Alces Flight Solo clusters on IBM SoftLayer

High Performance Cluster deployment on Private Cloud

Alces Flight - 2016.3

Mark Titorenko - Alces Flight Ltd

November 2016

Revision number	Release date	Release notes	Author(s)
1.1	December 2016	Updated release for 2016.3r2	Mark Titorenko

CONTENTS

THIS DOCUMENT AND INCLUDED ALCES FLIGHT LOGOS ARE COPYRIGHT 2016 ALCES FLIGHT LTD. OTHER PRODUCT NAMES, LOGOS, BRANDS AND OTHER TRADEMARKS REFERRED TO WITHIN THIS DOCUMENTATION, AS WELL AS OTHER PRODUCTS AND SERVICES ARE THE PROPERTY OF THEIR RESPECTIVE TRADEMARK HOLDERS. ALL RIGHTS RESERVED. THESE TRADEMARK HOLDERS ARE NOT AFFILIATED WITH ALCES FLIGHT, OUR PRODUCTS, OR OUR SERVICES, AND MAY NOT SPONSOR OR ENDORSE OUR MATERIALS.

THIS MATERIAL IS DESIGNED TO ASSIST CAPABLE USERS TO CUSTOMISE OUR SOFTWARE PRODUCTS, EXTENDING THEM TO PERFORM OPERATIONS BEYOND ORIGINAL DESIGN PARAMETERS. EXAMPLES OF POSSIBLE EXTENSIONS ARE PROVIDED TO HELP CUSTOMERS REALISE THE FUTURE POTENTIAL OF THE SOFTWARE AND ARE PROVIDED WITHOUT WARRANTY AND ARE NOT SUPPORTED OR GUARANTEED BY ALCES FLIGHT LTD OR LOCAL SOFTWARE RESELLERS. THIS DOCUMENTATION HAS BEEN CREATED TO INFORM AND SHARE KNOWLEDGE WITH CUSTOMERS FOR REFERENCE PURPOSES ONLY; SOFTWARE FEATURES AND PERFORMANCE ARE NOT GUARANTEED. THIS DOCUMENTATION IS NOT DESIGNED AS A STAND-ALONE TRAINING TOOL – EXAMPLE COMMANDS AND SYNTAX ARE INTENDED TO DEMONSTRATE FUNCTIONALITY IN A TRAINING ENVIRONMENT AND MAY CAUSE DATA LOSS IF EXECUTED ON LIVE SYSTEMS. REMEMBER: ALWAYS TAKE BACKUPS OF ANY VALUABLE DATA. THIS DOCUMENTATION IS PROVIDED “AS IS” AND WITHOUT ANY WARRANTY; WITHOUT EVEN THE IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. NO COMPUTER SYSTEM CAN BE COMPLETELY SECURE - FOLLOW GOOD SECURITY ADVICE AND ALWAYS HAVE YOUR COMPUTER SYSTEMS CHECKED BY COMPETENT SECURITY PROFESSIONALS BEFORE BEING USED WITH LIVE DATA. PLEASE SEE THE EULA INCLUDED WITH THE PROVIDED SOFTWARE PACKAGES FOR FULL USAGE TERMS AND CONDITIONS. WE WELCOME CUSTOMER FEEDBACK AND SUGGESTIONS FOR FUTURE ENHANCEMENTS - PLEASE VISIT THE COMMUNITY SUPPORT SITE AT WWW.ALCES-FLIGHT.COM.

ALCES FLIGHT COMPUTE IS FREE SOFTWARE PUBLISHED UNDER THE TERMS OF THE GNU AFFERO GENERAL PUBLIC LICENSE AS PUBLISHED BY THE FREE SOFTWARE FOUNDATION, EITHER VERSION 3 OF THE LICENSE, OR (AT YOUR OPTION) ANY LATER VERSION. SOFTWARE IS MADE AVAILABLE IN THE HOPE THAT IT WILL BE USEFUL, BUT WITHOUT ANY WARRANTY; WITHOUT EVEN THE IMPLIED WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. SEE THE GNU AFFERO GENERAL PUBLIC LICENSE FOR MORE DETAILS (HTTP://WWW.GNU.ORG/LICENSES/). A COPY OF THE GNU AFFERO GENERAL PUBLIC LICENSE IS DISTRIBUTED ALONG WITH THIS PRODUCT. FOR MORE INFORMATION ON ALCES FLIGHT, PLEASE VISIT: HTTP://WWW.ALCES-FLIGHT.COM/. PLEASE SUPPORT SOFTWARE DEVELOPERS WHEREVER THEY WORK – IF YOU USE OPEN-SOURCE SOFTWARE, PLEASE CONSIDER CONTRIBUTING TO THE MAINTAINING ORGANISATION OR PROJECT, AND CREDITING THEIR WORK AS PART OF YOUR RESEARCH AND PUBLICATIONS.

