoy@bsc.es