cloudera[®] Impala HA with F5 BIG-IP



Important Notice

© 2010-2016 Cloudera, Inc. All rights reserved.

Cloudera, the Cloudera logo, Cloudera Impala, Impala, and any other product or service names or slogans contained in this document, except as otherwise disclaimed, are trademarks of Cloudera and its suppliers or licensors, and may not be copied, imitated or used, in whole or in part, without the prior written permission of Cloudera or the applicable trademark holder.

Hadoop and the Hadoop elephant logo are trademarks of the Apache Software Foundation. All other trademarks, registered trademarks, product names and company names or logos mentioned in this document are the property of their respective owners to any products, services, processes or other information, by trade name, trademark, manufacturer, supplier or otherwise does not constitute or imply endorsement, sponsorship or recommendation thereof by us.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Cloudera.

Cloudera may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Cloudera, the furnishing of this document does not give you any license to these patents, trademarks copyrights, or other intellectual property.

The information in this document is subject to change without notice. Cloudera shall not be liable for any damages resulting from technical errors or omissions which may be present in this document, or from use of this document.

Cloudera, Inc. 1001 Page Mill Road, Building 2 Palo Alto, CA 94304-1008 info@cloudera.com US: 1-888-789-1488 Intl: 1-650-843-0595 www.cloudera.com

Release Information Date: 20160414

Table of Contents

Overview	1
Instructions	1
Prerequisites	1
Basics	1
Create the Nodes	2
Create the Pools	4
Port 21000	4
Port 21050	7
Create the Virtual Servers	8
Port 21000	8
Port 21050	11
Configure Impala	12
Verification	13
Local Traffic Network Map	13
JDBC	14
ODBC	14
Impala Shell	14
Hue	15
Supplemental Configuration	15
Supplemental Configuration Increasing Idle Timeouts	15 15
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile	15 15 16
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile	15 15 16 18
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile	15 16 18 19
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile	15 16 18 19 21
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos	15 16 18 19 21 22
Supplemental Configuration. Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL.	15 16 18 19 21 22 22
Supplemental Configuration. Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL.	15 16 18 19 21 22 22 23
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough	15 16 18 19 21 22 22 23 25
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload	15 16 18 19 21 22 22 23 25 26
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload Verification	15 16 16 18 21 21 22 23 25 26 26
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload Verification	15 16 18 19 21 22 23 25 26 26 26
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload Verification Known Issues Backend nodes visible in Hue	15 16 18 19 21 22 22 25 26 26 26 27
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload Verification Known Issues Backend nodes visible in Hue Error 104: Connection reset by peer	15 16 16 18 19 21 22 22 23 26 26 26 26 27 27 28
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Protocol Profile Create a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL Client/Server SSL TLS/SSL Passthrough TLS/SSL Offload Verification Known Issues Backend nodes visible in Hue Error 104: Connection reset by peer TTransportException, Could not start SASL	15 16 18 19 21 22 23 25 26 26 26 27 27 28 29
Supplemental Configuration Increasing Idle Timeouts Create a Custom Protocol Profile Apply a Custom Persistence Profile Apply a Custom Persistence Profile Kerberos TLS/SSL. Client/Server SSL. TLS/SSL Passthrough TLS/SSL Offload Verification Known Issues Backend nodes visible in Hue Error 104: Connection reset by peer TTransportException, Could not start SASL	15 16 18 19 21 22 22 23 26 26 26 26 27 27 27 29 29 29

Overview

This guide walks you through configuring an <u>F5 BIG-IP</u> to manage client connection traffic to <u>Apache Impala</u> (incubating) traffic using <u>Local Traffic Manager</u> (LTM), providing high availability and protecting against Impala daemon failures. Step-by-step instructions are provided to configure LTM to work with ODBC, JDBC, impala-shell, and Hue, both with and without TLS/SSL enabled.

Known issues and limitations are discussed.

This guide does not address performance tuning of Impala or LTM.

We find that users often use the terms F5, BIG-IP, and LTM interchangeably. This usage is not accurate and can lead to confusion. Think of the BIG-IP as the physical hardware appliance sitting in the datacenter, running a specially crafted operating system to manage network traffic. Local Traffic Manager (LTM) is one of the modules running on that OS, providing traffic optimization, load balancing, and offloading.

The vast majority of this guide refers to LTM. Some of the configuration described -- such as loading certificates -- affects the BIG-IP as a whole, not just the LTM. A few operations pertain to Cloudera Manager and the cluster hosts themselves.

Instructions

Prerequisites

- An F5 BIG-IP configured with functional network interfaces, trunks, VLANs, and routes required to pass traffic
- A self-IP added to BIG-IP networking for the Impala services, or available ports on an existing one
- An existing cluster running Impala
- A fully qualified domain name (FQDN) that resolves to the self-IP

Basic understanding of <u>general load balancing concepts</u> is recommended, especially before making changes to production systems. Links to the <u>F5 Glossary</u> and <u>BIG-IP LTM manual</u> are provided for F5-specific terms.

Basics

Impala uses two "external" TCP ports, each of which must be configured in a similar manner:

- 21000 Frontend (impala-shell, Beeswax, v1.2 of Cloudera ODBC)
- 21050 Frontend (Hue, JDBC, v2.0+ of Cloudera ODBC)

The Impala Daemon HTTP Server runs on TCP port 25000. You do not need to configure this port for load balancing; users should go to the host directly.

The StateStore and Catalog server ports are also listed as external, but only run on a single host. Do not configure these services for load balancing.

There are four steps when configuring an Impala service on an LTM:

- 1. Create the <u>nodes</u>.
- 2. Create the pools.
- 3. Create the virtual servers.
- 4. <u>Configure Impala to be used</u> through a proxy for high availability.

Depending on your need for TLS/SSL, additional configuration may be required. Once you have the basics, see <u>Supplemental Configuration</u>.

Create the Nodes

For each cluster host that is targeted by the load balancer, you create a node record on the load balancer.

Local Traffic > Nodes > Node List > Create (green plus)



Assign each node a descriptive label and IP address. Although you can use a fully qualified domain name (FQDN), Cloudera recommends using the IP address of the node instead of relying on name resolution.

