Troubleshooting SFTP server issues
Last updated by | Veena Pachauri | Mar 8, 2023 at 11:59 PM PST

Contents

[SFTP] Hit SFTP server throttling

201031- [SFTP] Hit SFTP server throttling			ASC Onboard	DiagnosticsConnectorIssueByActivityId
Monday, December 18, 2017 11:22 AM				

SME	
Symptoms	Server response does not contain SSH protocol identification
Cause	ADF will create multiple connection to download from customer's sftp server in parallel. Sometimes it will hit sftp server throttling and lead to copy fail. Different server will return different error when hit throttling. Below are some known error messages indicate it could be a throttling issue.

Resolution	First need to verify this is a throttling issue. Please let customer specify max concurrent connection of SFTP dataset to 1 and re-run the Copy. If copy can pass, it means the issue is caused by throttling.
	    In this case, ADF will only establish 1 connection to download file, thus the throughput will be low.
	Second, if customer has concern of the low throughput, please let customer contact sftp admin to increase the concurrent connection count limit or add to whitelist for below IP:
	    1. If customer is using Managed IR, add Azure DataCenter IP range.
	       If customer don't want to add such large list of IP ranges into sftp server whitelist, suggest customer to install and self-hosted IR.
	    2. If customer is using Self-Hosted IR, add the machine IP which installed IR.

More Information	
Tags	SFTP,throttling

[SFTP]An existing connection was forcibly closed by the remote host

Symptom

Your pipeline with Copy activity which moves data from your SFTP server to Azure Storage failed with following error:

Failure happened on 'Sink' side. ErrorCode=SftpNetworkIssue,'Type=Microsoft.DataTransfer.Common.Shared.HybridDeliveryException,Message=Meet network issue when connect to Sftp server '172.16.104.126', SocketErrorCode: 'ConnectionReset'.,Source=Microsoft.DataTransfer.ClientLibrary.SftpConnector,''Type=System.Net.Sockets.SocketException,Message=An existing connection was forcibly closed by the remote host,Source=Renci.SshNet,

Cause

One of the reason can be following:

This error was caused because the Copy activity was run many times inside a ForEach loop, which then created many connections. And the SFTP server closed the connection because of too many connection requests sent from Azure Data Factory.

Resolution

Optimized the pipeline structure to reduce the connections created to your SFTP server.

Reference

120080523000234

[SFTP]Failed to connect to Sftp server

Issue

Failed to connect to Sftp server 'sftp.preceda.com.au'. Server response does not contain SSH protocol identification.

Root Cause

Too many attempts to block ip address.

Solution

1, Added ip address back and work as expected.

2, Please check configuration for the "Max concurrent connections" property. The recommendation made was to manually set this to 1 as well as enabling retries in order to reduce the connection numbers.

3, You should better engage the SFTP vendor team to see whether possible to raise the connection count from server side.

Reference

120080325000035

SFTP concurrent connections run issue

To many concurrent connections runs: sftp server cx owns can't handle too many concurrent connections, here is the chart for two good runs where we see relatively low connection requests, around 60 connections in 10s

```
CustomLogEvent
// failed
// | where ActivityId in ('ae4b3506-c439-4d66-ac59-77d2a0e8d868','1a08c4b9-33e1-4435-89a6-7aaa576971a7')
// succeeded
| where ActivityId in ('a7766807-1131-4b7c-b947-3e5a100f8f79', '93a21b62-872e-4497-be07-8f9d9a59e036')
| where TraceMessage contains_cs 'Sftp'
| as hint.materialized=true T
// | where Message !has_cs 'exception' | reduce by Message
| parse Message with * 'Message=' ErrorMessage ',' *
| extend Action = case
                (
                    Message has_cs 'Connect to Sftp ',
                    'connect',
                    Message has_cs 'Host Key Received from the SFTP server.',
                    'connected',
                    Message has_cs 'Try Reconnect.',
                    'retry',
                    Message has_cs 'exception',
                    'exception',
                    'NA'
                )
| summarize count() by Action, bin(PreciseTimeStamp, 10s)
| where Action != 'NA'
| render timechart
```


Similarly, we use the same(above) query to chart two failed runs on the same day. The blue dots jump to **300 concurrent connections**, and then green line (connection failure) starts to become outstanding. Red line shows successful connections. Based on the red line (successful connection) it looks to me the sftp server can handle less than 180 connections in 10s.
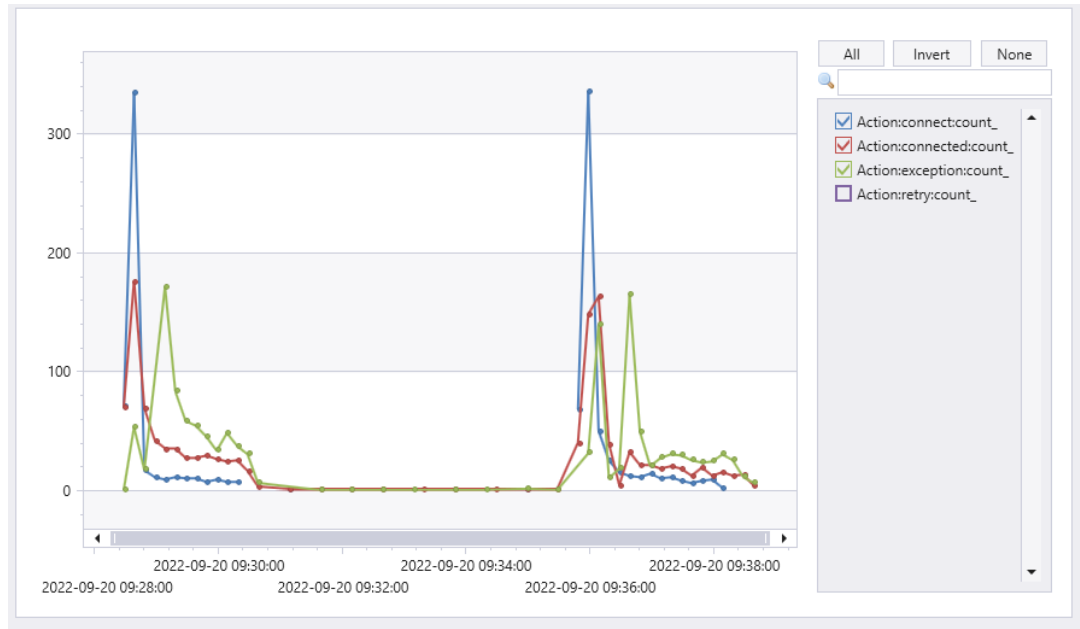
Solution:

Scenario when, SFTP server is throttled. Try to reduce the concurrent connection in copy activity or configure the SFTP server to allow more concurrent threads or scale up the server.cx should reduce the parallelism or configure sftp server to allow more concurrent connections. Increasing memory size will not help.

Reference:

2209190030001956

Created with Microsoft OneNote 2016.

How good have you found this content?