Making SSIS a Clustered Service

posted on Wednesday, July 06, 2005 2:27 PM by Knight_Reign

http://www.sqljunkies.com/WebLog/knight_reign/archive/2005/07/06/16015.aspx

One of the requests that we see very often is how to run SSIS on a cluster. The goal of clustering is either load-balancing or failover (restarting a process if the machine it is on dies). SSIS today is not cluster aware: the SSIS runtime is installed on a single machine, and operates from that machine. Somewhat confusingly, the SSIS service can be made to run on a cluster. But the service does not run or schedule SSIS packages --- it only collects monitoring information from them. So making the service clustered doesn't accomplish the goal most people are setting out to do.

As a consequence, we recommend that people do not cluster SSIS (or the SSIS service). Given a clustered hardware environment, install SSIS on each server that you intend to use it on, and schedule jobs separately on each of those machines. You can use the checkpoint restartability feature to recover from failed package executions. ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== ===== =====

First of all, this blog entry is wholly the work of Ranjeeta Nanda and Nick Berezansky who are two of our top notch testers on the Integration Services team.

They have outlined the steps to make SSIS a clustered service. This is interesting stuff.

Thanks for doing this Ranjeeta and Nick.

How to make SSIS a clustered service?

Prerequisite:

You have already installed a cluster with one or more nodes; you have installed Yukon on the cluster.

You have installed all the non-virtualized components (in this case SSIS) on all available nodes.

There are two ways to configure SSIS as a clustered service:

1) In same group as SQL Server (Preferred)

2) In different group from SQL Server

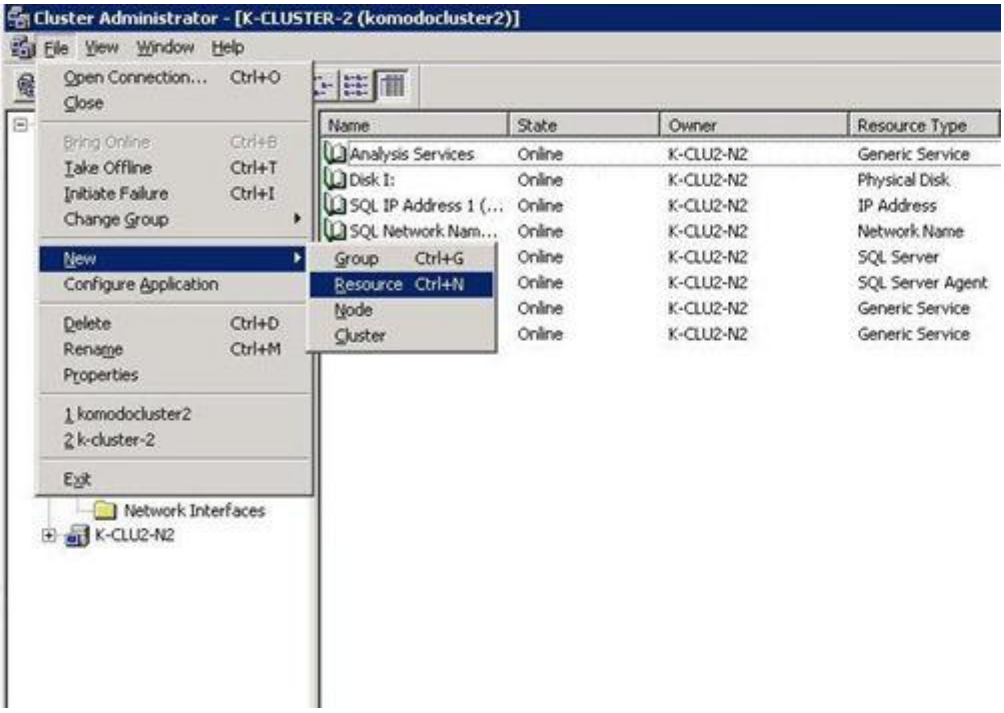
IN SAME GROUP (Preferred):