• •	BIG-IP® - f5-eng.m	tv.cloud ×		Lega
⊢ ⊣	C 🛛 Attps://f5-eng	g.sjc.cloudera.com/xui/#	☆ 😔 🗖 🧿 !	=} 🚺 😳
Hostna IP Add	rne: f5-eng.mtv.cloudera.com ress: 172.17.190.10	Date: Mar 23, 2016 User: ale: Time: 11:46 AM (PDT) Role: Res	xm Partition: Common Partition:	Cog out
R	ONLINE (ACTIVE) Standalone Provisioning Warning			
Mai	n Help Abou	t Local Traffic » Nodes : No	ode List » New Node	
SI	tatistics	General Properties		
iA 🖏	pps	Name	nightly56-2	
🕄 w	lizards	Description		
S D	NS	Address	Address CDDN 172.31.8.124	
i Lo	ocal Traffic	Configuration		I
	Network Map	Health Monitors	Node Default	
	Virtual Servers	Ratio	1	
	Policies	Connection Limit	0	
	Profiles	Connection Rate Limit	0	
	iRules)		
	Pools	Cancel Repeat Finished	i)	
	Nodes	19		
	Monitors	\odot		
	Traffic Class	⊙		
	Address Translation			

Repeat until all of your Impala nodes are defined.

C A https://		×									
	5-eng.sjc.cl	oudera.	com/xui/#						2		{=}
Hostname: f5-eng.mtv.clouder IP Address: 172.17.190.10	a.com Date Time	s: Mar 23, s: 11:48 A	2016 User M (PDT) Role	: alexm : Resource Administrat					Partition: Co	mmon C	Log ou
ONLINE (ACTIV Standalone Provisioning W	E) aming										
Main Help	About	Local	Traffic » Nodes	: Node List	1						
Statistics		☆ -	Node List	Default Monitor	Statistics						
iApps											
		nightly			Search Reset	Search		[Create
Wizards			Status A Nam	e		Application	Address	≑ FQDN	Ephemeral General General	Description	Partition / Pa
			nightly:	56-2			172.31.8.124		No		Common
			o nightiyi	50-3			172.31.8.120		No		Common
Local Traffic		Enable	Dicable Fo	Do-4	1		172.31.6.130		NO		Common
Network Map		Enable		de offine j Delete							
Virtual Servers											
Policies	10										
Profiles											
iRules											
IRules Pools											
iRules Pools Nodes	4 										
iRules Pools Nodes Monitors	+ + •										
iRules Pools Nodes Monitors Traffic Class	, • •										

Create the Pools

A pool represents a collection of hosts running a service to be balanced according to a particular mechanism. Impala has two frontend services, each of which need a distinct pool.

Port 21000

Local Traffic > Pools > Pool List > Create (green plus)

	ng.mtv.clou	× //										
> C 🛛 🔒 https://f5	-eng.sjc.c	loudera.c	om/xui/#						23	5 0 5	{=}	U
Hostname: f5-eng.mtv.cloudera.c P Address: 172.17.190.10	om Da Tin	te: Mar 23, 2 ne: 11:48 AM	016 Us I (PDT) Ro	er: alexm le: Resource Administra					Partition: Co	mmon	Log	out
ONLINE (ACTIVE) Standalone Provisioning War	ning											
Main Help	About	Local T	raffic » Node	as : Node List	1							
Statistics		₩ ~ 1	Node List	Default Monitor	Statistics							
iApps		nightly			Search R	eset Search					Crea	te
Wizards			Status 🔺 Na	me		Application	+ Address	≑ FQDN	Ephemeral	Description	+ Partition / F	Path
		0	nightl	ly56-2			172.31.8.124		No		Common	
DNS		0	nightl	ly56-3			172.31.8.126		No		Common	
📋 Local Traffic		0	nightl	ly56-4			172.31.8.130		No		Common	
Network Map		Enable	Disable	orce Offline Delete.)							
Virtual Servers	16											
Policies												
Profiles	÷											
iRules												
Pools		Pool List	0									
Nodes	-	Statistics		2								
Nodes	0											
Monitors	U.S.											
Monitors Traffic Class	0											

Use the following settings for the pool.

Configuration (Basic)

Name: pool-nightly56-impala-21000 Description: impalad frontend - impala-shell, beeswax, v1.2 ODBC Health Monitors: tcp

Resources

Load Balancing Method: Least Connections (node)

🕨 😐 🥥 🕼 BIG-IP® - f5-eng.mt	tv.cloud ×		Leg
÷ → C 🕼 https://f5-eng	.sjc.cloudera.com/xui/	숬 😔 🗔 🔾 🛤 🞑	۳
Hostname: f5-eng.mtv.cloudera.com IP Address: 172.17.190.10	Date: Mar 23, 2016 User: alex Tirne: 1:50 PM (PDT) Role: Res	an Partition: Common 3	out
ONLINE (ACTIVE) Standaione Provisioning Warning			
Main Help About	Local Traffic » Pools : Poo	ol List » New Pool	
Statistics	Configuration: Basic	•	
👸 iApps	Name	pool-nightly56-impala-21000	
Wizards	Description	impalad frontend - impala-shell, beeswax, v1.2 ODBC	
🕥 dns	Hadith Manifere	Active Available //Common https://443 top <<< https://doi.org/10.1016/j.j.doi.0016/	
Local Traffic	Health Monitors	inband tcp_half_open udp	
Virtual Servers	Resources		
Policies	Load Balancing Method	Least Connections (node)	
Profiles	Priority Group Activation	Disabled	
iRules	•	New Node New FORN Node Node List	
Pools	•	Address: nightly56-2 (172.31.8.124)	
Nodes	→ 1	Service Port: 21000 Select \$	
Monitors	Now Momborn	Add	
Traffic Class	New Members		
Address Translation	→		
Acceleration		Edit Delete	

In the **New Members** section:

- 1. Click the **Node List** radio button.
- 2. Select the first node, assign the **Service Port** as 21000.
- 3. Click Add.
- 4. Repeat for each Impala node.

