BlobEventsTrigger - Delayed pipeline execution

Last updated by | Jackie Huang | Jan 4, 2022 at 12:24 AM PST

Contents

Issue

Customer raised an issue that the pipeline execution is delayed. Pipeline triggered by BlobEventsTrigger.

BlobTrigger created time: 2020-08-20 08:**08**:17

Pipeline execution time: 2020-08-20 08:**29**:25 (delayed by approx 20 mins)

Investigation

1. Query the **TriggerRuns** table to get the blob trigger creation time and execution time:

```
cluster('adfcus').database('AzureDataFactory').TriggerRuns | union cluster('adfneu').database('AzureDataF
| where PreciseTimeStamp >= ago(7d)
| where dataFactoryName =~ "<dataFactoryName>"
| where triggeredPipelines has "<PipelineRunId>" //PipelineRunId
| extend blobTime = split(triggerEvent,' - ',1)[0]
| extend delay = PreciseTimeStamp - blobTime
//| where triggerId == "<TriggerRunId>" //TriggerRunId
//| where triggerName =~ "<TriggerName>"
| project delay,blobTime, PreciseTimeStamp, triggerEvent, dataFactoryName, triggerName, triggerId, trigge
| sort by PreciseTimeStamp asc
```

delay	blobTime	PreciseTimeStamp	triggerEvent	dataFactoryName	triggerName	triggerId	triggerType	status
00:28:12.7644664	2020-09-04T10:18:32	2020-09-04 10:46:44.7644664	BlobEventsTrigger - 2020-09-04T10:18:32	[redacted]	za_downloads_batch_tr	c8769627-61f6-4729-aea7-b50f58077380	BlobEventsTrigger	Succeeded
00:20:20.4618218	2020-09-04T10:26:32	2020-09-04 10:46:52.4618218	BlobEventsTrigger - 2020-09-04T10:26:32	[redacted]	za_downloads_batch_tr	4dc3c65b-2f71-4a70-8c8f-19f3a3277362	BlobEventsTrigger	Succeeded
00:00:16.9457725	2020-09-06T19:18:51	2020-09-06 19:19:07.9457725	BlobEventsTrigger - 2020-09-06T19:18:51	[redacted]	za_downloads_batch_tr	ad4dc8c3-37f9-435b-9d63-6ac7221289d1	BlobEventsTrigger	Succeeded
00:00:13.0923786	2020-09-04T07:44:31	2020-09-04 07:44:44.0923786	BlobEventsTrigger - 2020-09-04T07:44:31	[redacted]	za_downloads_batch_tr	14056ca1-1d19-4a99-9a13-bdc4fd90fb61	BlobEventsTrigger	Succeeded
00:00:11.9523366	2020-09-04T07:50:31	2020-09-04 07:50:42.9523366	BlobEventsTrigger - 2020-09-04T07:50:31	[redacted]	za_downloads_batch_tr	fc6b7c35-c44b-4e08-a7cc-ab169966069f	BlobEventsTrigger	Succeeded
00:00:11.8266775	2020-09-06T21:00:52	2020-09-06 21:01:03.8266775	BlobEventsTrigger - 2020-09-06T21:00:52	[redacted]	za_downloads_batch_tr	91d42e94-605c-4ee2-8931-d6c35f60e112	BlobEventsTrigger	Succeeded

   In this the PreciseTimeStamp is the time when the trigger was executed. triggerEvent is the time of the blob trigger creation. Hence, we notice there's a delay in the Pipeline Execution.

   NOTE: *Event triggers are not for real time and this should not be the expectation.*

2. Query the **AdfTraceEvent** with the triggerId:

```
cluster('adfcus').database('AzureDataFactory').AdfTraceEvent | union cluster('adfneu').database('AzureDat
| where env_time >= ago(7d)
| where TraceCorrelationId == "<triggerId>" // triggerId from TriggerRuns table
| where ComponentId == "PipelineManager"
| where Message has "InvokeEventsTriggerWorkflowAsync"
| project env_time, env_cloud_name, env_cloud_role, env_cloud_deploymentUnit, ComponentId, Message,TraceC
| sort by env_time asc
```

We noticed the following error:

Skip InvokeEventsTriggerWorkflowAsync, **Pipeline execution throttled**. If the concurrency value is set for a pipeline, a maximum of 100 runs will be queued for execution.

Root Cause

This issue can be caused due to:

1. Throttling by the LogicApps when there are large number of events received from EventTrigger. Refer the Query the rate at which the triggers are being executed section in this article.

   If you notice the rate of EventTrigger are really high, causing the LogicApps to throttle the requests, then it's a design behavior in case of EventTrigger. When we receive a large number of BlobEvents we queue them and process them in the batches of 100. This is done to prevent throttling. If customer wants to have a better performance they should think about breaking the blob trigger in to multiple triggers. Multiple approaches can be considered:

   - If the files are going in different folder separate triggers can be created base on folder name.
   - Trigger can also be created based on the pattern in which file names can be ending.