ACKNOWLEDGEMENTS:

THE CREATION OF THIS WHITEPAPER WAS SUPPORTED BY AN IBM GLOBAL ENTREPRENEUR GRANT WHICH FUNDED THE NECESSARY TESTING RESOURCES ON THE IBM SOFTLAYER PLATFORM. WE WOULD LIKE TO THANK IBM FOR THEIR ASSISTANCE IN SUPPORTING THIS ACTIVITY, AND HOPE THAT THE PROVIDED DOCUMENTATION IS USEFUL TO THEIR CUSTOMERS IN DELIVERING HPC CLUSTER SOLUTIONS ON SOFTLAYER.

Overview

High Performance Computing (HPC) clusters provide an excellent mechanism for processing large numbers of compute jobs in a short amount of time. By allowing users to deploy an HPC cluster on a range of Cloud platforms, Alces Flight enables customers to choose the environment which suits their workloads. This choice may encompass freedom to choose platforms based on a range of factors, including:

- Locality; the platform might be locally hosted a the customer's facility, or held remotely from them

- Performance; the required compute and storage resources can be selected to suit the workload. This includes jobs which may incorporate many different stages during completion, and a mixture of different types of compute and storage processing elements.

- Budget; platform providers may deliver services with different cost models, or even enable jobs to be run for free at point of use.

- Features; some jobs may require very basic features, while other workflows can be optimized via the use of specialist features, such as additional cloud services, dedicated or specialized hardware, or availability of compatible software tools.

- Availability; customer can choose a platform that offers additional availability, redundancy and resilience, in combination with the other factors in this list. Some platforms may be generally available all the time, while others may optimize other factors with the time when jobs are run.

There are many different platform choices available, and no single solution is correct for every use case. This whitepaper details the use of Alces Flight on the IBM SoftLayer environment – a private cloud facility managed and maintained for customers by IBM. SoftLayer has many different solutions available to customers, including both bare-metal servers and virtual servers with a range of high-bandwidth networking options.

Intended audience

This white paper is primarily intended for use by users and administrators who are already utilizing HPC clusters for their research and scientific computing requirements. The paper is not a detailed configuration guide - but does advise best practices for particular use cases and scenarios. This paper assumes that you have an existing IBM SoftLayer account set up, or are able to set up an account in order to support the steps described. For further information on accessing IBM SoftLayer, contact their sales team at the following URL:

http://www.softlayer.com/

What is a High Performance Compute cluster?

A High Performance Compute (HPC) cluster is typically formed using a number of compute nodes which are loosely coupled to allow compute jobs to be run which are larger than the resources available on any individual compute node. HPC clusters typically deliver a cost-effective way to run both high-performance and high-throughput workloads using commodity computer hardware, managed by a batch scheduler system that efficiently organizes and processes user jobs.

HPC clusters typically run a modern Linux operating system, and are deployed using high-performance networking components and storage facilities to aid efficient workload processing. Most clusters employ one or more *master* nodes that are responsible for management facilities, and a larger number of *compute* nodes which process the user workload. Depending on the software applications used by users, special purpose hardware may also be included such as graphical rendering hardware, multi-core processing devices and accelerated interconnect technologies to improve the scalability of jobs. Clusters can scale from a handful of nodes to many thousands of connected computers – the optimum size of a cluster depends on many factors, not just the budget available for implementation.

What is Alces Flight Solo?

Once you have hardware available to create your cluster, many different software packages can be integrated together to create an HPC cluster. The software selected defines the features of your cluster, and the workloads which it can be used to process. Whichever team is responsible for maintaining the hardware typically have a set of software to keep the platform running – this may include the server operating system, storage and interconnect drivers, monitoring and management software needed for that team to do their job efficiently. Modern cloud platforms such as IBM SoftLayer, Amazon Web Services, Microsoft Azure and Google Cloud Platform include in-house software to provide these functions; for bare-metal hardware, Alces provides the *Alces Symphony* software toolkit to deliver platform management services.

Alces Flight is a software HPC environment designed to be compatible with a range of different platform providers. It comes complete with large library of software applications compiled as the *Alces Gridware* project, and provides common HPC facilities such as a choice of job schedulers and data management tools. Flight has been designed to compatible with both private cloud and public cloud when complemented by the software tools provided by the platform itself – or when deployed with the *Alces Symphony* toolkit in a bare-metal environment.

Alces Flight Solo is a single-user cluster environment, designed for use by researchers, scientists and developers using HPC clusters in their day-to-day jobs. Flight Solo is designed to enable individuals to quickly and easily deploy HPC environments wherever they're needed, delivering optimized applications ready to use on a chosen cloud platform. Once workload processing is completed, the ephemeral cluster can be easily dissolved by the researcher, with resources returned to the cloud provider. This method of working helps to leverage some of the benefits of

cloud platforms, such as paying for resources only when they are required, instead of requiring a large capital investment to be able to process HPC workloads.