BIG-10- 15-0			
• → C ▲ https://f5	-eng.sjc.c	loudera.com/xui/	(1) 🖾 🐨 🚺 🤟
lostname: f5-eng.mtv.cloudera.o P Address: 172.17.190.10	com Dai Tim	te: Mar 23, 2016 User: alexin ne: 1:51 PM (PDT) Role: Resol	n Partition: Common 🗘 Log out
ONLINE (ACTIVE) Standalone Provisioning War) ning		
Main Help	About	Local Traffic » Pools : Pool	I List » New Pool
Statistics		Configuration: Basic \$	
iApps		Name	pool-nightly56-impala-21000
Wizards		Description	impalad frontend - impala-shell, beeswax, v1.2 ODBC
DNS		Health Monitors	Active Available /Common tcp << Available https: 443 https. head 15 inband
Network Map			<pre>>>> tcp_half_open udp</pre>
Virtual Servers	F	Resources	
Policies	÷	Load Balancing Method	Least Connections (node)
Profiles	÷	Priority Group Activation	Disabled
iRules) - F		New Mede ONew FODM Mede Chief
Pools	F		Address: nightly56-4 (172.31.8.130)
Nodes	÷.		Service Port: 21000 Select \$
Monitors	÷	blow Mambara	Add
Traffic Class	÷	New Members	R:1 P:0 C:0 nightly56-2 172.31.8.124 (21000 R:1 P:0 C:0 nightly56-3 172.31.8.126 (21000
Address Translation	÷		R:1 P:0 C:0 nightly56-4 172.31.8.130 :21000

Click Finished.

Follow the same process for the next pool.

Port 21050

Local Traffic > Pools > Pool List > Create (green plus)

Use the following settings for the pool.

Configuration (Basic)

Name: pool-nightly56-impala-21050 Description: impalad frontend - JDBC, v2.0+ of Cloudera ODBC Health Monitors: tcp

Resources Load Balancing Method: Least Connections (node)

In the New Members section

- 1. Click the **Node List** radio button.
- 2. Select the first node, assign the **Service Port** as 21050.
- 3. Click Add.
- 4. Repeat for each Impala node.

Click Finished.

You now see the two pools, each with three members.



Create the Virtual Servers

A Virtual Server is the client-facing side of the load balancer—the IP and port that the client connects to for a particular service. Virtual Servers are backed by one or more pools of backend nodes; in most cases, the client is unaware of the backend nodes.

Port 21000

Local Traffic > Virtual Servers > Virtual Server List > Create (green plus)

	eng.mtv.clou	× \					Le
• -> C' 🔒 https://f8	s-eng.sjc.c	loudera.com/xui/		2 0 k	0	{=}	•
Hostname: f5-eng.mtv.cloudera P Address: 172.17.190.10	.com Da Tin	le: Mar 24, 2016 User: alexm le: 7:35 AM (POT) Role: Resource Administ	trator	Partition: Comm	on	Log o	ut
CONLINE (ACTIVE Standalone Provisioning Wa	E) ming						
Main Help	About	Local Traffic » Pools : Pool List					
Statistics		🔅 🗸 Pool List Statistics					
JApps		nightly	Search Reset Search			Creat	e
Wizards		Status A Name		Application	Members	+ Partition / P	ath
D DUG		pool-nightly56-impala-21	000		3	Common	
DNS		pool-nightly56-impala-21	050		3	Common	
Local Traffic		Delete					
Network Map							
Virtual Servers		Virtuai Server List 🛛 🔂					
Virtual Servers Policies	Р. 1-	Virtual Server List O Virtual Address List					
Virtual Servers Policies Profiles	+ + 	Virtual Server List Virtual Address List Statistics					
Virtual Servers Policies Profiles iRules	+ 	Virtual Server List Virtual Address List Statistics					
Virtual Servers Policies Profiles iRules Pools	4 	Virtual Server List Virtual Address List Statistics					
Virtual Servers Policies Profiles iRules Pools Nodes	4 	Virtual Server List Virtual Address List Statistics					
Virtual Servers Policies Profiles iRules Pools Nodes Monitors		Virtual Server List Virtual Address List Statistics					
Virtual Servers Policies Profiles iRules Pools Nodes Monitors Traffic Class	* * * * *	Virtual Server List Virtual Address List Statistics					

Here, you set all the required properties (denoted with blue bars to the left of the field name) as well as some optional properties.

General Properties

Name: vs-dev-impala-21000 Type: Performance (Layer 4) Destination Address: <Self IP> Service Port: 21000

The Self IP is the address that clients communicate with. Your FQDN should resolve to this address.

You can change the service port, but if you leave it as the default for Impala, you do not need to make changes to the clients.

BIG-IP® - f5-eng.mtv.clou	< × \	Leç
> C 🕼 https://f5-eng.sjc.c	loudera.com/xui/	☆ 😔 🔜 🧿 🔝 🔘
Hostname: f5-eng.mtv.cloudera.com Da IP Address: 172.17.190.10 Tin	te: Mar 24, 2016 User: alexm ne: 8:04 AM (PDT) Role: Resource	Administrator Partition: Common 🗘
ONLINE (ACTIVE) Standalone Provisioning Warning		
Main Help About	Local Traffic » Virtual Servers	: Virtual Server List » New Virtual Server
Statistics	General Properties	
iApps	Name	vs-dev-impala-21000
Wizards	Description	
	Туре	Performance (Layer 4) \$
	Source Address	
Local Traffic	Destination Address	10.17.84.12
Network Map	Service Port	21000 Other: \$
Virtual Servers	Notify Status to Virtual Address	
Profiles	State	Enabled \$
iRules	Configuration: Basic \$	
Pools	Protocol	(TCP +)
Nodes	Protocol Profile (Client)	fastL4 🔶
Monitors 📀	HTTP Profile	None
Traffic Class	VLAN and Tunnel Traffic	All VLANs and Tunnels \$
Address Translation	Source Address Translation	Auto Map 1

In addition, you must configure Source Address Translation and the default pool.

Configuration (Basic)

Source Address Translation: Auto Map

Resources

Default Pool: pool-nightly56-impala-21000

- → G 5	k https://f5-eng.sjc.	cloudera.com/xui/	
Profiles	•	State	Enabled \$
iRules	÷	Configuration: Basic \$	
Pools	Þ	Protocol	(TCP ¢)
Nodes	Þ	Protocol Profile (Client)	fastL4 ♦
Monitors	÷	HTTP Profile	None
Traffic Cla	ss 🕀	VLAN and Tunnel Traffic	All VLANs and Tunnels \$
Address T	ranslation +	Source Address Translation	Auto Map ¢
Acceleration	1	Acceleration: Basic \$	
Access Poli	cy	Rate Class	None
		SPDY Profile	None \$
Device Mana	agement	Basauraaa	
Network		Resources	Enabled Available
System		iRules	(<
		Default Pool +	pool-nightly56-impala-21000
		Default Persistence Profile	None
		Fallback Persistence Profile	None
		Cancel Repeat Finished	

3. Click Finished.

Port 21050

Local Traffic > Virtual Servers > Virtual Server List > Create (green plus)

General Properties

Name: vs-dev-impala-21050 Type: Performance (Layer 4) Destination Address: <Self IP> Service Port: 21050

Configuration (Basic)

Source Address Translation: Auto Map

Hue requires persistent (or "sticky") sessions, meaning its requests need to be serviced by the same node when possible. So in addition to the default pool, you also configure a Persistence Profile for this Virtual Server. Without persistent sessions, Hue can be disconnected from long-running queries.

