

FailoverCluster-report

by Lu Han

FILE	12041_LU_HAN_FAILOVERCLUSTER-REPORT_181263_1663882019.DOCX (4.55M)		
TIME SUBMITTED	01-MAY-2019 11:52PM (UTC+0100)	WORD COUNT	2837
SUBMISSION ID	106033506	CHARACTER COUNT	15055

Contents

1. What Is Failover Cluster	2
1.1 Introduction	2
1.2 Key technology	2
SAN	2
Heartbeat	3
Quorum	3
1.3 Cluster Terminology	3
1.4 RAID	3
RAID level	4
Summarize	5
Application of RAID technology:	5
2. Build a failover cluster	5
2.1 Network Topology	5
IP address allocation	6
2.2 Virtual machine preparation	6
Install hypervisor platform	6
Create virtual machines and configure it	8
Reasons for choosing a private cloud	9
Install Window Server 2016 on the virtual machine	9
Select a VM network	10
Add a new virtual hard disk and the iSCSI target role (Cardiff)	10
Create a new iSCSI initiator (Swansea)	12
Run the iSCSI initiator	13
Clone Swansea to create a new node – Bridgend	14
2.3 Failover Cluster	14
Failover Cluster installation	14
Failover Cluster configuration	15
Install the DHCP role on both nodes	16
2.4 Test the Failover cluster	16
3. Expand Discussion	17
Network Load Balanced (NLB) Cluster	17
SPI	18
IaaS	19
PaaS	19
SaaS	19
References	19

1. What Is Failover Cluster

1.1 Introduction

Failover Clustering is a service provided by Microsoft Corporation for ensure uptime for mission-critical applications. A Windows failover cluster is a group of computers with independent capabilities, and the cluster server passes Physical cables and dual-machine software work together to enhance the usability of applications and services.

Failover cluster requires [1]:

- 2 – 8 Servers.
- 2+ Switches to connect other components together.
- 2 SAN trays with:
 - Data storage, in RAID 10 (or 5/6).
 - Quorum, in RAID 1.

1.2 Key technology

The most critical technologies in Windows Server Failover Cluster are SAN, Heartbeats, and Quorum Disks.

SAN

SAN (Storage Area Network) is a high-speed network dedicated to storage operations, usually independent of a computer local area network (LAN). The SAN connects the host to the storage device and provides a dedicated communication channel for any host and any storage device on it. The SAN separates the storage device from the server and implements storage resource sharing at the server level. SANs bring channel technology and network technology to the storage environment, providing a new type of network storage solution that meets both throughput, availability, reliability, scalability, and manageability.

FC-SAN

A fiber channel storage area network (FC-SAN) is a system that enables multiple servers to access network storage devices by using fiber channel.

IP-SAN

The IP-SAN (IP storage) communication channel uses IP channels to connect servers to storage devices. In addition to the fastest development standard iSCSI, there are other standards such as FCIP and iFCP. Like Fiber Channel, IP storage is interchangeable. But unlike Fiber Channel, IP networks are mature and there is no interoperability issue. In IT industry, there are many network management software and service available by IP address [2].

IP storage overcome the limits of geographic distance between different servers. How far IP can extend and how far storage can extend is almost an epoch-making revolution that is ideal for remote backup of existing critical data.

Heartbeat

Detecting hardware failure in a node is a tricky issue. The solution is that all nodes send out a “heartbeat” – a small packet of data which essentially says “I’m alive” to the other nodes in every 1.2 seconds. If 5 sequential heartbeats are missed, the nodes take action and re-assess the cluster configuration. Once 1 or more unconnectable nodes have been detected, and the quorum voting process occurs, where possible, and if the cluster is still quorate, the work being performed will be redistributed. As active nodes, it continuously sends information to the rest of the nodes in the cluster. These messages are called heartbeat information. A private network for transmitting heartbeat information is called a heartbeat line. If the active node fails, the backup node does not receive heartbeat information within the specified period, then the failover will occur, and the backup node will take over the cluster and provide external services.

Quorum

Quorum can effectively define how many device failures a cluster can sustain before the cluster itself fails. The “quorum” device is just an area of storage which holds some basic cluster configuration. This is called the “*disk witness*”. It is often distributed which is the SAN model with an extra “quorum” disk. Sometimes hardware failure isn’t the cause of a cluster failing. Imagine a situation in which the network fails, effectively splitting the cluster in half. To decide which half continues running the cluster tasks, we need quorum because the cluster can’t differentiate between node failure and network failure. Quorum is a voting system which ensures that, if something goes wrong, at most one section of the cluster will continue running the cluster jobs.

There are four main approaches to defining quorum configuration [1]:

- **Node majority.** Each node gets 1 vote. Allows failure of half of the nodes, but no more. Normal for clusters with odd numbers of nodes.
- **Node and disk majority.** Each node and the disk witness get 1 vote. Allows failure of half of the nodes, if the disk witness remains online. Normal for clusters with even numbers of nodes.
- **Node and file share majority.** Same as disk majority, using a file share.
- **No majority, disk only.** Only the disk has a vote. Allows failure of all but the node but turns the disk into the single point of failure.

1.3 Cluster Terminology

Nodes: Physical computers that build clusters

Resources: IP addresses, disks, server applications, etc.