The Alces Flight Enterprise product is also available to build persistent, multi-user clusters on different cloud platforms. Capable of being managed by a traditional infrastructure team, a Flight Enterprise cluster delivers the same benefits of *Gridware* application delivery with the ability to scale your HPC environment on demand. This whitepaper focuses on delivery of the Flight Solo product for individual researchers – for more information on the Flight Enterprise product, please visit the Alces Flight webpage at the URL below:

http://alces-flight.com

How do I build an HPC cluster on IBM SoftLayer?

Before starting to build your cluster, you need to decide what hardware you want to use. Contact the IBM SoftLayer team for help choosing which servers and storage configurations are available in the data centre where you want to run your workload. Alternatively, customers can also purchase directly from the IBM SoftLayer website, at the URL below:

http://www.softlayer.com/

You will want to think carefully about the equipment that you choose to build your cluster, as this will have implications on the performance you can achieve for your workload. The server types you'll need are:

- One login node; the login node is the most important server in the cluster, acting as your access point and shared filesystem host. Choose a node with the resilience you need, along with the CPU, memory and disk resources suitable for your workload. Login nodes can be physical or virtual; they need a minimum of two network interfaces, at least two CPU cores, and 4GB RAM.

- Compute nodes; most clusters have at least two compute nodes – the number and size of your nodes will depend on the jobs you want to run. Flight will automatically detect the configuration of your compute nodes, and setup the cluster job scheduler to take any limitations into account. Compute nodes can be physical or virtual; they need a minimum of one network interface and 2GB RAM. For job scheduling purposes, it can be beneficial to deploy all your compute nodes as the same type of hardware – however, this is not a requirement.

- Networking; the primary interface of all your login and compute nodes will be deployed on the same private SoftLayer VLAN. The secondary interface of your login node will be configured on the public SoftLayer VLAN, and provide access to the cluster for your users. The performance of your network interfaces will depend on the type of hardware selected for your login and compute nodes. In order for Flight to function correctly, it must have outbound internet access via the public VLAN.

How do I keep my data safe?

This whitepaper provides the steps required to build your own HPC cluster using the IBM Software platform. Good security practices are followed throughout, including the use of SSH public key access instead of setting passwords, and configuring a network firewall on the cluster login node. However – users are responsible for their own clusters, and ensuring that their data is safe from both loss and unauthorized access.

Alces Flight Solo contains a range of data-management tools which you can use to copy your data to secure storage once your workflow has been completed. For more information, see the product documentation at the URL below:

http://docs.alces-flight.com/en/stable/databasics/data_basics.html

Contact your IBM SoftLayer sales team for advice on backup and disaster recovery options, and for assistance in securing your environment.

Building your cluster – prerequisites

Before starting to build your cluster, order the login and compute nodes in your IBM SoftLayer platform. In the examples below, we have selected the following hardware:

- 1 x login node:
  - Supermicro server with E3-1270-V3-Quadcore CPU
  - 8GB RAM
  - 1TB of RAID1 configured onboard storage
  - Dual 1Gb network interfaces
    - Interface eth0 connected to private VLAN
    - Interface eth1 connected to public VLAN

- 2 x compute nodes:
  - Supermicro server with E3-1270-V3-Quadcore CPU
  - 8GB RAM
  - 1TB onboard storage
  - 1Gb network interface
    - Interface eth0 connected to private VLAN

Next, import the *Flight Compute* provisioning script into your IBM SoftLayer environment – as shown in appendix B to this whitepaper. Please see the following URL for information on how to import a provisioning script to your IBM SoftLayer platform:

https://knowledgelayer.softlayer.com/topic/provisioning-scripts

In order to inform your cluster nodes how they are to be configured, your servers must be deployed with two SSH keys to identify their configuration parameters. The first key is a "***Cluster***" key which allows the provisioning script to select the set of configuration to be applied. The second key is a "***Role***" key which identifies whether they should function as a login (master) node or compute (slave) node.

We have supplied three SSH keys in appendix A of this whitepaper:

- An example ***Cluster*** key which will configure a standard cluster
- A ***Role*** key identifying the function of the server as a login node, "Master"
- A ***Role*** key identifying the function of the server as a compute node, "Slave"

In order to run Flight on SoftLayer, you will need to import these SSH keys to allow Flight to identify the purpose of each of the deployed servers. Users will also need to have a user SSH key configured in SoftLayer to allow them to login to the cluster environment. To add a new SSH key, follow the instructions at the following URL:

https://knowledgelayer.softlayer.com/procedure/ssh-keys-0

Building your Flight Compute Solo cluster

Once your platform has been deployed, follow the steps below to deploy Alces Flight Solo and build your HPC environment.

