
stackspace-docs Documentation

Release 0.1.1

B23 LLC

January 11, 2017

1	Create AWS User Creds	3
1.1	Background	3
1.2	Creating an IAM User in your AWS Account and Saving Credentials:	3
1.3	Attaching Amazon Security Policies to the User	5
1.4	Start Launching Stacks!	6
2	Launch a Space	7
2.1	First: Logon to Stackspace	7
2.2	Second: Launch a Space (AWS Virtual Private Cloud)	8
2.3	Spin it up, Shut it down	8
2.4	Problems, Questions, Concerns, Feedback?	9
3	Launch a Stack	11
3.1	R-IDE Stack	11
3.2	Zeppelin Stack	13
3.3	Spark Stack	15
3.4	Sparklyr (R on Spark) Stack	17
3.5	ELK Stack	20
3.6	H2O Stack	22
3.7	Ambari Stack	24
3.8	Hadoop and Yarn Stack	26
3.9	Kafka Stack	28
3.10	Streamsets Stack	30
3.11	Nifi Stack	32
4	Using Stackspace Stacks	35
5	How To: Connect R with H2O in the Cloud Using Stackspace	37
5.1	Option 1: Quick Start for H2O and R Studio both on Stackspace	37
5.2	Data Ingest	37
5.3	Option 2: Quick Start for H2O on Stackspace and R on your Local Machine	38
5.4	Extra notes about H2O in Stackspace	38
6	EasyIngest	41
6.1	EasyIngest with Stackspace	41
7	EasyAnalyze	43
7.1	How to use EasyAnalyze	43

8	FAQ	45
8.1	General Stackspace Questions	45
8.2	Stackspace Pricing Questions	46
8.3	Spaces	46
8.4	Stacks	47
8.5	EasyIngest	48

Launch your AWS infrastructure at <https://beta.stackspace.io/>



Create AWS User Creds

1.1 Background

Stackspace goes beyond the AWS Security Best Practices in our deployments. We lock down Amazon Web Services as much as possible while allowing you to use all the Stackspace capabilities. In the following user guide, we will walk you through creating a user, getting the user credentials and giving the user Full Amazon EC2, Full Amazon IAM, and Full Amazon S3 permissions. These permissions are needed to successfully spin up your AWS resources, create your stack, and ingest and egress data. If you need help with custom AWS policies, we are glad to walk you through creating them as well. When you are completed with this process, you will be able to enter the AWS Credentials in the **NEW SPACE** Page and create your **SPACE**.

1.2 Creating an IAM User in your AWS Account and Saving Credentials:

1. Sign in to the AWS Management Console and open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the navigation pane, click **[Users]** and then click **[Create New Users]**. (If you are a new AWS user you may need to click **[Get Started]** first.)
3. Enter the user name for the user you want to create – we used “Stackuser” for this guide.
4. Select “Generate an access key for each user” radio button. Click **[Create]**. (See Figure 1)
5. A page appears that enables you to display and download the Credentials for the new user. To save the access keys for the new user or users, click **[Download Credentials]**. (See Figure 2)
6. Keep these keys in a safe place. We will use AWS ACCESS KEY and AWS SECRET KEY.

The screenshot shows the 'Create User' wizard in the AWS IAM console. The top navigation bar includes the AWS logo, 'AWS', 'Services', 'S3', 'EC2', 'VPC', 'Edit', a search bar, 'Global', and 'Support'. The left sidebar has a 'Create User' link. The main content area is titled 'Enter User Names:'. It contains five numbered input fields. The first field, labeled '1.', contains the text 'Stackuser'. Below the fields is the text 'Maximum 64 characters each'. A checkbox labeled 'Generate an access key for each user' is checked and highlighted with a red box. Below this checkbox is explanatory text: 'Users need access keys to make secure REST or Query protocol requests to AWS service APIs. For users who need access to the AWS Management Console, create a password in the Users panel after completing this wizard.' At the bottom right, there are 'Cancel' and 'Create' buttons, with the 'Create' button highlighted by a red box.

Figure 1

The screenshot shows the second step of the 'Create User' wizard. The top navigation bar is the same as in Figure 1. The left sidebar has a 'Create User' link. The main content area displays a success message: 'Your 1 User(s) have been created successfully. This is the last time these User security credentials will be available for download. You can manage and recreate these credentials any time.' Below this message is a link 'Hide User Security Credentials'. A yellow box contains the user details for 'Stackuser': 'Access Key ID:' followed by a masked value and 'Secret Access Key:' followed by a masked value. At the bottom right, there are 'Close' and 'Download Credentials' buttons, with the 'Download Credentials' button highlighted by a red box.

Figure 2

1.3 Attaching Amazon Security Policies to the User

1. Sign in to the AWS Management Console and open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the navigation pane, click **[Policies]**.
3. In the search box type **[ec2]**
4. In the list of policies select **[AmazonEC2FullAccess]**
5. Click **[Policy Actions]**, then click **[Attach]**. (See Figure 3)
6. Select the User you created and click **[Attach Policy]**. (See Figure 4)
7. Repeat steps 3-6, although this time in the search box type **[s3]**, then select **[AmazonS3FullAccess]** and attach that policy.
8. Repeat steps 3-6, in the search box type **[iam]**, then select **[IAMFullAccess]** and attach that policy.

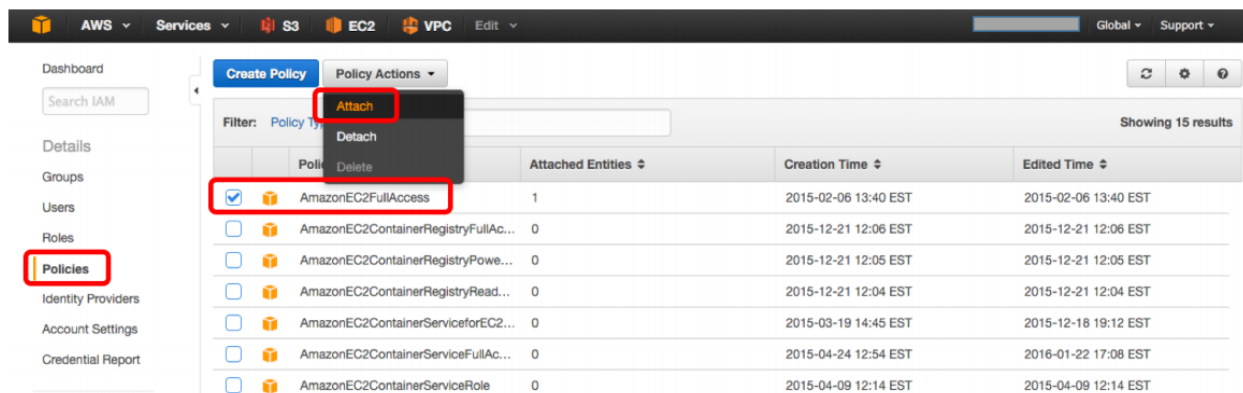


Figure 3

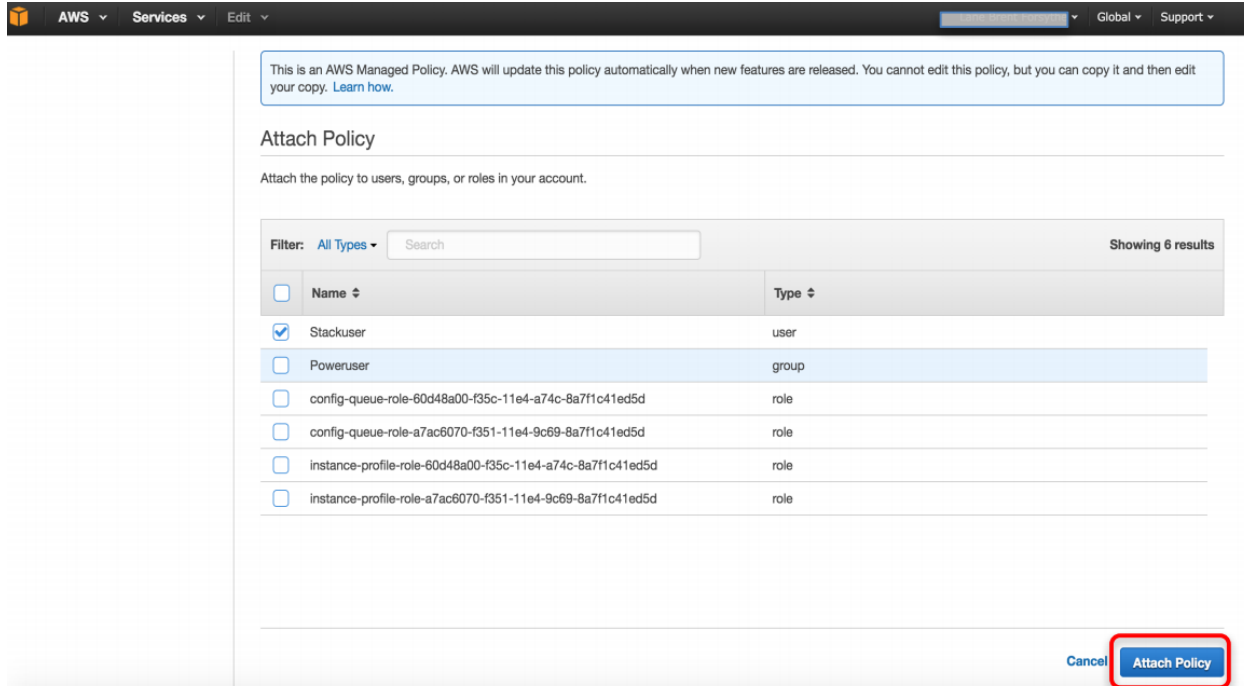


Figure 4

1.4 Start Launching Stacks!

You are now ready to take your Access Keys and use Stackspace to launch your stacks! If you have questions, need support, or find a bug in using Stackspace please contact us via the Stackspace Slack channel and someone will contact you as soon as possible.