1. Open the Cluster Administrator.

2. In the console tree, double-click the Groups folder

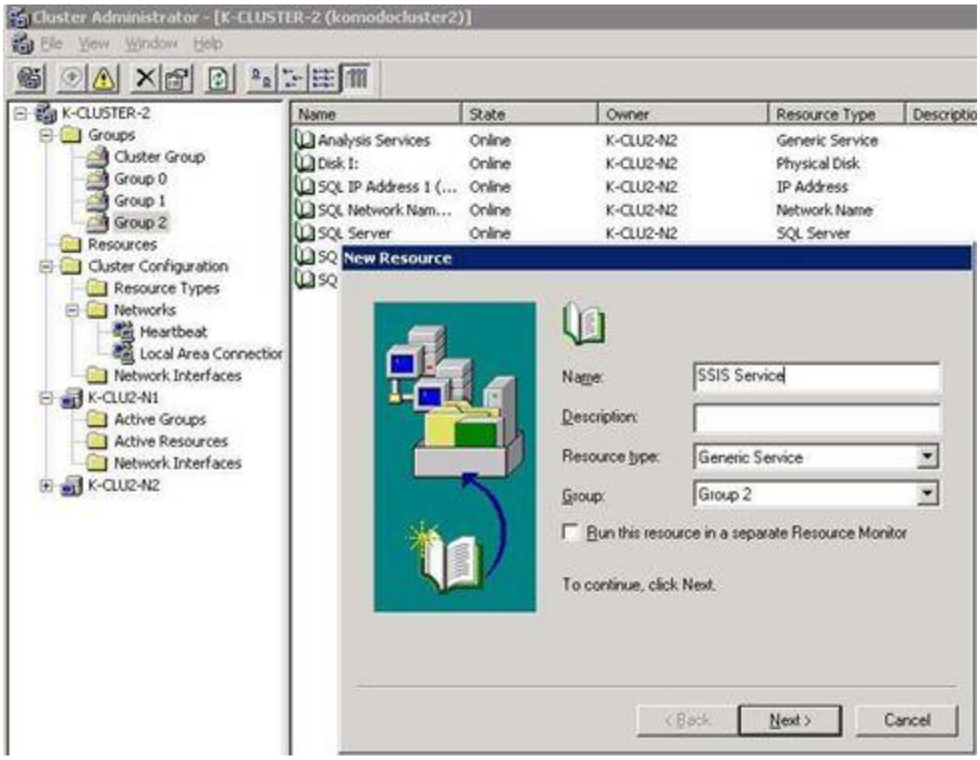
3. In the details pane, click the group to which SQL Server belongs

4. On the File menu, point to New and then click Resource.



Name	State	Owner	Resource Type
Analysis Services	Online	K-CLU2-N2	Generic Service
Disk I:	Online	K-CLU2-N2	Physical Disk
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address
SQL Network Nam...	Online	K-CLU2-N2	Network Name
	Online	K-CLU2-N2	SQL Server
	Online	K-CLU2-N2	SQL Server Agent
	Online	K-CLU2-N2	Generic Service
	Online	K-CLU2-N2	Generic Service

5. In the new Resource Wizard that appears, type a Name and Choose “Generic Service” as the Service Type, and then click Next.



Cluster Administrator - [K-CLUSTER-2 (komodocluster2)]

Name	State	Owner	Resource Type	Descriptio
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	

New Resource

Name: SSIS Service

Description:

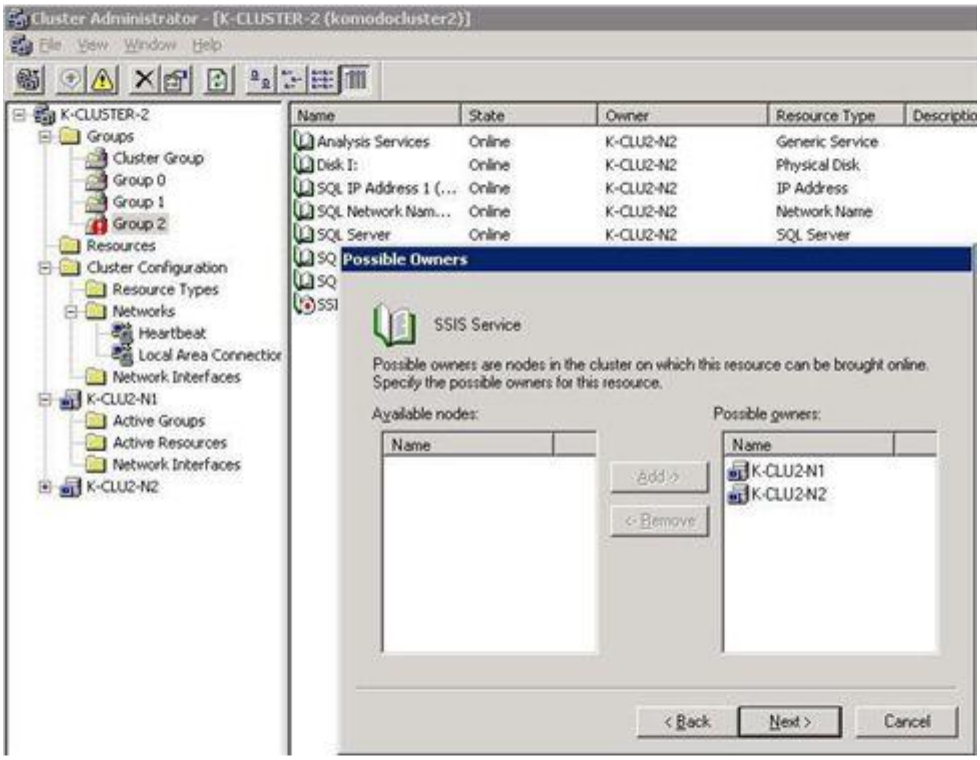
Resource type: Generic Service

Group: Group 2

□ Run this resource in a separate Resource Monitor

To continue, click Next.

6. Add or remove the nodes of the cluster as the possible owners of the Resource, and then click Next.



Cluster Administrator - [K-CLUSTER-2 (komodocluster2)]

Name	State	Owner	Resource Type	Descriptio
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	

Possible Owners

SSIS Service

Possible owners are nodes in the cluster on which this resource can be brought online. Specify the possible owners for this resource.

Available nodes:

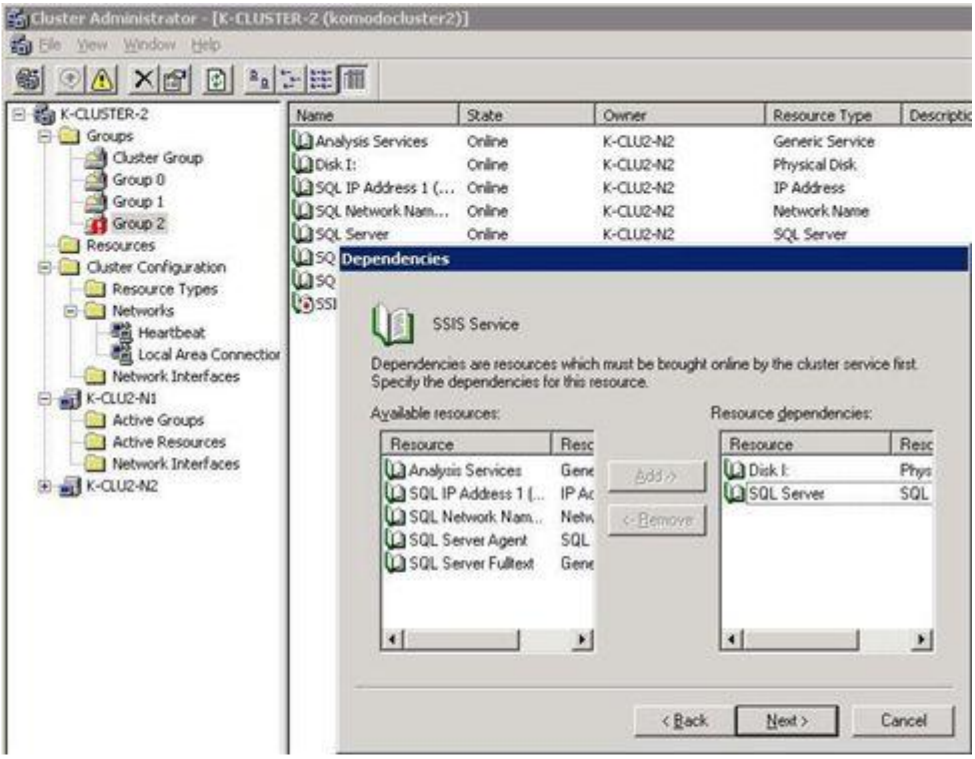
Name

Possible owners:

Name
K-CLU2-N1
K-CLU2-N2

7. To add dependencies, under Available resources, click a resource, and then click Add. In this case, we want the shared disk which holds the packages, and SQL Server to

come online before SSIS Service is brought online. Click on Next after you have chosen all the dependencies.



Name	State	Owner	Resource Type	Descriptic
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	

Dependencies

SSIS Service

Dependencies are resources which must be brought online by the cluster service first. Specify the dependencies for this resource.

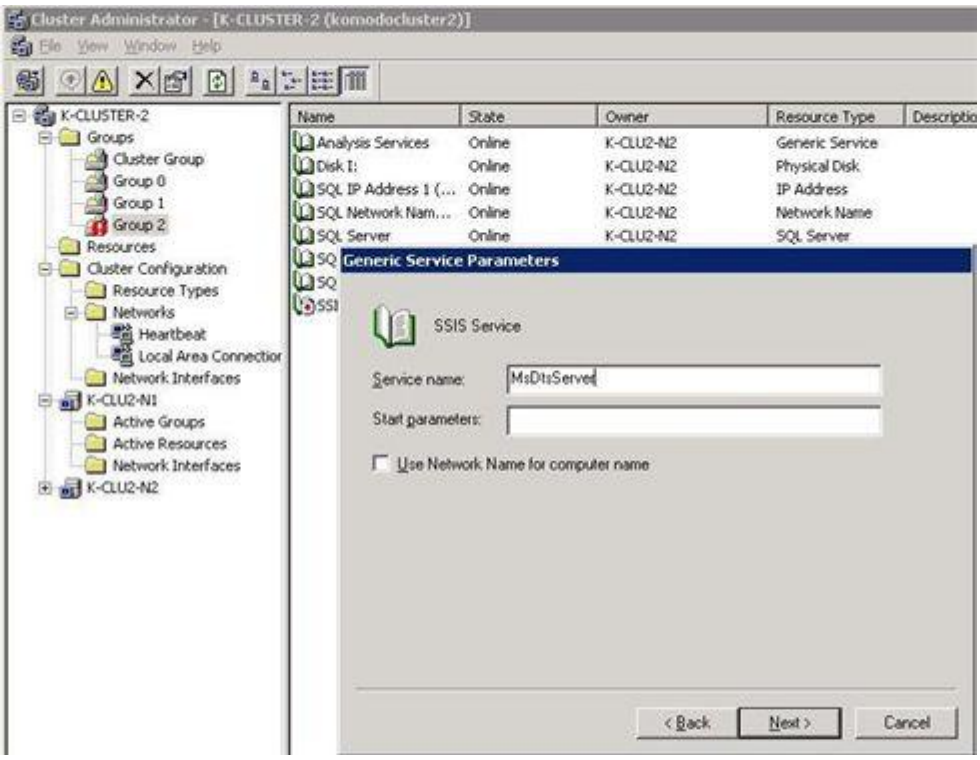
Available resources:

Resource	Resc
Analysis Services	Gene
SQL IP Address 1 (...	IP Ac
SQL Network Nam...	Netw
SQL Server Agent	SQL
SQL Server Fulltext	Gene

Resource dependencies:

Resource	Resc
Disk I:	Phys
SQL Server	SQL

8. Set Resource properties in the Resource Parameters dialog box. For SSIS, the name is MsDtsServer, enter this in the Service Name box, and click Next.