Step 1 – Choosing the provisioned operating system

Alces Flight Solo has been tested with the IBM SoftLayer CentOS7 operating system. When your servers are ordered, or if you choose to redeploy your servers once they are available, choose the “CentOS7.0 64-bit” operating system option.

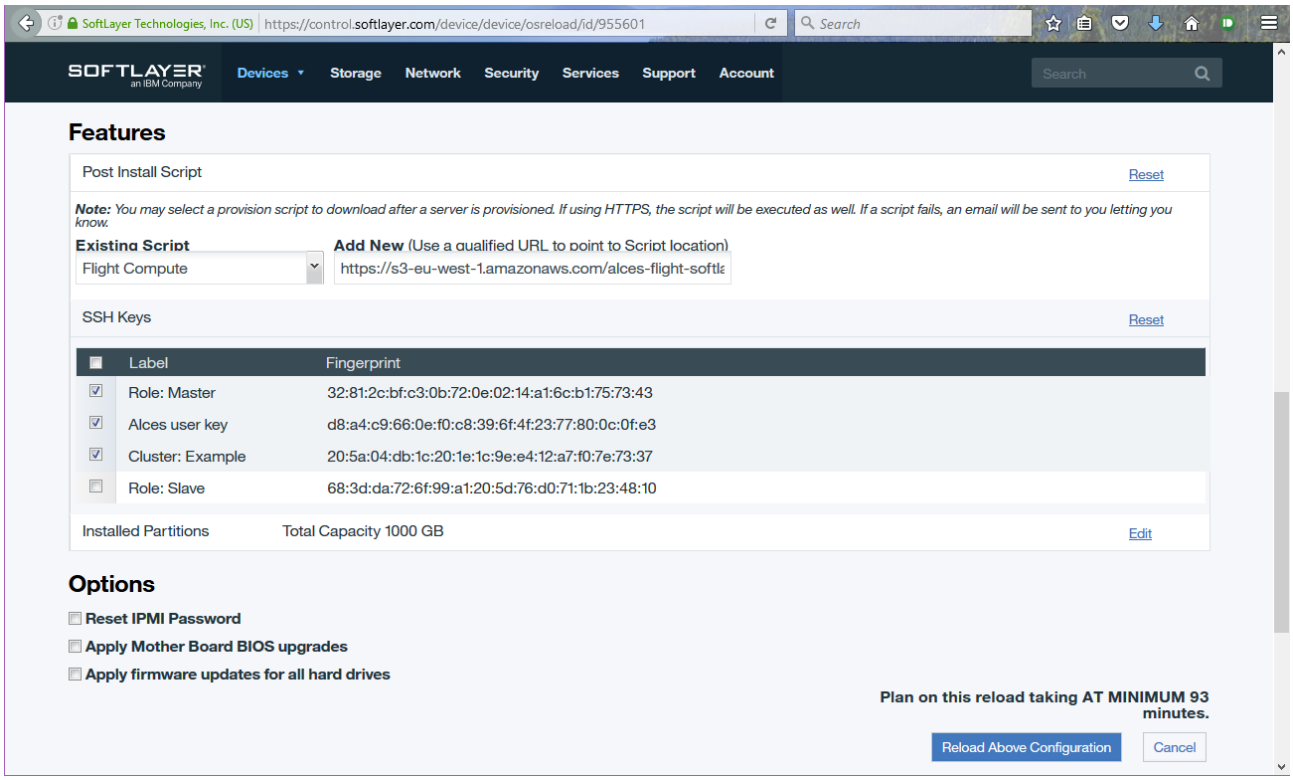
Select the ***Alces Flight Solo*** provisioning script to run when your provisioned operating system is installed. The same provisioning script can be used for both cluster login and compute nodes. The URL for the provisioning script is included in appendix B to this whitepaper.

Step 2 – Building the login node

The cluster login node should be built using the ***Flight Compute*** provisioning script enabled in the step above. Select the following SSH keys for your login node:

- The *“**Role: Master**”* SSH key, to enable this server to be built as a login node
- The "***Cluster: Example***" SSH key, to identify the configuration set for your cluster

- A user public SSH key which you can use to login to your cluster once it is deployed; you should ensure that you have access to the private SSH key before deploying your login node.