Launch a Space

2.1 First: Logon to Stackspace

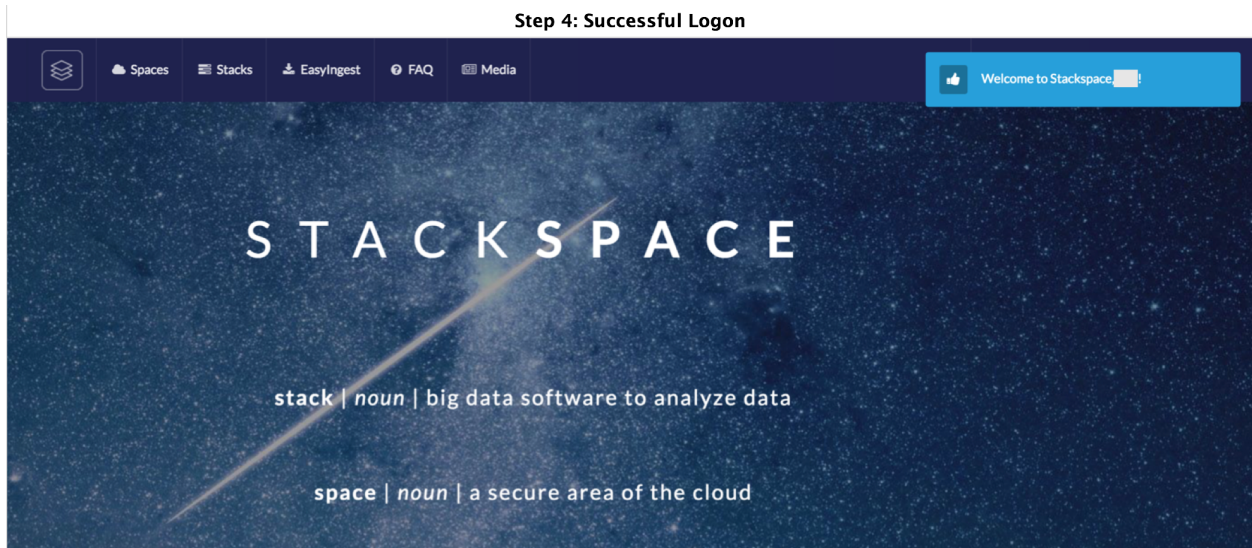
Currently Stackspace leverages the power and security of OAuth via social and collaboration tools like Facebook, Github, Google, Slack, and Twitter.

1. Ensure that you are logged into one of your social media/collaboration accounts.

Step 1: Logon via Social Media



2. Select the provider you would like to use for OAuth authentication into Stackspace.
3. Click on any 'Authorization' Icons that may appear to allow you to login.
4. Once you are logged in you will see a "Welcome to Stackspace, <Your Name>" flag in the upper right corner of the browser.



2.2 Second: Launch a Space (AWS Virtual Private Cloud)

Currently the only cloud provider offered by Stackspace is AWS. You will need to create an AWS ‘Space’ before you can launch a ‘Stack’.

To Launch a Space:

1. Navigate to the Spaces tab in Stackspace and hit the green ‘New Space’ button.
2. In the input form, give your space a name - it can be anything as long as it only contains letters and numbers.
3. Enter your AWS credentials (Access Key and Secret Access Key).
4. Select the AWS region you would like to launch in.
5. To ensure the Security Group IP address is populated, make sure ad blockers, such as adblock plus or ublock origin, are turned off in your browser.
6. Hit the blue ‘Create’ button and wait for your space to reach ‘Running’ state (usually takes 4-5 minutes).
7. To see the ‘Status’ of your Space look in the box under the Stack name. The ‘Status’ icon should be teal. The ‘Status’ icon will show ‘running’ and turn green when the Space is ready.
8. Your Space is a locked down AWS Virtual Private Cloud with a Bastion Host and Real-time monitoring capability (click on the orange tab to access).

2.3 Spin it up, Shut it down

Stackspace is designed to make Big Data automation fast and easy so you can terminate a cluster as soon as you’re done with it and pay AWS no more than you need to. Remember however that once a stack is terminated, all the work you produced inside is gone, so be sure to download any files or workspace objects you wish to save. Once you’re done working, hit the ‘Terminate’ button on your space, wait for it to die, and log off. You must terminate all stacks before attempting to terminate the space.

2.4 Problems, Questions, Concerns, Feedback?

Stackspace would love to hear from you. Email us at [info AT stackspace DOT io](mailto:info@stackspace.io)

Launch a Stack

3.1 R-IDE Stack

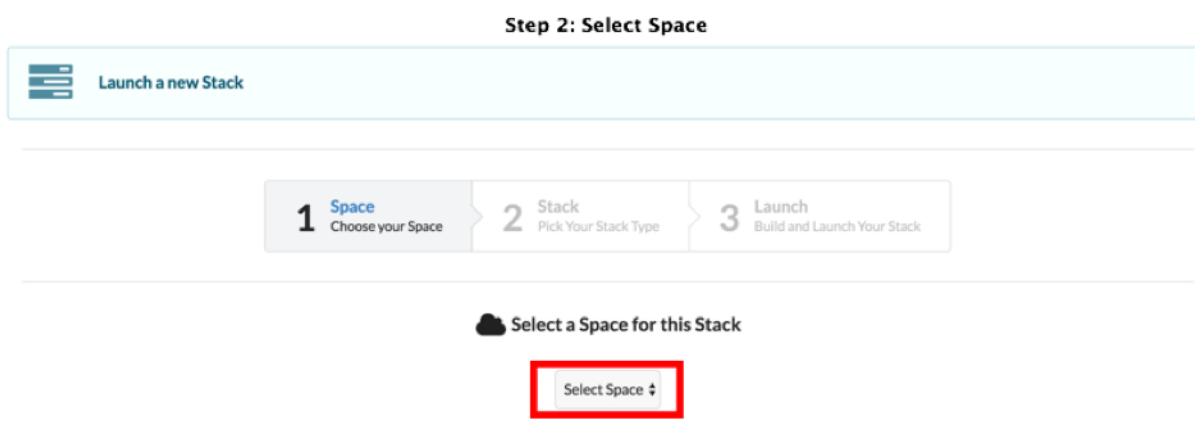
3.1.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.1.2 Second: Launch an R-IDE (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just R IDE, so feel free to explore any and all of the stack offerings. To Launch a Stack:

1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button.
2. Open the 'Select Space' dropdown and choose the space you just launched.



3. You will now see all the stacks available on Stackspace. The R IDE is located under the 'Analysis and Visualization' column. Find it and click on the card.

Stack Name

Number of rstudio nodes

Select an instance type

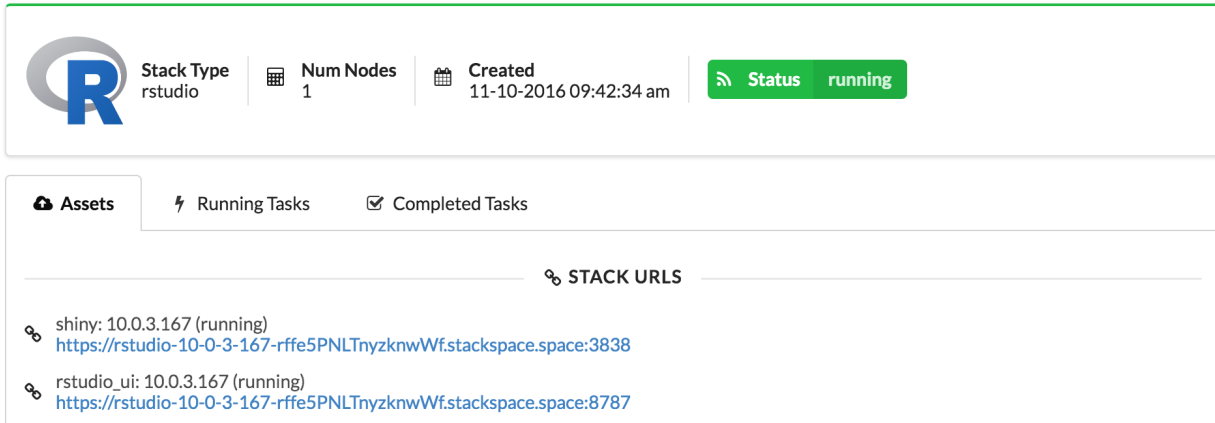
Select a volume size (between 8GB and 500GB)?

4. Name the R-IDE stack anything you want (same rules as space naming apply).
5. Select the instance type you would like R IDE to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
6. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
7. Enable any Optional Configuration settings needed for enhancing stack functionality. View instructions for all optional configurations in the R stack [here_](#).
8. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.1.3 Third: Use the R-IDE Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.



The screenshot shows the Stackspace interface for a running stack. At the top, there's a header bar with the following information:

- Stack Type:** rstudio
- Num Nodes:** 1
- Created:** 11-10-2016 09:42:34 am
- Status:** running (indicated by a green button)

Below the header, there are tabs for **Assets**, **Running Tasks**, and **Completed Tasks**. Under the **Running Tasks** tab, there's a section titled **STACK URLS** with two entries:

- shiny: 10.0.3.167 (running)
<https://rstudio-10-0-3-167-rffe5PNLTnyzknwWf.stackspace.space:3838>
- rstudio_ui: 10.0.3.167 (running)
<https://rstudio-10-0-3-167-rffe5PNLTnyzknwWf.stackspace.space:8787>

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.2 Zeppelin Stack

3.2.1 First: Make sure that you have launched a Space.

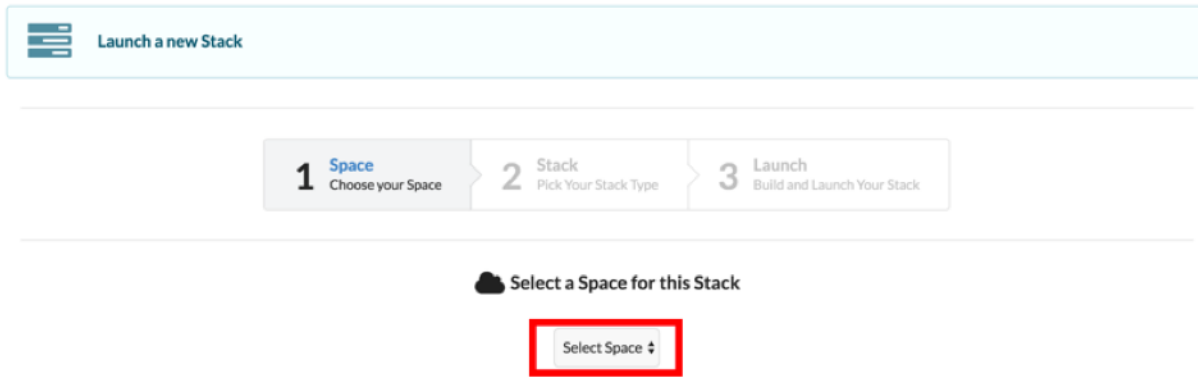
1. Please see the Space launch documentation.

3.2.2 Second: Launch a Zeppelin (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Apache Zeppelin (Zeppelin), so feel free to explore any and all of the stack offerings. To Launch a Stack:

1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button.
2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space



The screenshot shows the 'Step 2: Select Space' interface. At the top, there's a button labeled 'Launch a new Stack'. Below this, there's a progress bar with three steps:

- 1 Space**: Choose your Space
- 2 Stack**: Pick Your Stack Type
- 3 Launch**: Build and Launch Your Stack

Below the progress bar, there's a section titled 'Select a Space for this Stack' with a dropdown menu labeled 'Select Space'.