Name	State	Owner	Resource Type	Descriptio
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	

Generic Service Parameters

SSIS Service

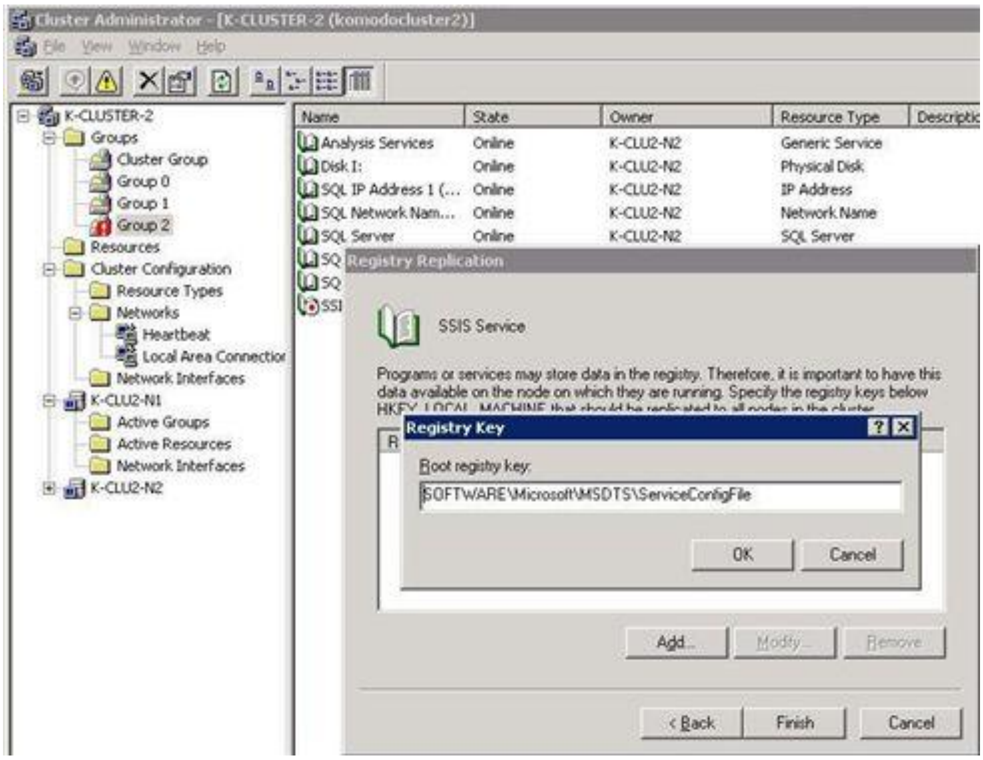
Service name: MsDtsServer

Start parameters:

□ Use Network Name for computer name

9. In the Registry key setting, we need to add the key for the configuration file needed for SSIS Service. This file needs to be at a shared disk which is in the same group as SSIS service, and fails over to the next node along with SSIS service. In this case, we use the registry key

