

Intermediate : Configuring Tungsten Replicator to Extract from Amazon RDS



Topics

In this short course we will:

- Review the Replicator
- Explain how we can extract from Amazon RDS
- Review prerequisites required to extract from Amazon RDS
- Discuss installation parameters (tungsten.ini)
- Walk through an installation
- Recap key resources and tools

Course Prerequisite Learning

- Basics: Introduction to Tungsten Replicator
- Visit www.continuent.com/videos/ to watch this recording



Extracting from RDS using Tungsten Replicator



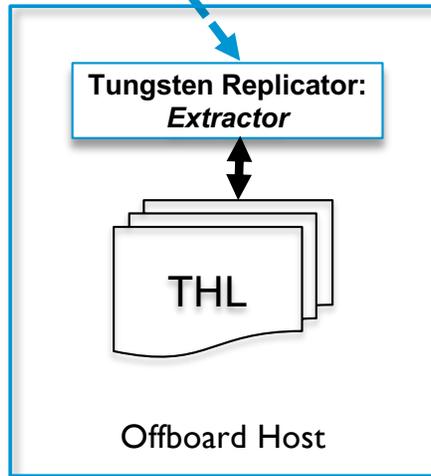
Extraction from RDS

- When extracting from MySQL, Tungsten Replicator usually reads the MySQL binary logs directly from the filesystem
- Amazon RDS does NOT provide filesystem access to the binary logs, nor any access to the RDS host itself
- However, we can configure two aspects of Tungsten Replicator to extract from RDS:
 - Direct mode – allows us to read from the binary logs from the RDS instance over the network, rather than accessing them on the filesystem
 - Unprivileged user support – allows the user to access and update information within RDS without requiring SUPER privileges which are unavailable in RDS
- Tungsten Replicator can then connect to the RDS instance as a regular MySQL client and request the binary logs
- Note that by using the above features, we can extract from any remote MySQL database. For example, the Tungsten Replicator Extractor can be installed on a secondary server, and extract from a remote MySQL database without having to install Tungsten Replicator directly on the database server



Tungsten Replicator

Amazon RDS



Slave Replicator:
Applier



Slave Replicator:
Applier



Slave Replicator:
Applier



Slave Replicator:
Applier



Slave Replicator:
Applier



Prerequisites



Prerequisites (specific to RDS)

- RDS Instance
 - Note IP/Endpoint
 - Also note user name and password for RDS
- Ensure that security group configuration allows access from the offboard host to the RDS instance
- Configure `binlog retention hours` in RDS to allow enough time for Tungsten Replicator to read the binary logs before they are purged. For instance, in RDS:

```
call mysql.rds_set_configuration('binlog retention hours', 24);
```
- Configure `max_allowed_packet_size` in RDS. This should be large enough to read entire transactions from the RDS instance.
- Configure the tungsten user in RDS:
 - `GRANT ALL ON tungsten_service.* to tungsten@'%' IDENTIFIED BY 'secret';`
 - `GRANT SELECT ON *.* TO tungsten@'%' IDENTIFIED BY 'secret';`
 - `GRANT REPLICATION SLAVE ON *.* TO tungsten@'%' IDENTIFIED BY 'secret';`
- RDS by default is configured for MIXED-mode replication. Depending on your topology, you may have to change this to ROW-based replication.



Prerequisites (standard Replicator host)

- Review docs
 - <http://docs.continuent.com/tungsten-replicator-5.3/prerequisite.html>
- Host Prerequisites
 - OS user
 - /etc/hosts
 - sudoers and ssh
 - Ruby
 - Java
- Network Prerequisites
 - Review port requirements
- MySQL Prerequisites
 - my.cnf settings
 - tungsten user accounts