3. You will now see all the stacks available on Stackspace. Zeppelin is located under the ‘Analysis and Visualization’ column. Find it and click on the card.

Steps 4–9: Launch Zeppelin

Stack Name

Number of master nodes

Number of worker nodes

Select an instance type

Select a volume size (between 8GB and 500GB)?

Optional Configuration Fields

Name of S3 bucket you have access to persist notebooks - current must be in us-east1 region (example, my-s3-notebooks)

Launch

4. Name the Zeppelin stack anything you want (same rules as space naming apply).
5. Select the number of worker nodes for your Zeppelin Stack. The number of workers may vary with the amount of data required, or the complexity of the algorithms within the Zeppelin notebooks.
6. Select the instance type you would like Zeppelin to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you’re unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of ‘8GB’ works for most small projects, but you may need to alter this configuration depending on your needs.
8. You have the option to set the S3 bucket in which to import/export your Zeppelin notebooks. The S3 bucket must be in the same region as your Space. For this example the S3 bucket must be in us-east-1, aka US Standard.
9. Hit the blue ‘Launch’ button and wait for your stack to reach ‘running’ state (usually takes 5-10 minutes).
10. You can monitor the status of your Stack launch by clicking on the ‘Current Task’ and ‘Completed Tasks’ icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you’ve registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.2.3 Third: Use your Zeppelin Stack

1. When the Stack launch is complete the ‘Status’ icon will change to green and switch from ‘launching’ to ‘running’.
2. Click on the ‘Assets’ tab.

- The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs

Stack Type zeppelin | Num Nodes 4 | Created 09-26-2016 11:46:56 am | Status running

Assets | Current Tasks | Completed Tasks

STACK URLs

- zeppelin: 10.0.5.220 (running)
<https://worker-10-0-5-220.FduvkmQG2aDaczfb9.stackspace.space:8080>
- spark_history: 10.0.5.75 (running)
<https://master-10-0-5-75.FduvkmQG2aDaczfb9.stackspace.space:18080>
- spark_master: 10.0.5.75 (running)
<https://master-10-0-5-75.FduvkmQG2aDaczfb9.stackspace.space:8080>
- hadoop_yarn: 10.0.5.75 (running)
<https://master-10-0-5-75.FduvkmQG2aDaczfb9.stackspace.space:8088>
- hadoop_namenode: 10.0.5.75 (running)
<https://master-10-0-5-75.FduvkmQG2aDaczfb9.stackspace.space:50070>

- If a login and password are required, the default login is 'admin' and password 'admin'.

3.3 Spark Stack

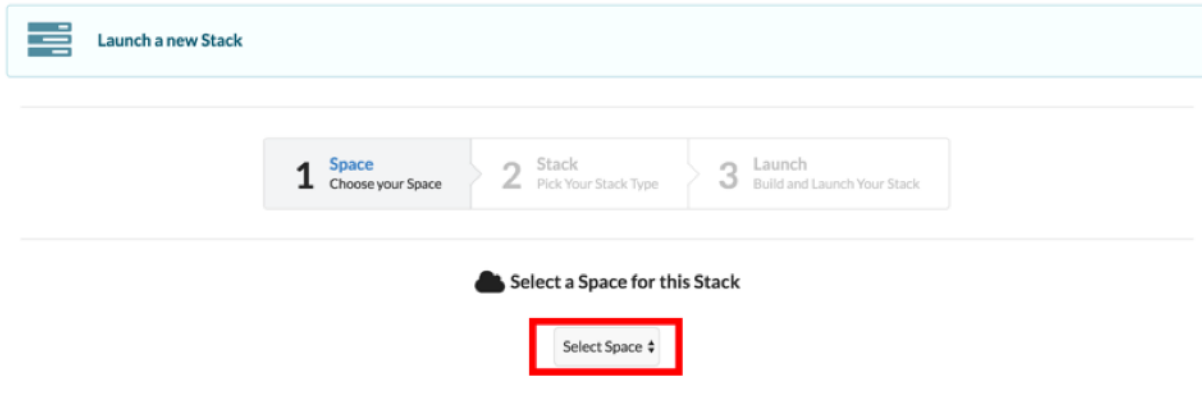
3.3.1 First: Make sure that you have launched a Space.

- Please see the Space launch documentation.

3.3.2 Second: Launch Spark (Stack)

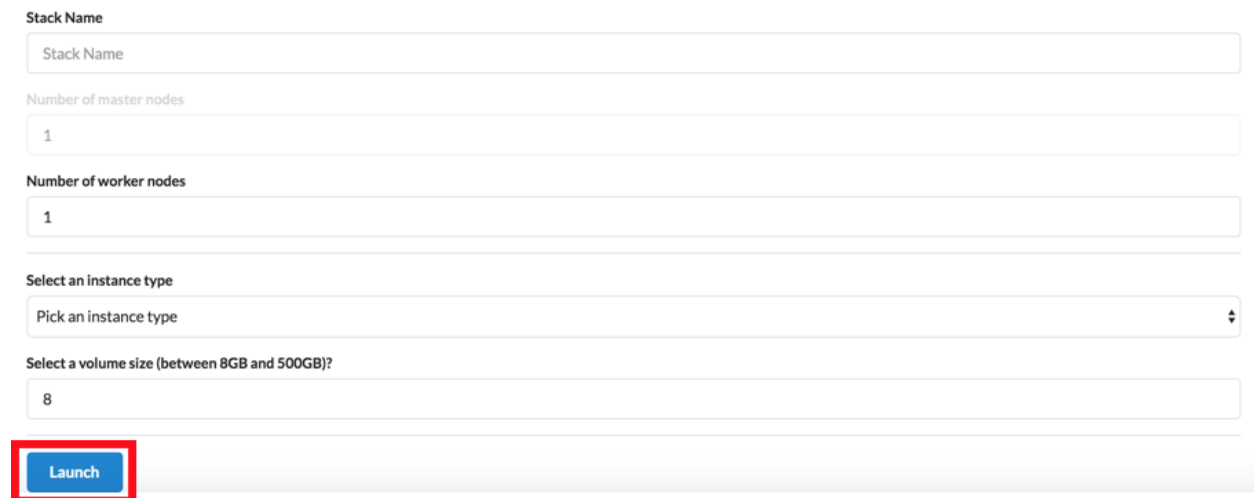
Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Apache Spark (Spark), so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space



3. You will now see all the stacks available on Stackpace. Spark is located under the ‘Processing and Access’ column. Find it and click on the card.

Steps 4–8: Launch Spark



4. Name the Spark stack anything you want (same rules as space naming apply).
5. Enter the number of worker nodes. This number may vary depending on the size of your data to process and the complexity of the Spark applications to run.
6. Select the instance type you would like Spark to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you’re unhappy with the configuration. Stackpace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of ‘8GB’ works for most small projects, but you may need to alter this configuration depending on your needs.
8. Hit the blue ‘Launch’ button and wait for your stack to reach ‘running’ state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the ‘Current Task’ and ‘Completed Tasks’ icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackpace representative may reach out to you if you’ve registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackpace and they will do their best to fix the issue as soon as possible.

3.3.3 Third: Use your Spark Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs

The screenshot shows the Stackspace interface for a Spark stack. At the top, the stack is identified as 'Spark' with a star icon. It shows 'Stack Type: spark', 'Num Nodes: 3', and 'Created: 09-26-2016 10:06:20 am'. A green status bar indicates the stack is 'Status: running'. Below this, the 'Assets' tab is selected, showing 'Current Tasks' and 'Completed Tasks'. The 'STACK URLs' section lists four URLs for different components: spark_master, hadoop_yarn, hadoop_namenode, and spark_history. The first URL, <https://master-10-0-5-176.gPBh5sDynoyPWehgQ.stackspace.space:8080>, is highlighted with a red box.

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.4 Sparklyr (R on Spark) Stack

3.4.1 First: Make sure that you have launched a Space.

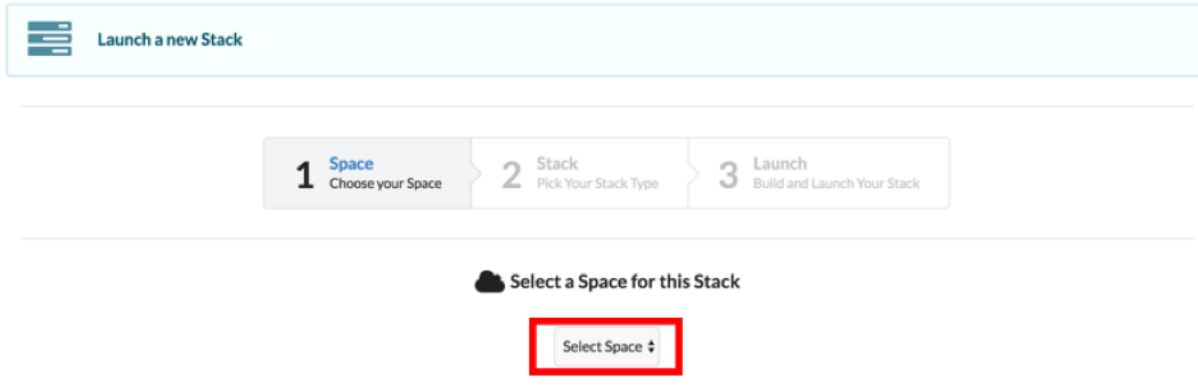
1. Please see the Space launch documentation.