SOFTWARE\Microsoft\MSDTS\ServiceConfigFile

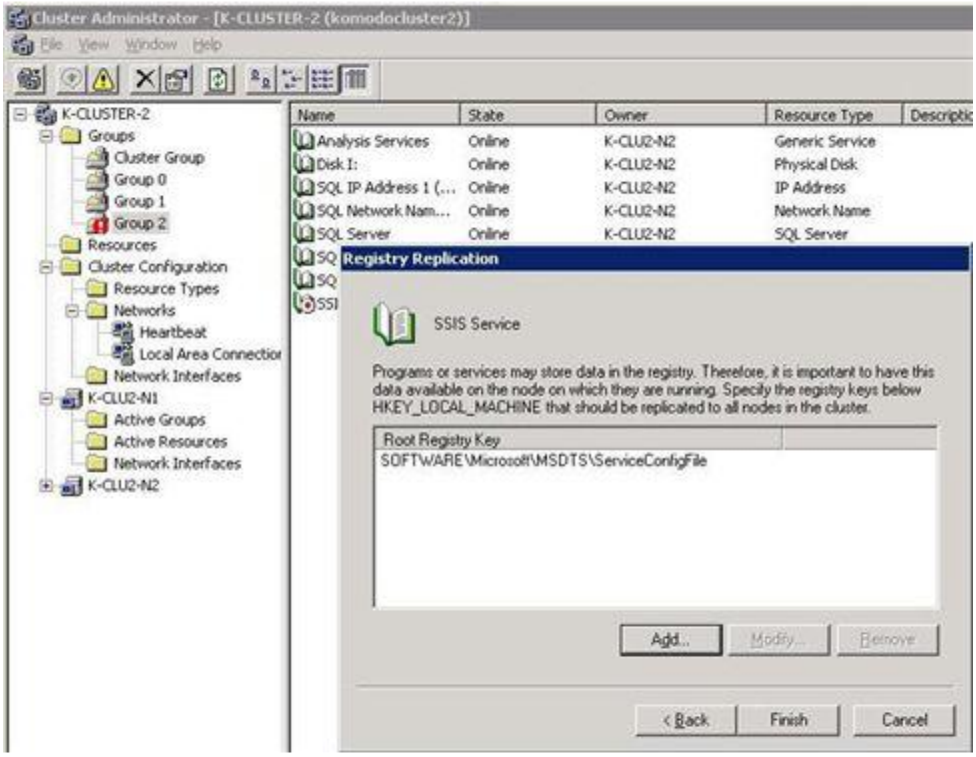


Name	State	Owner	Resource Type	Descriptic
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	
SQ				
SQ				
SSI				

Programs or services may store data in the registry. Therefore, it is important to have this data available on the node on which they are running. Specify the registry keys below HKEY_LOCAL_MACHINE that should be replicated to all nodes in the cluster.

Root registry key: SOFTWARE\Microsoft\MSDTS\ServiceConfigFile

10. Finally click Ok and then click Finish.



Name	State	Owner	Resource Type	Descriptic
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	
SQ				
SQ				
SSI				

SSIS Service

Programs or services may store data in the registry. Therefore, it is important to have this data available on the node on which they are running. Specify the registry keys below HKEY_LOCAL_MACHINE that should be replicated to all nodes in the cluster.

Root Registry Key
SOFTWARE\Microsoft\MSDTS\ServiceConfigFile

11. The Service should get added successfully.

12. In addition to adding SSIS Service to the Cluster, we also need to set up the config file and Package store properly so that they are accessible to all nodes in case of the failure.

There is a config file at the following location:

% ProgramFiles%\Microsoft SQL Server\90\DTS\Binn\MsDtsSrvr.ini.xml

Copy this file over to the Shared Disk I:\

Also create a folder called Packages in the Shared Disk I:\

Set ACL on the folder to List Folders+Write for built-in users

Now open the file I:\MsDtsSrvr.ini.xml

The content of the file looks like this:

true

MSDB

File System

..\Packages

You need to change it to the following (the changes are in Bold). K-CLUSTER2-SQL1 is the name of the virtual sql server which is in the same group, in this case. We need to use the virtual server name instead of "." for localhost. The Shared packages path has also been updated to use the shared folder.