Resource Status: Mainly refers to whether resources are currently online or offline.

Resource Group: The smallest unit of failover.

1.4 RAID

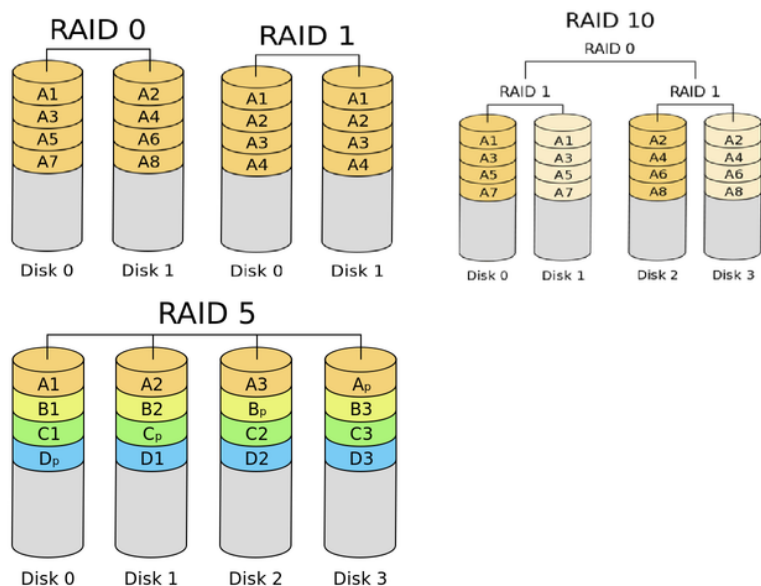
Redundant Arrays of Independent Drives (RAID) is composed of a large number of independent disks, which are combined into a huge disk group. By using RAID, the data is cut into a number of segments, which are stored on each hard disk.

The RAID can also take advantage of the concept of Parity Check. When any RAID in the array fails, the

data can still be read. When the data is reconstructed, the data is calculated and re-inserted into the new hard disk.

There are a lot of different ways a RAID may be configured [1].

RAID level



RAID 0 (striped)

RAID 0 is strange, because it isn't actually redundant. When we write data to a RAID 0 configuration, it's evenly split across each disk. It is extremely risky – if any disk in a RAID 0 configuration fails, the entire RAID fails. If you don't have a backup, your data is permanently lost, and cannot be recovered from the remaining functional disk(s).

RAID 1 (mirrored)

All data written to the array is written to all disks in the array simultaneously. Sometimes it is considered as the opposite of RAID 0 because it is absolute redundancy. Array rebuilds unlikely to result in data loss for large RAID 1 arrays.

RAID 10 (striped + mirrored)

This applies mirroring AND striping. It is good combination of performance and redundancy which contains most of the benefits of RAID 0 and RAID1. It Requires a minimum of 4 disks, and must always have a multiple of two. Failure of 2 disks can cause the array to fail, while it may tolerate the failure of up to 50% while it is very luck-dependent on which disks fail.

RAID 5

RAID 5 Uses a form of striping with an added (distributed) "parity" block which puts stress on a single disk. Parity information can be used to re-calculate the data on a failed disk, in combination with the other data. The parity value itself must be calculated, making RAID 5 more computationally expensive than the RAID levels we've looked at before it. It actually performs slightly worse than a single disk for write. Large numbers of disks are extremely inadvisable for this array type.

RAID 6 is same as RAID 5, but with two disks worth of distributed parity information.

Summarize

	Disk space usage	Read speed	Write speed	Redundancy
RAID 0	100%	N* single disk	N* single disk	None
RAID 1	1/N	Can only be read on one disk, depending on the faster disk	Reasonably well with the number of disks	100%
RAID 10	Between RAID 0 and RAID 1	due to striping	due to mirroring and/or striping	One disk in a pair of mirror disks can be used
RAID 5	(N-1)/N	(N-1)* single-disk	Poor write speed than a single disk	Only one disk is allowed to be damaged

Application of RAID technology:

DAS--Direct Access Storage Device

DAS was previously used on large and medium-sized machines. The latest form of using a DAS that includes a hard disk device on a PC is RAID. "Direct access" means that the time to access all data is the same.

NAS--Network Attached Storage

A special purpose server with an embedded software system that provides file sharing services across a variety of system platforms over the network.

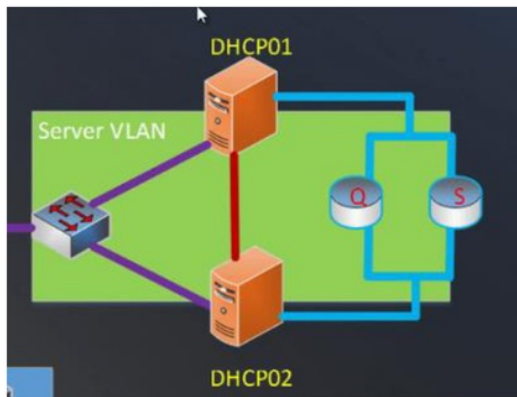
SAN--Storage Area Networks

That is what we talked about before.

