## Technical Prerequisites/Requirements

1. You will be required to create a database on SQL Server and run in the ‘Seed Code Sample Request.sql’ file to create your database and populate it with data. You may have to move the file locations in the Create Database section.

1. Any scripts provided must be written using TSQL and/or Powershell;
2. The scripts must be built to support a database deployed using SQL Server 2012 Enterprise . If you do not have access to this, please download and install a trial version from Microsoft. This is required so you can use the enterprise feature set available for partitioning and indexing.
3. Each answer should be implemented in separate file.
4. You must provide tests to highlight how you know your solution is correct, each test will be documented in a separate file ( and could comprise of simple list of items to check, sql scripts, table of expected results or any other method you wish to use ).
5. The source code with all its dependencies must be published into a Git repository created/owned by you at either BitBucket or GitHub, and you should give access to our user “NephilaCodeReview” and jweisbecker so we can access that repository. When submitting the code to the on-line repository, please publish all of your commits (deep check-in), so we can have some understanding on how you organize your work;
6. Your code will be reviewed by our senior developer team within Nephila. Our developers should be able to open the files in SQL Studio Enterprise Manager or SSDT hosted in Visual Studio. They should be able to run the tests without having to setup anything on the environment, code or database, with the exception of setting the connection string that points to the database.
7. Your response should also consider scripts to migrate the data, from that provided to the final proposed solution. You should also consider, the possibility of errors, transaction handling/rollback etc.

We are specifically looking for elegant, maintainable and well documented code. Inclusion of new technology features will be considered a bonus if it does not compromise the simplicity of the solution.

## Business Requirements

### Background

To help the developer understand this exercise, consider the following scenario:

You have purchased your dream home on the beach in the state of Florida, USA on January 1st 2013.

When you purchased the home, you also purchased home insurance from the local ***insurance company*** named ‘***X’***.

Other home owners in the state of Florida have also bought home insurance from this local ***insurance company***.

On August 15th, 2013, you are watching the ***Weather Channel*** and see an important news flash that a hurricane is approaching Florida. You know if the hurricane makes landfall in Florida, you and other home owners will probably experience some damage to your homes due to the potentially very strong gale force winds and possible flooding. You do your best to secure your home and family.

The hurricane passes and you and your family are safe and well. Then when you step outside to survey the damage you realize the very strong winds blew over a tree and damaged your garage, car and other property. You immediately call your ***insurance company*** and advise them that you will be making a ***claim*** against your ***insurance policy*** to repair the damage. The ***insurance company*** in due course will need to make a payment to you based on your ***claim.*** You will then use this payment to repair the damage.

### Project Outline

We have been tasked with helping our company upgrade a small application that they use to manage/query several data-sets associated with ***Insurance Company ‘X’***.

***Insurance company*** ***‘X’*** sent us two ***event-set*** files. These ***event-set*** files represent this ***insurance company’s*** projected financial losses should an ***event*** (i.e. a natural catastrophe) take place. These files contain a variety of scenarios for us to consider/review.

The data you will be working with consists of the following:

* geographical ***regions***
* natural catastrophe ***perils***
* two ***event-sets*** ( each contains an ***event*** file and an ***event loss footprint*** file )

A glossary and entity relationship diagram for this database is included in the appendix.

The company faces several issues with the data included in the event-set files. They project this data is going to grow dramatically over the next year. They also would like to leverage some of the new Velocity features available in SQL Server 2012 but they are concerned about the effort vs. benefit required to implement these. They have asked you to perform a technical spike to see if this is worth pursuing or if there will be insufficient return on their investment.

Goals  
As part of this project, you have been tasked to help the team with a number of problems they are facing. They have asked if you can help them with any the problems they face.