true

MSDB

K-CLUSTER2-SQL1

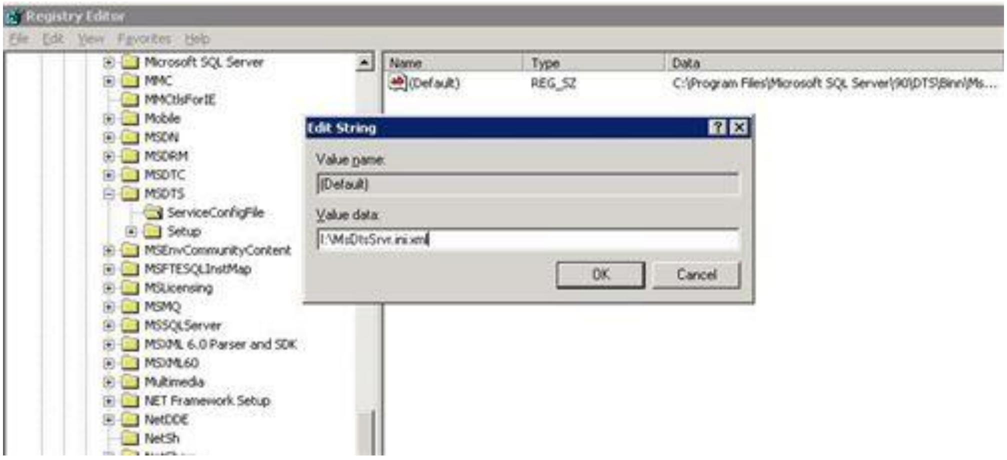
File System

I:\Packages

13. The last thing we need to do is to update the Registry Key setting so that the correct config file is used. The registry key is the one we specified before:

HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\MSDTS\ServiceConfigFile

Its value need to change so that the config file in the Shared disk is used.

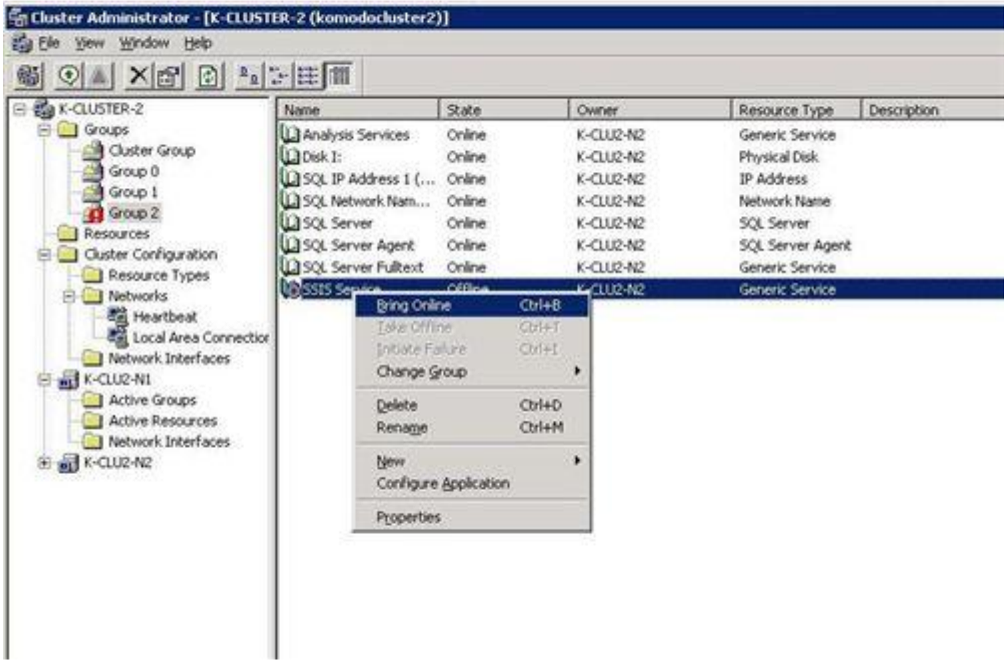


Name	Type	Data
(Default)	REG_SZ	C:\Program Files\Microsoft SQL Server\90\DTS\Binn\Ms...

Value name: (Default)

Value data: I:\MsDtsSrvr.ini.xml

13. Once we have performed all the above steps, Right click on the SSIS Service in the cluster administrator window, and bring it online.