2. Build a failover cluster

2.1 Network Topology

Here is the network topology of our Failover Cluster. In this Cluster, there are two networks. One is used to connect to public network, and one is used to communicate between cluster nodes, commonly known as "heartbeat connections."



- Purple line: public network
- Red line: heartbeat network
- Q disk: Qurum
- S disk: Data

IP address allocation

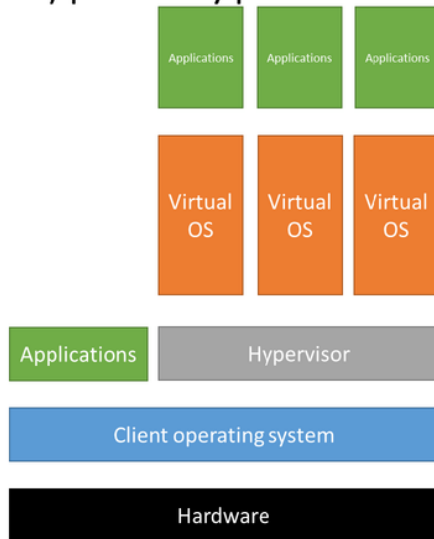
Server Name	Public	Heartbeat	Role
Cardiff	192.168.0.1/24		Domain Controller (Domain name: nwtrader.msft)
Swansea	192.168.0.2/24	192.168.11.2/24	Node 1
Bridgend	192.168.0.4/24	192.168.11.3/24	Node 2
Lab01-01	192.168.0.3/24		Client

2.2 Virtual machine preparation

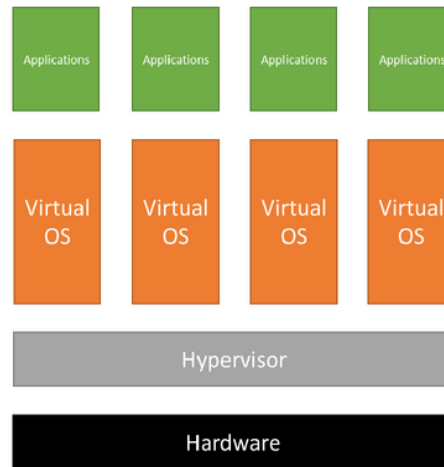
Install hypervisor platform.

Hypervisor also called VMM (virtual machine monitor). This is essentially an operating system inside an operating system. What we used is Microsoft Hyper-V. There are two type of hypervisor.

Type 2 hypervisor



Type 1 hypervisor



Type 1 is relatively easy to use which is not beyond the capabilities of less technical users. It allows the client operating system to run on “bare metal” which must navigate both the type 2 hypervisor and the client OS to get access to the physical hardware. It is generally used by users who want to use two or more operating systems simultaneously, but do not need good performance from the non-client OSes. It is not used for cloud computing.

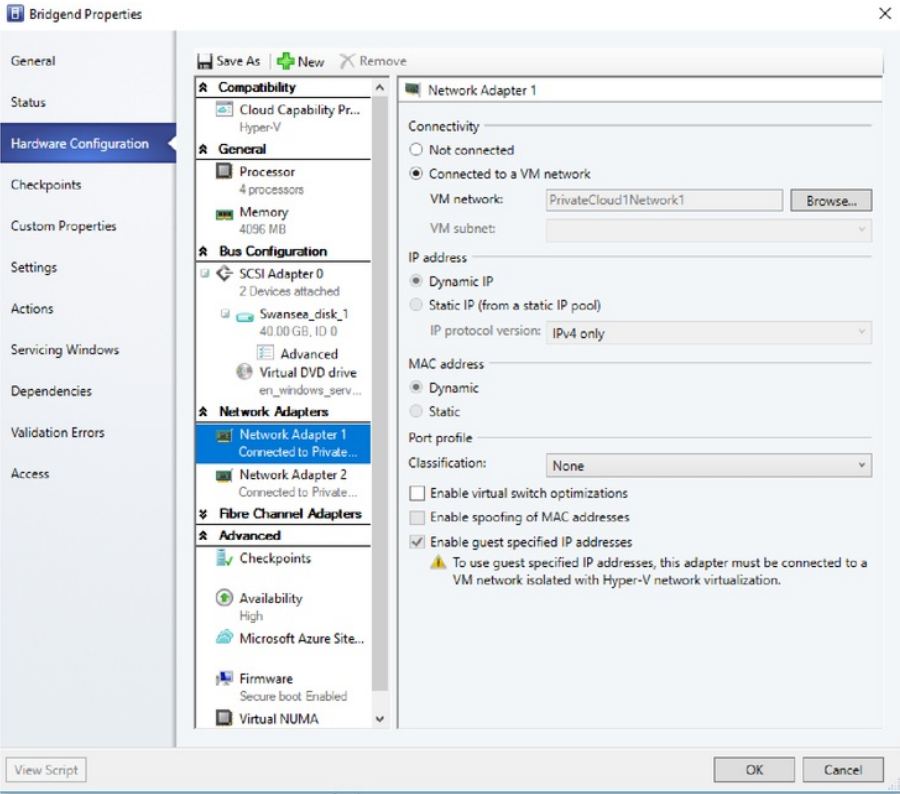
Type 2 hypervisor is much more difficult to use with far more advanced in the technology required. It is used for cloud implementations.

