

How long the recovery PostgreSQL server

Last updated by | Hamza Aqel | Feb 2, 2022 at 3:08 AM PST

Please don't modify or move as this is part of GT , please contact haaqel@microsoft.com if needed

Use this Kusto query to check the recovery process, how long it took:

```
MonRdmsPgSqlSandbox
| where originalEventTimestamp > datetime(yyyy-mm-dd hh:mm:ss) // REPLACE THIS WITH 5 MINS BEFORE SERVER RESTART
| where originalEventTimestamp < datetime(yyyy-mm-dd hh:mm:ss) // REPLACE THIS WITH 5 MINS AFTER SERVER IS READ
| where LogicalServerName == "pg-server-name"
| where text startswith "Loaded package: [PostgreSQL]"
or text contains "CheckpointPG"
or text contains "WaitOneWithLog"
or text contains "archive"
or text contains "BACKEND_TYPE_ALL"
or text contains "exit"
or text contains "recovering "
or text startswith "SbsProvider::IsSbsXioStorage: Is SBS"
or text startswith "SbsProvider::IsSbsPfsStorage: Is SBS_PFS"
or text contains "PostgresLauncher: DoCheckpointPG: PG checkpoint in progress"
or text contains "PostgresLauncher: DoCheckpointPG: Time taken in PG checkpoint is"
or text contains "PostgresLauncher: Warning: DoCheckpointPG: PG checkpoint failed"
or text contains "PostgresLauncher: DoCheckpointSBS: SBS checkpoint in progress"
or text contains "PostgresLauncher: DoCheckpointSBS: Time taken in SBS checkpoint is"
or text contains "PostgresLauncher: ShutdownWithSFIntegration: SIGINT has been sent"
or text contains "PostgresLauncher: ShutdownWithSFIntegration: SIGQUIT has been sent"
or text contains "Parallel redo is started for database 'SBS'"
or text contains "Parallel redo is shutdown for database 'SBS'"
or text contains "Recovery completed for database SBS"
or text contains "database system was not properly shut down"
or text contains "entering standby mode"
or text contains "redo starts at"
or text contains "consistent recovery state reached at"
or text contains "database system is ready to accept read only connections"
or text contains "FATAL: could not connect to the primary server"
or text contains "started streaming WAL from primary"
or text contains "trigger file found: " // PromoteTriggerFile.txt
or text contains "redo done at"
or text contains "checkpoint starting: end-of-recovery immediate"
or text contains "database startup complete in"
or text contains "database system is ready to accept connections"
or text contains "SqlSocketDuplicateProvider::DB engine is ready"
or text contains "Listen to port" and text contains "at checkpoint"
or tolower(text) contains "role" and tolower(text) contains "orcasql" and tolower(text) !contains "Wait on"
| order by originalEventTimestamp asc
| project originalEventTimestamp, NodeName, process_id, text=trim("((\r)?(\n)?)*", text)
```

originalEventTimestamp	nodeName	processId	text
2020-07-10 09:00:07.1669643	DB.15	110112	OrcasqlInstanceProvider: FabricChangeRole (currentRole = Primary, newRole = ActiveSecondary, dataLossNumber = 01D62D4FCE9F4164, configurationNumber = 0000000500000000) start
2020-07-10 09:00:07.1669643	DB.15	110112	OrcasqlInstanceProvider: Write to pipe newRole = 4
2020-07-10 09:19:08.3653293	DB.15	110112	PostgresLauncher: ShutdownWithSfIntegration: SIGQUIT has been sent to PG process by Orcas PG Shared Memory. pid=36, name=36, PgKillByMem return=0
2020-07-10 09:19:58.0898331	DB.15	110112	OrcasqlInstanceProvider: FabricChangeRole (currentRole = Primary, newRole = ActiveSecondary) failed with 80004005.
2020-07-10 09:19:58.2484054	DB.15	110112	OrcasqlInstanceProvider: FabricChangeRole (currentRole = Primary, newRole = ActiveSecondary, dataLossNumber = 01D62D4FCE9F4164, configurationNumber = 0000000500000000) start
2020-07-10 09:19:58.2484188	DB.15	110112	OrcasqlInstanceProvider: Write to pipe newRole = 4
2020-07-10 09:19:58.2484339	DB.15	110112	OrcasqlInstanceProvider: FabricChangeRole (currentRole = Primary, newRole = ActiveSecondary) failed with 80004005.
2020-07-10 09:20:02.6724846	DB.20	84568	OrcasqlInstanceProvider: FabricChangeRole (currentRole = ActiveSecondary, newRole = Primary, dataLossNumber = 01D62D4FCE9F4164, configurationNumber = 0000000500000000) start
2020-07-10 09:20:02.6724964	DB.20	84568	OrcasqlInstanceProvider: Write to pipe newRole = 2
2020-07-10 09:20:17.5314538	DB.20	84568	2020-07-10 09:20:17.53 spid370s Parallel redo is started for database 'SBS' with worker pool size [8].
2020-07-10 09:20:22.4139703	DB.20	84568	2020-07-10 09:20:22.40 spid370s Recovery completed for database SBS (database ID 5) in 5 second(s) (analysis 59 ms, redo 4590 ms, undo 123 ms [system undo 0 ms, regular undo 0 ms].) This is an informational message only. No user action required.
2020-07-10 09:20:22.4213580	DB.20	84568	2020-07-10 09:20:22.41 spid370s Parallel redo is shutdown for database 'SBS' with worker pool size [8].
2020-07-10 09:20:35.2692644	DB.20	84568	2020-07-10 09:20:35 UTC-5f0832e1.2c-LOG: database system was not properly shut down; automatic recovery in progress
2020-07-10 09:20:35.4104776	DB.20	84568	2020-07-10 09:20:35 UTC-5f0832e1.2c-LOG: redo starts at 39/7004310
2020-07-10 09:20:35.6562245	DB.20	84568	2020-07-10 09:20:35 UTC-5f0832e1.2c-LOG: redo done at 39/7005598
2020-07-10 09:20:35.8990268	DB.20	84568	2020-07-10 09:20:35 UTC-5f0832e1.2c-LOG: database startup complete in 2 seconds, startup began 163 seconds after last stop
2020-07-10 09:20:36.3224330	DB.20	84568	OrcasqlInstanceProvider: FabricChangeRole Database port open signaled.
2020-07-10 09:20:36.3224420	DB.20	84568	OrcasqlInstanceProvider: FabricChangeRole (currentRole = ActiveSecondary, newRole = Primary) finished, duration 33655 ms
2020-07-10 09:20:36.3224483	DB.20	84568	OrcasqlInstanceProvider: FabricChangeRole unavailability duration 33655 ms
2020-07-10 09:20:36.3224673	DB.20	84568	SqlSocketDuplicateProvider:DB engine is ready.
2020-07-10 09:20:36.3476323	DB.20	84568	2020-07-10 09:20:36 UTC-5f0832de.20-LOG: database system is ready to accept connections

Canned RCA (please edit this if needed and read it carefully before sharing with customers

In database systems when the server is restarted it needs to go through database recovery to ensure data consistency. Time to complete database recovery depends on how much data was generated by the workload and how far along is the database checkpoint that flushes these changes to the disk. As we continuously improve our service to improve your experience we are looking at ways to better handle checkpoints that needs longer time to complete.

Mitigation: The Azure Postgres server is back to normal and consistent state after restart. Auto-mitigated.

****How good have you found this content?***

