



---

# Database-as-Service

---

*Last generated: February 21, 2018*

© 2018 Splice Machine, Inc. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.

# Table of Contents

## Database-as-Service Product

- » [Welcome!](#) ..... 0
- » [About our DBaaS](#) ..... 0
- » [Service Overview](#) ..... 0
- » [User Interface Overview](#)..... 0
- » [Release Notes](#)..... 0

## Cloud Manager Guide

- » [Introduction](#)..... 0
- » [Registering with Cloud Manager](#) ..... 0
- » [Logging into the Cloud Manager](#) ..... 0
- » [Creating Your Cluster](#) ..... 0
- » [Exploring Your Dashboard](#) ..... 0
- » [Managing a Cluster](#) ..... 0
- » [Managing Your Account](#) ..... 0
- » [Reviewing Event Notifications](#)..... 0

## Using Zeppelin

- » [Introduction](#)..... 0
- » [Getting Started with Zeppelin](#) ..... 0
- » [Zeppelin Usage Notes](#)..... 0
- » [A Simple Tutorial](#) ..... 0

## Splice Machine Home

- » [Splice Machine Home Page](#)..... 1
- » [Splice Machine Product Page](#) ..... 1
- » [Splice Machine Solutions](#) ..... 1
- » [Splice Machine Support](#) ..... 1

- » Splice Machine Resources..... 1
- » Splice Machine Community..... 1

# Welcome to the Splice Machine Database Service!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation for our [Managed Database Service in the Cloud](#), which includes Release 2.7 of the Splice Machine Database.

## Getting Started With our Database Service

Getting started with our database is as simple as can be; just follow these steps, and you can be up and running in less than an hour:

- 1**

**LOG INTO  
SPLICE  
MACHINE**  
**(page 19)**

→ Log in directly, or use your Google or Amazon ID.
- 2**

**CREATE  
DATABASE  
CLUSTER**  
**(page 20)**

→ Adjust 4 sliders for your processing and storage needs; your database cluster is ready within 15 minutes.
- 3**

**LOAD YOUR  
DATA (page**  
**52)**

→ Copy data to S3, then perform a fast import. Time required varies with dataset size. Our [Zeppelin Simple Example \(page 52\)](#) provides a quick example.
- 4**

**QUERY AND  
UPDATE  
YOUR  
DATABASE**  
**(page 44)**

→ Use Zeppelin notebooks to quickly update, query, and display results graphically, without coding.

## Next Steps

Easy next steps you can take to become more proficient with your new database system:

- » Our [About the Splice Machine Database Service \(page 5\)](#) topic introduces this edition of Splice Machine and links to main documentation pages related to the service.
- » Spend some time learning more about [creating and using Zeppelin notebooks \(page 44\)](#), which you can use to prepare and run SQL DDL and DML, stored procedures, Java, Scala, and Python and Spark-SQL programs with Splice Machine data, all without writing code.
- » Spend a few minutes with our [Cloud Manager Interface \(page 13\)](#), which you can use to modify your cluster configuration, administer your account, set up events, and review database usage.
- » Check this documentation web for best practices, usage tips, developer guides, and reference material

# About the Splice Machine Database Service

With the *Splice Machine Cloud Manager*, configuring a new cluster is as easy as using a few sliders to set compute units for OLTP and OLAP processing, allocate storage, and schedule backup frequency and retention. Splice Machine does the rest. You can seamlessly scale out from gigabytes to petabytes of data when needs or data volumes change, and the same configurator adds or subtracts resources dynamically. You pay only for what you use. You can then:

- » Power your applications on a scale-out, ANSI SQL database
- » Power apps with simultaneous OLAP & OLTP workloads
- » Ingest millions of records and process thousands of transactions in nanoseconds
- » Elastically scale resources as needed
- » We've got you covered – availability, backups, monitoring and alerts

## Service Configuration

You only need to understand a few key concepts to configure your service:

- » A **Splice Unit** is a measure of processing work; one unit currently translates (approximately) to 2 virtual CPUs and 16 GB of memory. When you provision a new Splice Machine cluster, you can select the number of Splice Units you want to use for OLAP and OLTP workloads. The minimum number of Splice Units required for your cluster changes when you update the amount of data you want to access or the amount processing power you want to use.
- » The space allocated for your **Internal Dataset**, which is data that you're storing within your database. Note that as this size increases, the number of Splice Units required (especially OLTP Splice Units) can also increase.
- » The space allocated for your **External Dataset**, which is data stored externally that you can access from your database using features such as external tables and VTI. Note that as this size increases, the number of OLAP Splice Units required can also increase.

Use our [Database-Service documentation \(page 3\)](#) to quickly walk through provisioning your database cluster, loading your data, and querying your data in notebooks, all in less than an hour. Once you're ready, our documentation offers:

## Available Tools

In addition to easy connectivity with almost any Business Intelligence tool, Splice Machine includes:

- » An integrated *Zeppelin Notebook* interface. Zeppelin notebooks are like text documents, but with code that can be executed and of which the output can be rendered as tables, reports and beautiful graphs. This enables you to prepare and run SQL DDL and DML, stored procedures, Java, Scala, and Python and Spark-SQL programs with Splice Machine data. Splice Machine comes pre-configured with a set of notebooks to get started, load data and see examples of the work that can be done with the RDBMS.
- » Our JDBC and ODBC drivers allow you to connect third-party business intelligence tools to your database.
- » You can also take advantage of machine learning, streaming, and other services that you can access from our predesigned notebooks, your own notebooks, or code written by your developers.

## Learn More

Our documentation provides:

- » Complete descriptions of our [Cloud Manager dashboard \(page 13\)](#)
- » Numerous [Tutorials \(page 0\)](#) about connecting with other tools, using various programming languages with Splice Machine, ingesting data efficiently, and so on.
- » An introduction to [Using Zeppelin with Splice Machine \(page 43\)](#)
- » A wealth of [Developer's Guide information \(page 0\)](#) and our [SQL Reference Manual \(page 0\)](#)

You can visit our company web site to learn about what our [Cloud-Managed Database-as-Service \(DBaaS\)](#) (<https://www.splicemachine.com/get-started/>) can do for your company.



# Service Overview

Our Database Service is a subscription-based service, hosted in the Cloud. We take care of managing your cluster services, and you can focus on working with our scalable, dual-engine database.

**This is a DB-Service-Only topic!** [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

## Service Availability

Splice Machine’s target Service availability commitment is 99.9% per calendar month, excluding scheduled downtime. You can expect the following:

- » Splice Machine will deliver product updates with minimal, scheduled downtime.
- » Splice Machine can recover your database from a stored backup after receiving your request to do so.
- » Splice Machine can resize your cluster with minimal downtime.

## Support for Your Service

Splice Machine provides two support options, as shown in the following table:

Support Type	Pricing	Support Feature	Description	Details
Standard Support	Free	Coverage Hours	Monday-Friday 9am-6pm Pacific time	(subject to local holidays)
		System Impaired	Significant issues with speed, quality, or funtionality of Service.	< 12 business hours

Support Type	Pricing	Support Feature	Description	Details
		Other Issues	General queries and guidance requests	< 24 business hours
<b>Business Support</b>	As per contract	Coverage Hours	24 hours a day, 7 days a week, 365 days a year	
	Includes SLA	Production System Down	Complete loss of Service on Production cluster.	< 1 hour
		Production System Impaired	Significant issues with speed, quality, or functionality of Service on Production cluster.	< 4 hours
		Production System Down	Significant issues with speed, quality, or functionality of Service on non-production cluster.	< 12 hours
		Other Issues	General queries and guidance requests	< 24 business hours

## Service Level Agreement (SLA)

Our *business support agreement* includes a *Service Level Agreement (SLA)* that specifies our commitment to a target Service availability of 99.9% per calendar year, excluding scheduled downtimes.

## Service Terms

Subscription fees are payable monthly in advance on the 1st of the month, pro-rated for any partial months. We'll charge your credit card or withdraw payment by ACH on the first of each month, until your service is cancelled. See your license agreement for more details.

# Database Service User Interface

**This is a DB-Service-Only topic!** [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

In addition to our database, the Splice Machine Database Service includes all of the tools you need to create your cluster, load data into your database, query and manipulate your database, and create visual representations of your query results, as described here:

UI Component	Description
Dashboard	The <a href="#">Splice Machine Dashboard (page 13)</a> or <b>Cloud Manager</b> is your entry point to your Database Service. You can register and log into your account here, as well as accessing the other managers described in this table.
Cluster Manager	Use the <a href="#">Cluster Manager (page 20)</a> to create new clusters and to monitor the health of your clusters.
Notebooks Manager	<a href="#">Apache Zeppelin notebooks (page 43)</a> make it easy to query your database and apply various visualizations to the results of your queries. We've created several notebooks that will help you to quickly become productive and to see how easy it is to create your own notebooks.
Database Console	The <a href="#">Database Console (page 0)</a> is a browser-based tool that you can use to monitor database queries on your cluster in real time. The Console UI allows you to see the Spark queries that are currently running in Splice Machine on your cluster, and to then drill down into each job to see the current progress of the queries, and to identify any potential bottlenecks. If you see something amiss, you can also terminate a query.
Events Manager	Our <b>Events Manager</b> allows you to examine events that have occurred on your cluster.

UI Component	Description
Account Manager	The <a href="#">Splice Machine Account Manager (page 36)</a> is where you manage your users, your profile, and your billing information.

# Release Notes for the Splice Machine Database-as-a-Service Product

**This is a DB-Service-Only topic!** [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This topic includes any release notes that are specific to the Splice Machine *Database-as-Service* product, in these sections:

- » [Features Not Yet Available \(page 11\)](#)
- » [Current Limitations \(page 11\)](#)
- » [Important Notes \(page 12\)](#)

Most of the information about changes in the Splice Machine database that underlies this product are found in the [Splice Machine database \(page 0\)](#) release notes.

## Features Not Yet Available

These features are not yet available, but will be very soon:

- » TLS is not yet enabled for JDBC connections.
- » VPC Settings are not yet enabled but will be in a near future release.
- » You currently cannot cancel queries that are running through Zeppelin or JDBC tools; you can use the Spark User Interface to cancel Spark queries.

## Current Limitations

These limitations exist in this release, and will be removed in the near future:

- » On a JDBC connection, individual queries or actions will time out after one hour; you can run long-running queries within a Zeppelin notebook.
- » Updating of CPU, Memory, and Disk usage graphs for clusters is currently limited: the updates are happening only intermittently.

## Important Notes

These are important notes about issues you need to be aware of when using our Database Service:

- » The timestamps displayed in Zeppelin will be different than the timestamps you see in the Splice Machine Spark User Interface, depending upon your time zone.
- » Although Splice Machine backs up your database regularly, it does not back up your Zeppelin Notebook changes; please export your Notebooks regularly if you make changes.

# Splice Machine Cloud Manager

This guide helps you to get registered with and start using the Splice Machine Cloud Manager




You can initiate a chat session with one of our support specialists by clicking this chat icon, which you'll see on Cloud Manager screens.