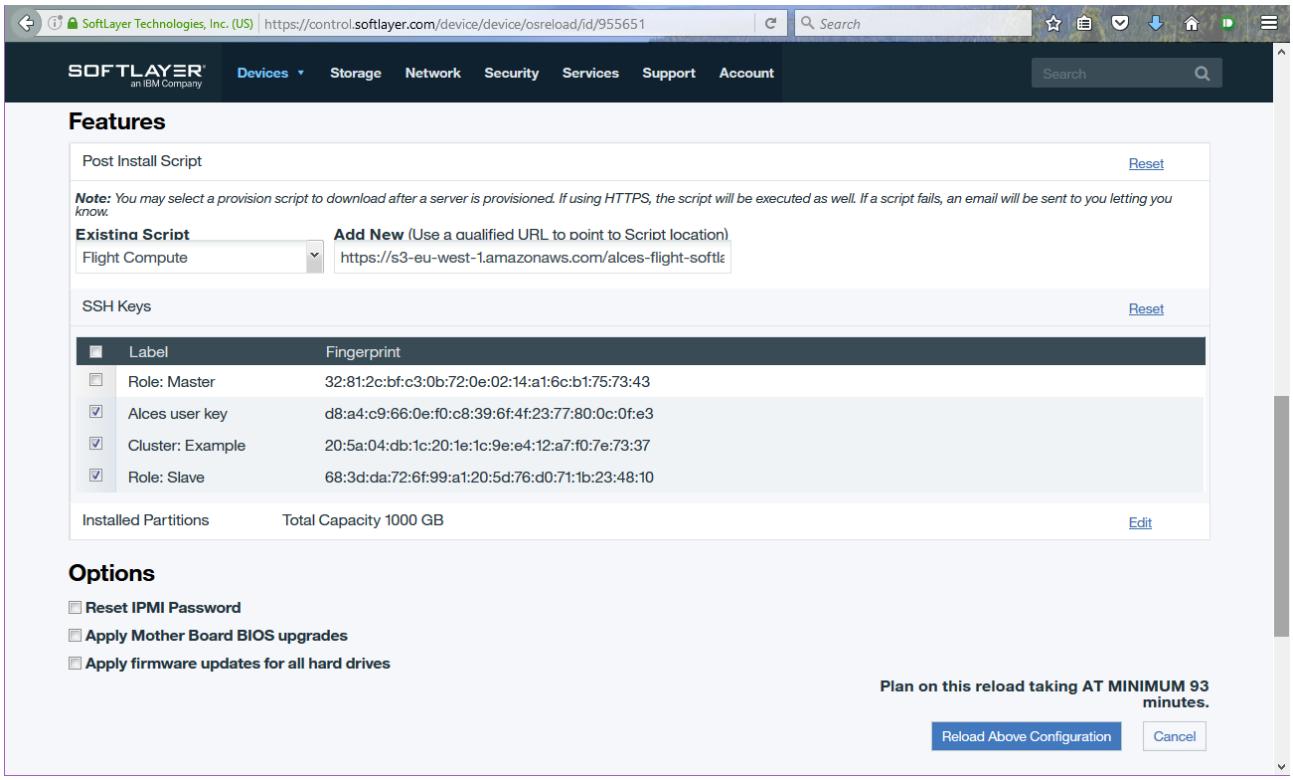


Start the OS deployment once your selection is complete – the time taken to complete the OS reload will depend on the performance of the server selected. Do not deploy any compute nodes until the OS deployment on your login node has been completed.

Step 3 – Building cluster compute nodes

Once the cluster login node has been built, users can move on to building their cluster compute nodes. Each compute node should be deployed with the same OS revision, on the same private VLAN, and use the same Alces Flight Compute provisioning script as the cluster login node. Select the following SSH keys for all your cluster compute nodes:

- The ***"Role: Slave"*** SSH key, to enable this server to be built as a compute node
- The "***Cluster: Example***" SSH key, to identify the configuration set for your cluster
- The same user public SSH key which you used to build your cluster login node



Start the OS deployment once your selection is complete – the time taken to complete the OS reload will depend on the performance of the server selected, but it is possible to deploy all your compute nodes in parallel. Do not deploy any compute nodes until the OS deployment on your login node has been completed. Deploying several compute nodes at the same time is possible – the time taken to configure your compute nodes will depend on the performance of your chosen hardware and networks. Compute nodes will automatically be registered in your cluster job scheduler once they are ready for use.

Alces Flight Solo will automatically configure your compute nodes to join the cluster services provided by the existing login node discovered on their private VLAN. If you want to move an existing compute node between Flight Compute clusters, they should be powered off, changed to the private VLAN of the new login node host, before being redeployed as described above.