3.4.2 Second: Launch Sparklyr (Stack)

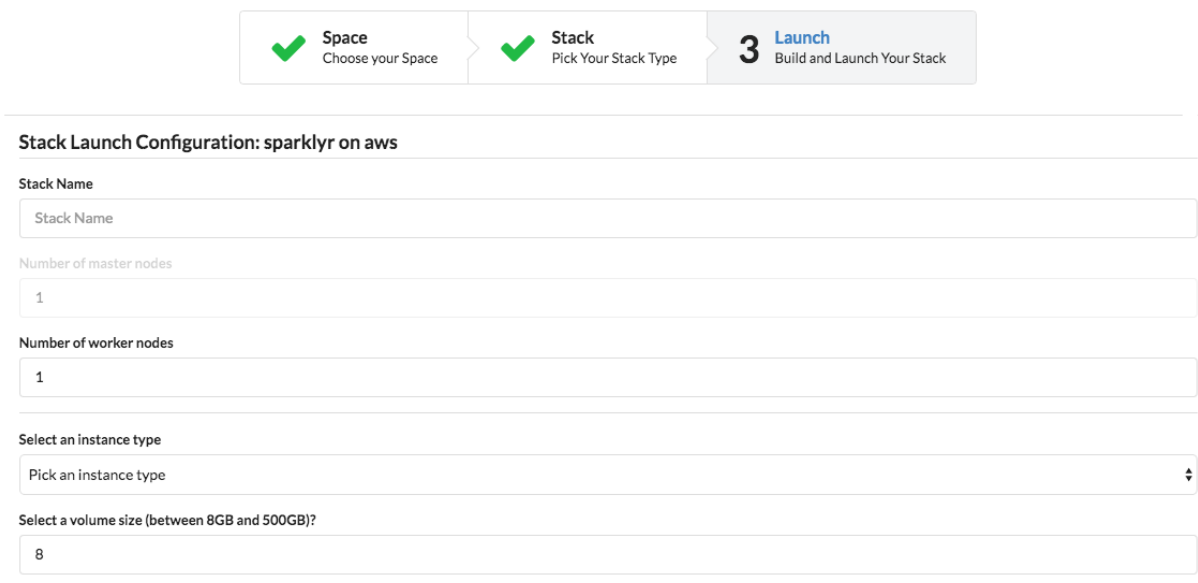
Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just R on Spark, so feel free to explore any and all of the stack offerings. To Launch a Stack:

1. Navigate to the Stacks tab in Stackspace and hit the green ‘New Stack’ button.
2. Open the ‘Select Space’ dropdown and choose the space you just launched.

Step 2: Select Space



3. You will now see all the stacks available on Stackspace. Sparklyr is located under the ‘Processing & Access’ column. Find it and click on the card.



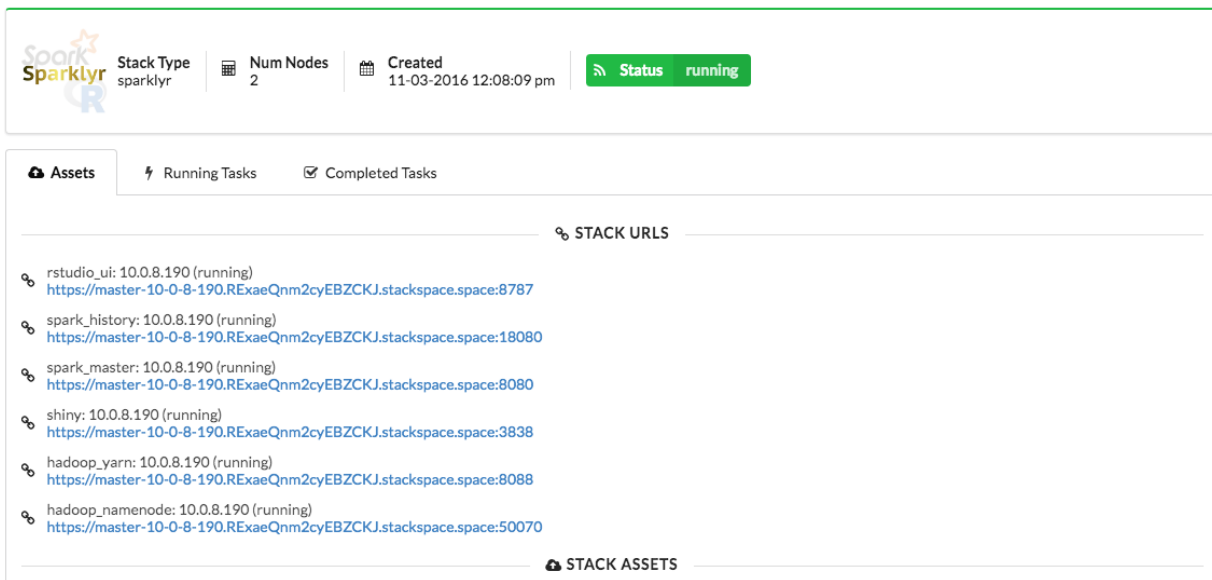
4. Name the Sparklyr stack anything you want (same rules as space naming apply).
5. Select the instance type you would like Sparklyr to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you’re unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
6. Set the EBS volume size (similar story to the instance type). Default of ‘8GB’ works for most small projects, but you may need to alter this configuration depending on your needs.
7. Enable any Optional Configuration settings needed for enhancing stack functionality. View instructions for all optional configurations in the Sparklyr stack [here](#).
8. Hit the blue ‘Launch’ button and wait for the stack to reach ‘running’ state (usually takes 17-20 minutes).

9. You can monitor the status of your Stack launch by clicking on the ‘Current Task’ and ‘Completed Tasks’ icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you’ve registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.4.3 Third: Use your Sparklyr Stack

1. When the Stack launch is complete the ‘Status’ icon will change to green and switch from ‘launching’ to ‘running’.
2. Click on the ‘Assets’ tab.
3. The top section within the ‘Assets’ tab is ‘Stack URLs’. You can access the Stack capabilities by clicking on the links embedded in the URLs.

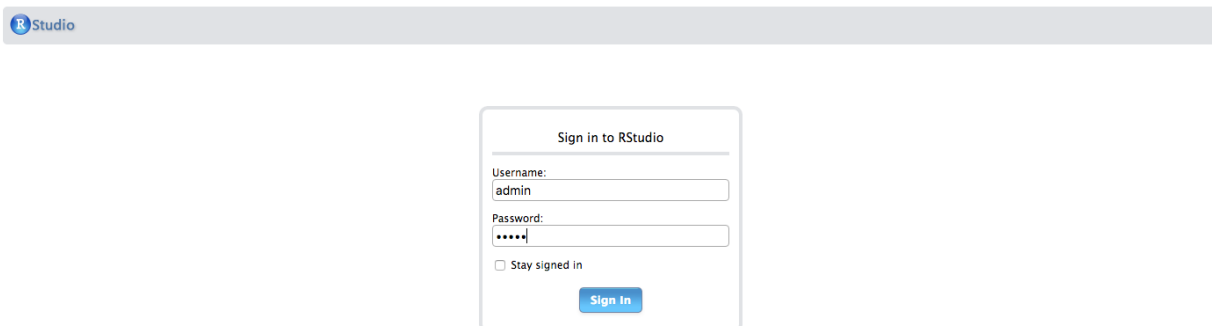


The screenshot shows the Sparklyr Stack interface. At the top, there's a header with the Sparklyr logo, 'Stack Type sparklyr', 'Num Nodes 2', 'Created 11-03-2016 12:08:09 pm', and a green 'Status running' button. Below this is a tabbed interface with 'Assets', 'Running Tasks', and 'Completed Tasks'. The 'Assets' tab is active, showing a section titled 'STACK URLs' with a list of services and their URLs:

- rstudio_ui: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:8787>
- spark_history: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:18080>
- spark_master: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:8080>
- shiny: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:3838>
- hadoop_yarn: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:8088>
- hadoop_namenode: 10.0.8.190 (running) <https://master-10-0-8-190.RExaeQnm2cyEBZCKJ.stackspace.space:50070>

 Below the URLs is a section titled 'STACK ASSETS'.

4. If a login and password are required for access to any asset URL, the default login is ‘admin’ and password ‘admin’.



The screenshot shows the RStudio login form. It has a title 'Sign in to RStudio'. There are two input fields: 'Username:' with the value 'admin' and 'Password:' with masked characters '•••••'. Below these is a checkbox labeled 'Stay signed in' which is unchecked. At the bottom is a blue 'Sign In' button.

3.5 ELK Stack


3.5.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.5.2 Second: Launch ELK (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just ELK, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space

 Launch a new Stack

1 Space
Choose your Space

2 Stack
Pick Your Stack Type

3 Launch
Build and Launch Your Stack

Select a Space for this Stack

Select Space ▾

3. You will now see all the stacks available on Stackspace. ELK is located under the 'Processing and Access' column. Find it and click on the card.

Steps 4-8: Launch ELK

Stack Name

Number of kibana nodes

Number of elasticsearch nodes

Select an instance type

Select a volume size (between 8GB and 500GB)?

Launch

4. Name the ELK stack anything you want (same rules as space naming apply).
5. Enter the number of Elasticsearch nodes desired within the ELK cluster. The number may vary depending on the size of data to ingest within ELK.

6. Select the instance type you would like ELK to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
8. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.5.3 Third: Use your ELK Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs

The screenshot shows the Stackspace interface for an ELK stack. At the top, there's a header with the Elasticsearch logo, 'Stack Type elk', 'Num Nodes 2', 'Created 09-26-2016 11:46:20 am', and a green 'Status running' button. Below this, there are three tabs: 'Assets' (selected), 'Current Tasks', and 'Completed Tasks'. Under the 'Assets' tab, there's a section titled 'STACK URLs'. The first entry is 'kibana: 10.0.3.175 (running)' with a URL 'https://kibana-10-0-3-175.FduvkmQG2aDaczb9.stackspace.space:5601' highlighted by a red box. Other entries include 'elasticsearch_rest: 10.0.4.37 (running)', 'elasticsearch_transport: 10.0.4.37 (running)', 'elasticsearch_marvel: 10.0.4.37 (running)', and 'elasticsearch_head: 10.0.4.37 (running)', each with their respective URLs.

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.6 H2O Stack


3.6.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.6.2 Second: Launch an H2O (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just H2O, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.


Step 2: Select Space

 Launch a new Stack

1 Space
Choose your Space

2 Stack
Pick Your Stack Type

3 Launch
Build and Launch Your Stack

 Select a Space for this Stack

Select Space ▾

3. You will now see all the stacks available on Stackspace. H2O is located under the 'Analysis and Visualization' column. Find it and click on the card.

Steps 4-8: Launch H2O

Stack Name

Number of h2o nodes

Select an instance type

Select a volume size (between 8GB and 500GB)?

Launch

4. Name the H2O stack anything you want (same rules as space naming apply).
5. Enter the number of H2O nodes you would like in your H2O Stack. This may vary depending on the size of your data.