Name	State	Owner	Resource Type	Description
Analysis Services	Online	K-CLU2-N2	Generic Service	
Disk I:	Online	K-CLU2-N2	Physical Disk	
SQL IP Address 1 (...	Online	K-CLU2-N2	IP Address	
SQL Network Nam...	Online	K-CLU2-N2	Network Name	
SQL Server	Online	K-CLU2-N2	SQL Server	
SQL Server Agent	Online	K-CLU2-N2	SQL Server Agent	
SQL Server Fulltext	Online	K-CLU2-N2	Generic Service	
SSIS Service	Offline	K-CLU2-N2	Generic Service	

14. Now, you can connect to Integration services from a client machine.

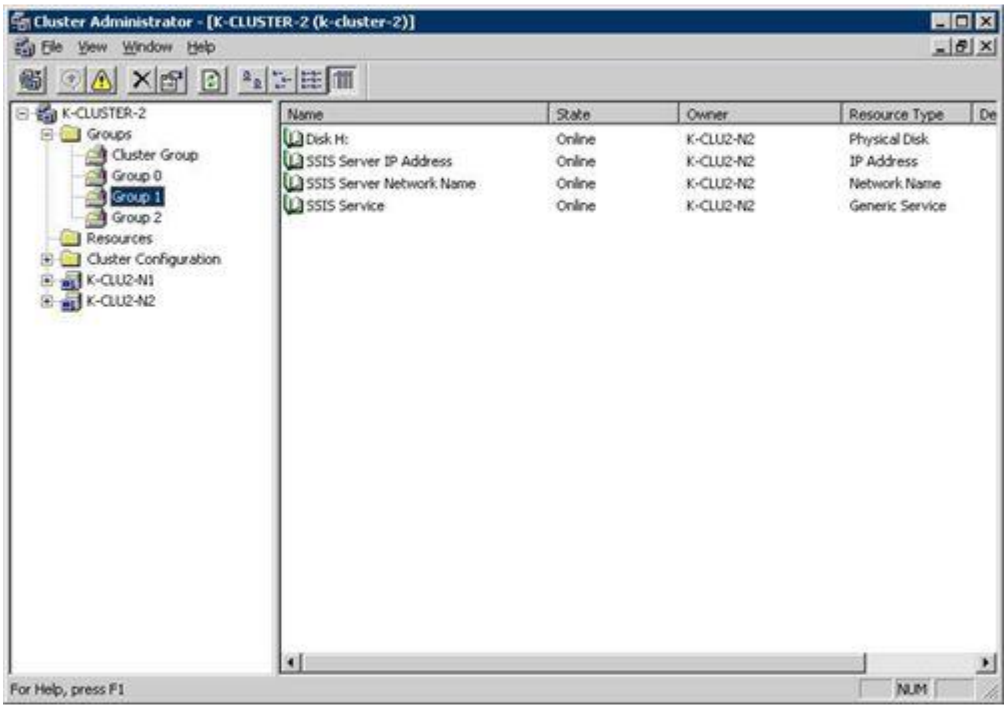
IN DIFFERENT GROUP:

1. You can also configure SSIS as a clustered Service that belongs to a different Group from SQL Server.

2. In the Cluster Administrator, go to a group other than the SQL Server group.

3. Add and configure IP Address, Network Name and Generic service resources as below

4. The shared packages and config file need to be on the Shared disk of this group (in this case H:\)



Name	State	Owner	Resource Type	De
Disk H:	Online	K-CLU2-N2	Physical Disk	
SSIS Server IP Address	Online	K-CLU2-N2	IP Address	
SSIS Server Network Name	Online	K-CLU2-N2	Network Name	
SSIS Service	Online	K-CLU2-N2	Generic Service	

5. Bring the group online

SIDE NOTE

Configuring SSIS Service to run in a group separate from SQL Server has some pros and cons

Pros:

- Allows SSIS Server to run an a different node so SSIS and SQL Server won’t compete for CPU time
- Since most of SSIS Server functionality (except MSDB store) doesn’t depend on a SQL Server it is reasonable to run it in a separate group
- Once SSIS Server runs in a separate group it can be moved to another node faster.

Cons:

- When UI, SSIS Server and SQL Server runs on 3 different machines MSDB Node is unavailable. This is IDW15 design that may or may not be changed.
- When SSIS and SQL Server run on the same node operations against MSDB may be faster and generate less network traffic.

Hopefully this will help folks when trying to set up SSIS on clustered machines.

Thanks,

Ranjeeta Nanda

Nick Berezansky

Universe.Earth.Software.Microsoft.SQLServer.IS.KirkHaselden