Acting as the dba,

* 1. Implement a partitioning strategy in TSQL for the event-set / event-loss tables ( t00001/t00002 ). The strategy / considerations here is based on your assessment of the type of data included in the sample. There are many possible answers to this question and the candidate should review the data in order to try and derive ideas.
  2. Propose 3 examples in TSQL of how a csv could be loaded/removed from these tables when (a) is in place without using SSIS
  3. Propose alternatives ( TSQL examples optional ) for indexing the event-set / event-loss tables ( t00001/t00002 ) describing the pro/cons of each after implementing (a).
  4. Discuss how (c ) impacts (a) including considerations for performance, disk i/o and memory.
  5. Demonstrate using TSQL, how you could monitor the data spread across partitions
  6. Discuss how you could monitor the effectiveness of any indexing implemented in 1.c. using the tools available to you within SQL Management Studio, TSQL.

1. Acting as a developer,
   1. You have been asked to provide a SET based TSQL query to present the hierarchy represented in tables geo\_code (t031) and geo\_region (t030). Provide 2 options to do this. Explain the pros/cons of each method ( if any ).
   2. You have been asked to provide some TSQL code to split a CSV string without using a cursor, how would you do this.

CSV string = ‘a,b,c,d,e,f,g,NULL,h,,I’

1. Acting as the dba, a developer on another project approaches you for advice. He wants a user in his database to access data on another database (on the same server). He is concerned about security and allowing his user access to the underlying tables etc.
   1. What options are available to the dba to manage / facilitate this. You have been told by the developer the user wants access to some of the data in the metarisk table (t011) but not all.
   2. Repeat question (5.a) but consider his database/user is located on a remote server and you are trying to do this as easily as possible (with no concerns regarding performance).
   3. Repeat question (5.a) again but your customer is now concerned about performance vs. ease of implementation. Describe what (if any) options are available to you as a dba to tackle this and highlight any things you believe you need to make the customer aware of ( complexity, ease of maintenance, cost ) when proposing each option. Ideally you would provide a table showing your proposals highlighting pros, cons of each.
2. Acting as a developer, you have been asked to offer a solution that will track a list of tasks ( that will be scheduled, executed and completed ). This list of tasks will be processed by a multi-threaded application written by the development team.
   1. What considerations/advice could you give to this team to handle race conditions, blocking.
   2. You are free to propose any solutions available within the enterprise toolset.

### Presentation

On completion of the exercises, you will be required to present your solution to the technical team.

## Glossary

***Reinsurance Company*** – A company providing insurance to other insurance companies so they can minimize/diversify their own insurance risk.

***Reinsurance Policy*** – An arrangement between an ***insurance company*** and the ***reinsurance company*** to provide them ***insurance***.

***Event*** – A theoretical event that is caused by a natural catastrophe that happens at some point in time i.e. hurricane, winter storm, flood;

***Event Loss Footprint*** – A list of losses by geographical location for each ***event*** that an insurance company suffered as a result of an ***event;***

***Event Simulation*** – The output of a Monte Carlo simulation to generate a sequence of ***events;***

***Event-set*** – A list of ***event simulations*** where each simulation in the ***event-set*** contains 0..n ***events.*** Each ***event*** has an e***vent loss footprint***

***Peril*** – The type of ***event*** that occurred i.e. hurricane, thunderstorm, flood. ***Peril*** is hierarchical, i.e. Wind includes ***perils*** such as tornado, hurricane;

***Region*** – The geographic area where an ***event*** occurred and/or experienced a loss. ***Region*** is hierarchical, i.e. North America includes the United States / Canada;

***Meta-Region*** – A list of ***regions*** grouped together to represent a larger region/cluster of regions;

***Meta-Peril*** – A list of ***perils*** grouped together to represent a larger peril/cluster of perils. i.e. Earthquake contains fires that occur after an earthquake, the ground shake and tsunami;

***Meta-Risk*** – A list of ***Meta-Regions*** and ***Meta-Perils*** grouped together to represent larger sets of perils and regions. i.e. United States Wind, World Wide All Perils;

## Entity Relationship Diagram