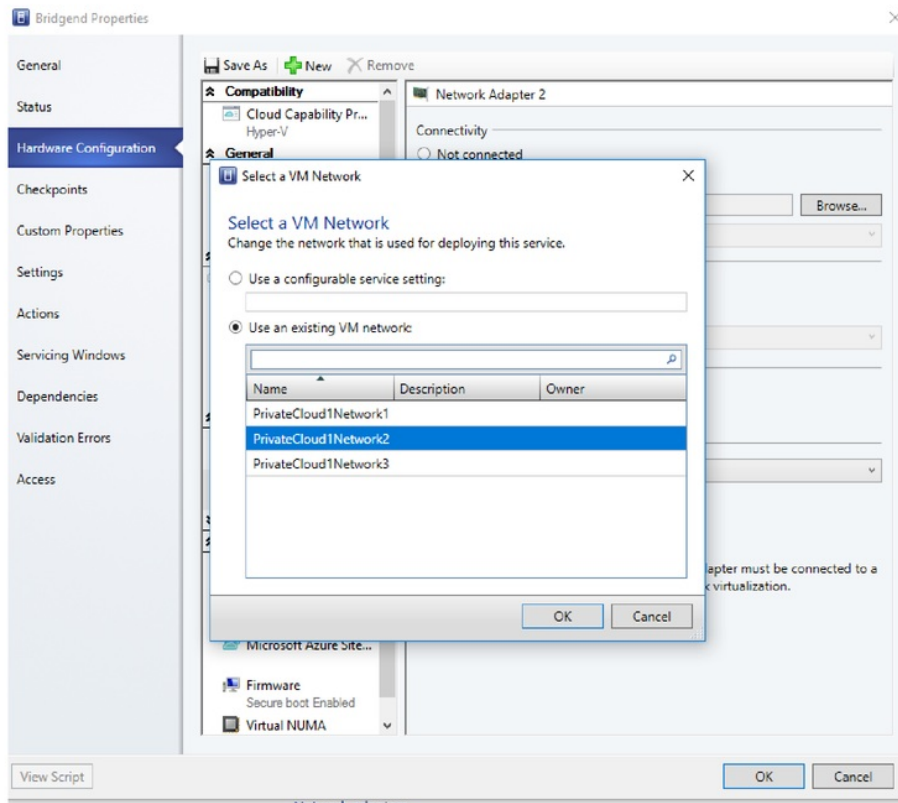
Create virtual machines and configure it

We create three virtual machines first, named them Cardiff, Swansea and Lab01-01.

Name	Status	Virtual Machine State	Host
Cardiff	Running	Running	
Lab01-01	Running	Running	
Swansea	Running	Running	

When configuring a virtual machine, it's worth noting that we chose to deploy the virtual machine to a private cloud.





Reasons for choosing a private cloud

Private clouds within the enterprise can provide:

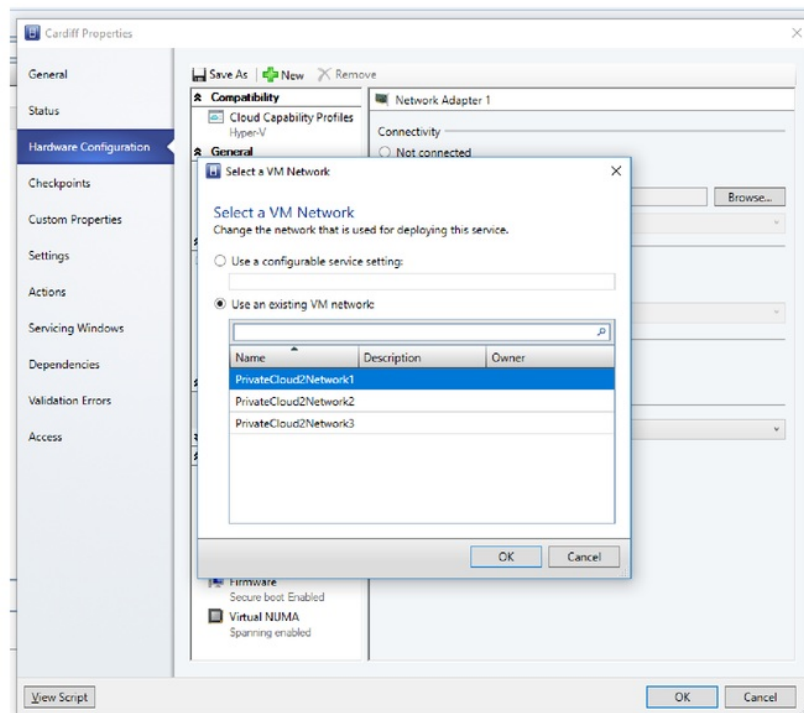
1. Possible increased security and performance control
2. Better test environment control than public clouds
3. Cloud-specific components
4. Better threshold level control than public clouds

Organizations want to test more efficient data center failover and scalability algorithms than are used by public cloud providers. Cloud providers do not allow organizations to test the algorithm of the provider. The private cloud provides the organization with control over the failover algorithm.

Install Window Server 2016 on the virtual machine

We install Windows Server 2016 as virtual machine in each operation system. Windows Server is the server operating system of Windows launched by Microsoft on April 24, 2003.