Here are the topics included in this guide:

Topic	Description
<a href="#">Navigating Your Dashboard (page 29)</a>	Your <b>Dashboard</b> is the entry point to the Splice Machine Cloud Manager. From here, you can create new clusters, access existing clusters, manage your account, review notifications, update your profile, and log out.
<a href="#">Registering (page 17)</a>	Shows you how to complete the first step of using your database service: registering as a user of the Splice Machine Cloud Manager.
<a href="#">Logging In (page 19)</a>	Shows you how to log into the Splice Machine Cloud Manager once you've registered.
<a href="#">Creating a New Cluster (page 20)</a>	Follow this steps in this topic to quickly become productive with your clustered database. In only a few minutes, you'll have your cluster up and running, and will be able to load and work with your data.

Topic	Description
<a href="#">Loading Your Data (page 0)</a>	<p>You can load your data into Splice Machine from an AWS S3 bucket:</p> <ul style="list-style-type: none"><li>» If you don't yet know how to create an S3 bucket or upload data to a bucket, please check our <a href="#">Uploading Data to an S3 Bucket (page 0)</a> tutorial.</li><li>» You may need to configure IAM permissions to allow Splice Machine to access your bucket; see our <a href="#">Configuring an S3 Bucket for Splice Machine Access (page 0)</a> tutorial.</li><li>» Once you've got your data in a bucket, you can follow our <a href="#">Importing Data Tutorial (page 0)</a> to load that data into Splice Machine.</li></ul> <p>Also note that the S3 directory you specify for the log of record import issues (the <i>bad record</i> directory), must be write-accessible using the same AWS credentials that apply to the input directory, i.e. the bad record directory must allow the same access key/secret key pair.</p>



Topic	Description
<a href="#">Using the DB Console (page 0)</a>	<p>The <b>Splice Machine Database Console</b> is a browser-based tool that you can use to monitor database queries on your cluster in real time. The Console UI allows you to see the Spark queries that are currently running in Splice Machine on your cluster, and to then drill down into each job to see the current progress of the queries, and to identify any potential bottlenecks. If you see something amiss, you can also terminate a query.</p> <p>The DB Console is available for all Splice Machine products; you access the Splice DB Console for the Database-as-Service product by clicking the <b>DB Console</b> link in your Cluster Management dashboard, or by following the link sent to you by Splice Machine when your cluster was originally created.</p> <div data-bbox="602 804 1232 1201" style="border: 1px solid #ccc; border-radius: 10px; padding: 10px; margin: 10px 0;">  <p>The <b>DB Console</b> link in your Cluster Management dashboard allows you to monitor Splice Machine jobs running on your cluster. If you are running non-Splice jobs, you'll need to use a different console to monitor them; you'll find a link to this <b>External Spark Console</b> in the bottom left corner of your cluster dashboard.</p> </div>
<a href="#">Using the Notebooks Manager (page 43)</a>	<p>One of the great features of our database service is the ease with which you can use Apache Zeppelin notebooks to interact with your database. Our <b>Notebooks Manager</b> provides convenient access to your notebooks, and to information about using Zeppelin.</p> <p>This tutorial walks you through using a notebook created by Splice Machine that:</p> <ul style="list-style-type: none"> <li>» Creates a schema and the tables in your database to store the TPC-H-1 benchmarks data.</li> <li>» Loads the data from an S3 bucket into your database.</li> <li>» Runs any or all of the TPC-H-1 queries</li> </ul>

Topic	Description
<a href="#"><u>Managing Your Account (page 36)</u></a>	Use our Account Manager to manage your company profile, billing, and users.
<a href="#"><u>Reviewing Event Notifications. (page 41)</u></a>	Use our Events Manager to review notification messages that have been sent to your account.

# Cloud Manager User Registration

This page describes the Splice Machine Cloud Manager registration page.

The registration screen is straightforward and familiar:

**splice**  
MACHINE

Set up Splice Machine in minutes and start powering your applications on a scale-out, ANSI SQL database.

- + Full ANSI SQL database
- + Power apps with simultaneous OLAP & OLTP workloads
- + Ingest millions of records and process thousands of transactions in seconds
- + Scale up or scale down when you need it
- + We've got you covered - availability, backups, monitoring and alerts

**Let's get started...**

**First Name**  
Enter First Name

**Last Name**  
Enter Last Name

**Email**  
Enter Email

**Confirm Email**  
Re-enter Email

**Password**  
Enter Password

**Confirm Password**  
Re-enter Password

**Register** >

**Or register with...**

**Google**

**Amazon**

You can use your Google or Amazon account information to register, or you can manually register by following these steps:

1. Enter your first and last name.
2. Enter and then confirm the email address you want to use for logging into the Splice Machine Cloud Manager.
3. Enter and then confirm the password you want to use for logging into the Splice Machine Cloud Manager.
4. Click the **Register** button.

Once you've successfully registered, you'll land on the *Create New Cluster* screen; we've created a tutorial to quickly walk you through [Creating a Cluster \(page 20\)](#) topic.

# Logging in to the Splice Machine Cloud Manager

This page describes the Splice Machine Cloud Manager Login page.



Set up Splice Machine in minutes and start powering your applications on a scale-out, ANSI SQL database.

- + Full ANSI SQL database
- + Power apps with simultaneous OLAP & OLTP workloads
- + Ingest millions of records and process thousands of transactions in seconds
- + Scale up or scale down when you need it
- + We've got you covered - availability, backups, monitoring and alerts

Login

Email

Password

Register

Login >

G Login with Google

a Login with Amazon

[Forgot password?](#)

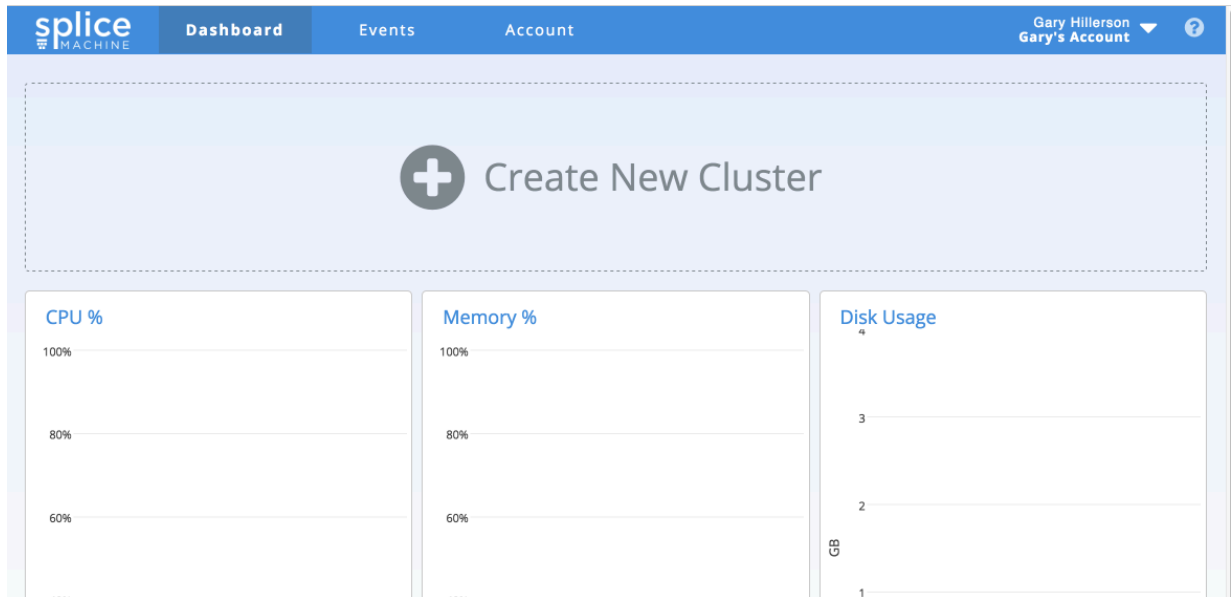
If you've already registered with the Splice Machine Cloud Manager, you can enter your registered email address and password, and click the **Login** button; alternatively, you can log in using your Google or Amazon credentials.

If you've not yet registered with the Splice Machine Cloud Manager, you'll need to register, which will also log you in. Click the **Register** button and fill in the [registration form \(page 17\)](#).

After successfully logging in, you'll land on your dashboard page.

# Creating a New Splice Machine Cluster

When you first visit your new Splice Machine Cloud Manager dashboard, you'll see the initial dashboard view, which prompts you to create a new cluster:



Click **Create New Cluster** to start the process of provisioning your Splice Machine cluster. You'll then need to:

1. [Configure Cluster Parameters \(page 20\)](#) for data sizing, cluster power, and backup frequency.
2. [Configure Cluster Access \(page 25\)](#) for your users.
3. [Set Up Payment \(page 26\)](#) for your Splice Machine cluster.
4. Start Using Splice Machine!

## Configure Cluster Parameters

You use the **Create New Cluster** screen to provision your cluster:

spliceMACHINE

DashboardEventsAccount

Account

Create New Cluster

SpliceDocs1

Region: us-east-1

Data Sizing

Internal Dataset (TB)

0.1 TB

5

10

25

50

100 TB

0.1

External Dataset (TB)

0 TB

50

100

250

500

1 PB

0

Cluster Power

OLTP Splice Units

4

25

50

100

250

500

600

4

OLAP Splice Units

4

25

50

100

250

500

600

4

Backup Frequency

Frequency:

Daily

Start Window (UTC):

02:30

Backups to Keep:

1

Estimated Benchmarks

Single Record Lookup (ms)

20ms

Import Rate (record/sec)

40,000

TPC-C 100 Users (TpmC)

4,000

TPC-H Query #2 (sec)

25

Estimated Costs (Monthly)

Storage

\$85.00

Compute

\$3,558.00

Total

\$3,643.00

\* Does not include monthly I/O costs.

Reset

Back

Next

If you have subscribed to Splice Machine via the AWS Marketplace, your costs will be estimated on an hourly basis instead of a monthly basis:

Estimated Costs (Hourly)

Storage

\$0.09

Compute


\$4.00

Total

\$4.09

\* Does not include monthly I/O costs.

## Screen Help

Many of the components of the [Create Cluster](#) screen, like most of our Cloud Manager screens, include small information buttons  that you can click to display a small pop-up that describes the components.

## About the Cluster Parameters

You'll notice several sliders that you can adjust to modify the configuration of your cluster. As you move these sliders, you'll see how the estimated monthly costs for your cluster change. Here are explanations of the adjustments you can make to your cluster provisioning:

**NOTE:** Note that you can come back and modify your cluster configuration in the future, so you're not stuck forever with your initial settings.

	Cluster Name	Supply whatever name you want for your Splice Machine cluster.
	Region	You can select in which AWS region your cluster will reside by clicking the previously selected region name, which drops down a list of choices.
Data Sizing	Internal Dataset (TB)	Move the slider to modify your estimate of how large your database will be.  Internal Dataset is the amount of data that you will be storing within your Splice Machine database.
	External Dataset (TB)	Move the slider to modify your estimate of how large your external data will be.  External Dataset is the amount of data that you will be accessing from external data sources, using features such as external tables and our virtual table interface.
Cluster Power	OLTP Splice Units	Move the slider to modify your estimate of how much processing power you need for transactional query processing. More OLTP units means more servers in your cluster.