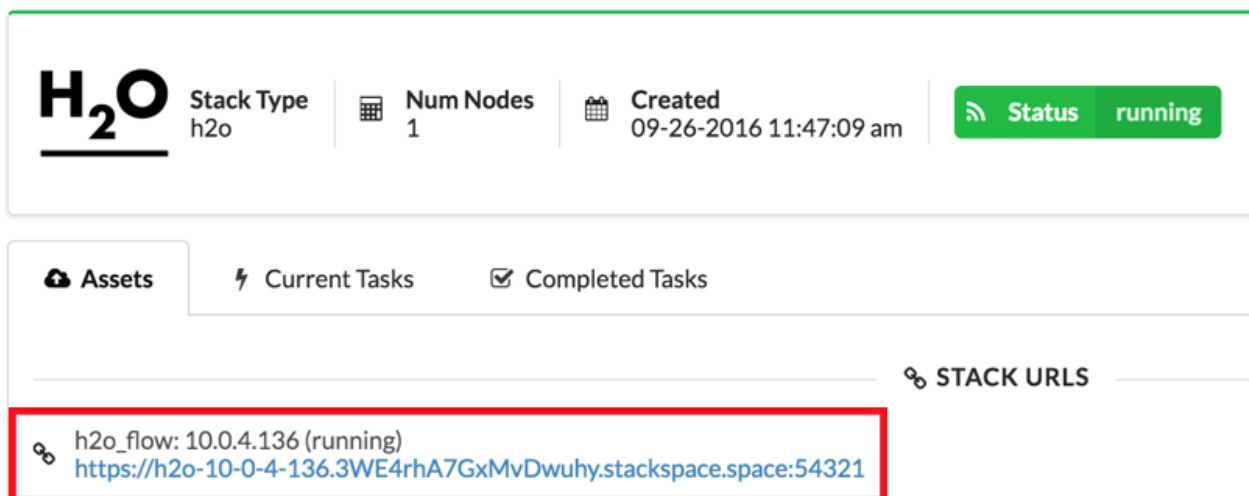
6. Select the instance type you would like H2O to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
8. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.6.3 Third: Use your H2O Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs



The screenshot shows the Stackspace interface for an H2O stack. At the top, there's a header bar with the H2O logo, 'Stack Type h2o', 'Num Nodes 1', 'Created 09-26-2016 11:47:09 am', and a green 'Status running' button. Below this is a tabbed interface with 'Assets', 'Current Tasks', and 'Completed Tasks'. The 'Assets' tab is selected. Under 'Assets', there's a section titled 'STACK URLs'. A red box highlights the first entry: 'h2o_flow: 10.0.4.136 (running)' followed by a blue hyperlink 'https://h2o-10-0-4-136.3WE4rhA7GxMvDwuhy.stackspace.space:54321'.

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.7 Ambari Stack


3.7.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.7.2 Second: Launch Ambari (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Ambari, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.


Step 2: Select Space

 Launch a new Stack

1 Space
Choose your Space

2 Stack
Pick Your Stack Type

3 Launch
Build and Launch Your Stack

 Select a Space for this Stack

Select Space ▾

3. You will now see all the stacks available on Stackspace. Ambari is located under the 'Processing and Access' column. Find it and click on the card.

Steps 4-8: Launch Ambari

Stack Name

Number of ambari_master nodes

Number of hadoop_master nodes

Number of hadoop_slave nodes

Select an instance type

Select a volume size (between 8GB and 500GB)?

Launch

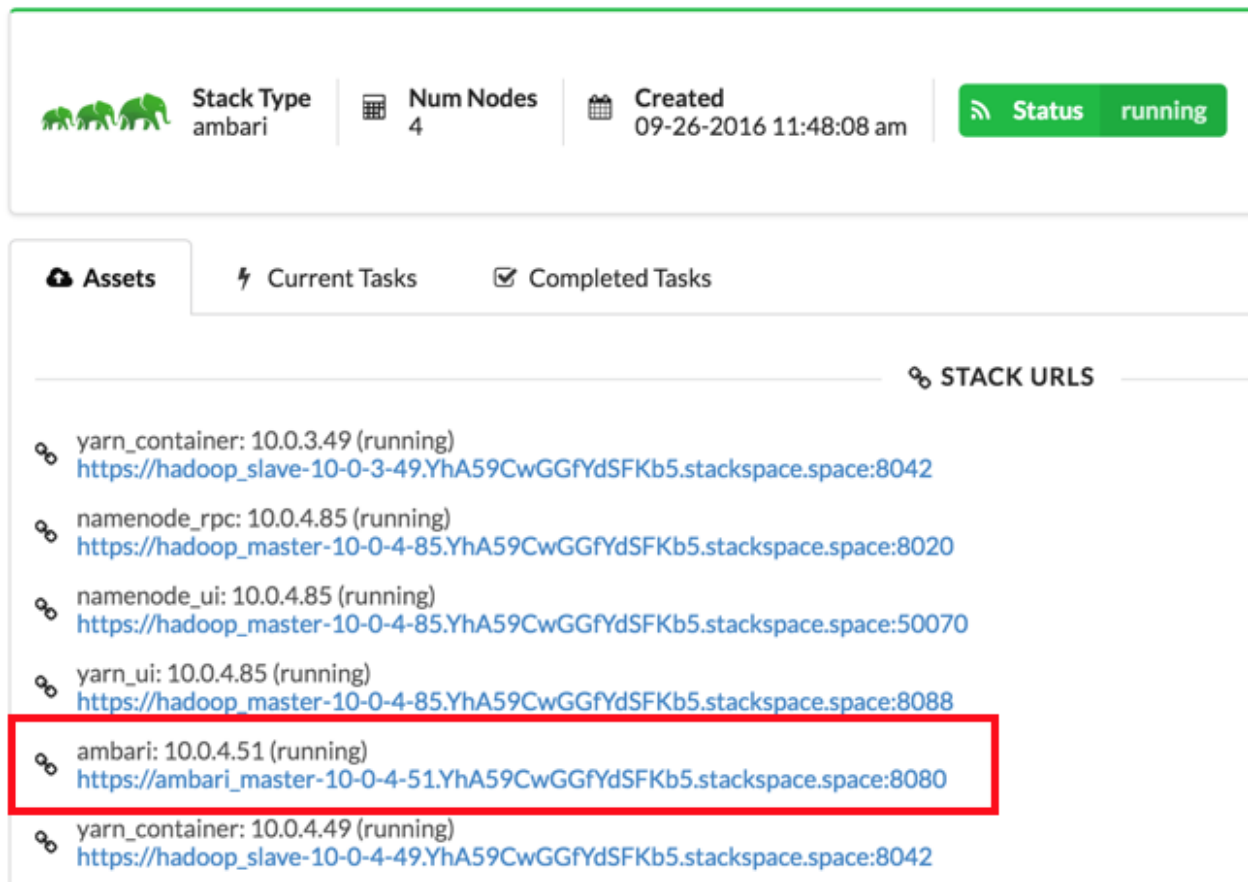
4. Name the Ambari stack anything you want (same rules as space naming apply).

5. Enter the number of Hadoop slave nodes. This will vary depending on the size of your data or the complexity of the jobs to be run within the cluster.
6. Select the instance type you would like Ambari to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
8. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.7.3 Third: Use your Ambari Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs


The screenshot displays the Stackspace interface for a stack named 'ambari'. At the top, there are four tabs: 'Assets', 'Current Tasks', 'Completed Tasks', and 'STACK URLs'. The 'STACK URLs' tab is selected. Below the tabs, a list of stack components and their URLs is shown. The 'ambari' component is highlighted with a red box. The list includes:

- yarn_container: 10.0.3.49 (running) https://hadoop_slave-10-0-3-49.YhA59CwGGfYdSFKb5.stackspace.space:8042
- namenode_rpc: 10.0.4.85 (running) https://hadoop_master-10-0-4-85.YhA59CwGGfYdSFKb5.stackspace.space:8020
- namenode_ui: 10.0.4.85 (running) https://hadoop_master-10-0-4-85.YhA59CwGGfYdSFKb5.stackspace.space:50070
- yarn_ui: 10.0.4.85 (running) https://hadoop_master-10-0-4-85.YhA59CwGGfYdSFKb5.stackspace.space:8088
- ambari: 10.0.4.51 (running) https://ambari_master-10-0-4-51.YhA59CwGGfYdSFKb5.stackspace.space:8080**
- yarn_container: 10.0.4.49 (running) https://hadoop_slave-10-0-4-49.YhA59CwGGfYdSFKb5.stackspace.space:8042

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.8 Hadoop and Yarn Stack

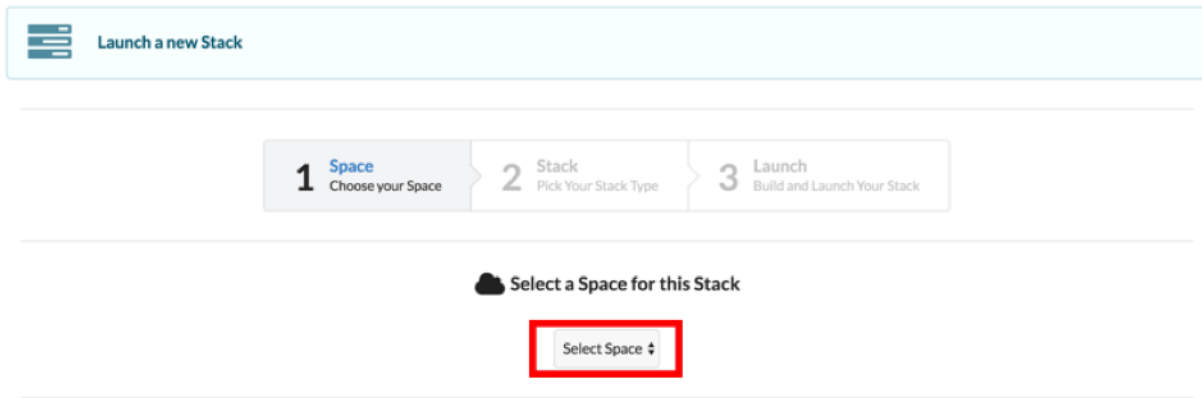
3.8.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.8.2 Second: Launch Hadoop and Yarn (Stack)

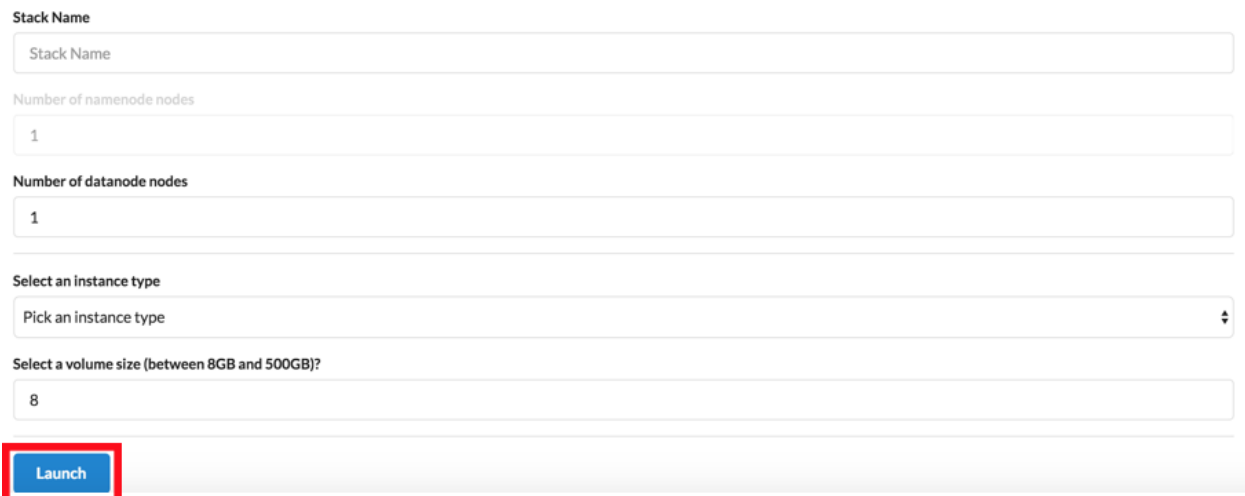
Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Hadoop and Yarn, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space



3. You will now see all the stacks available on Stackspace. The Hadoop and Yarn stack is located under the 'Processing and Access' column. Find it and click on the card.

