julian.rueger@gwdg.de

Julian Rüger

The Clustershell

A thousand nodes but only one admin?!

Table of contents

1 Introduction

2 The Clustershell

3 Exercise

Learning Objectives

- Understand the difficulty of managing many computers remotely
- Use the clustershell to
  - execute commands on remote nodes
  - copy files between cluster-manager and remote nodes

Real World Problems

- You need to restart the slurmd on nodes `n100–n200`
- `systemctl restart slurmd` (?)

How would you do it?

The Clustershell

- A shell tool to execute commands on many nodes in parallel
- Set up on management node
- Uses ssh to connect to remote nodes

The Clustershell

Execute commands on many nodes at once

- `clush -w n[100-200] 'systemctl restart slurmd'`
- `clush -w n[100-200] -x n[110-120] 'systemctl restart slurmd'`

Gather information with binned responses

- `clush -w n[100-200] 'ls /some/path | wc -l'`
- `clush -b -w n[100-200] 'ls /some/path | wc -l'`

The Clustershell

Copy files to remote nodes

- `clush -w n[100-200] --copy /some/file`

Copy files from remote nodes back to the cluster-manager

- `clush -w n[100-200] --rcopy /some/file --dest /where/to/collect`

Use memorable group names to refer to node sets

- `clush -w @compute ...`
- `clush -w @racks:rack1 -x n150 ...`

Exercise

- Install the clustershell on your cluster
- Use the clustershell
- Explore the possibilities of the nodeset utility tool