	OLAP Splice Units	Move the slider to modify your estimate of how much processing power you need for analytical query processing. More OLAP units means more executors.
Backup Frequency	Frequency	<p>Select how frequently you want Splice Machine to back up your data. You can select Hourly, Daily, or Weekly; each selection displays additional timing and retention options:</p> <p>Hourly:</p> <div><p><b>Backup Frequency</b></p><p>Frequency: <input type="text" value="Hourly"/> Every: <input type="text" value="1"/> Start Window (UTC): <input type="text" value="02:30"/> Backups to Keep: <input type="text" value="1"/></p><p>Hours</p></div> <p>Daily:</p> <div><p><b>Backup Frequency</b></p><p>Frequency: <input type="text" value="Daily"/> Start Window (UTC): <input type="text" value="02:30"/> Backups to Keep: <input type="text" value="1"/></p></div> <p>Weekly:</p> <div><p><b>Backup Frequency</b></p><p>Frequency: <input type="text" value="Weekly"/> Day of Week: <input type="text" value="Monday"/> Start Window (UTC): <input type="text" value="02:30"/> Backups to Keep: <input type="text" value="1"/></p></div>



A *Splice Unit* is a measure of processing work; one unit currently translates (approximately) to 2 virtual CPUs and 16 GB of memory.

## Modifying Cluster Parameters

We recommend that you spend a few minutes experimenting with modifying the cluster parameters; you'll notice that as you increase various values, the estimated monthly cost of your cluster changes.

When you're satisfied with your cluster configuration parameters, click the **Next** button to set up access to your cluster.

You'll notice that when you increase some values, Splice Machine may indicate that the current setting for a parameter clashes with a change that you've made. For example, in the following image, we have increased the **Internal Dataset** size to 20 TB, and as a result the **Cluster Power** values are no longer adequate to support that large a dataset, as indicated by the striping:

**Create New Cluster**

SpliceDocs1

Region: us-east-1

**Data Sizing**

Internal Dataset (TB): 20.1

External Dataset (TB): 0

**Cluster Power**

OLTP Splice Units: 4

OLAP Splice Units: 4

**Backup Frequency**

Frequency: Daily

Start Window (UTC): 02:30

Backups to Keep: 1

**Estimated Benchmarks**

Single Record Lookup (ms)	20ms
Import Rate (record/sec)	40,000
TPC-C 100 Users (TpmC)	4,000
TPC-H Query #2 (sec)	25

**Estimated Costs (Monthly)**

Storage	\$17,065.00
Compute	\$3,558.00
<b>Total</b>	<b>\$20,623.00</b>

\* Does not include monthly I/O costs.

Reset Back Next >

Splice Machine will not allow you to create your cluster if any of your values clash. You can click the vertical bar at the end of the striping to instantly set the parameter to the required value.

**NOTE:** If you don't correct the required setting and attempt to advance to the **Next** screen, you'll see an error message and will be unable to advance until you do correct it.

## Configure Cluster Access

Once you've configured your cluster, click the **Next** button to display the **Cluster Access** screen. The following image includes displays of the pop-up help information displays for the different access methods:

The screenshot shows the 'Create New Cluster' interface for Splice Machine. The top navigation bar includes 'Dashboard', 'Events', and 'Account'. The user is logged in as 'Gary Hillerson' with 'Gary's Account'.

**Create New Cluster**

**SpliceDocs1**

**VPC Setup**

☐ Client VPC connectivity required

Account ID:

VPC ID (Accepter):

**IAM S3 Access**

☐ S3 connectivity

S3 Access Key:

S3 Secret Key:

**Virtual Private Cloud Configuration**

If applicable, enter your VPC information below to allow the cluster to connect. If you don't have this information ready now, you can always add it later.

**AWS S3 Configuration**

Important: follow [instructions](#) to set up IAM privileges for Splice to read from and write to select S3 folders. If you don't have this information ready now, you can always add it later.

**Configuration Summary**

Internal Dataset (TB)	0.1
External Dataset (TB)	0
OLTP Splice Units	4
OLAP Splice Units	4
Backups	daily
Region	us-east-1

**Splice Machine Authorization Code**

Contact us at [websales@splicemachine.com](mailto:websales@splicemachine.com) to see if you qualify for a Splice Authorization Code.

**Authorization Code**

**Confirmation**

☒ I read and accept the [terms and conditions](#).

\* An email will be sent to you once cluster/database creation is complete. This email will include instructions for how to connect to your database.

You can set your cluster up for access to your Amazon Virtual Private Cloud (VPC) access by selecting the Client VPC connectivity required option and providing your VPC account ID.

You need to configure AWS Identity and Access Management (IAM) for your cluster to allow Splice Machine to access selected S3 folders; this is described in our [Configuring an S3 bucket for Splice Machine Access \(page 0\)](#) tutorial.

For more information about Amazon VPC, see <https://aws.amazon.com/vpc/> (<https://aws.amazon.com/vpc/>).

For more information about Amazon IAM, see <https://aws.amazon.com/iam/> (<https://aws.amazon.com/iam/>).

After setting up any access methods, please confirm that you accept our terms and conditions, then click the **Launch** button, which will take you to the **Payment** screen, unless you’ve subscribed to Splice Machine from the Amazon Marketplace or have already set up a payment method for your account.

## Set Up Payment

When you click the **Launch button**, then one of these actions happens:

- » If you subscribed to Splice Machine via the AWS Marketplace, or you already have a payment method set up on your account, you’ll land on your dashboard and will be notified when your cluster has been initialized.
- » If you don’t yet have a payment method set up, you’ll land on the **Payment** screen, in which you can elect to use on of three payment methods:

Credit Card

ACH

Authorization Code

Name on Card

Credit Card Number

4242 4242 4242 4242

MM / YY

CVC

Cancel

Submit >

Credit Card	<div><div>Credit Card</div><div>ACH</div><div>Authorization Code</div></div> <div><div><div>Name on Card</div><div>My Name</div></div><div><div>Credit Card Number</div><div><div><div>VISA</div></div>4242 5678 9012 3456<div>09 / 17</div><div>999</div><div>99999</div></div></div></div> <div><div>Cancel</div><div>Submit &gt;</div></div>
-------------	---

ACH Electronic Transfer	<div> <a href="#">Credit Card</a> <a href="#">ACH</a> <a href="#">Authorization Code</a> </div> <div> <p>TODO: Implement</p> <div> <a href="#">Cancel</a> <a href="#">Submit &gt;</a> </div> </div>
Authorization Code	<div> <a href="#">Credit Card</a> <a href="#">ACH</a> <a href="#">Authorization Code</a> </div> <div> <p><b>Authorization Code:</b></p> <input type="text" value="3345-8888-9078-ABCF-4321"/> <div> <a href="#">Cancel</a> <a href="#">Submit &gt;</a> </div> </div>

## Modifying Payment Information

If you ever need to change your Splice Machine payment information, you can update it in the **Billing Activity** tab of the **Account** screen; just click the **Update** button to revisit the **Payment** screen:

The screenshot shows the Splice Machine web interface. The top navigation bar includes 'Dashboard', 'Events', and 'Account'. The 'Account' section is active, showing tabs for 'Billing Activity', 'Users', and 'Company Information'. Under 'Billing Activity', the 'Billing Source' section displays a VISA card ending in 4242, expiring 9/2017. A red box highlights the 'Update' button next to the card details. Below this is a 'Recent History' bar chart showing a single data point for April. At the bottom, an 'Invoices' table lists an invoice for 'MyTrialCluster' dated April 4, 2017, with a total of \$2,588.40.

Invoice	Date	Total
MyTrialCluster	April 2017 April 4, 2017	\$2,588.40

**NOTE:** If you've purchased Splice Machine through Amazon Marketplace, change your billing credentials in the Marketplace instead.

## Start Using Your Database!

After your cluster spins up, which typically requires about 10 minutes, you can load your data into your Splice Machine database and start running queries.

The easiest way to get going with your new database is to use our [Zeppelin Notebook interface \(page 43\)](#), with which you can quickly run queries and generate different visualizations of your results, all without writing any code. We've provided a number of useful Zeppelin tutorials, including one that walks you through setting up a schema, creating tables, loading data, and then running queries.

Note that your data must be in an AWS S3 bucket before you can import it into your Splice Machine database:

- » If you don't yet know how to create an S3 bucket or upload data to a bucket, please check our [Uploading Data to an S3 Bucket \(page 0\)](#) tutorial.
- » You may need to configure IAM permissions to allow Splice Machine to access your bucket; see our [Configuring an S3 Bucket for Splice Machine Access \(page 0\)](#) tutorial.
- » Once you've got your data in a bucket, you can follow our [Importing Data Tutorial \(page 0\)](#) to load that data into Splice Machine.

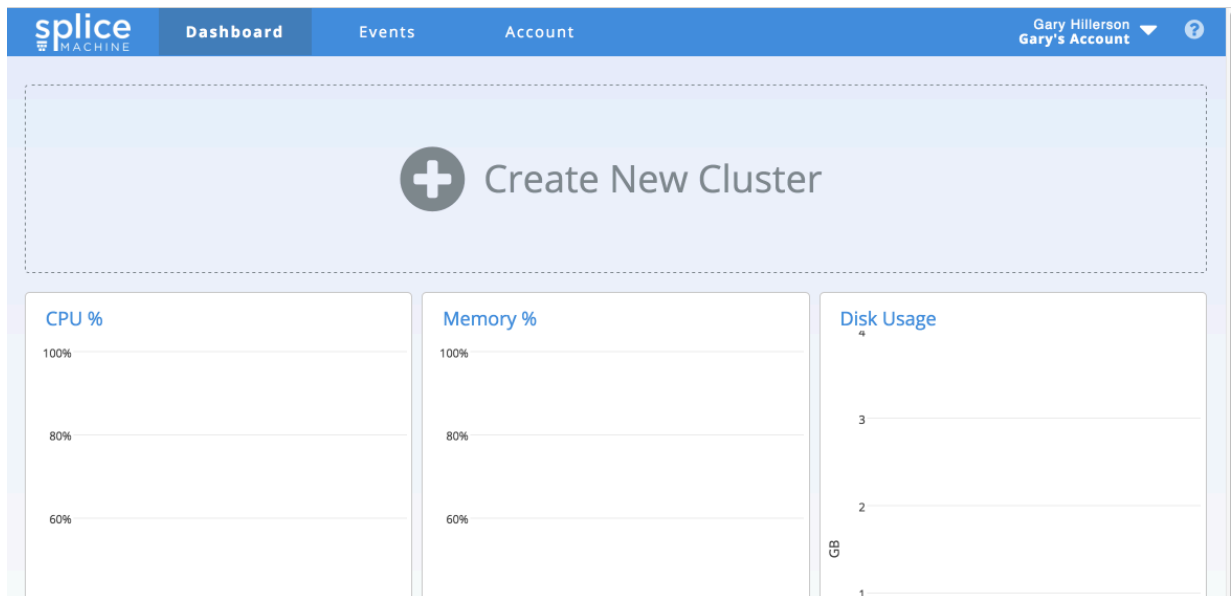
# Exploring Your Cloud Manager Dashboard

This topic describes the actions you can initiate from your Splice Machine dashboard, which include:

- » [Creating a New Cluster \(page 29\)](#)
- » [Viewing and Managing Your Clusters \(page 29\)](#)

## Creating a New Cluster

Click the large **Create New Cluster** button to start the process of creating a new cluster. This process, which requires just a few minutes, is described in detail in our [Creating a New Cluster \(page 20\)](#) topic.



## Viewing and Managing Your Clusters

When you create a new cluster, you land back on this Dashboard screen, which shows you the status of your current clusters. In the following image, we've created our first Splice Machine cluster, which is currently *Initializing*.

spliceMACHINE

DashboardEventsAccount

Gary HillersonGary's Account?