If you do not use Hue, leave Default Persistence Profile set to None.

Resources

Default Pool: pool-nightly56-impala-21050

Default Persistence Profile: source_addr

Click Finished.

Configure Impala

The following guidance mirrors the documentation for using a load balancer with Impala.

In Cloudera Manager, navigate to the Impala service, select the **Configuration** pane, then search for "balancer" to find the **Impala Daemons Load Balancer** parameter. The load balancer should be specified in *host:port* format, where host is your virtual server's FQDN and port. These values are used by Cloudera Manager and are also passed to Hue.

In the example, the self-IP FQDN is f5-demo-12.sjc.cloudera.com and the TCP port required is 21050. Type f5-demo-12.sjc.cloudera.com:21050 in the field.

● ● / (È BIG-IP®)	×C	IMPALA-1 - Cloudera Mana ×					Le
· → C 🗋 nightly56-1.gce	.cloudera.co	m:7180/cmf/services/13/co	nfig#filterfreeText=balancer		☆ ⊗ 5⊐	• =>	0 👳
oudera manager				🖀 🔓 Sear	rch (Hotkey: /)	Support	admin
	notion - A	udite Charte – Reekun	- Administration -				
Clusters + Hosts Diagno	DSUCS A	udits Charts + Backup					
IMPALA-1 (Cluste	r 1)				Apr	il 8, 2016, 1:0	5 PM PDT
tatus Instances Configu	ration Cor	mmands Audits Querie	s Charts Library Best Practices	s Web UI 🔻	Quick Links -	[Actions -
Configuration IMPALA-1 o	n Cluster 1 🔻	Selected Filters: × bal	ancer Switch to th	e classic layout	Role Groups	History and	Rollback
Filters	Clear	Reason for change	Save Changes	Edited Value			
♥ SEARCH	Clear						
balancer						Show All Des	scriptions
• STATUS		Impala Daemons Load	Impala Daemon Default Group C				8
B Error	0	Balancer	5-demo-12.sjc.cloudera.com:2105	iO			-
A Warning	0						
Edited	1	Suppress Parameter	Impala Daemon Default Group 🗌				8
Non-default	1	Validation: Impala					
Has Overrides	0	Balancer					
♥ SCOPE							
IMPALA-1 (Service-Wide)	0						
Impala Catalog Server	0						
Impala Daemon	2						
Impala Llama ApplicationMaster	0						

If the **Impala Daemons Load Balancer** parameter is specified and Kerberos is enabled, Cloudera Manager adds a principal for 'impala/<load_balancer_host>@<realm>' to the keytab for all Impala daemons. No additional configuration is required for Kerberos.

Click Save.

Note: If you are using Hue, you must restart the Hue service for the change to take effect.

Verification

Local Traffic Network Map

After everything is created, view the **Network Map** to see if everything is mapped correctly (**Local Traffic > Network Map**). You should see two virtual servers (on ports 21000 and 21050), each with a pool of backend Impala nodes. In this configuration, the port of the Virtual Server should match the ports of the pool nodes.



Throughout verification and operation, you can watch connection statistics in the F5 UI.

Statistics > Module Statistics > Local Traffic

Display Options Statistics Type: Pools Data Format: Normalized

If you have many pools, enter a search term to limit the number of pools displayed.

impal	а		Search Reset Search	B	its	Pac	kets	Co	onnections	
	Status	Pool/Member	Partition / Path	In	Out	In	Out	Current	Maximum	Total
	0	pool-nightly56-impala-21000	Common	15.6K	32.7K	22	18	0	1	1
	0	nightly56-2:21000	Common	0	0	0	0	0	0	0
	0	nightly56-3:21000	Common	15.6K	32.7K	22	18	0	1	1
	0	nightly56-4:21000	Common	0	0	0	0	0	0	0
	0	pool-nightly56-impala-21050	Common	187.1K	753.6K	258	190	1	5	5
	0	nightly56-2:21050	Common	7.5K	1.7K	10	4	0	2	2
	0	nightly56-3:21050	Common	3.7K	896	5	2	0	1	1
	0	nightly56-4:21050	Common	175.8K	750.9K	243	184	1	2	2
Rese	ht.									

JDBC

You can verify JDBC operation using IntelliJ and ImpalaJDBC (or other client) by connecting to jdbc:impala://f5-demo-12.sjc.cloudera.com:21050/ and running queries.

Impala 4.1 - test

Database: Impala/2.4.0-cdh5.6.x Driver name: ImpalaJDBC Driver version: 02.05.29.1048 JDBC version: 4.1 Case sensitivity: MIXED (quoted: EXACT)

Connection successful

ODBC

ODBC can be verified by a client that utilizes the Simba Impala ODBC driver.

Impala Shell

Launch three terminal sessions using screen or tmux, and then launch impala-shell in each:

OK

```
> impala-shell -i f5-demo-12.sjc.cloudera.com
Starting Impala Shell without Kerberos authentication
Connected to 10.17.84.12:21000
Server version: impalad version 2.4.0-cdh5.6.x RELEASE (build
85c0772d5455fe4ee5fe1d5fa39d162ad3c9e52f)
*****
                                           *************************************
Welcome to the Impala shell. Copyright (c) 2015 Cloudera, Inc. All rights reserved.
(Impala Shell v2.4.0-cdh5.6.x (85c0772) built on Mon Mar 21 07:08:54 PDT 2016)
Run the PROFILE command after a query has finished to see a comprehensive summary
of all the performance and diagnostic information that Impala gathered for that
query. Be warned, it can be very long!
[f5-demo-12.sjc.cloudera.com:21000] > show tables;
Query: show tables
+----+
name
+----+
customers
sample_07
| sample_08 |
+---+
Fetched 3 row(s) in 0.50s
[f5-demo-12.sjc.cloudera.com:21000] > quit;
```

You should see one client connected to each pool member.

Hue

Hue inherits the **Impala Daemons Load Balancer** setting <u>when set within Cloudera Manager</u> (as described above). No other configuration is required, although Hue must be restarted for the setting to take effect.

