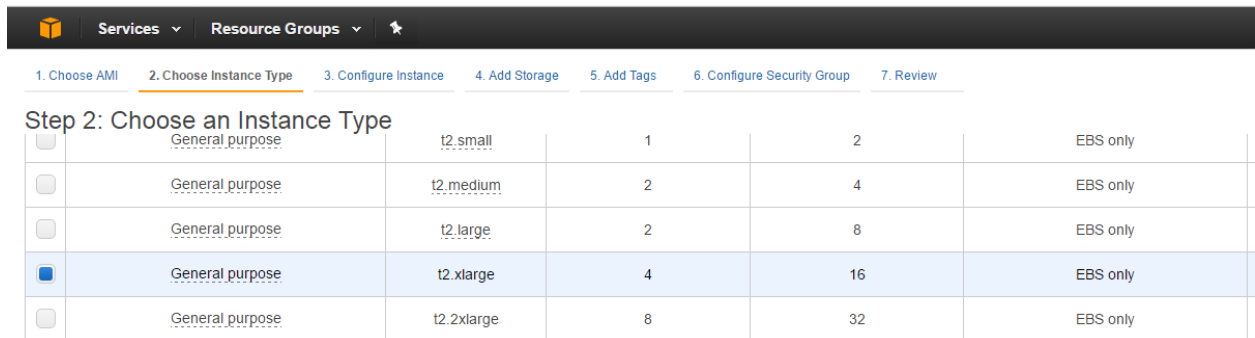


All code has been created in AWS. Requirements are to spin up a Windows EC2 server from an AMI image for all ETL and Analytics development via Pentaho EE (installed on the server). For the back-end database, a SQL Server Express Cloud instance has been used. The EC2 AMI as well as the SQL Server Express snapshot have been made public via AWS (see instructions below).

## Setting up Windows EC2 Server

To find a shared public AMI using the console

1. Open the Amazon EC2 console at <https://console.aws.amazon.com/ec2/>.
2. In the navigation pane, choose **AMIs**.
3. In the first filter, choose **Public images**. To granulate your search, choose the Search bar and use the filter options provided in the menu.
4. Search for the public image (AMI ID: **ami-9e68f2fe**)
5. Once you have found the image, right click and select to launch.
6. Select the following EC2 instance type.



<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only
<input type="checkbox"/>	General purpose	t2.medium	2	4	EBS only
<input type="checkbox"/>	General purpose	t2.large	2	8	EBS only
<input checked="" type="checkbox"/>	General purpose	t2.xlarge	4	16	EBS only
<input type="checkbox"/>	General purpose	t2.2xlarge	8	32	EBS only

## Setting up SQL Server Express Instance:

1. Open the Amazon RDS console at <https://console.aws.amazon.com/rds/>.
2. In the navigation pane, choose **Snapshots**
3. In the first filter, choose **All Public Snapshots**.
4. Search for the snapshot **transpoc-reporting-20170508**

5. Once you have found the image, right click and select to launch.

6. Set the options as shown below:

Resource Groups ▾ 🔖

◀ Restore DB Instance

You are creating a new DB Instance from a source DB Instance at a specified time. This new DB Instance will have the default DB Security Group and DB Parameter Groups.

Instance Specifications

DB Engine

sqlserver-ex ▾

License Model

license-included ▾

DB Instance Class

db.t2.micro — 1 vCPU, 1 GiB RAM ▾

Storage Type

General Purpose (SSD) ▾

Scaling storage

after launching a DB Instance is currently not supported for SQL Server. You may want to provision storage based on anticipated future storage growth.

⚠

Provisioning less than 100 GB of General Purpose (SSD) storage for high throughput workloads could result in higher latencies upon exhaustion of the initial General Purpose (SSD) IO credit balance. [Click here](#) for more details.

Settings

DB Snapshot ID

transpoc-reporting-20170508

DB Instance Identifier\*

transpoc-reporting

DB Instance identifier. This is the unique key that identifies a DB Instance. This parameter isn't case sensitive. (e.g. mydbinstance)

Network & Security

VPC\*

Default VPC (vpc-116d5d76) ▾

Subnet Group

default ▾

Publicly Accessible

Yes ▾

Availability Zone

us-west-2c ▾

### Microsoft SQL Server Windows Authentication

Select a directory in which you want to allow authorized domain users to authenticate with this SQL Server instance using Windows Authentication.

Directory  ↻

[Create a new Directory](#)

By selecting a directory and continuing with database instance creation you authorize Amazon RDS to create the IAM role necessary for using Windows Authentication

### Database Options

Database Port

Option Group

Copy Tags To Snapshots ☐

### Maintenance

Auto Minor Version Upgrade

7. Click Restore DB Instance when ready.
8. Record the ARN of your SQL Server DB instance once ready as this will be needed when trying to connect.

## Environment Configuration/Setup:

As a note, Pentaho will already be installed on the EC2 Windows instance. Also, all Kettle and Data Analysis/Visualization files will already be present as well following the Image restore.

The base code for the Pentaho transformations and jobs are located on the EC2 image in the following directory: C:\TransPOC

These base code files are also included in the GitHub repository

here: [https://github.com/reddraider/TransPOC\\_1/tree/master/Pentaho\\_Code/Kettle\\_Code](https://github.com/reddraider/TransPOC_1/tree/master/Pentaho_Code/Kettle_Code)

The kettle.properties file is located

here: [https://github.com/reddraider/TransPOC\\_1/tree/master/Pentaho\\_Code/Config\\_Files](https://github.com/reddraider/TransPOC_1/tree/master/Pentaho_Code/Config_Files)

This file will have to be modified with the SQL Server Express ARN as well as the NREL API Key. This file will need to be placed in the C:\Users\Administrator\.kettle EC2 directory (can overwrite existing file). This can be done by copying/pasting the file via RDP once you have modified it. Modify the following attributes in the kettle.properties file:

```
#####
#SQL NREL Fuel Station Database
#Connection_URL=jdbc:sqlserver://reporting.caylwz8kpbxv.us-west-
2.rds.amazonaws.com:1433;databaseName=reporting;selectMethod=cursor;
Connection_URL=can copy the string above and modify the database ARN connection
Driver_Class_Name=com.microsoft.sqlserver.jdbc.SQLServerDriver
User=reddraider
```

Password=

##NREL URL and API KEY

AgileBIDatabase=AgileBI

NREL\_API=[http://api.data.gov/nrel/alt-fuel-stations/v1.json?api\\_key=](http://api.data.gov/nrel/alt-fuel-stations/v1.json?api_key=)

API\_Key=

#####

## Pentaho Analytics:

All reporting files exist within EC2 instance by going to the URL <http://localhost:8080/pentaho/Home>

User: admin

Pass: password

These files can also be found on Github

here: [https://github.com/reddraider/TransPOC\\_1/tree/master/Pentaho\\_Code/Reports](https://github.com/reddraider/TransPOC_1/tree/master/Pentaho_Code/Reports)

## Database Creation/DDL

The SQL Server Express database creation DDL is located

here: [https://github.com/reddraider/TransPOC\\_1/tree/master/Pentaho\\_Code/DDL](https://github.com/reddraider/TransPOC_1/tree/master/Pentaho_Code/DDL)

Please let me know if you have any questions.