Clusters

RefreshCreate Cluster

Cluster Name	OLTP (Splice units)	OLAP (Splice units)	Storage (TB)	Creation Date	Status
SpliceDocs1	4	4	0.1	6/14/2017 - 17:29	Initializing

Once your new cluster is initialized, its status changes to *Active*, and you receive an email message from Splice Machine notifying you that your cluster is ready. At that point, you can click the cluster name (e.g. *SpliceDocs1*) to use the *Cluster Management* screen for that cluster.

**NOTE:** If you have multiple clusters associated with your account, each will be listed in your dashboard. Simply click or tap a cluster name to jump to its management screen.



# Managing a Cluster

This topic describes the actions you can initiate from the **Cluster Management** screen, which include:

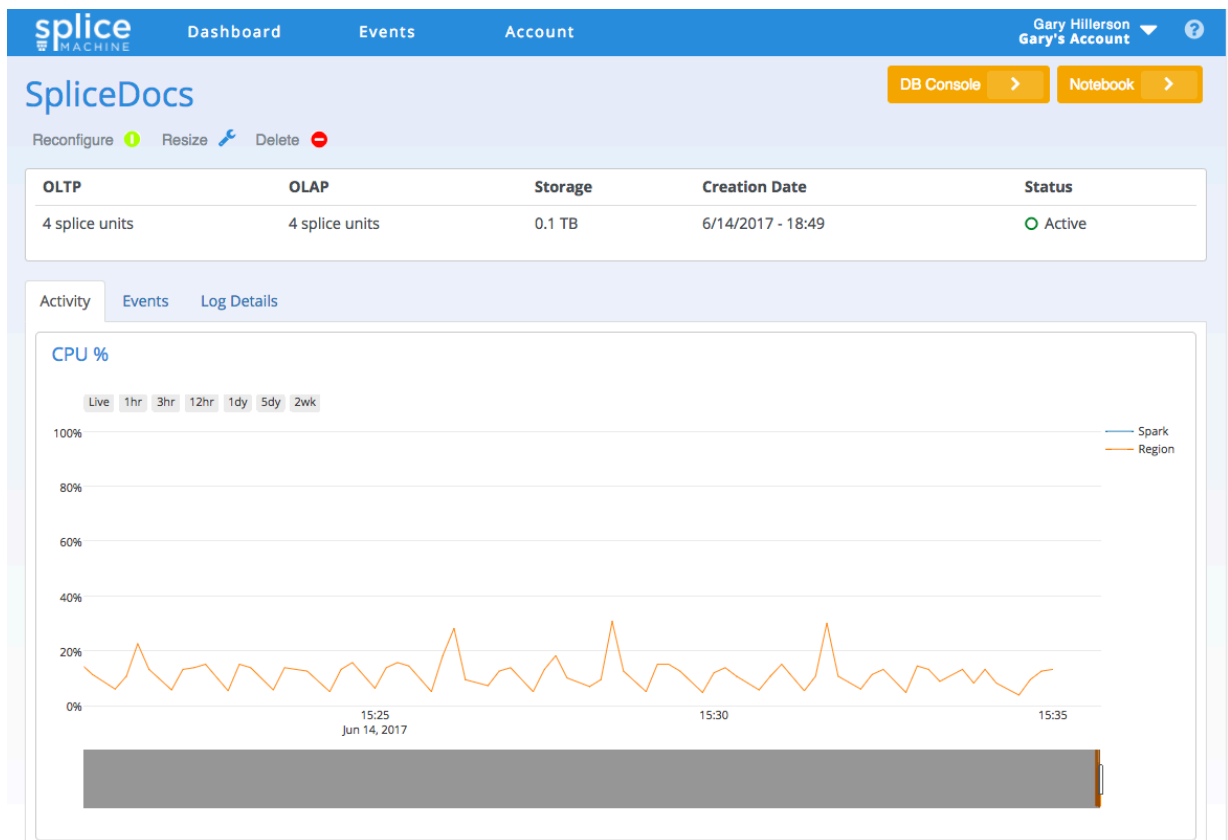
- » [Viewing the cluster's CPU, Disk, and Memory Usage \(page 0\)](#)
- » [Reconfiguring cluster access \(page 32\)](#)
- » [Resizing your cluster \(page 33\)](#)
- » [Deleting your cluster \(page 34\)](#)
- » [Using Zeppelin and the Database Console with your cluster \(page 0\)](#)

## The Cluster Management Screen

You can access the **Cluster Management** screen for any cluster in your Dashboard by simply clicking the name of the cluster.

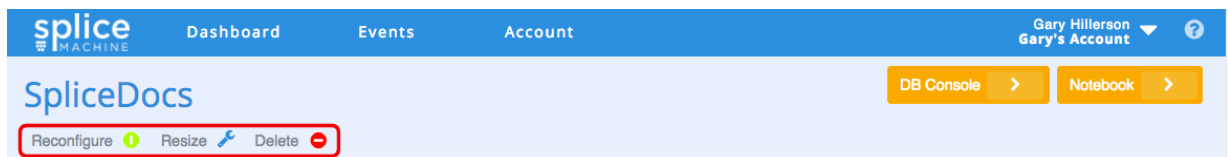
This screen displays information about the cluster, and includes three panels that display a graph of resource usage over time. There are three similar resource usage graphs displayed:

- » CPU Percentage Used (shown in the following image)
- » Disk Usage
- » Memory Usage



## Modifying Your Cluster

Your dashboard features three cluster modification choices; you can [Reconfiguring \(page 0\)](#), [Resizing \(page 0\)](#), or [Deleting \(page 0\)](#) your cluster by clicking one of the buttons near the top-left of your Dashboard screen:



## Reconfiguring Your Cluster

To reconfigure your cluster, click the **Reconfigure** button; the cluster reconfiguration screen displays:

## Reconfigure Cluster SpliceDocs2

### VPC Setup

☐ Client VPC connectivity required

Account ID:

VPC ID (Acceptor):

### IAM S3 Access

☐ S3 connectivity

S3 Access Key:

S3 Secret Key:

Cancel

Reconfigure >

You can modify your VPC and/or IAM configuration information in this screen. Once you've entered your new information, click the **Reconfigure** button to update your configuration and return to your Cluster Management screen.

## Resizing Your Cluster

To resize your cluster, click the **Resize** button. The Resize Cluster screen, which is pretty much identical to the Create New Cluster screen, displays:

The dashboard for cluster **SpliceDocs914** includes the following sections:

- Data Sizing:** Sliders for Internal Dataset (0.1 TB to 100 TB, set to 0.1) and External Dataset (0 TB to 1 PB, set to 0).
- Cluster Power:** Sliders for OLTP Splice Units (4 to 600, set to 4) and OLAP Splice Units (4 to 600, set to 4).
- Backup Frequency:** Frequency set to Daily, Start Window (UTC) set to 02:30:00, and Backups to Keep set to 1.
- Estimated Benchmarks:**
  - Single Record Lookup (ms): 20ms
  - Import Rate (record/sec): 40,000
  - TPC-C 100 Users (TpmC): 4,000
  - TPC-H Query #2 (sec): 25
- Estimated Costs (Hourly):**
  - Storage: \$0.09
  - Compute: \$4.00
  - Total: \$4.09**

Buttons at the bottom include **Reset**, **Back**, and **Resize** (highlighted in orange). A note states: "Note - resizing a cluster can take some time. We will notify you by email once your cluster is resized."

You can adjust your cluster parameters, then click the **Resize** button, which will return you to your Cluster Management screen, where you'll see the status of your cluster set to *Updating*.

## Deleting Your Cluster

To delete your cluster, click the **Delete** button. You'll be asked to confirm the deletion:

**SpliceDocs914**

Are you sure you want to delete this cluster?

Enter the word "DELETE" in the input field below to confirm.

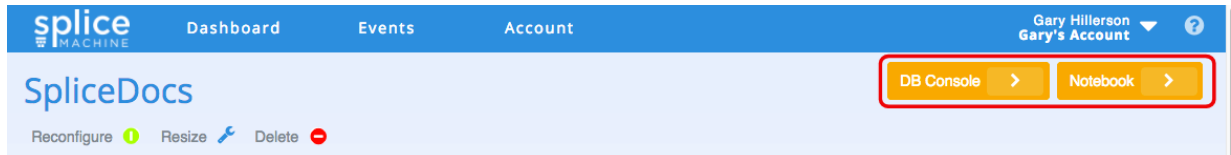
Confirm Code

**Close** **Delete** (with right arrow)

After confirming that you want to do so, your cluster will be deleted, and its status in your dashboard will show as *Deleted*.

## Working With Your Database

Your dashboard includes buttons with which you can access two different means of working with your database; you can launch the **DB Console** or the Apache Zeppelin **Notebook** interface by clicking one of the buttons near the upper-right corner of the Dashboard screen:

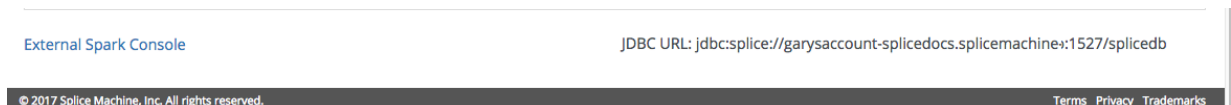


Click the **DB Console** button to land on our [Database Console \(page 0\)](#) interface.

Click the **Notebook** button to land on our [Zeppelin Notebook \(page 43\)](#) interface.

## Connecting to Your Database with JDBC

At the very bottom of the Cluster Management screen, below the three graph panels, is the URL you can use for a JDBC connection with your database service:



The same JDBC link was emailed to you in the message from Splice Machine letting you know that your cluster creation was successful.

# Managing Your Splice Machine Account

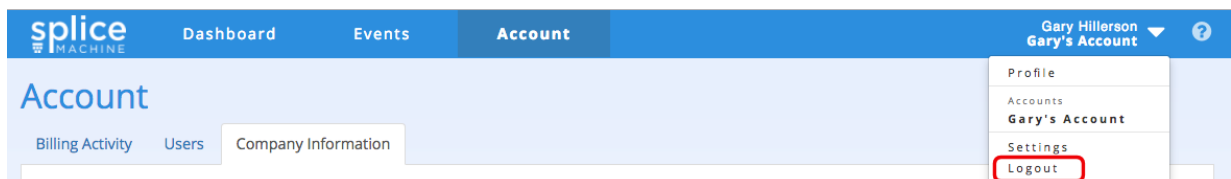
This is a DB-Service-Only topic! [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This topic describes the actions you can perform from the Account tab and Account drop-down in your Dashboard, which include:

- » [Logging Out of Your Account \(page 36\)](#)
- » [Reviewing and Updating Your Billing Information \(page 36\)](#)
- » [Viewing and Updating Your User Profile and Password \(page 37\)](#)
- » [Viewing and Adding Users \(page 39\)](#)
- » [Reviewing and Updating Your Company Information \(page 39\)](#)

## Logging Out of Your Account

To log out of your Cloud Manager account, click the Account Drop-down arrow in the upper-right of your dashboard screen, and select **Logout**:



You'll be logged out and will land back on the Splice Machine Cloud Manager [Login](#) page.