Select a VM network

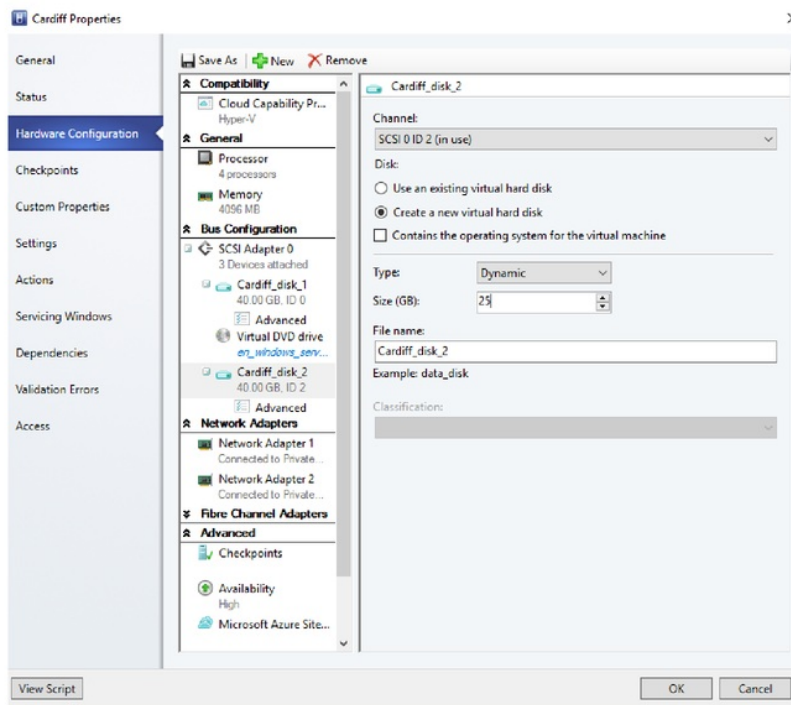


This is an existing VM network which is derive from the private cloud.

Remote volume management must be allowed on the node server or the shared folder cannot be created within the cluster. The method of operation is to allow remote volume management on the network location where the public network is located. It is also possible to disable the firewall directly.

Add a new virtual hard disk and the iSCSI target role (Cardiff)

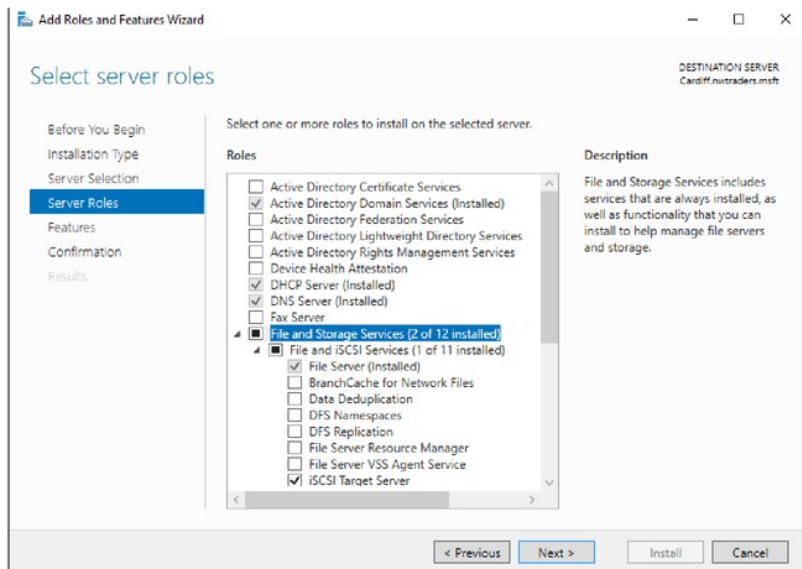
Cardiff acts as a domain controller. So shared storage resources are located in Cardiff to connect to the disk over the Internet using the iSCSI protocol to support storage for clustered Swansea and Bridgend nodes.



iSCSI (Internet Small Computer System Interface), also known as IP-SAN, is a storage technology based on the Internet and SCSI-3 protocol, proposed by the IETF, and became the official standard on February 11, 2003.