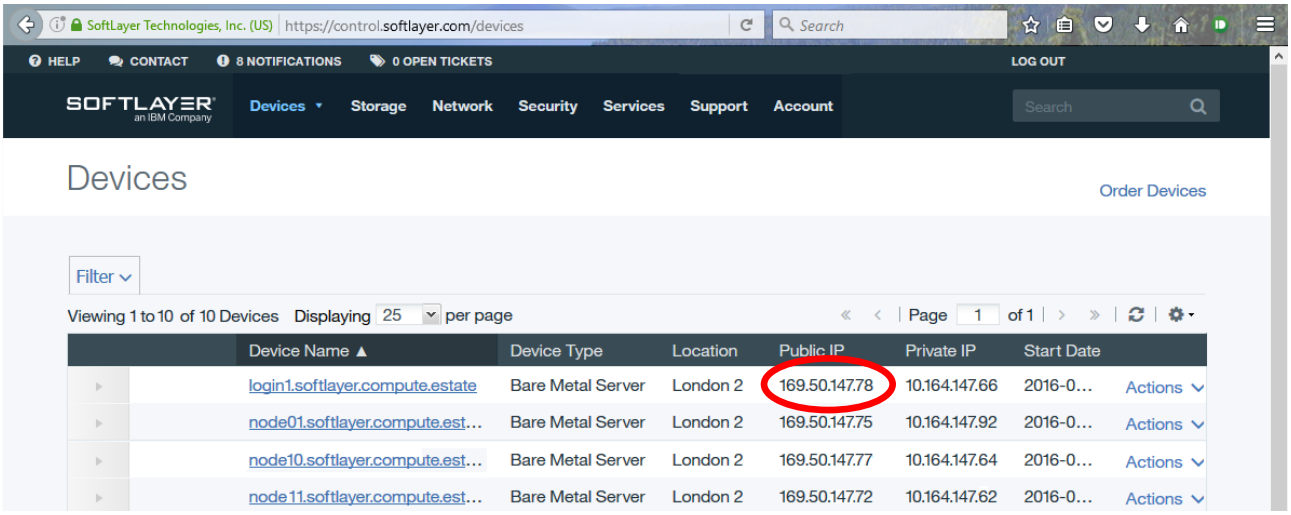
Accessing your Flight Solo cluster

Once your cluster has been deployed, a user will be able to login with the user SSH key provided during the steps above. Use a compatible SSH client to access your cluster login node using the public IP address allocated by SoftLayer. For information on accessing your Flight Solo cluster, review the product documentation at the following URL:

http://docs.alces-flight.com/en/stable/overview/whatisit.html#prerequisites

The IBM SoftLayer platform automatically configures the provided user public SSH key for the **root** account on provisioned nodes. In addition, Flight creates a standard user login with the name ***alces*** which users should use for running workloads and installing software. Users should login to their cluster as the ***alces*** user, using their SSH private key which matches the public key provided during deployment.

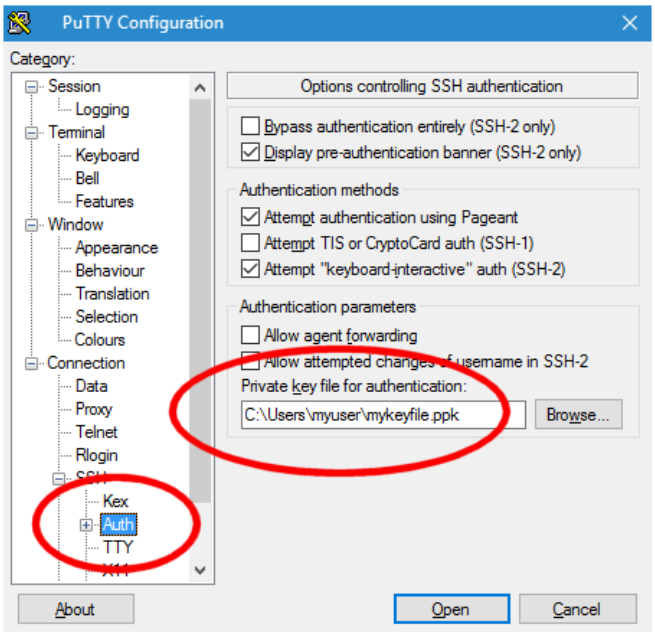
The public IP address for your cluster login node will be reported by the IBM SoftLayer web portal – from the ***Devices*** menu, choose ***Device List*** to show your purchased resources, and find the public IP address of your login node.