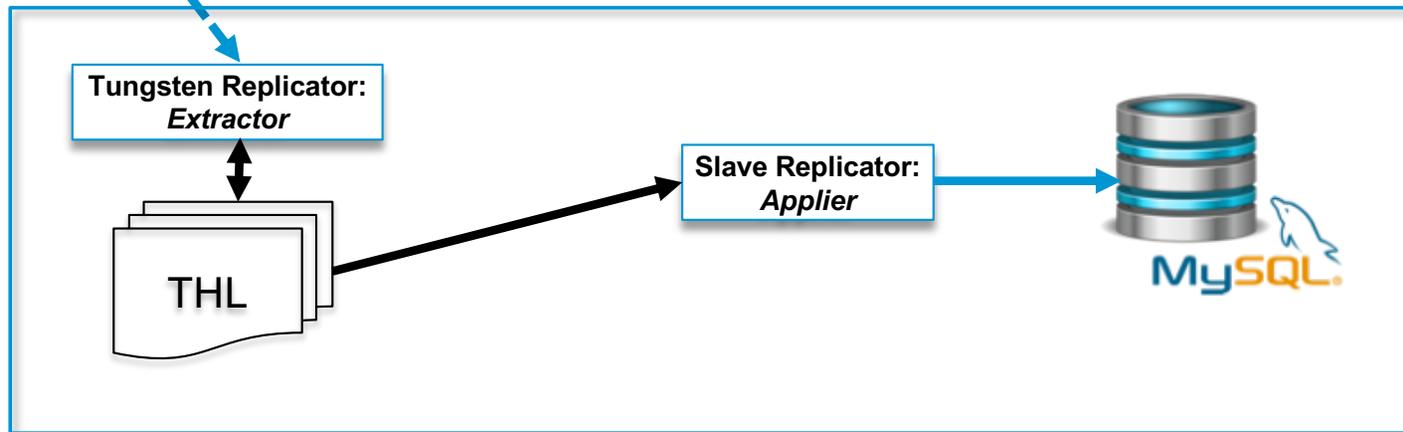
RDS/Direct Parameters for tungsten.ini

- `topology=direct`
 - Replication topology, use “direct” for reading directly from a MySQL server
- `master=db1`
 - The master is the host that will be serving the THL to the appliers, typically it is the offboard host where the replicator extractor is installed
- `direct-datasource-host=rds_endpoint.us-east-1.rds.amazonaws.com`
 - The RDS endpoint
- `direct-datasource-port=3306`
 - The port in RDS where the replicator will connect to read binary logs. Typically 3306
- `direct-datasource-user=rds_user`
 - RDS user name with privileges to read the binary logs (ie tungsten)
- `direct-datasource-password=rds_password`
 - Password for the rds user above
- `privileged-master=false`
 - Disable privileged updates, which are not available in RDS (i.e. require SUPER privileges)



RDS to Local MySQL Demo

Amazon RDS



Command Line Tools & Resources



Tools : trepctl

- “trepctl status” can be run from any node to view the status of the local replicator
- “trepctl status -r 3” will show status output refreshed every 3 second until CTRL+C
- “trepctl qs” provides a quick summary overview of the local replicator
- “trepctl perf” provides deeper diagnostics of the different stages in the replicators

```
$ trepctl qs
State: east Online for 21.069s, running for 45.654s
Latency: 0.837s from DB commit time on dbl into THL
          21.839s since last database commit
Sequence: 1 last applied, 0 transactions behind (0-1 stored) estimate 0.00s before synchronization
```



Log Files

- The `/opt/continuent/service_logs/` directory contains both text files and symbolic links.
- Links in the `service_logs` directory go to following subdirectory:
 - `/opt/continuent/tungsten/tungsten-replicator/log/`

```
tungsten@db1:/opt/continuent/service_logs $ ll
total 116
lrwxrwxrwx 1 tungsten tungsten      60 Jan 06 09:52 trepsvc.log -> /opt/continuent/tungsten/tungsten-replicator/log/trepsvc.log
```



Next Steps

- If you are interested in knowing more about the software and would like to try it out for yourself, please contact our sales team who will be able to take you through the details and setup a POC – sales@continuent.com
- Read the documentation at <http://docs.continuent.com/tungsten-replicator-5.3/index.html>
- Follow us on Twitter @Continuent or www.facebook.com/Continuent
- Subscribe to our Tungsten University YouTube channel! <http://tinyurl.com/TungstenUni>
- Visit the events calendar for upcoming Webinars and Training sessions at www.continuent.com/events/



For more information, contact us:

Eero Teerikorpi

Founder, CEO

eero.teerikorpi@continuent.com

MC Brown

VP Products

mc.brown@continuent.com

Eric Stone

COO

eric.stone@continuent.com

Matthew Lang

Director, Professional Services Americas

matthew.lang@continuent.com

