

Outsourcer Oracle Source Guide

<http://www.PivotalGuru.com>

Document Date 2013-10-11

Oracle Source	3
Memory.....	3
Fetch Size	3
XMN.....	4
Entropy.....	4

Oracle Source

Oracle database is a supported source for Outsourcer but it comes with challenges that Microsoft SQL Server does not have. For both sources, Outsourcer relies on the JDBC thin driver to make connections but for Oracle, more tuning is often needed.

The two main areas to understand are Memory and Entropy.

Error Type	Error Message	Description	Resolution
Memory	Protocol violation	The XMX is too low for the number of rows retrieved per fetch multiplied by the number of bytes per row.	Increase the XMX value and/or decrease the oFetchSize.
Entropy	Connection reset	The Entropy pool has been depleted.	Install the rng-tools and start a background process that maintains the random pool.

Memory

Outsourcer uses Java to make connections to Oracle and the Oracle JDBC driver requires far more memory than the SQL Server JDBC driver. Each setting has an impact to memory with a consideration for performance.

You will get an error indicating “java.sql.SQLException: Protocol violation” when you need to make an adjustment to either the Fetch Size or the XMX value.

Fetch Size

The fetch size or the “defaultRowPrefetch” value in JDBC terminology, is the number of rows to retrieve on each fetch. The default value is 10 which makes sense for the typical way in which Oracle database is used. It uses just a little bit of memory as it only gets 10 rows at a time. However, this is very slow when extracting large amounts of data because it requires more round trips to the database.

The default value defined in Outsourcer for the fetch size is 40,000. It is configured by adjusting environment variable stored in os.variables with the name “oFetchSize” or by adjusting this value in the User Interface.

Adjusting this value higher will help increase performance but at a cost of requiring more memory allocated by the JVM. If you get the error, “java.sql.SQLException: Protocol violation”, then you can try decreasing the oFetchSize and re-running the job.

XXM

This is the maximum memory pool allocation for each job execution. Outsourcer creates a thread for each Job and each thread gets allocated a memory pool that can grow as large as the XXM value. The default value is 256M and setting an environment variable in your .bashrc file and then restarting Outsourcer.

Example:

```
export XXM=1024m
```

Be careful adjusting this value because Outsourcer can consume this much memory per thread, which is determined by the max_jobs value found in os.variables.

Entropy

Entropy in computers is the randomness used by cryptography or other applications that require random data. Oracle's JDBC driver uses entropy to when making a connection and it relies on the host operating system to do this. Often times, this "feature" causes problems with Oracle's JDBC driver and subsequently, you will get this error: "java.sql.SQLException: Io exception: Connection reset".

The error happens when the JDBC driver executes several requests from /dev/random for a random number and it depletes the pool. This is when the error happens because there aren't enough bytes in the pool to satisfy the driver.

Outsourcer uses "file:///dev/urandom" to use the non-blocking version of random, which usually does satisfy the needs of the Oracle JDBC driver. However, there are times when the entropy pool is still not large enough and you will still get the "Connection reset" error.

To solve this, use the rng-tools which can be installed on the Master server from rpm on Linux or using yum. Next, edit /etc/sysconfig/rngd and add:

```
EXTRAOPTIONS="-r /dev/urandom -o /dev/urandom -t 1"
```

Start rngd:

```
/etc/init.d/rngd start
```

Make sure it starts again on reboot:

```
chkconfig rngd on
```

And you can watch entropy. It should be maxing out around 4000 and staying above 1000.

```
watch -n 1 cat /proc/sys/kernel/random/entropy_avail
```