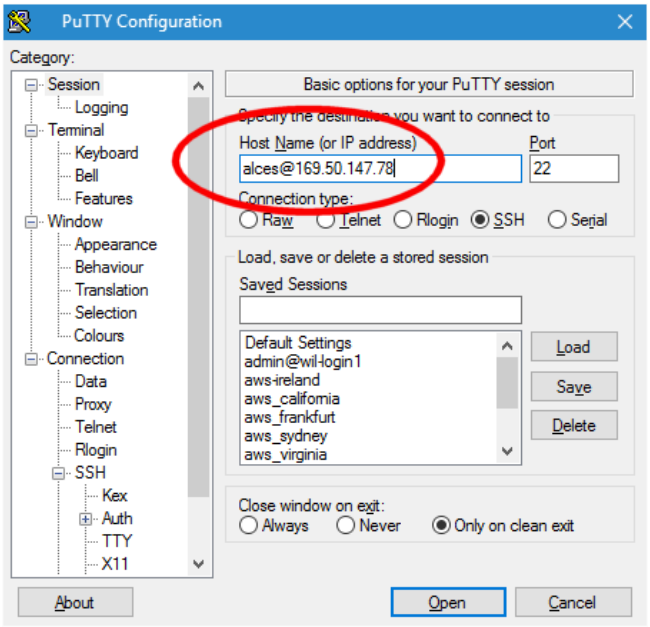


From a Linux or Mac client, use the **`ssh`** command to log in to your cluster login node as the **alces** user, providing your public SSH key; e.g.

```
ssh -i mykeyfile.pub alces@169.50.147.78
```

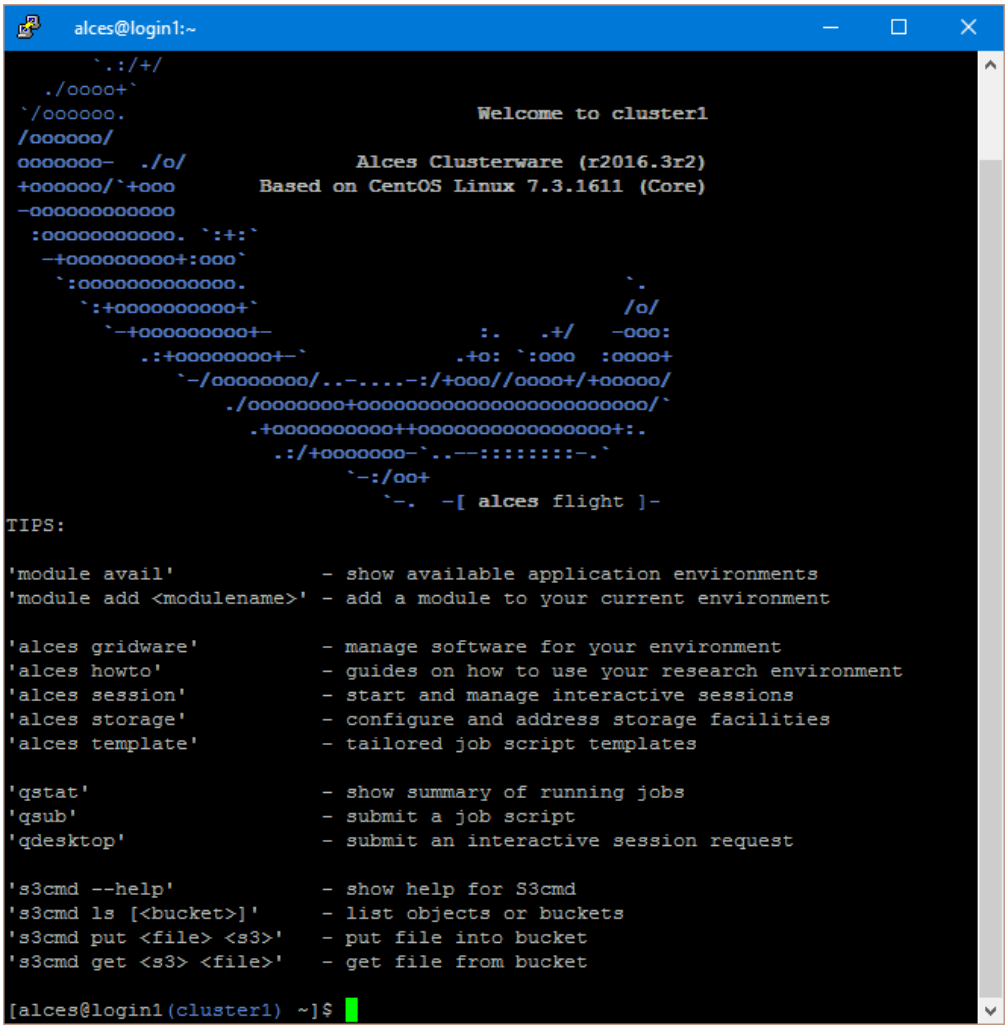
From a Windows client, use an SSH client utility such as PuTTY, or OpenSSH. To add your public SSH key file, select the ***Connection*** category on the left-hand side of the configuration page, choose the ***SSH*** sub-menu, then choose your public SSH under the ***Auth*** menu item.





Enter your login node public IP address in the ***Host Name*** box and click on the ***Open*** button to connect to your cluster.

You will be logged into your cluster as the ***alces*** user; you have full *sudo* access to all cluster nodes, and can install gridware software packages as required.