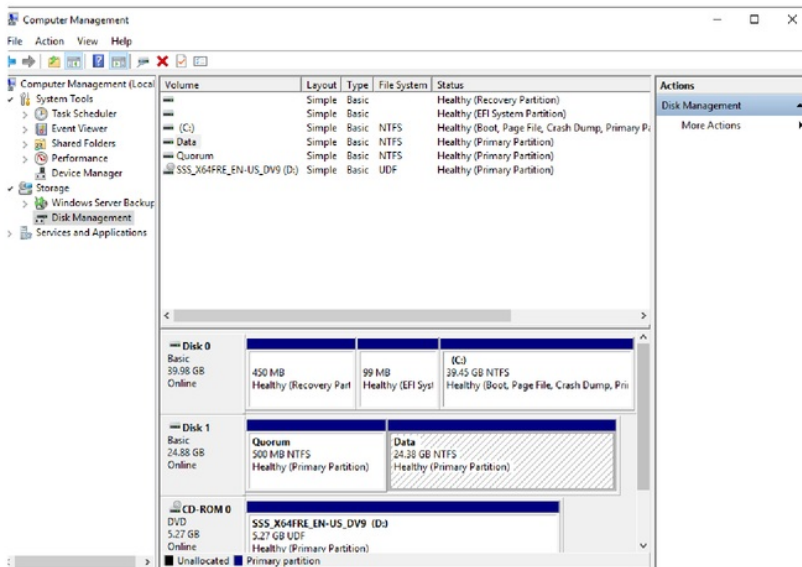
iSCSI uses the TCP/IP protocol (typically using TCP ports 860 and 3260). Essentially, iSCSI allows two hosts to negotiate with each other over an IP network and then exchange SCSI commands. In this way, iSCSI creates a storage area network (SAN) by emulating a common high-performance local storage bus with a wide area network. Without some SAN protocols, iSCSI does not require a dedicated cable; it can run on existing switching and IP infrastructure. However, if you do not use a dedicated network or subnet (LAN or VLAN), the deployment performance of the iSCSI SAN can be severely degraded [3].

The iSCSI model adopts the client/server mode. The client is called the initiator, and the server is called the target. When the data is transmitted, the initiator sends a request to the target, and the target responds to the response.



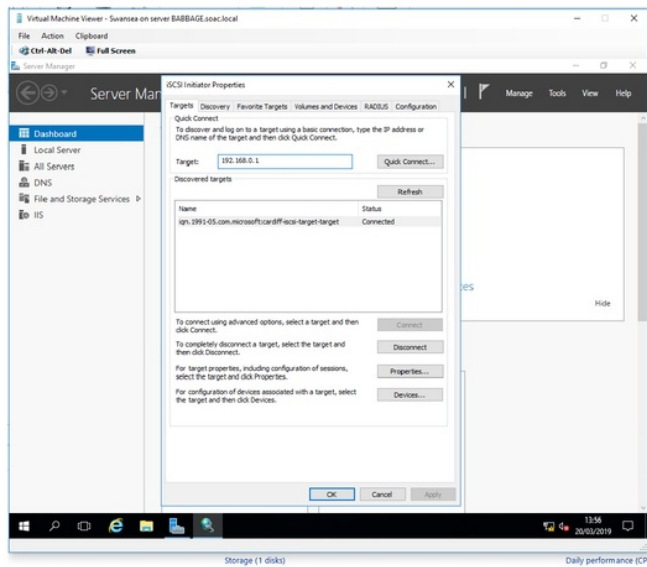
This disk is partitioned to include the quorum disk 500mb and the data disk that utilizes all remaining space. Format both with NTFS.

NTFS is a disk format designed specifically for the management security features of network which is supported by the Microsoft Windows NT kernel family of operating systems.



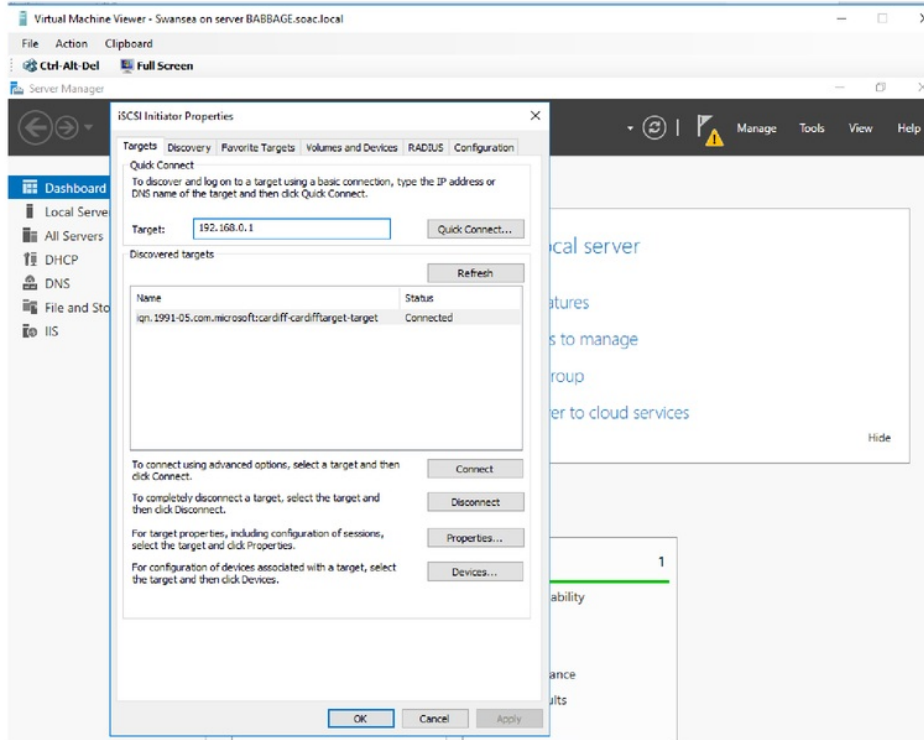
Create a new iSCSI initiator (Swansea)

It is aimed to connect shared storage on the domain controller.