Steps 4–8: Launch Hadoop and Yarn



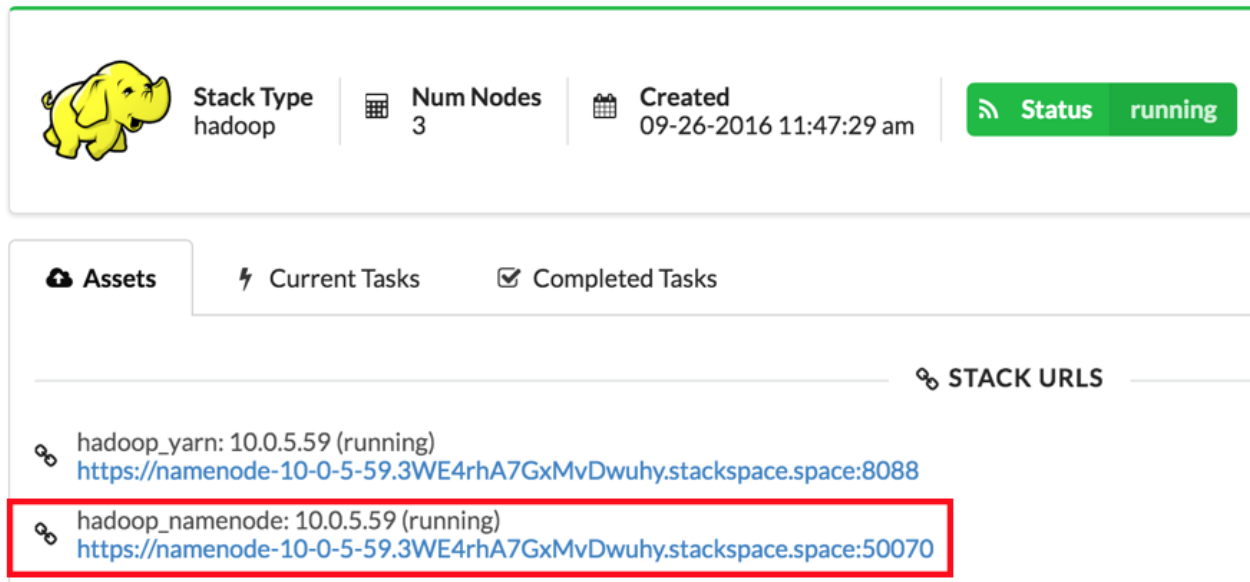
4. Name the Hadoop and Yarn stack anything you want (same rules as space naming apply).
5. Enter the number of desired datanodes. This number may change depending on the size of your data.
6. Select the instance type you would like Hadoop and Yarn to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
7. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
8. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
9. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.8.3 Third: Use your Hadoop and Yarn Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs



The screenshot shows the Stackspace interface for a Hadoop stack. At the top, there's a header with a yellow elephant icon, 'Stack Type: hadoop', 'Num Nodes: 3', 'Created: 09-26-2016 11:47:29 am', and a green 'Status running' button. Below this is a tabbed interface with 'Assets', 'Current Tasks', and 'Completed Tasks'. The 'Assets' tab is selected, and within it, the 'STACK URLs' section is visible. It contains two entries: 'hadoop_yarn: 10.0.5.59 (running)' with a URL 'https://namenode-10-0-5-59.3WE4rhA7GxMvDwuhy.stackspace.space:8088' and 'hadoop_namenode: 10.0.5.59 (running)' with a URL 'https://namenode-10-0-5-59.3WE4rhA7GxMvDwuhy.stackspace.space:50070'. The second entry is highlighted with a red rectangular box.

4. If a login and password are required, the default login is 'admin' and password 'admin'.

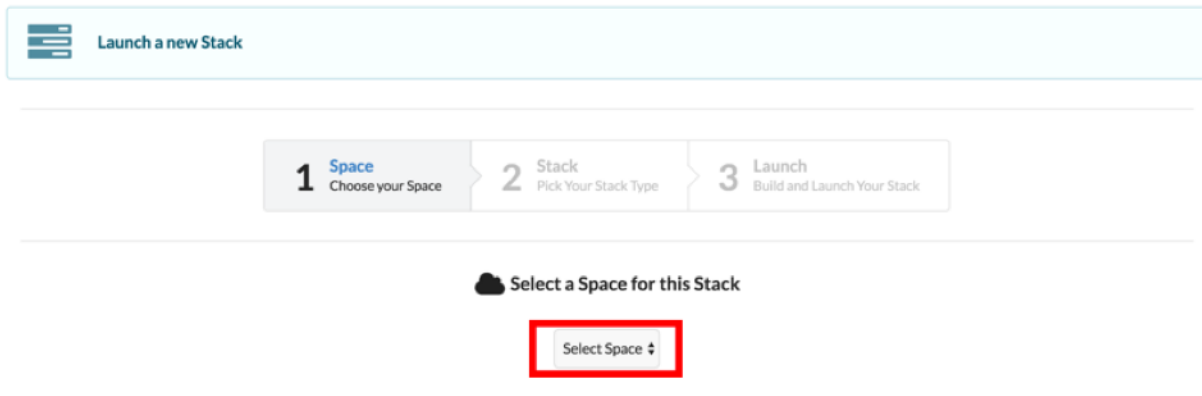
3.9 Kafka Stack

3.9.1 First: Make sure that you have launched a Space.

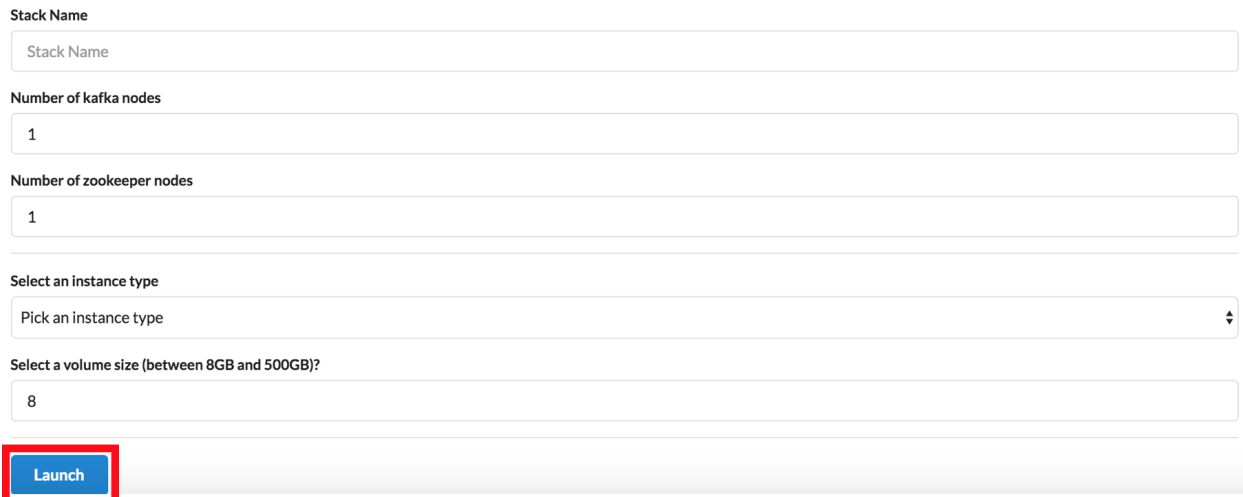
1. Please see the Space launch documentation.

3.9.2 Second: Launch a Kafka (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Kafka, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space


3. You will now see all the stacks available on Stackspace. The Kafka stack is located under the ‘Flow and Ingest’ column. Find it and click on the card.

Steps 4–9: Launch Kafka


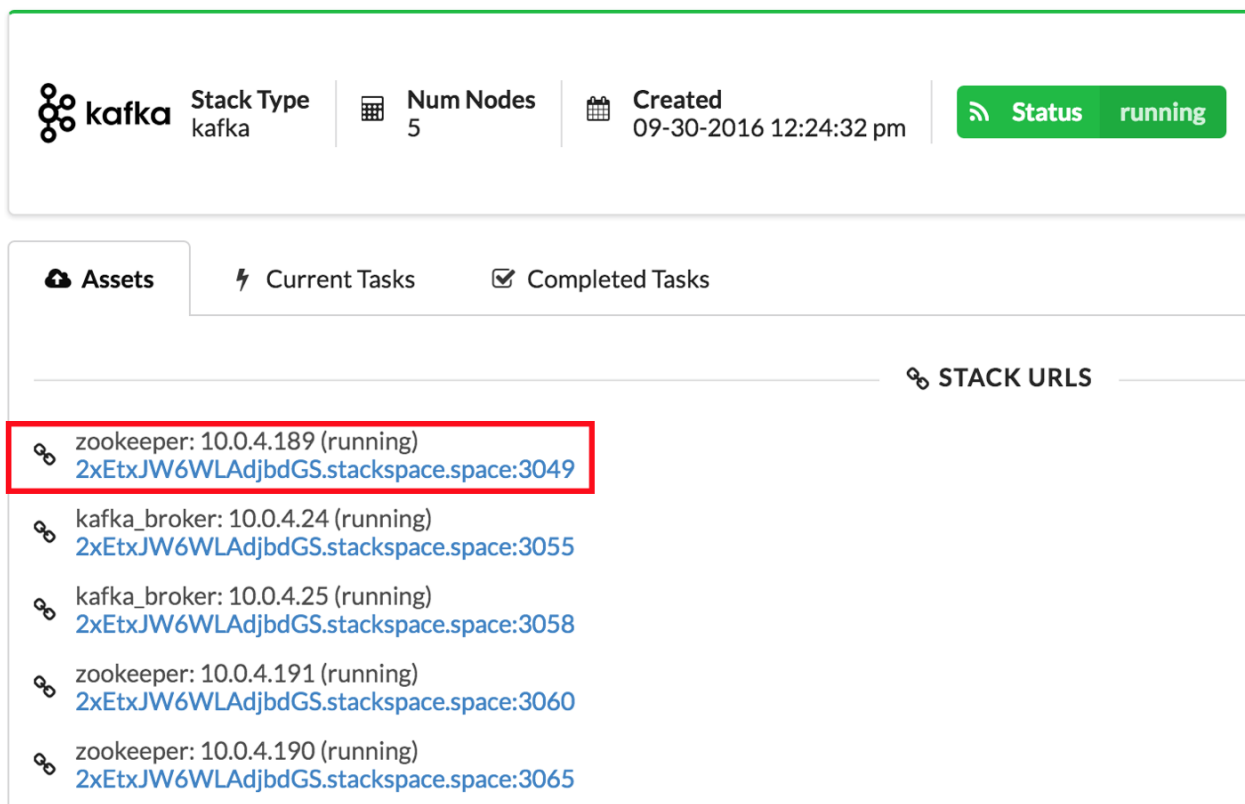
4. Name the Kafka stack anything you want (same rules as space naming apply).
5. Enter the number of desired Kafka nodes.
6. Enter the number of desired Zookeeper nodes. To create a deployment that can tolerate the failure of n machines, you should create $2n+1$ Zookeeper nodes.
7. Select the instance type you would like Kafka and Zookeeper to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you’re unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
8. Set your EBS volume size (similar story to the instance type). Default of ‘8GB’ works for most small projects, but you may need to alter this configuration depending on your needs.
9. Hit the blue ‘Launch’ button and wait for your stack to reach ‘running’ state (usually takes 5-10 minutes).
10. You can monitor the status of your Stack launch by clicking on the ‘Current Task’ and ‘Completed Tasks’ icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.9.3 Third: Use your Kafka Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. Use the links to harness the Stack capabilities.