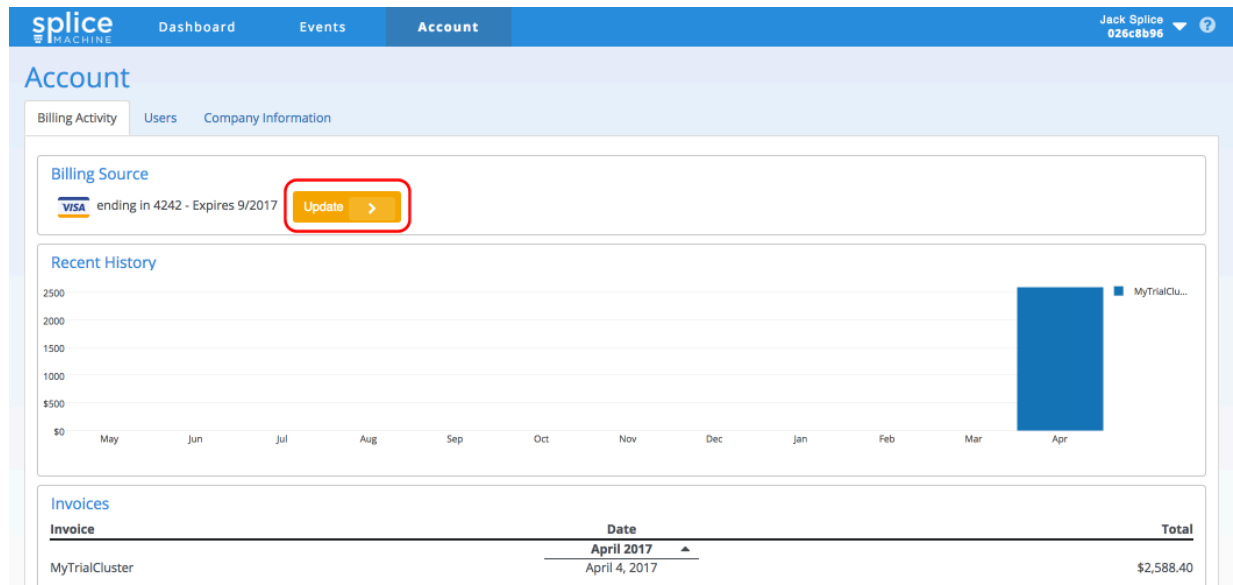
## Reviewing and Updating Your Billing Information



If you subscribed to Splice Machine via the AWS Marketplace, your billing is handled by AWS, not Splice Machine. Your *Account Management* screen will not contain a **Billing Activity** tab; this section does not apply to you.

To display billing information for your account, select the **Billing Activity** tab in a Cloud Manager screen. You can see billing details for each month of each year that your account has been alive. You can also hover over one of the bars representing a cluster to see exactly how much that cluster cost in a month (as shown for July in the image below).

If you have provisioned more than one cluster in your account, each cluster is shown in a different color in the billing detail graphic, as shown below.



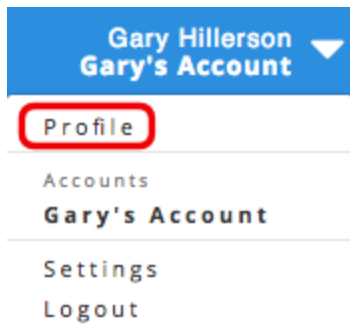
To update your payment source, click the **Update** button.

## Prorated Monthly Billing

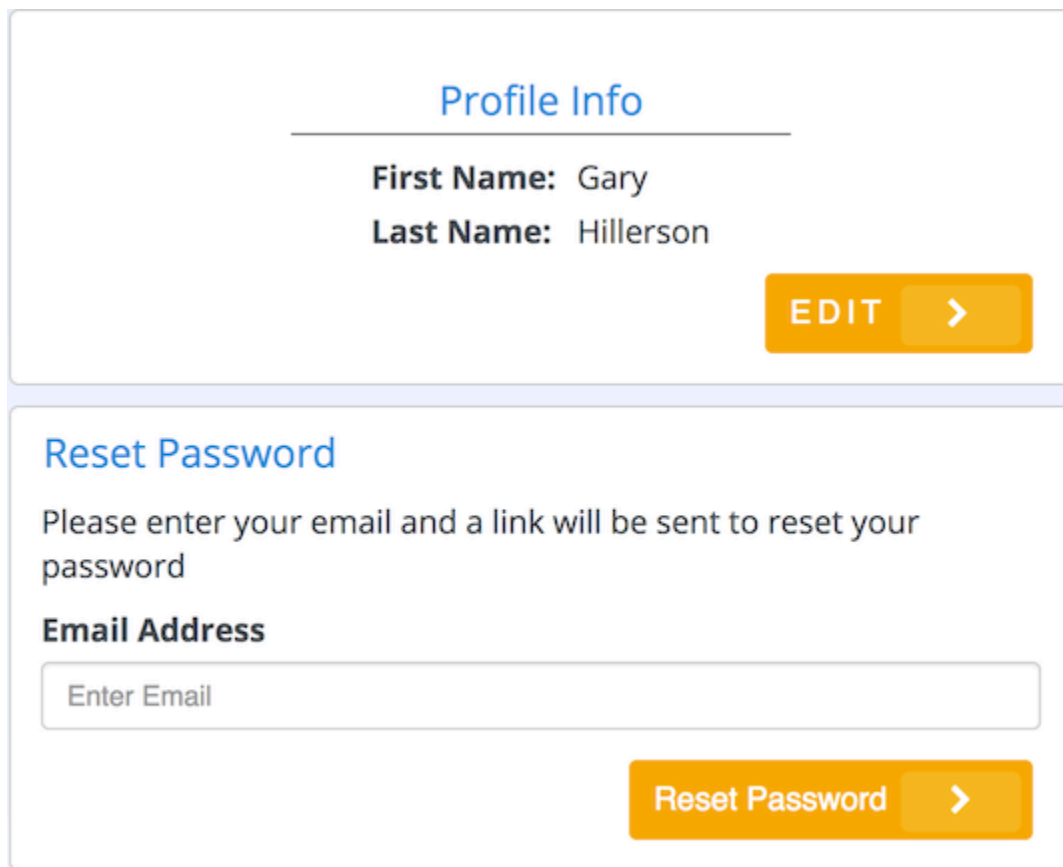
Splice Machine bills for our database service on a prorated monthly basis; any adjustments for deleting or downsizing your cluster(s) are applied to future bills or cluster purchases.

## Viewing and Updating Your User Profile and Password

You can review or edit your profile information by selecting **Profile** from the click the Account Drop-down:



The **Profile** screen displays:

The Profile screen displays two main sections. The first section, titled 'Profile Info', shows the user's first name as 'Gary' and last name as 'Hillerson'. To the right of this information is an orange 'EDIT' button with a right-pointing arrow. The second section, titled 'Reset Password', contains a text prompt: 'Please enter your email and a link will be sent to reset your password'. Below this is a label 'Email Address' followed by a text input field with the placeholder text 'Enter Email'. At the bottom right of this section is an orange 'Reset Password' button with a right-pointing arrow.

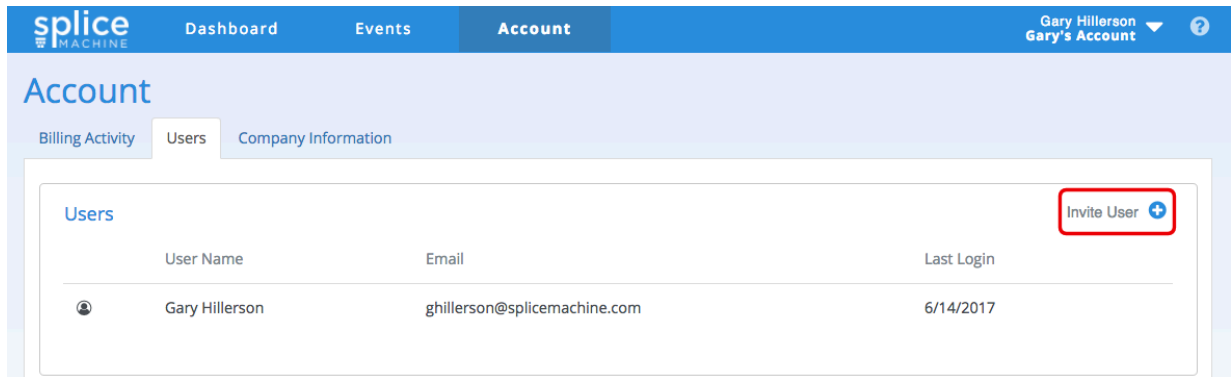
You can edit your name information by clicking the **EDIT** button in the Profile Info panel.

You can reset your account password by entering your email address and then clicking the **Reset Password** button. You'll receive an email from Splice Machine that contains a link you can use to reset your password.

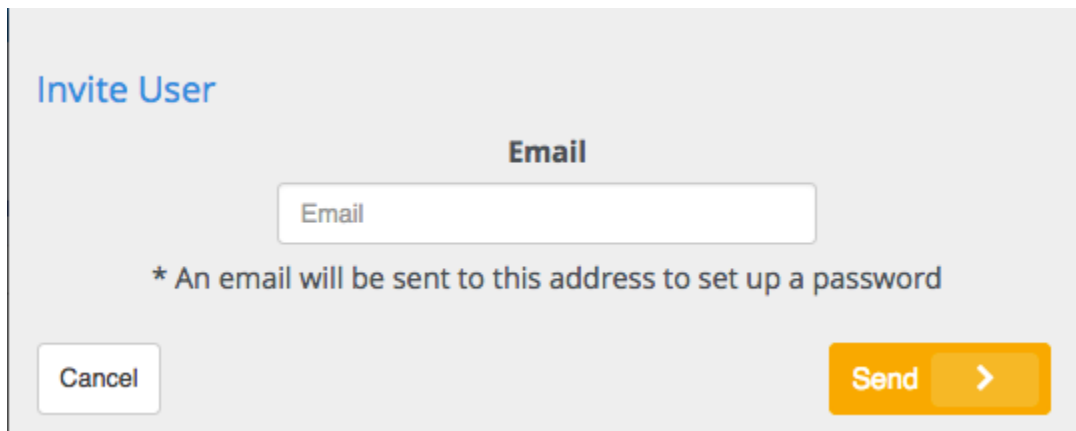


## Viewing and Adding Users

To display the names and log-in information for the users of your database service, select the **Users** tab in your Cloud Manager screen. The *Users* screen displays:

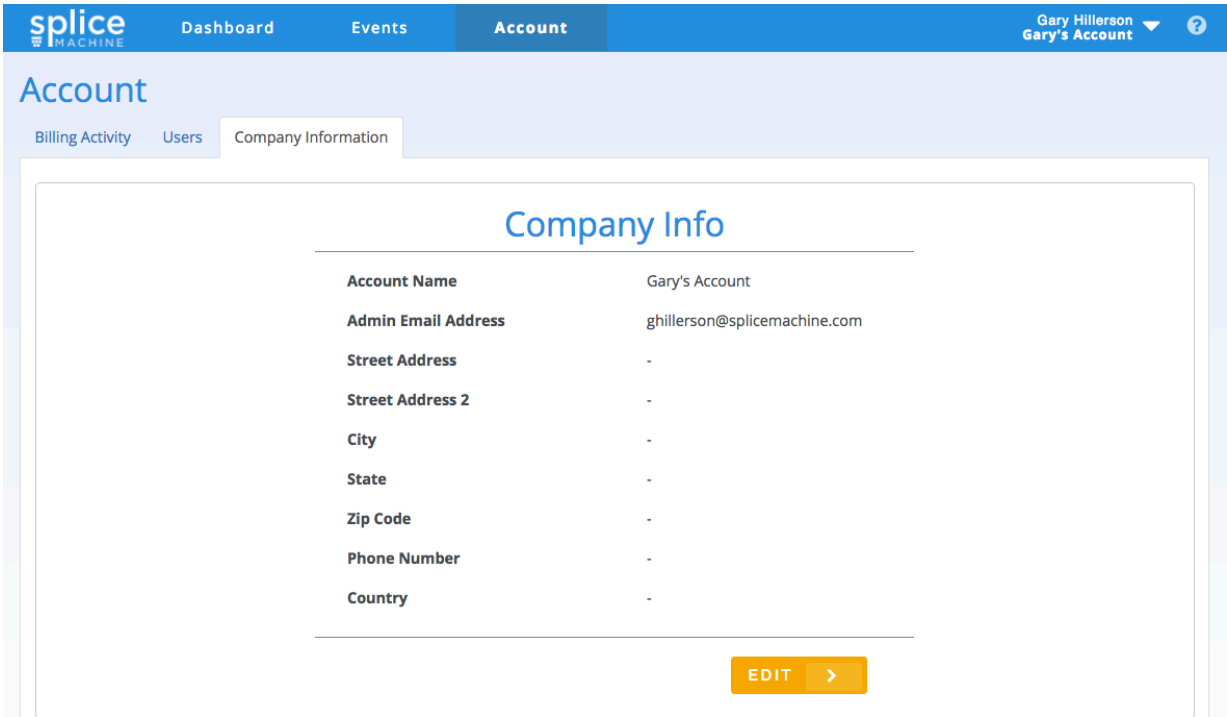


To add another user, click the **Invite User +** button in the *Users* screen. Then enter the new user's email address in the *Invite User* screen and click the **Send** button. We'll send an email inviting that person to set up a password to access your database.



## Reviewing and Updating Your Company Information

To display the company information associated with your account, select the **Users** tab in your Cloud Manager screen. The *Company Information* screen displays:



To edit the company information associated click the **Edit** button.

# Managing Your Event Notifications

This is a DB-Service-Only topic! [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This topic describes the Splice Machine Events Manager, which allows you to examine notification messages sent to your cluster.

Here's a screenshot of a partially populated **Events Manager screen**:

Date	Detail
7/25/2017 - 20:07:45	Daily backup completed at 2017/07/18 12:35:02AM
7/25/2017 - 20:07:21	Daily backup is scheduled for tonight at 2017/07/18 12:00:00AM
7/25/2017 - 20:07:03	Daily backup completed at 2017/07/17 12:33:59AM
7/25/2017 - 20:06:38	Daily backup is scheduled for tonight at 2017/07/17 12:00:00AM
7/25/2017 - 20:06:24	Daily backup completed at 2017/07/16 12:34:51AM
7/25/2017 - 20:06:11	Daily backup is scheduled for tonight at 2017/07/16 12:00:00AM

You can initiate these actions in the Events Manager:

- » Display messages for one specific cluster, or all of your clusters; in the screenshot above, events are displayed for the cluster named *GaryDocs2*.
- » Filter which messages are displayed; enter filter criteria, then click the **Filter** button. You can filter on:
  - » A start date.
  - » An end date.
  - » A keyword or exact phrase.

You can filter on a start date or end date on its own, or combine them together to specify a date range. You can also combine a date or date-range filter with a keyword filter to find only events that meet the combined criteria.

- » You can click the **Clear Filter** button to clear any filters and display all of your notification messages.
- » Click the < (Prev), > (Next), << (First), or >> (Last) buttons to move through multiple screenfuls of messages.
- » Click the arrow to the right of a message to display the full or shortened version of the message.

# Using Zeppelin Notebooks

This is a DB-Service-Only topic! [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This guide helps you to get started with using Zeppelin notebooks to interact with your Splice Machine database.

Topic	Description
<a href="#">Getting Started with Zeppelin (page 44)</a>	Introduces you to using Zeppelin with your Splice Machine database.
<a href="#">Zeppelin Usage Notes (page 50)</a>	Specific usage notes for creating Zeppelin notebooks to use with Splice Machine..
<a href="#">A Simple Tutorial (page 52)</a>	Walks you through a quick and simple tutorial that shows you how to use Zeppelin notebooks to load and query data, and apply different visualizations to the results.

# Getting Started with Zeppelin

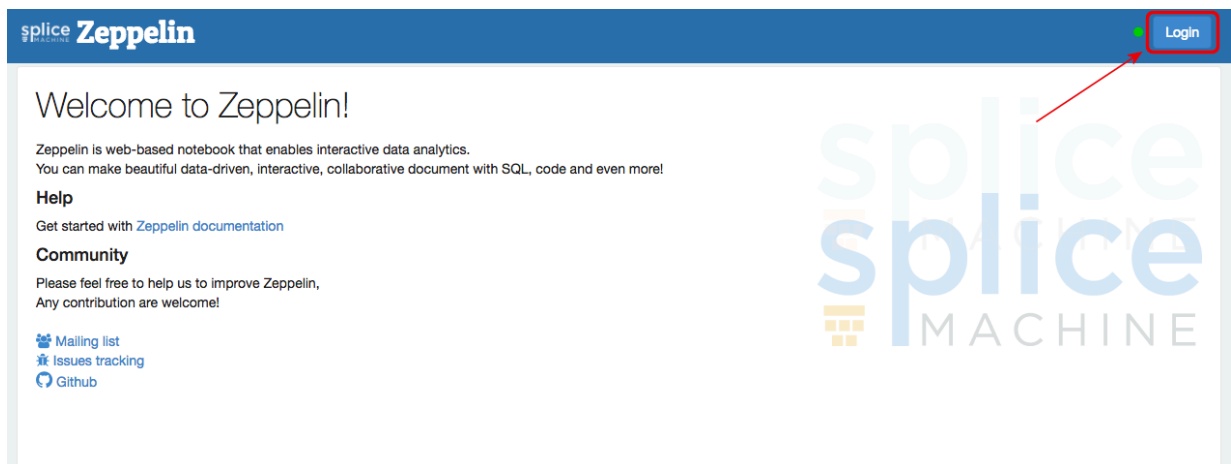
This is a DB-Service-Only topic! [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This topic helps you to get started with using Zeppelin with your Splice Machine database.

**NOTE:** We strongly encourage you to visit the [Zeppelin documentation site \(https://zeppelin.apache.org/docs/\)](https://zeppelin.apache.org/docs/) to learn about creating, modifying, and running your own Zeppelin notebooks.

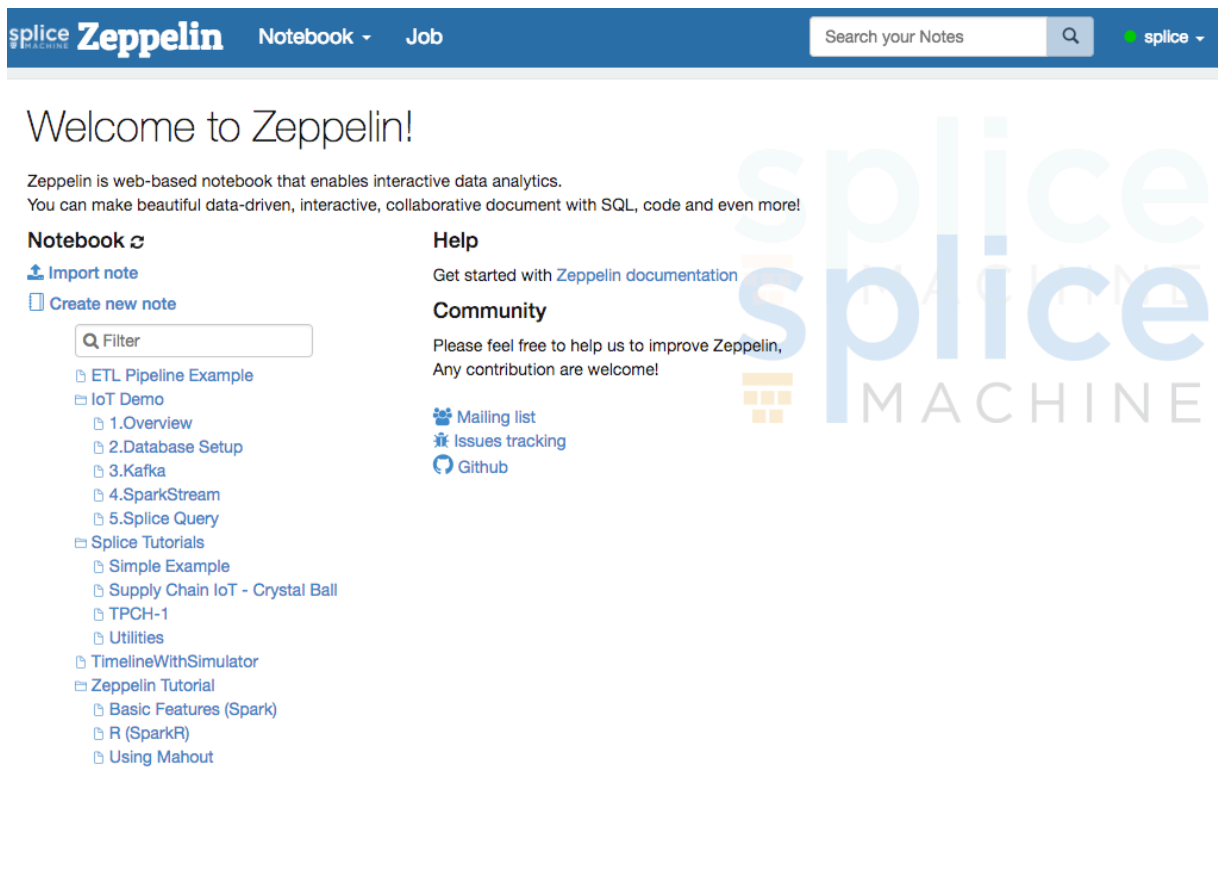
## The Zeppelin Dashboard

When you click the **Notebook** button in your Cluster Management dashboard, you land on the Zeppelin welcome page. To start using Zeppelin with your database service, you need to log in to your database by clicking the **Login** button.



Use the same user ID and password to log into Zeppelin as you use to log into your database.

When you log into Zeppelin for your database, you'll land on the Zeppelin dashboard, which displays the list of available notebooks. As you can see, notebooks can be organized in folders.



splice MACHINE Zeppelin Notebook Job

Search your Notes splice

## Welcome to Zeppelin!

Zeppelin is web-based notebook that enables interactive data analytics.  
You can make beautiful data-driven, interactive, collaborative document with SQL, code and even more!

### Notebook

Import note  
Create new note

Filter

- ETL Pipeline Example
- IoT Demo
  - 1. Overview
  - 2. Database Setup
  - 3. Kafka
  - 4. SparkStream
  - 5. Splice Query
- Splice Tutorials
  - Simple Example
  - Supply Chain IoT - Crystal Ball
  - TPCH-1
  - Utilities
- TimelineWithSimulator
- Zeppelin Tutorial
  - Basic Features (Spark)
  - R (SparkR)
  - Using Mahout

### Help

Get started with [Zeppelin documentation](#)

### Community

Please feel free to help us to improve Zeppelin,  
Any contribution are welcome!