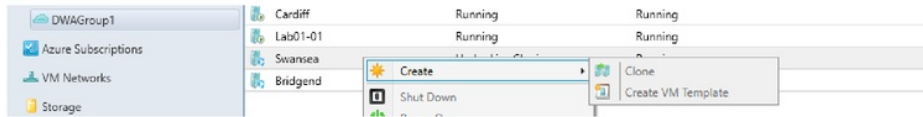


Run the iSCSI initiator

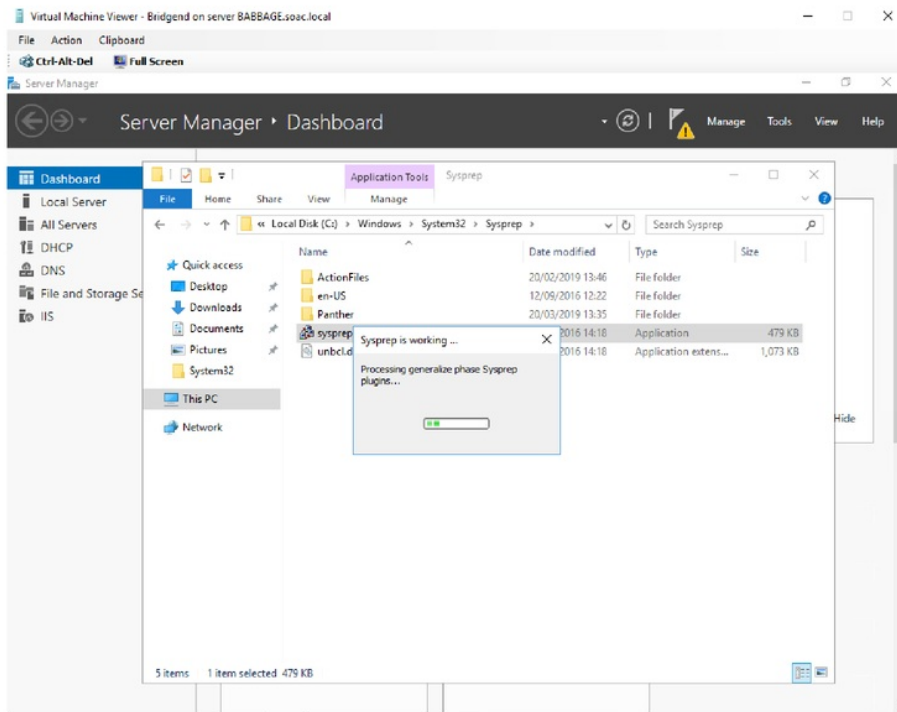
Discover the IP address of the domain controller, connect to it, then auto discover the targets.



Clone Swansea to create a new node – Bridgend.



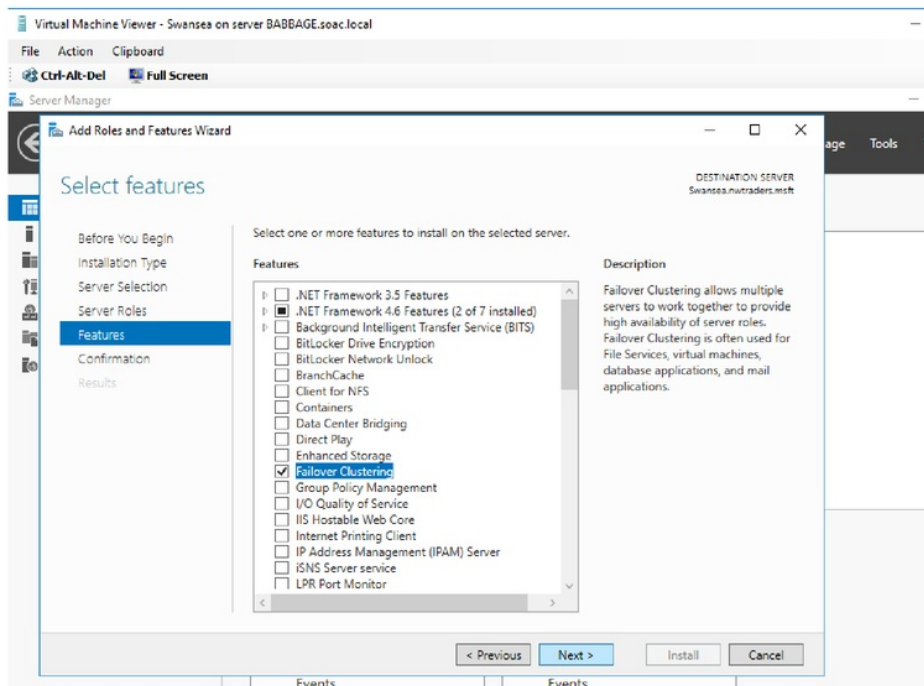
Open Bridgend and run sysprep to generate new ip address. (windows->system32->sysprep)