Step 3: Stack URLs



The screenshot displays the Stackspace interface for a Kafka stack. At the top, the stack is identified as 'kafka' with a 'Stack Type' of 'kafka', 'Num Nodes' of 5, and a 'Created' timestamp of '09-30-2016 12:24:32 pm'. A green 'Status' button indicates the stack is 'running'. Below this, the 'Assets' tab is selected, showing 'Current Tasks' and 'Completed Tasks'. The 'STACK URLs' section lists several service URLs, with the first one, 'zookeeper: 10.0.4.189 (running) 2xEtxJW6WLAdjbdGS.stackspace.space:3049', highlighted by a red box.

Service	IP Address	Status	URL
zookeeper	10.0.4.189	running	2xEtxJW6WLAdjbdGS.stackspace.space:3049
kafka_broker	10.0.4.24	running	2xEtxJW6WLAdjbdGS.stackspace.space:3055
kafka_broker	10.0.4.25	running	2xEtxJW6WLAdjbdGS.stackspace.space:3058
zookeeper	10.0.4.191	running	2xEtxJW6WLAdjbdGS.stackspace.space:3060
zookeeper	10.0.4.190	running	2xEtxJW6WLAdjbdGS.stackspace.space:3065

3.10 Streamsets Stack


3.10.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.10.2 Second: Launch Streamsets (Stack)

Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Streamsets, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.


Step 2: Select Space

 **Launch a new Stack**

1 Space
Choose your Space

2 Stack
Pick Your Stack Type

3 Launch
Build and Launch Your Stack

 **Select a Space for this Stack**

Select Space ▾

- You will now see all the stacks available on Stackspace. The Streamsets stack is located under the 'Flow and Ingest' column. Find it and click on the card.

Steps 4-7: Launch Streamsets

Stack Name

Number of datacollector nodes

Select an instance type
 ▾

Select a volume size (between 8GB and 500GB)?

Launch

- Name the Streamsets stack anything you want (same rules as space naming apply).
- Select the instance type you would like Streamsets to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
- Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
- Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
- You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

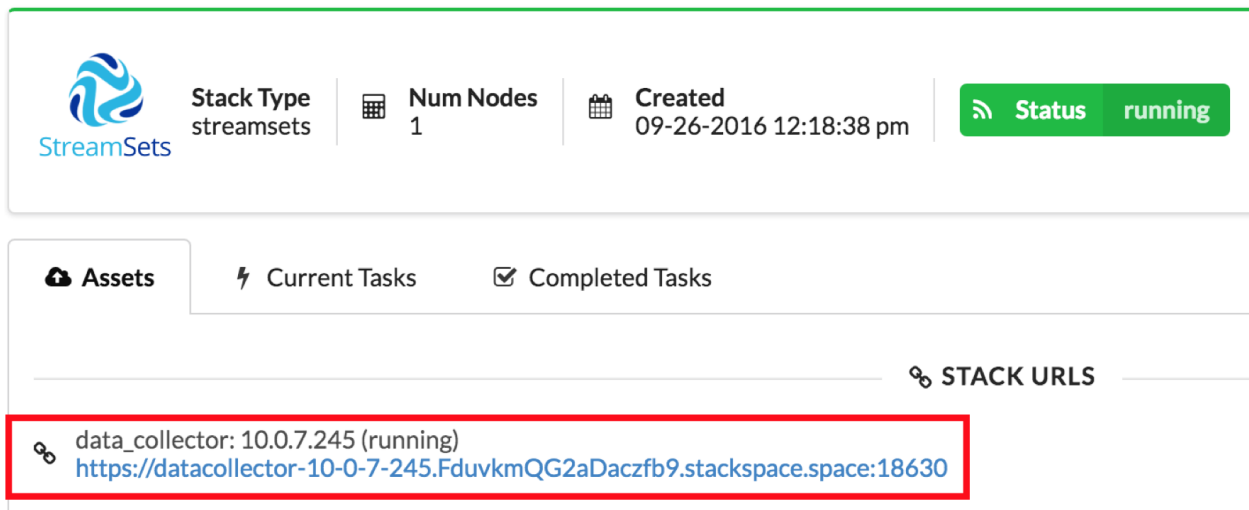
Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach

out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.10.3 Third: Use your Streamsets Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs



The screenshot shows the Stackspace interface. At the top, there's a header bar with the StreamSets logo, 'Stack Type streamsets', 'Num Nodes 1', 'Created 09-26-2016 12:18:38 pm', and a green 'Status running' button. Below this is a tabbed interface with 'Assets', 'Current Tasks', and 'Completed Tasks'. The 'Assets' tab is selected. Underneath, there's a section titled 'STACK URLs' with a red box highlighting a link: 'data_collector: 10.0.7.245 (running) https://datacollector-10-0-7-245.FduvkmQG2aDaczfb9.stackspace.space:18630'.

4. If a login and password are required, the default login is 'admin' and password 'admin'.

3.11 Nifi Stack

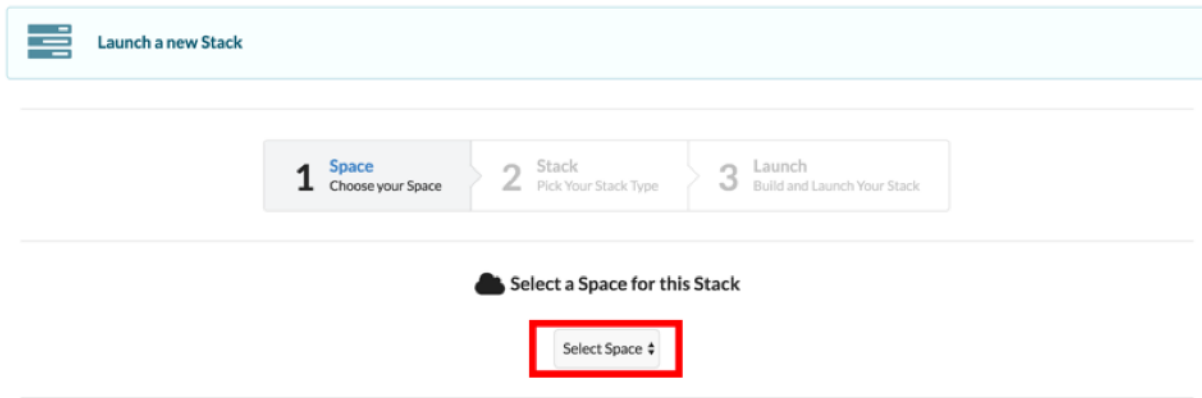
3.11.1 First: Make sure that you have launched a Space.

1. Please see the Space launch documentation.

3.11.2 Second: Launch Nifi (Stack)

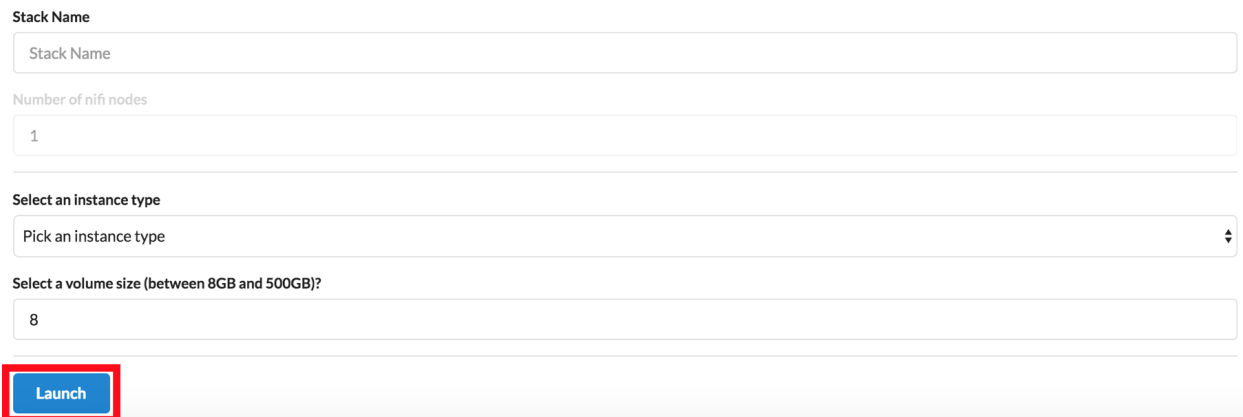
Once you have a running space, you're all set to launch any number of stacks in your secure Stackspace environment. These steps should work for just about any stack available, not just Nifi, so feel free to explore any and all of the stack offerings. To Launch a Stack: 1. Navigate to the Stacks tab in Stackspace and hit the green 'New Stack' button. 2. Open the 'Select Space' dropdown and choose the space you just launched.

Step 2: Select Space



3. You will now see all the stacks available on Stackspace. The Nifi stack is located under the 'Flow and Ingest' column. Find it and click on the card.