Hue makes a connection and uses it for a period of time. It communicates to the impalad to determine the real host (the real hostname is visible on the Session tab of the <u>Impala query page</u>). After a few minutes of inactivity, LTM reaps the connection; when a fresh query is requested, Hue makes a new connection.

If you enabled persistence, you should see subsequent Hue query requests (and incrementing traffic) for a single pool.

Supplemental Configuration

Now that you are familiar with the basics, here are other configuration options that may be relevant to your installation.

Increasing Idle Timeouts

By default the LTM reaps idle connections after five minutes and persistence is limited to three minutes. For longrunning queries, that might not be ideal. The following procedures set the effective idle timeout to one hour. In general, you set the value to the maximum query duration of the workload. Keep in mind that network resources on the LTM are shared, and that Impala requirements must be weighed against the requirements of other LTM users.

Create a Custom Protocol Profile

By default, a Performance (Layer 4) Virtual Server uses the fastL4 protocol profile. If you just configure that profile to have a longer idle timeout, all other users of the profile are also affected. Instead, you create a new profile based on fastL4.

Local Traffic > Profiles > Protocol > Fast L4 > Create (green plus)



General Properties

Name: fastL4_1h_idle_timeout Parent Profile: fastL4

Because new profiles inherit all the properties of the parent profile, you should avoid changing the configuration on the default profiles. For each setting that you want to override, check the box to the right and make modifications.

<u>Settings</u>

Idle Timeout: Specify, 3600

•••	BIG-IP® - f5-eng.m	tv.cloud ×		Le
· -> 1	C https://f5-eng	.sjc.cloudera.com/xui/	F 🔂 😔 🖵 🤇) 🔚 (=) 🚺 🤨
Hostname P Address	a: f5-eng.mtv.cloudera.com as: 172.17.190.10	Date: Mar 30, 2016 User: alexm Time: 7:44 AM (PDT) Role: Resource	e Administrator Partition: Common	Cog out
6	ONLINE (ACTIVE) Standalone Provisioning Warning			
Main	Help Abou	Local Traffic » Profiles : Prot	ocol : Fast L4 » New Fast L4 Profile	
🚽 Stati:	istics	General Properties		
iApp:	15	Name	fastL4_1h_idle_tim	
📋 Wiza	ards	Parent Profile	fastL4 \$	
		Settings		Custom 🗌
~		Reset on Timeout	C Enabled	
Loca	al Traffic	Reassemble IP Fragments		0
Ne	etwork Map	Idle Timeout	Specify \$ 3600 seconds	
Vi	irtual Servers	TCP Handshake Timeout	Specify ¢ 5 seconds	
	Policies	Maximum Segment Size	Disabled A	
	Profiles	Override		
1	iRules	, PVA Offload Dynamic	Enabled \$	
1	Pools	PVA Dynamic Client Packets	1 Packets	0
1	Nodes	PVA Dynamic Server Packets	0 Packets	0
M	Ionitors	IP ToS to Client	Pass Through \$	
Tr	raffic Class	IP ToS to Server	Pass Through 1	
Ac	ddress Translation	Link OoS to Client	Dass Through A	0
		Link Qos to Client	Pass mough -	

Scroll down and click **Finished**.

The new profile should appear.

🖲 🤍 🌀 🖉 🕼 🖉 🖉 🖉	ng.mtv.cloud	×									Le
⊢ → C 🔒 https://f5	-eng.sjc.cl	oudera.com/xui/							₩2 8		{=} 🚺 🤠
Hostname: f5-eng.mtv.cloudera. IP Address: 172.17.190.10	com Dati Tim	в: Mar 30, 2016 L e: 7:46 AM (PDT) F	lser: alexn ole: Reso	n urce Administrat	or				Partition:	Common S	Log out
ONLINE (ACTIVE Standalone Provisioning War) ning										
Main Help	About	Local Traffic » Pr	ofiles : Pr	otocol : Fast I	4						-
Statistics		👳 🗸 Services		Content		Persistence	Protocol	✓ SSL		- Authenticatio	n -
		Message Routing	- Other								
(B) IApps											
Wizards		·			Search						Create
		✓ A Name							Application	Parent Profile	+ Partition / Path
		iii fastL4								(none)	Common
Local Traffic		fastL4_1h_idle_	timeout							fastL4	Common
Network Map		Delete									
Virtual Servers											
Policies	10										
Profiles											
iRules											
Pools	- F										
Nodes	1.5										
	0										
Monitors	0										
Monitors Traffic Class	•										

Apply a Custom Protocol Profile

Once the custom protocol profile has been created, configure the Virtual Server to use it instead of the default. Navigate to the Virtual Server you want to modify, and then make the following change. You must apply the profile to both port 21000 and 21050 Virtual Servers.

Configuration (Advanced)

Protocol Profile (Client): fastL4_1h_idle_timeout

BIG-IP® - f5-eng.mtv.clo	ud ×	Lega
$- \rightarrow C'$ \triangleq https://f5-eng.sjc.	cloudera.com/xui/	F 🕁 🥝 🗔 🐚 🚺 🧔
Hostname: f5-eng.mtv.cloudera.com D IP Address: 172.17.190.10 Ti	ate: Mar 30, 2016 User: alexm me: 7:51 AM (PDT) Role: Resource	e Administrator Partition: Common 🗘 🕍 Log out
ONLINE (ACTIVE) Standalone Provisioning Warning		
Main Help About	Local Traffic » Virtual Servers	: Virtual Server List » vs-dev-impala-21000
Statistics	🚓 👻 Properties Reso	ources Statistics 🗵
IApps	General Properties	
Wizards	Name	vs-dev-impala-21000
	Partition / Path	Common
DNS	Description	
🛐 Local Traffic	Туре	Performance (Layer 4) \$
Network Map	Source Address	0.0.0.0/0
Virtual Servers	Destination Address	10.17.84.12
Policies	Service Port	21000 Other: \$
Profiles	Notify Status to Virtual Address	
iRules	Availability	Available (Enabled) - The virtual server is available
Pools	Syncookie Status	Off
Nodes	State	Enabled \$
Monitors 📀		
Traffic Class 📀	Configuration: Advanced \$	
Address Translation	Protocol	
	Protocol Profile (Client)	fastL4_1h_idle_timeout

Scroll down and click Update.

Repeat for each Virtual Server you want to have an extended idle timeout.