- Mailing list
- Issues tracking
- Github

Splice Machine has already created a number of useful notebooks; we suggest that you try running some of them to get a feel for what Zeppelin can do: click a notebook name, and you'll land on the notebook page in Zeppelin. From there, you can run all or portions of the notebook, modify its content, and create new notebooks. Our next topic, [A Simple Tutorial \(page 52\)](#), uses the our **Simple Example** tutorial.

## First Notebook Run: Save Interpreter Bindings

The first time that you run any Zeppelin notebook, you need to bind any interpreters needed by the notebook. For our tutorials, these are preconfigured for you; all you need to do is click the Save button:

splice **Zeppelin**

Notebook ▾Job

Search your Notes 🔍

admin ▾

Splice Tutorials / Simple Ex...

▶

⌕

📖

📝

📄

📥

👤

📄

⌕

Head ▾

🗑️

🕒

⌂

⚙️

🔒

default ▾

Settings

Interpreter binding

Bind interpreter for this note. Click to Bind/Unbind interpreter. Drag and drop to reorder interpreters. The first interpreter on the list becomes default. To create/remove interpreters, go to [Interpreter](#) menu.

🔄 spark %spark (default), %sql, %dep, %pyspark, %r

🔄 md %md

🔄 angular %angular

🔄 python %python, %python.sql, %python.conda, %python.docker

🔄 plsql %plsql

🔄 splicemachine %splicemachine

Save

Cancel


**NOTE:** If you neglect to save its bindings, the notebook will not run. And again: you only need to do this one time for each notebook that you run.

## The Zeppelin Note Toolbar










Zeppelin displays a toolbar at the top of each note that provides convenient access to a number of options:



The following table describes the toolbar buttons:

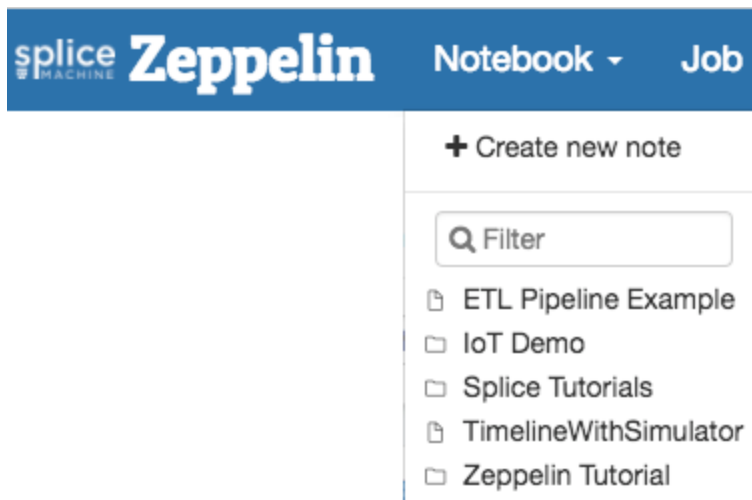
	Executes all of the paragraphs in the note, in display-order sequence.
---	--



	Shows or hides the code sections of the paragraphs in the note.
	Shows or hides the result sections of the paragraphs in the note.
	Clears the result sections of the paragraphs in the note.
	Clones the current note.
	Exports the current note in JSON format. <b>NOTE:</b> The code and result sections of all paragraphs are exported; you might want to clear your results before exporting a note.
	Switches between personal and collaboration modes.
	Commits changes that you've made to the content of the current note (and allows you to add a commit note).
Head ▼	Displays the revision you're currently viewing, and lets you select from available revisions.
	Deletes the note.
	Schedules execution of the note, using CRON syntax.

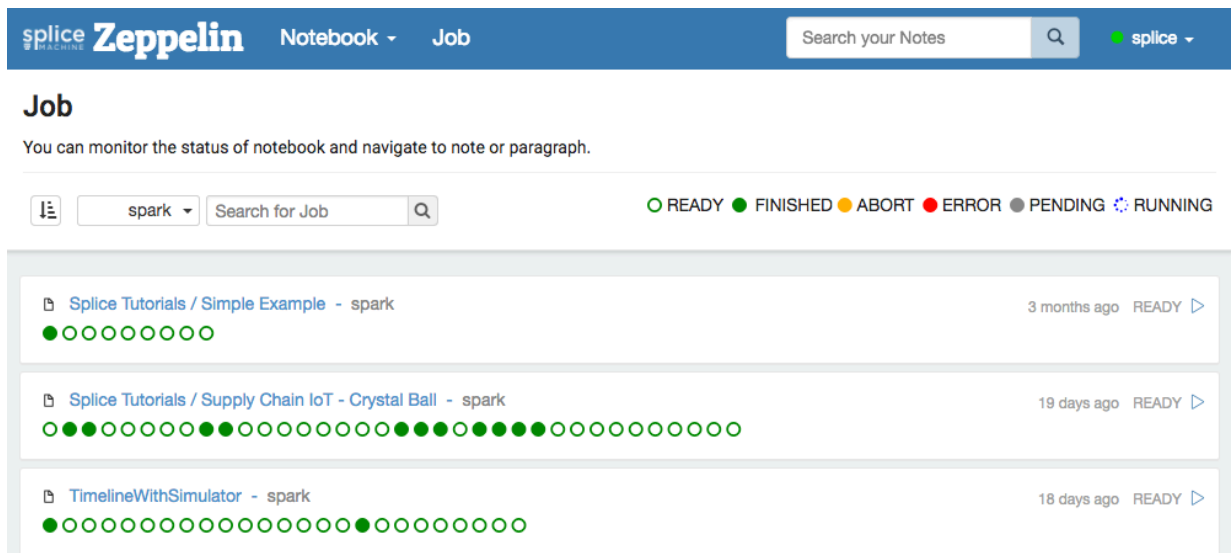
## The Zeppelin Drop-Down Menu

When you're working in Zeppelin, you can quickly jump to another notebook or create a new note by clicking the **Zeppelin** drop-down menu:



## Monitoring Job Status

You can monitor the status of any Zeppelin notebook job(s) running in your cluster by clicking the **Job** button at the top of the Zeppelin screen. This displays a list of the notebook jobs that are running and have run on your cluster.



From the **Job** screen, you can:

- » Monitor all jobs associated with your account.
- » Filter which jobs are displayed.
- » Search for notebooks.
- » Start, Pause, or Terminate a running job.

» Click a notebook job name to navigate to that notebook.

## Creating Notebooks

Be sure to view our [Usage Notes \(page 50\)](#) page for important information about creating Zeppelin notebooks to use with Splice Machine.

# Zeppelin Usage Notes

**This is a DB-Service-Only topic!** [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This page currently contains exactly one tip about using Zeppelin with Splice Machine; this will grow into a loose collection of tips over time.

## Use Full Classpath!

If you're coding a Zeppelin notebook in Java, you must specify the full class of imported classes, such as `java.sql.Timestamp`; otherwise, an error occurs.

For example, this generates an error:

```
%spark
import java.util.Date
import java.sql.
{Connection, Timestamp}
classOfTimestamp
classOffoo
val tt = Timestamp.valueOf("2261-12-31 00:00:00")
class foo extends Object { val xx: Timestamp = Timestamp.valueOf("2261-12-31 00:0
0:00") }

import java.util.Date
import java.sql.{Connection, Timestamp}
res12: Classjava.sql.Timestamp = class java.sql.Timestamp
res13: Classfoo = class foo
tt: java.sql.Timestamp = 2261-12-31 00:00:00.0
<console>:13: error: not found: type Timestamp
val xx: Timestamp = Timestamp.valueOf("2261-12-31 00:00:00")
^
<console>:13: error: not found: value Timestamp
val xx: Timestamp = Timestamp.valueOf("2261-12-31 00:00:00")
^
ERROR
```

The error is resolved by specifying the full classpath:

```
%spark
import java.util.Date
import java.sql.
{Connection, Timestamp}
classOfTimestamp
classOffoo
val tt = Timestamp.valueOf("2261-12-31 00:00:00")
class foo extends Object { val xx: java.sql.Timestamp = java.sql.Timestamp.valueOf("2
261-12-31 00:00:00") }

import java.util.Date
import java.sql.{Connection, Timestamp}
res14: Classjava.sql.Timestamp = class java.sql.Timestamp
res15: Classfoo = class foo
tt: java.sql.Timestamp = 2261-12-31 00:00:00.0
defined class foo
FINISHED
```

# A Simple Zeppelin Tutorial

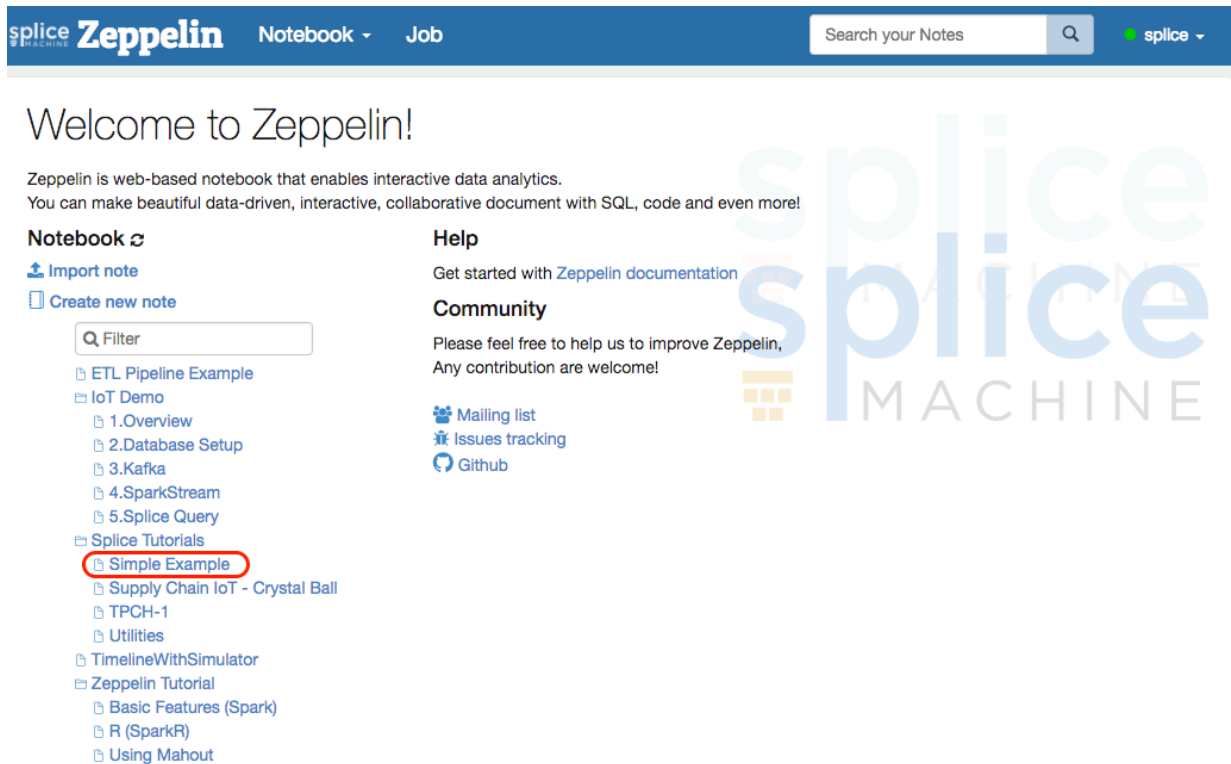
This is a DB-Service-Only topic! [Learn about our products \(https://www.splicemachine.com/get-started/\)](https://www.splicemachine.com/get-started/)

This topic walks you through using a very simple Zeppelin notebook, to help you learn about using Zeppelin with Splice Machine.