2. LogicApp workflows getting stuck causing couple of events getting delayed, and then processing the queued events quickly when it eventually runs. This is why the customer saw a delay and followed by the Pipelines getting triggered quickly. This is an issue with a partner service (LogicApps) that can happen occasionally, we are tracking this with them.

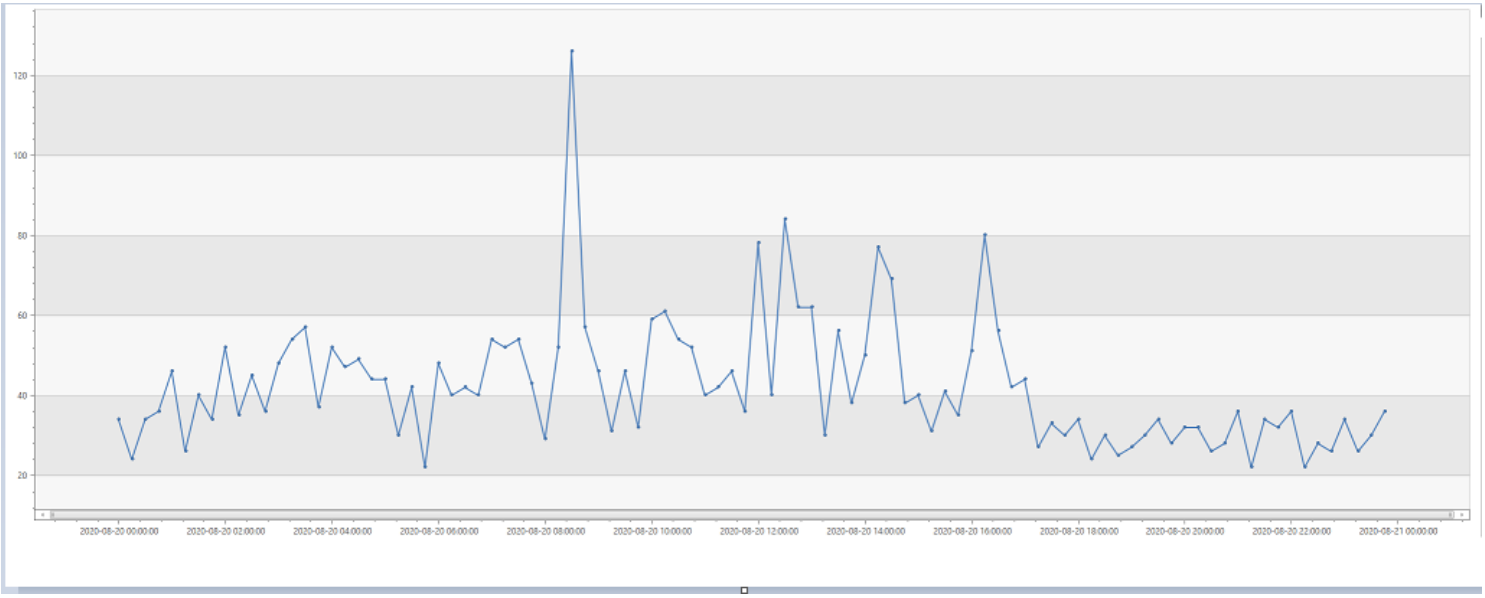
   If the customer has downstream issues when the pipelines get triggered in parallel, we suggest for them to add concurrency of 1 to the pipeline so that any runs triggered simultaneously will go into queued state and wait for the previous run to complete.

   NOTE: PG has already released a fix to resolve this issue, if customer is still seeing delayed event trigger execution, raise an Icm with the Orchestration PG.

3. Check with the customer if the EventTrigger was deactivated state and reactivated recently? If yes, than we cannot purge the events for a trigger. Even if the files were added 2 months back. As soon as the trigger will be activated again, the events queued will start getting processed. If the customer does not want the old events to get processed. Please stop this trigger and use a new trigger with new name for new events. Refer: Icm link

Query the rate at which the triggers are being executed

```
cluster('adfcus.kusto.windows.net').database('AzureDataFactory').TriggerRuns
| union cluster('adfneu.kusto.windows.net').database('AzureDataFactory').TriggerRuns
| where PreciseTimeStamp >= datetime(2020-08-20 00:00:00) and PreciseTimeStamp < datetime(2020-08-21 00:00:00)
| where subscriptionId =~ "xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"
| where triggerName   =~ "<triggerName>"
| project  PreciseTimeStamp, ['time'], triggerId, groupId, dependencyStatus, correlationId, dataFactoryId , da
| summarize count() by bin(['time'], 1h)
| render timechart
```



Additional Information:

- **Icm References:** Icm Link
- **Author:** vimals
- **Reviewer:** vimals
- **Keywords:**

How good have you found this content?