Create a Custom Persistence Profile

Create custom persistence profiles in the same way as protocol profiles.

Local Traffic > Profiles > Persistence > Create (green plus)



General Properties

Name: source_addr_1h_idle_timeout Persistence Type: Source Address Affinity Parent Profile: source_addr

Because new profiles inherit all the properties of the parent profile, you should avoid changing the configuration on the default profiles. For each setting that you want to override, check the box to the right and make modifications.

Settings

Timeout: Specify, 3600

→ C A https://f5-eng.sjc.cl	oudera.com/xui/		F 🔂 😔 🖬 🔾	(=) 🖸 😳
Hostname: f5-eng.mtv.cloudera.com Date P Address: 172.17.190.10 Time	:: Mar 30, 2016 User: alexm :: 8:10 AM (PDT) Role: Resourc	e Administrator	Partition: Common	C Log out
ONLINE (ACTIVE) Standalone Provisioning Warning				
Main Help About	Local Traffic » Profiles : Pers	Istence >> New Persistence Profile		
J Statistics	General Properties			
Apps IApps	Name	source_addr_1h_ic		
Wizards	Persistence Type	Source Address Affinity		
DNS	Parent Profile	(source_addr \$		
	Configuration			Custom 🗌
Di Local Traffic	Match Across Services	0		0
Network Map	Match Across Virtual Servers	0		0
Virtual Servers	Match Across Pools	0		0
Policies	Hash Algorithm	Default \$		
Profiles	Timeout	Specify \$ 3600 seconds		0
iRules	Mask	None 1		0
Pools	Man Provies			0
Nodes	Override Connection Limit			0
Monitors 💮	Stando Somosion Linit			
Traffic Class	Cancel Repeat Finished			

Click Finished.

Apply a Custom Persistence Profile

Navigate to the Virtual Server for Impala port 21050 (in the example, vs-nightly56-impala-21050). Click the **Resources** tab and make the following change.

Load Balancing

Default Persistence Profile: source_addr_1h_idle_timeout

Image: Second state of the se	udera.com/xuí/ Mar 30, 2016 User, alexm. 8:15 AM (PDT) Role: Resource Local Traffic >> Virtual Servers ☆ → Properties Reso Local Balancing Default Pool Default Persistence Profile Fallback Persistence Profile	e Administrator : Virtual Server List >> vs-dev-Impale-210 purces Statistics (2) (pool-nightly56-impale-21050 (2) (source_addr_1h_idle_timeout (2) None (2)	950	Partit	On: Common		out
Hostname: R6-eng.mtv.clouders.com IP Address: 172.17.190.10 IP Address: 172.17.190.10 I ONLINE Standalone Provisioning Warning Main Heip About II Statistics I Apps VWizards DNS Local Traffic Network Map Virtual Servers II	Mar 30, 2016 8:15 AM (PDT) User: alexm Role: Resource Local Traffic >> Virtual Servers ☆ ▼ Properties Reso Local Balancing Default Pool Default Persistence Profile Fallback Persistence Profile	e Administrator : Virtual Server List » vs-dev-impala-210 purces Statistics pool-nightly56-impala-21050 source_addr_1h_idle_timeout None	950	Parit	on: Common	Log	out
ONLINE (ACTIVE) Standalone Provisioning Warning Main Help About Statistics IApps Wizards DNS Local Traffic Network Map Virtual Servers	Local Traffic » Virtual Servers Properties Reso Load Balancing Default Pool Default Persistence Profile Fallback Persistence Profile	Virtual Server List >> vs-dev-Impala-210 purces Statistics pool-nightly56-impala-21050 source_addr_1h_idle_timeout None e	J50				
Main Help About Statistics Apps Wizards DNS Local Traffic Network Map Virtual Servers	Local Traffic » Virtual Servers Properties Reso Load Balancing Default Pool Default Persistence Profile Failback Persistence Profile	Virtual Server List » vs-dev-Impala-210 vurces Statistics pool-nightly56-impala-21050 pool-nightly56-impala-21050 None	050				
Statistics Apps Vizards DNS Local Traffic Network Map Virtual Servers I I I I I I I I I I I I I I I I I I	Properties Reso Coad Balancing Default Pool Default Persistence Profile Fallback Persistence Profile	purces Statistics () pool-nightly56-impala-21050 ¢ source_addr_1h_idle_timeout ¢ None ¢					
i Apps Vizards DNS Local Traffic Network Map Virtual Servers iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	Load Balancing Default Pool Default Persistence Profile Fallback Persistence Profile	pool-nightly56-impala-21050 source_addr_1h_idle_timeout None					
Virtual Servers	Default Pool Default Persistence Profile Fallback Persistence Profile	pool-nightly56-impala-21050 source_addr_1h_idle_timeout None					
DNS Clocal Traffic Network Map Virtual Servers	Default Persistence Profile Fallback Persistence Profile	source_addr_1h_idle_timeout \$ None \$					
Local Traffic Network Map Virtual Servers	Fallback Persistence Profile	None \$					
Virtual Servers							
Virtual Servers	Update						
Virtual Servers + iF							
	iRules					Manag	ge
Policies	Name						
Profiles •	No records to display.					 	
iRules •							
Pools F							
Nodes							
Monitors (+)							
Traffic Class							
Address Translation							

Click Update.

Kerberos

If you followed <u>Impala configuration instructions</u>, no additional configuration is required for Kerberos.

TLS/SSL

When TLS/SSL is enabled for Impala, the client application—whether impala-shell, Hue, or something else expects the certificate common name (CN) to match the hostname that it connected to. With no load balancer, the hostname and certificate CN are both that of the impalad instance. However, with a load balancer, the certificate presented by the impalad instance does not match the load balancer "front-end" hostname.

If you try to load-balance a TLS/SSL-enabled Impala installation without additional configuration, you see the following error when a client attempts to connect to the load balancer hostname:

```
Hostname we connected to "f5-demo-12.sjc.cloudera.com" doesn't match
certificate provided commonName "nightly57-kerberized-4.gce.cloudera.com"
(code THRIFTTRANSPORT): TTransportException(u'Hostname we connected to "f5-
demo-12.sjc.cloudera.com" doesn\'t match certificate provided commonName
"nightly57-kerberized-4.gce.cloudera.com"',)
```

You can configure an LTM in several ways to load-balance Impala:

- Client/Server SSL
- <u>TLS/SSL Passthrough</u>
- TLS/SSL Offload

Client/Server SSL

In this configuration, the LTM presents an SSL certificate to the client, decrypts the client request, then reencrypts the request before sending it to a backend impalad instance in the pool. This is often referred to as client/server SSL. At no point is the request or resulting payload unencrypted in transit. The client and server certificates can be managed separately, which can be convenient.