**NOTE:** Our [Getting Started with Zeppelin \(page 44\)](#) page provides a very brief overview of using Zeppelin; If you're new to Zeppelin, we strongly encourage you to visit the [Zeppelin documentation site \(https://zeppelin.apache.org/docs/\)](https://zeppelin.apache.org/docs/) to learn about creating, modifying, and running your own Zeppelin notebooks.

## Running the Tutorial Notebook

You can access this Zeppelin notebook by clicking the Basics (Spark) link under Zeppelin Tutorials on the Zeppelin Dashboard page:



Once you've opened the tutorial, you can run each step (each Zeppelin *paragraph*) by clicking the **Ready** button that you'll see on the right side of each paragraph. This example includes these steps:

- » Click the first **READY** button to create the schema and a table:

**Create a schema called EXAMPLE**

FINISHED ▶ ⌵ ⌶ ⌵

```
%splicemachine
CREATE SCHEMA EXAMPLE;
```

Query executed successfully. Affected rows : 0

Took 4 sec. Last updated by splice at September 14 2017, 10:09:30 AM.

**Create a table called LINEITEM**

FINISHED ▶ ⌵ ⌶ ⌵

```
%splicemachine
CREATE TABLE EXAMPLE.LINEITEM (
  L_ORDERKEY BIGINT NOT NULL,
  L_PARTKEY INTEGER NOT NULL,
  L_SUPPKEY INTEGER NOT NULL,
  L_LINENUMBER INTEGER NOT NULL,
  L_QUANTITY DECIMAL(15,2),
  L_EXTENDEDPRICE DECIMAL(15,2),
  L_DISCOUNT DECIMAL(15,2),
  L_TAX DECIMAL(15,2),
  L_RETURNFLAG VARCHAR(1),
  L_LINESTATUS VARCHAR(1),
  L_SHIPDATE DATE,
  L_COMMITDATE DATE,
  L_RECEIPTDATE DATE,
  L_SHIPINSTRUCT VARCHAR(25),
  L_SHIPMODE VARCHAR(10),
  L_COMMENT VARCHAR(44),
  PRIMARY KEY(L_ORDERKEY,L_LINENUMBER)
);
```

Query executed successfully. Affected rows : 0

Took 2 sec. Last updated by splice at September 14 2017, 10:09:37 AM.

- » Import data (in this case, TPC1 benchmark data) into the table, then verify the data load by counting the number of records in the table:

**Import Data**

READY ▶ ⌵ ⌶ ⌵

```
%splicemachine
call SYSCS_UTIL.IMPORT_DATA ('EXAMPLE', 'LINEITEM', null, 's3a://splice-benchmark-data/flat/TPCH/1/lineitem', 'I', null, null, null, null, 0, '/tmp'
, true, null);
```

Took 19 sec. Last updated by anonymous at April 24 2017, 6:56:51 AM.

**Count the number of records in the table**

READY ▶ ⌵ ⌶ ⌵

```
%splicemachine
SELECT COUNT(*) AS NUM_RECORDS FROM EXAMPLE.LINEITEM;
```

Took 1 sec. Last updated by anonymous at April 24 2017, 6:57:37 AM.

- » Create indexes on the table, and then run compaction on the data, which is always a good idea after updating a large number of records:

**Create some indexes**

READY ▶ ⌵ ⌶ ⌵

```
%splicemachine
create index EXAMPLE.L_SHIPDATE_IDX on EXAMPLE.LINEITEM(
  L_SHIPDATE,
  L_PARTKEY,
  L_SUPPKEY,
  L_EXTENDEDPRICE,
  L_DISCOUNT
);

create index EXAMPLE.L_PART_IDX on EXAMPLE.LINEITEM(
  L_PARTKEY,
  L_ORDERKEY,
  L_SUPPKEY,
  L_SHIPDATE,
  L_EXTENDEDPRICE,
  L_DISCOUNT,
  L_QUANTITY,
  L_SHIPMODE,
  L_SHIPINSTRUCT
);
```

Took 1 sec. Last updated by anonymous at April 24 2017, 6:56:52 AM.

**Run compaction**

READY ▶ ⌵ ⌶ ⌵

```
%splicemachine
call SYSCS_UTIL.SYSCS_PERFORM_MAJOR_COMPACTION_ON_SCHEMA('EXAMPLE');
```

Took 2 sec. Last updated by anonymous at April 24 2017, 6:56:54 AM.



» Collect statistics, to improve query planning, and then run a query:

Collect Statistics

```
%splice
analyze schema EXAMPLE;
```

READY

▶

⌵

⌵

⌵

Run a query

```
%splice
select
  l_returnflag,
  l_linestatus,
  sum(l_quantity) as sum_qty,
  sum(l_extendedprice) as sum_base_price,
  sum(l_extendedprice * (1 - l_discount)) as sum_disc_price,
  sum(l_extendedprice * (1 - l_discount) * (1 + l_tax)) as sum_charge,
  avg(l_quantity) as avg_qty,
  avg(l_extendedprice) as avg_price,
  avg(l_discount) as avg_disc,
  count(*) as count_order
from
  EXAMPLE.lineitem
where
  l_shipdate <= date({fn TIMESTAMPADD(SQL_TSI_DAY, -90, cast('1998-12-01 00:00:00' as timestamp))})
group by
  l_returnflag,
  l_linestatus
order by
  l_returnflag,
  l_linestatus
```

READY

▶

⌵

⌵

⌵

After the query runs, you can take advantage of Zeppelin's built-in visualization tools to display the query results in various graphical and tabular formats.

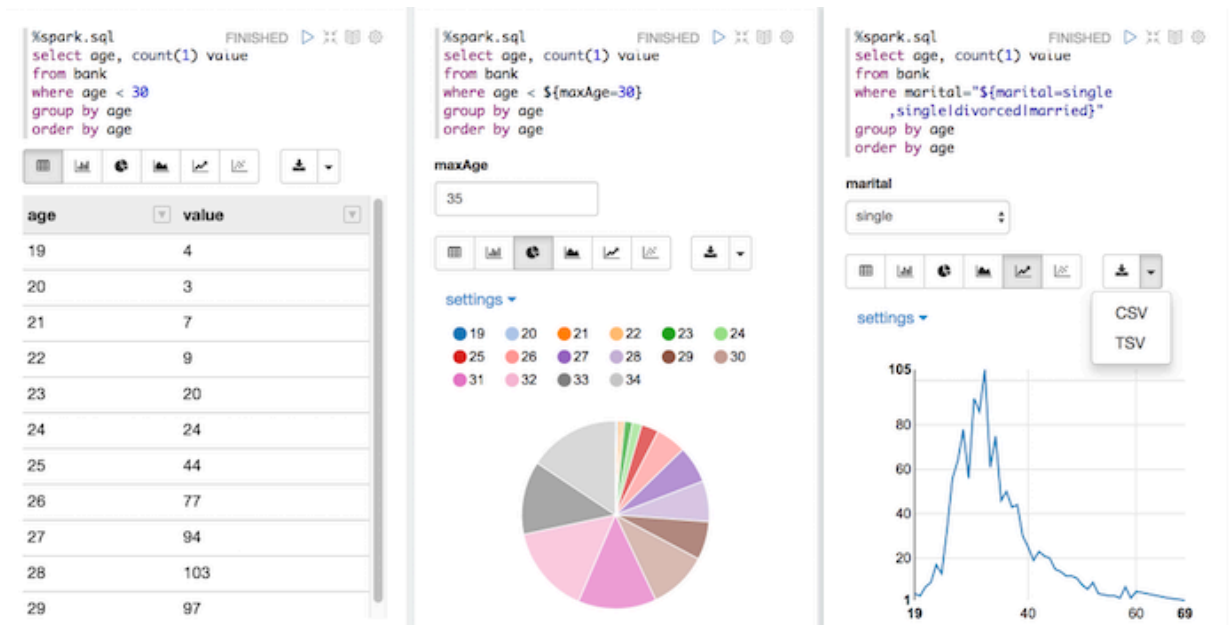
When you click the **READY** button, Zeppelin runs the paragraph that loads your data and subsequently displays the *Finished* message.

**NOTE:** If you see *Error* instead of *Finished*, it usually means that you've forgotten to set SpliceMachine interpreter as the default.

## Apply Different Visualizations to Your Results

Zeppelin provides a wealth of data visualization tools you can use. In the example below, we have modified the presentation of query results to use different visualizations by clicking different visualization icons in the output pane. You can define and modify the values of variables that you use in your queries; for example, the `maxAge` and `marital` values in the examples below:

Page 55 of 70



# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.

# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.

# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.



# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.

# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.

# Welcome to Splice Machine!

Welcome to Splice Machine, the database platform for adaptive applications that manage operational processes. This site contains documentation to how you use our products. Visit our company web site to [learn more about how our platform works](https://www.splicemachine.com/product/how-it-works/) (<https://www.splicemachine.com/product/how-it-works/>).

## Documentation Versions

This web contains the current customer documentation for **version 2.7** of Splice Machine.

If you're using an earlier version, you can find the documentation in these locations:

» [Splice Machine Version 2.5.1 Documentation. \(https://doc.splicemachine.com/2.5.1/index.html\)](https://doc.splicemachine.com/2.5.1/index.html)

## Splice Machine Editions

Splice Machine's database platform is available in several editions, all of which are documented in this web. Click an Edition name in the table below to navigate to the documentation home page for that edition.

Splice Machine Edition	Description
<a href="#">Cloud Edition (page 0)</a>	You can configure and deploy our <i>Managed Database Service in the Cloud Edition</i> in a matter of minutes: simply register, adjust a few configuration options for your anticipated data size, cluster power, and backup frequency, and you're up and running within 10-15 minutes.
<a href="#">Enterprise Edition (page 0)</a>	You can deploy the Enterprise Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks. This edition is a superset of our Community Edition, adding features such as backup and encryption that are required for enterprise users.

Splice Machine Edition	Description
<a href="#">Cluster Community Edition (page 0)</a>	You can deploy the free, open source Cluster Community Edition of our <i>On-Premise Database</i> on a cluster that is managed by Cloudera, MapR, or Hortonworks.
<a href="#">Standalone Community Edition (page 0)</a>	You can install and run the free Standalone Community Edition of our <i>On-Premise Database</i> on a computer running MacOS, Linux, or CentOS. This edition is a great way to quickly learn about and play with the many features of our database platform.



Our [Editions \(page 0\)](#) page compares the features available in each Splice Machine product edition.

## Finding Information

This documentation web include three main sections, each of which is accessed through the main menu at the top of each page:

- » The **Splice Machine** menu links to topics that apply to all of our database editions, including our Tutorials and our SQL Reference Manual.
- » The **DB-Service Only** menu links to topics that apply only to our Managed Database Service in the Cloud edition, including our Cloud Manager Guide and guides to using Zeppelin with Splice Machine.
- » The **On-Premise-Only** menu links to topics that apply only to the On-Premise editions of our database platform, including our Installation and Administrator's Guides.

Each topic page in this web also includes a sidebar that links to topics related to the one you are currently viewing. You can hide or show the sidebar by clicking the **Nav** button in the menu at the top of screen.

See the [Using our Documentation \(page 0\)](#) page to learn more about the content available in our documentation and how to easily navigate among our sections and topics.