Steps 4-7: Launch Nifi




4. Name the Nifi stack anything you want (same rules as space naming apply).
5. Select the instance type you would like Nifi to run on. Instance type selection does affect the AWS pricing. We recommend using t2.medium and above, but the needs of your project may require (or not require) a certain number of cores and RAM to operate efficiently, so this is something you may need to play around with. Remember that you can always terminate stacks and start new ones if you're unhappy with the configuration. Stackspace costs nothing to use, so feel free to launch and kill stacks until you find what you need.
6. Set your EBS volume size (similar story to the instance type). Default of '8GB' works for most small projects, but you may need to alter this configuration depending on your needs.
7. Hit the blue 'Launch' button and wait for your stack to reach 'running' state (usually takes 5-10 minutes).
8. You can monitor the status of your Stack launch by clicking on the 'Current Task' and 'Completed Tasks' icons.

Warning: Error State Sometimes stacks and spaces end up in an error state instead of running as they should. If this is the case for you, simply terminate the stack or space and try again. A Stackspace representative may reach out to you if you've registered with an email address. If the problem occurs multiple times, feel free to reach out to Stackspace and they will do their best to fix the issue as soon as possible.

3.11.3 Third: Use your Nifi Stack

1. When the Stack launch is complete the 'Status' icon will change to green and switch from 'launching' to 'running'.
2. Click on the 'Assets' tab.
3. The top section within the 'Assets' tab is 'Stack URLs'. You can access the Stack capabilities by clicking on the links embedded in the URLs.

Step 3: Stack URLs




Stack Type nifi	Num Nodes 1	Created 09-26-2016 12:18:59 pm	Status running
--------------------	----------------	-----------------------------------	-------------------

Assets

Current Tasks

Completed Tasks

STACK URLs

 nifi_ui: 10.0.7.203 (running)
<https://nifi-10-0-7-203.3WE4rhA7GxMvDwuhystackspace.space:8080/nifi>

4. If a login and password are required, the default login is 'admin' and password 'admin'.

Using Stackspace Stacks

How To: Connect R with H2O in the Cloud Using Stackpace

5.1 Option 1: Quick Start for H2O and R Studio both on Stackpace

To use and deploy R Studio and H2O in the cloud, follow these instructions:

1. Launch a [Space](#) and spin up a [H2O](#) and [RStudio](#) Stack in Stackpace.
2. Set up R Studio with the correct H2O library package
 - Install and load the ‘h2o’ package through the interface package manager.
 - Run in R:

```
install.packages("h2o")
library(h2o)
```

3. Initialize a connection between R and the H2O cluster
 - Instead of the full address and password set used in the local instructions above, you only need to pass in the private ip portion to the ip field here.
 - Run in R:

```
h2o.init(ip="[PrivateIP]", port=54321, startH2O = FALSE)
```

- If the connection is successful, you should see a success message and stats on your cluster.
- Stackpace is currently running H2O 3.10.0.8 (rel-turing). Access the [H2O R package documentation](#) for this release to learn about all the features available.

5.2 Data Ingest

5.2.1 Sending data to H2O from R

1. Run in R:

```
filepath = normalizePath("path-to-your-data") rData.hex = h2o.uploadFile(path=filepath, destination_f
```

2. Access the data in H2O Flow:
 - In the Flow UI, select Data > List All Frames (same as running the *getFrames* command)
 - You should see your data frame available with options to Inspect, Predict or Build Model.

5.2.2 Checking up on your H2O cloud status

1. Although running `h2o.init(...)` in R will print out the cluster and connection status report whenever you want to see it, this action is not recommended. Running the `h2o.init` command multiple times has been known to cause connection issues.
2. In R, try using `h2o.clusterInfo()` or in H2O Flow, select Admin > Cluster Status (same as running the `getCloud` command).

5.3 Option 2: Quick Start for H2O on Stackspace and R on your Local Machine

If you want to use R locally and H2O in the cloud, follow these instructions.

1. Launch a Space and spin up an [H2O Stack](#) in Stackspace:
 - Open the H2O Flow UI (admin/admin) and take note of the ip address your stack is running on.
 - If you have multiple nodes running you will have access to the H2O Flow UI at a private address for each of those nodes. You just need to open one. The nodes will be linked together in H2O to form a cloud.
2. Setup your local R or Rstudio program with the correct H2O library package
 - Run in R:

```
install.packages("h2o")
library(h2o)
```

- If you have a different version of the H2O R package library it is recommended that you remove or unlink that version from your workspace. The R package and H2O versions must match.
3. Initialize a connection between R and the H2O cluster
 - Within the stackspace environment, locate the private IP address of the H2O stack and the ID of the Space you have that stack running in. Alternatively, find the node URLs for your H2O stack. If you launched with multiple H2O nodes, just choose one to use as your ip connection.
 - Run in R:

```
h2o.init(ip="h2o-[PrivateIP].[StackspaceSpaceID].stackspace.space", port=54321, startH2O = FALSE, ht
```

- If the connection is successful, you should see a success message and stats on your cluster.
4. As long as the cluster stays healthy, you can now use H2O to scale R into the cloud
 - Note that H2O R package components have varied dramatically across releases, so older documentation and help forums may not be helpful.

5.4 Extra notes about H2O in Stackspace

- The smallest recommended stack size instance is the t2-small, but t2-medium stacks seem to be more reliable.
- If the cluster or any of the nodes go down, you'll need to power down, delete, and spin up a new H2O stack. None of your work will be retained.
- The status indicator of the H2O stack or nodes in the Stackspace dashboard once launched, will not accurately reflect the health of the H2O cluster. You will need to rely on the checkup methods listed above and not Stackpace for that information.

EasyIngest

6.1 EasyIngest with Stackspace

EasyIngest is ready to go in Stackspace, and its use is straight forward and simple to use. We are still working on the documentation although, so if you would like to use EasyIngest and have questions about how to use it, please email us at info@stackspace.io.

EasyAnalyze

7.1 How to use EasyAnalyze

EasyAnalyze is currently under construction. We do have a beta version of EasyAnalyze for R if you would like to try it out. If you have questions about EasyAnalyze email us at info@stackspace.io.

8.1 General Stackspace Questions

8.1.1 What is Stackspace?

Stackspace is the fastest and easiest way to launch your choice of big data tools to solve the largest number of data analytics challenges.

8.1.2 Who thought of Stackspace?

Stackspace was designed and built by data scientists for data scientists.

8.1.3 What can I do with Stackspace?

In a few mouse-clicks and a few minutes you can launch, configure, and leverage a variety of big data, analytics, & visualization tools using an intuitive interface to solve you hardest data science challenges, build big data applications, create full data pipelines, or IoT collection and analysis platforms.

8.1.4 Do I still have to wait for requested tools if I use Stackspace?

No more waiting for systems to be manually built and configured. Stackspace automates it for you.

8.1.5 How long do things take using Stackspace?

Using Stackspace, data can be securely ingested and analyzed automatically and in minutes.

8.1.6 Can I launch in the Cloud?

Don't just launch in the cloud, but launch securely and transparently, retaining control of your data all at the same time.

8.1.7 What can I expect to happen to my organization if it uses Stackspace?

You will free your data science and analytical teams from worrying about infrastructure and systems, and facilitate choice, experimentation, and results using Stackspace.

8.1.8 How do I communicate with Stackspace if I need support or help?

Please email us at info@stackspace.io if you have any questions, suggestions, or recommendations. If you would like to set up a more consistent and real-time line of communication we can set up a custom Slack channel for you.

8.1.9 What and where are your Terms and Conditions?

Please visit our terms and conditions page: <https://beta.stackspace.io/terms>

8.2 Stackspace Pricing Questions

8.2.1 How much extra does it cost to use Stackspace?

Zero. There is no additional fee to use Stackspace.

8.2.2 What about my Cloud resources?

Stackspace deploys to your existing Cloud accounts, so you pay for the Cloud resources the exact same way you would as if you were using the Cloud without Stackspace.

8.3 Spaces

8.3.1 What is a Space?

We call a Space, the “Stackspace Forcefield” and it composed of many capabilities that you should take for granted, but most IT Teams spend years creating and Stackspace does automatically in minutes. At its base a Space is your own secure software defined cloud within your infrastructure provider; defined with route tables, subnets, and firewalls already set up using best practices and codified lessons learned from hundreds of customers deployed on the cloud. Stackspace builds and manages your Space for you so you don’t have to waste hours or days doing it yourself. A Space has four functions, it acts as a DMZ to common cyberthreats, Enterprise Monitoring node, Data Flow node, and IP Traffic Log.

8.3.2 What do I need to create a Space?

Once you are logged in, you only need your AWS Credentials and to make up a name for your Space. Your AWS Credentials are your AWS Access Key ID (20 characters) and your Secret Access Key (40 characters).

8.3.3 How do I get my AWS Credentials?

At Stackspace we default to implementing the most secure possible security policies that will allow you to use all the capabilities of Stackspace. Please see the first option on the left, Creating AWS User Creds, which will walk you through the complete process.

8.3.4 When will I be able to build a Space in other cloud providers?

We anticipate adding more cloud providers to Stackspace in late 2016.

8.3.5 How can I monitor the health of my Space and Stack?

Stackspace has by default integrated Enterprise level monitoring. To access the monitoring UI simply click on the red “Real-Time Monitoring” icon in the Space Status box at the top of the Space screen. Through the full-featured interactive dashboard you can monitor every node in your Space and all Stacks within that Space.

8.4 Stacks

8.4.1 What is a Stack?

Once you have created your Space, you can create a Stack. A Stack is your choice of Big Data software tools and analytics. With a few clicks your choice of Big Data tools can be provisioned, SW installed, and configured and ready for use in minutes. For detailed information regarding Stack types see below.

8.4.2 What Stacks are currently available?

At this time we provide Stack types that will allow you to build a full scale comprehensive data pipeline from data ingest and storage to data analysis. The software Stacks include: RServer & Shiny, SparklyR, Apache Zeppelin, Apache Spark running on Yarn, Apache Hadoop with Yarn, Apache Ambari, Kafka, Streamsets, Nifi, Elasticsearch, and H2O.

8.4.3 What Stacks will you be making available?

The potential is limitless...we have a prioritized list based on customer requests and launch a new stack about once a month and recently launched a SparklyR within 2-days of its launch! If you have a preference for which Stacks you would like us to include in Stackspace, please let us know over your Stackspace Slack Channel or email us at: info@stackspace.io.

8.5 EasyIngest

8.5.1 What is EasyIngest?

EasyIngest is an automated way to take your data from its source and ingest it into your Stack so it is ready to analyze in minutes. No more waiting on your infrastructure personnel or data engineers to “prep” data for you. Stackspace EasyIngest can do it for you. For detailed information regarding custom ingest, please email us at info@stackspace.io.