Modify the configuration process by applying a Client and Server profile. To quote the BIG-IP LTM manual on <u>Managing SSL Traffic</u>:

A **Client profile** is a type of traffic profile that enables the BIG-IP system to accept and terminate any client requests that are sent by way of a fully SSL-encapsulated protocol. A **Server profile** is a type of profile that enables the BIG-IP system to initiate secure connections to a target web server.

To configure, perform the following tasks from the BIG-IP LTM documentation:

- 1. Install a key/certificate pair on the BIG-IP for terminating client-side secure connections.
- 2. Configure a client SSL profile (use the default server SSL profile).
- 3. Associate the profile with a virtual server.

These steps assume the pools and virtual servers have already been created following earlier guidance and that Impala has been configured for TLS/SSL.

Create/Import Certification and Key

Follow the BIG-IP documentation to <u>create or import an SSL certificate and corresponding key</u>. The certificate common name (CN) should match the load balancer hostname that you will use for Impala traffic.

If your certificate is self-signed, append the certificate to the Impala TLS/SSL CA Certificate file. You can find the file in Cloudera Manager by searching for **ssl_client_ca_certificate** in the Impala service configuration. The append is straightforward:

```
impalad-host$ cat certificate.crt >> /etc/certs/truststore.pem
```

If your certificate is signed by a CA already listed in the file, no further action is required. If not, append the client CA certificate to the file.

Create a Client SSL Profile

Local Traffic > Profiles > SSL > Client > Create (green plus)

Name: f5-demo-12 Parent Profile: clientssl

Configuration (Basic)

Click the small blue checkbox to the right, and then complete the following fields:

Certificate: f5-demo-12.sjc Key: f5-demo-12.sjc Chain: None

In the example, my certificate (**f5-demo-12.sjc** in the BIG-IP SSL Certificate List) is self-signed, so there is no certificate chain. If you have a certificate chain or passphrase, make the appropriate selections.

Click Add.



Scroll down and click Finished.

Modify the Virtual Servers

To use SSL profiles, you cannot use the Performance (Layer 4) type. You must change your existing Virtual Servers to use the Standard type and the SSL profile you just created.

General Properties

Type: Standard

Configuration: Basic

SSL Profile (Client): f5-demo-12 SSL Profile (Server): serverssl

🔴 🔍 🌑 🌀 BIG-IP® - f5-eng.mtv.cloud	×	Leç	gac
← → C 🔒 https://f5-eng.sjc.c	loudera.com/xui/	F 🔂 🥹 🔜 🔍 💟 🤨	=
Local Traffic	Туре	Standard	
Network Map	Source Address	0.0.0/0	
Virtual Servers	Destination Address	10.17.84.12	
Policies	Service Port	21050 Other: \$	
Profiles	Notify Status to Virtual Address		
iRules	Availability	Available (Enabled) - The virtual server is available	
Pools	Syncookie Status	Off	
Nodes	State	Enabled \$	
Monitors 📀	Configuration: Basic \$		1
Address Translation	Protocol	(TCP \$	
	Protocol Profile (Client)	(tcp •	
Acceleration	Protocol Profile (Server)	(Use Client Profile)	
Access Policy	HTTP Profile	None	
Device Management	FTP Profile	None 🛊	
Device management	RTSP Profile	None \$	
Network		Selected Available	
System	SSL Profile (Client)	f5-demo-12 < clientssl-insecure-compatible cloudera-com-wildcard-sha2 crypto-serve-default-clientssl wom-default-clientssl	
	SSL Profile (Server)	Selected Available //Common serverssl /	

Scroll down and click Update.

TLS/SSL Passthrough

In this configuration, TLS/SSL is terminated on the backend Impala instances; traffic is still encrypted end-to-end. No client-side SSL work is done by the LTM, and the encrypt/decrypt load on the backend hosts is not reduced.

- 1. Configure Impala with TLS/SSL.
- 2. Issue impalad certificates with a Subject Alternate Name (SAN) matching the frontend load balancer. See the example below.
- 3. Create pools and virtual servers as usual.
- 4. Configure Impala in Cloudera Manager as usual.

For example, if an impalad instance hostname is datanode05.sjc.cloudera.com, and the load balancer Virtual Server hostname is f5-impala.sjc.cloudera.com:

CN: datanode05.sjc.cloudera.com SAN: datanode05.sjc.cloudera.com f5-impala.sjc.cloudera.com Note: If the load balancer hostname is changed, *all* impalad certificates must be regenerated and redistributed, and the cluster must be restarted.

TLS/SSL Offload

In this configuration, TLS/SSL is terminated at the load balancer, and traffic between the backend Impala instances is unencrypted. Less overhead is incurred by the backend hosts because they do not have to encrypt or decrypt Impala traffic. This configuration presumes that cluster hosts reside on a trusted network and only external client-facing communication need to be encrypted in-transit. Traffic between Hue and the load balancer is also encrypted, however.

To configure:

- 1. Configure Impala without TLS/SSL.
- 2. Create pools.
- 3. Create/import an SSL certificate and key.
- 4. Create a Client SSL profile using the certificate and key.
- 5. Create a Standard Virtual Server with Client SSL profile; do not configure a Server SSL profile.
- 6. Configure Impala in Cloudera Manager as usual.
- Configure Hue to use SSL for Impala: in Cloudera Manager, add the following text in the Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini field (substituting the path to your truststore).

```
[impala]
[[ssl]]
enabled=true
validate=false
cacerts=/etc/certs/truststore.pem
```

8. If your SSL certificate is signed by a CA already listed in the **cacerts** truststore, no further action is required. If not, append the CA certificate to the truststore file (or the certificate itself if self-signed).

You must restart Impala and Hue for the changes to take effect.

Verification

To verify TLS/SSL + Kerberos configuration, you can use Hue or impala-shell. You kinit before launching impalashell:

Run the PROFILE command after a query has finished to see a comprehensive summary of all the performance and diagnostic information that Impala gathered for that query. Be warned, it can be very long! [f5-demo-12.sjc.cloudera.com:21000] > show tables; Query: show tables +----+ name T +----+ customers sample_07 sample_08 | +----+ Fetched 3 row(s) in 0.50s [f5-demo-12.sjc.cloudera.com:21000] > quit;

Known Issues

Backend nodes visible in Hue

Even when configured to utilize a load balancer, Impala clients are often aware of the backend impalad instance they are connected to. These hostnames and ports are visible in the Hue Sessions display and other places. This is normal and does not affect operation.