Running HPC workloads on your cluster

Once deployment is complete, users will be able to use their Flight Solo cluster by following the product documentation located at the following URL:

http://docs.alces-flight.com/en/stable/basics/basic_cluster_operation.html

To connect to the SWIFT compatible IBM SoftLayer object storage, use the credentials and service address provided when the storage was purchased. The example below shows a user connecting to the storage service located in the LON02 SoftLayer data-center:

```
[alces@login1(softlayerdemo) ~]$ alces storage enable swift
alces storage enable: enabled storage type: base/swift -> swift

[alces@login1(softlayerdemo) ~]$ alces storage configure mystorage swift
Display name [mystorage]: mystorage
Username: SLOS1023365-2:SL1023365
API key: **********************************
Authentication endpoint: https://lon02.objectstorage.softlayer.net/auth/v1.0/
alces storage configure: storage configuration complete

[alces@login1(softlayerdemo) ~]$ alces storage use mystorage
alces storage use: storage configuration 'mystorage' now set as default

[alces@login1(softlayerdemo) ~]$ alces storage ls
2016-09-15 10:54       DIR   foobar
2016-09-15 11:50       DIR   scripts
2016-09-15 10:00       DIR   test

[alces@login1(softlayerdemo) ~]$
```

Making sense of it all

Our testing has demonstrated how users can deploy their own Alces Flight Solo clusters using the IBM SoftLayer cloud platform. Clusters deployed with bare-metal compute nodes are capable of accessing maximum performance for compute jobs launched via the job-scheduler, without the overhead of a hypervisor-hosted environment.

For more information, we encourage users to launch their own environments to evaluate using their own IBM SoftLayer resources. Further documentation is provided at the URLs below which detail how to configure and use Alces Flight Solo clusters using the IBM SoftLayer platform for your scientific computing requirements.

https://docs.alces-flight.com/
https://www.softlayer.com/

Appendix A: Cluster role identification

Import the following SSH keys to enable cluster server role identification on the IBM SoftLayer platform:

The ***Cluster: Example*** SSH key:

```
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAABAQDg4OWEBX9LOewSAiWMEDCGf+IaiePQhTnHd4SZnJ9uUVlj8d
r1bhrWOO4+8A8adczhZwa9h5pTdIqOF1mtZBJgyG5FWW1PpC73OixL636D7X957ZM19f3V+OofuMHn
nAKrOEQwXpbzMGLhhwHTgztvyBbqeZlz6dIoomMg9WYR53Ucawk5k2saV0TJyZJHcTvFPgNwdR7+jQ
2Ua3VU5T3V26J1E5ftiyONGg5OFlPqOL2ogknvMoL/eZl+O8Cqy1Wfe+qO8eztpCRmuenECAgmvhT9
bDNyqKdrluK7ewvmUuxN4VGO6B699x5DnEzcYfIGQ52za6FBeOLDupO/Y05N Cluster: Example
```

The ***Role: Master*** SSH key:

```
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAABAQDjBDngEFCMgWkKhFQpR9aqclT68XErY0KXSXSxsWHP4bWde7
9CEZbOPxo1KWzs/eBrme6Bp+SkCXWedsRFq8Sh6/Qhh3KHFrvY4sj6nx+oztCcxRfLXNRruaZ0Q037
a2MIfTyksbA2LLceU0dTjCAH683AN7OlNg5w4HEPopUlX3oGnouIareiKQQNOVBMApJj2LWq3q5ABM
yFufl22jKPDxAguZES0RhMGJh4Q0hcIY6hqGjq2U5ei2gJ9faeZlm/JROUpTNfCvB/2tFXom7+HvPj
z5owZ4UOJVXzEDUV97VJXsoHn4f6GeXUSLfwgXr81WPuwBYmST0SFeZ17hwV Role: Master
```

The ***Role: Slave*** SSH key:

```
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAABAQDmUg4Ndt44xL7VQuR/sTHxYqDjYgfPCtspw2iPjjBY+gAJlF
1aZfcdIUUJMSpXfpx8tt4HhxKSYYbt2fdx972eeHhyAH1D8LRZRtHVJTLyvrc8ndh/1V7JOIUAdg9P
FTpQyr8He+w3/c4H6axlF7n0qhSdqLQNh4KSOmeOiwINm4sisfilRx5RUYFeTTQNh+Q7SIE9BsIoLf
/8O+tmGT0bzdBf5hY9R9i7akkW5M1K0r0po2l7UVBNXDZtKUcWfckzv89DfWhVLrYnmRUYD+6PmcVI
5xLHiQDRz6c2BAh0GxDiK/2SWe8OpI2Q0tQZwBi9p/7H6GMCxuY5G8nahMfF Role: Slave
```

Appendix B: Flight Compute provisioning script

Configure your IBM SoftLayer environment with the following URL to set up the Alces Flight Compute Solo provisioning script:

https://s3-eu-west-1.amazonaws.com/alces-flight-softlayer/provisioning/compute-solo.sh