BIG-IP® - f5-eng.mtv.cloud	×	IMPA	.A-1 - Clou	dera Mana × 🛃 Hue - Impala - Editor 🛛 ×					Leg	gacy
← → C	rberized	l-1.g	ce.cloude	rra.com:8889/notebook/editor?editor=14 🖒	3	0		{=}	•	≡
Sessions										×
Impala SQL										
ADDRESS										
nightly57-kerberized-3.gce.cloudera.com	:25000 🖸	3								
						× Clo	ose	C Recre	ate	
				^						
key (string) value (struct)	3								_	_
email preferences (struct <email formate<="" td=""><td></td><td>•</td><td>code</td><td>description</td><td>4</td><td>total_</td><td>emp</td><td>salary</td><td></td><td></td></email>		•	code	description	4	total_	emp	salary		
		1	00-000	All Occupations		13435	4250	40690		
		2	11-0000	Management occupations		60039	30	96150		
hbase		3	11-1011	Chief executives		29916	0	151370)	
🗅 solr		4	11-1021	General and operations managers		16554	10	103780)	
🗀 tmp		5	11-1031	Legislators		61110		33880		
🗀 user		6	11-2011	Advertising and promotions managers		36300	1	91100		
Documents		7	11-2021	Marketing managers		16524	0	113400	5	
G /		8	11-2022	Sales managers		32217	0	106790	5	
Empty directory		9	11-2031	Public relations managers		47210	1	97170		
		10	11-3011	Administrative services managers		23936	0	76370		
		11	11-3021	Computer and information systems managers		26499	0	113880)	
		12	11-3031	Financial managers		49430	0	106200	n	

Error 104: Connection reset by peer

During normal operation, an Impala client keeps a user session open to the backend impalad instance. If an impalad instance becomes unavailable unexpectedly, the Impala connection is lost.

Without a load balancer, you need to know the hostname of another impalad instance; changes to the client application may be required to reconnect to the new instance.

With a load balancer, no client reconfiguration or knowledge about other backend hosts is required, but you are not connected to a new impalad instance until a new request or query is made.

In impala-shell, you see the following:

[f5-demo-12.sjc.cloudera.com:21050] > show tables; Connection lost, reconnecting... Socket error 104: Connection reset by peer

To reestablish a session in impala-shell, CONNECT to the load balancer hostname and port.

[Not connected] > connect f5-demo-12.sjc.cloudera.com:21050; Connected to f5-demo-12.sjc.cloudera.com:21050 Server version: impalad version 2.5.0-cdh5.7.1 RELEASE (build f1464330fcc33b3709490a67a7ad1241ee983a3c) [f5-demo-12.sjc.cloudera.com:21050] >

In Hue, the error looks like this:

IGBIG-IP® - f5-eng.mtv.cloud	× \ C	IMPAI	LA-1 - Clou	dera Mans × (f) Hue - Impala - Editor ×				Lega
← → C	rberized	d-1.g	ce.cloude	ra.com:8889/notebook/editor?editor=14 🖒 😔 🖁			=}	•
HUE 🖀 Query Editors 🗸	Notebo	ooks	Data Bro	owsers v Workflows v Search Security v			■ 0 % \	0
💡 Impala Untitled Impala	Query	A	dd a des	cription [Errno 104] Connection reset by peer				
■ □ ⁴					277s defa	ult 🔻		
JQL		1	SELECT *					
🕻 🛢 default		2	FROM san	ple_07;				
Tables (4) Q Z	•	4						
m customers								
addresses (map <string,struct<street_1:st key (string)</string,struct<street_1:st 								
value (struct)								
id (int)		•	code	description	total_en	np 🍦	salary	
5/ 0		1	00-000	All Occupations	1343542	250	40690	
		2	11-0000	Management occupations	6003930	D	96150	
accumulo		3	11-1011	Chief executives	299160		151370)
		4	11-1021	General and operations managers	1655410	D	103780	,
🗅 tmp		5	11-1031	Legislators	61110		33880	
🗅 user		6	11-2011	Advertising and promotions managers	36300		91100	
ocuments		7	11 2021	Marketing managers	165240		113400	
۵ <i>۱</i>		1	11-2021		105240		113400	
mpty directory		8	11-2022	Sales managers	322170		106790	1
		9	11-2031	Public relations managers	47210		97170	
		10	11-3011	Administrative services managers	239360		76370	
		11	11-3021	Computer and information systems managers	264990		113880	
		12	11-3031	Financial managers	484390		106200)

In Hue, the your next query or Impala action results in a fresh connection being established with an operational impalad instance.

TTransportException, Could not start SASL

When TLS/SSL + Kerberos is enabled, you see the following error if you have not restarted Impala after configuring the load balancer host and port in Cloudera Manager. You see this even if you have a valid Kerberos ticket as seen through klist.

```
# impala-shell -k --ssl -i f5-demo-12.sjc.cloudera.com:21051
Starting Impala Shell using Kerberos authentication
Using service name 'impala'
SSL is enabled. Impala server certificates will NOT be verified (set --ca_cert
to change)
Error connecting: TTransportException, Could not start SASL: Error in
sasl_client_start (-1) SASL(-1): generic failure: GSSAPI Error: Unspecified
GSS failure. Minor code may provide more information (Server
krbtgt/SJC.CLOUDERA.COM@GCE.CLOUDERA.COM not found in Kerberos database)
```

Hue

To spread the load across multiple Hue instances, you can <u>configure Hue high availability</u>.

References

Cloudera Documentation

- Using Impala through a Proxy for High Availability
- User a Load Balancer with Impala
- Ports Used by Impala

F5 Documentation

- Glossary and Terms
- Manual Chapter: <u>Configuring Load Balancing Pools</u>
- <u>SOL14163</u>: Overview of BIG-IP virtual server types (11.x)
- Manual Chapter: <u>Session Persistence Profiles</u>
- Manual Chapter: Managing SSL Traffic
- <u>SOL14620</u>: Managing SSL certificates for BIG-IP systems using the Configuration utility
- Technical Article: <u>LTM: Action on Service Down</u>