2.3 Failover Cluster

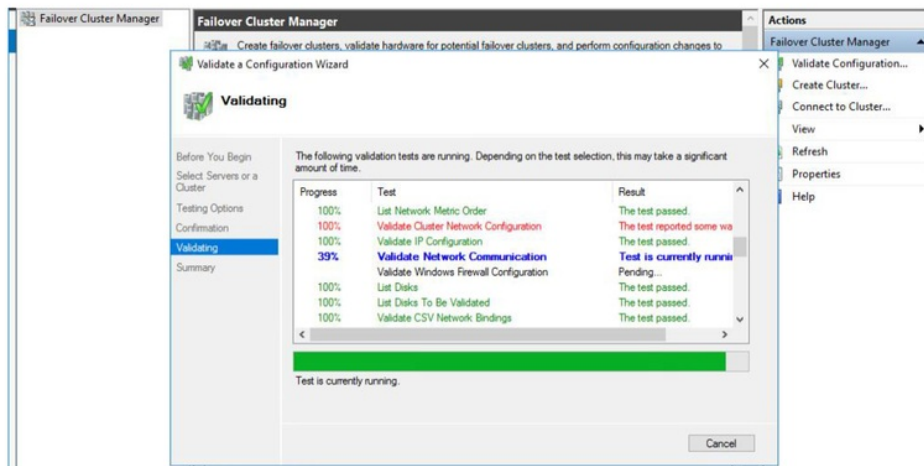
Failover Cluster installation

Add the failover clustering role on both nodes.

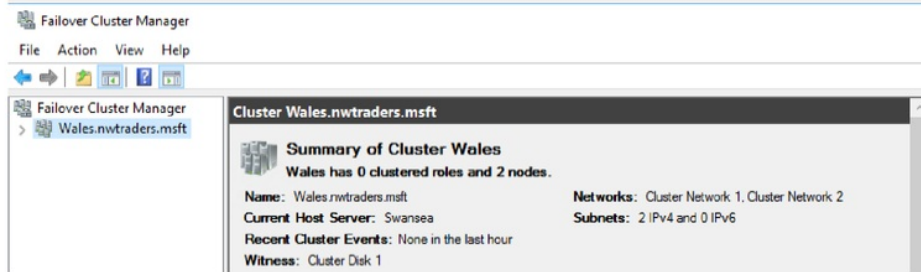


Failover Cluster configuration

Once complete, open the failover cluster manager. Combine Swansea and Bridgend. Add the servers and validate in either Swansea or Bridgend. Add both nodes and validate.

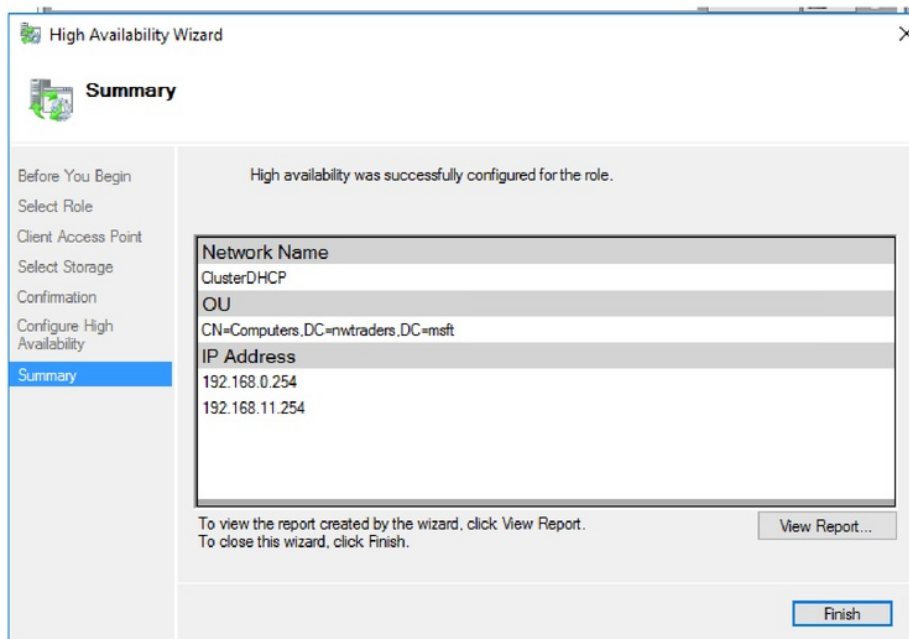


After this validates successfully, name the cluster as “Wales” and give it an IP address (192.168.0.254).



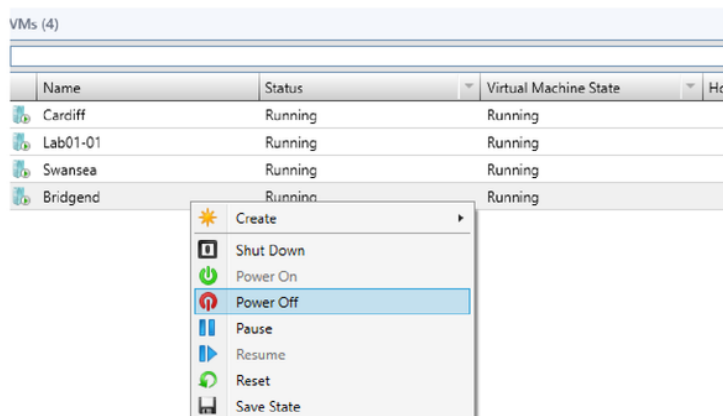
Install the DHCP role on both nodes

After installation, configure the failover DHCP servers to define a scope of IP on both nodes Swansea and Bridgend.

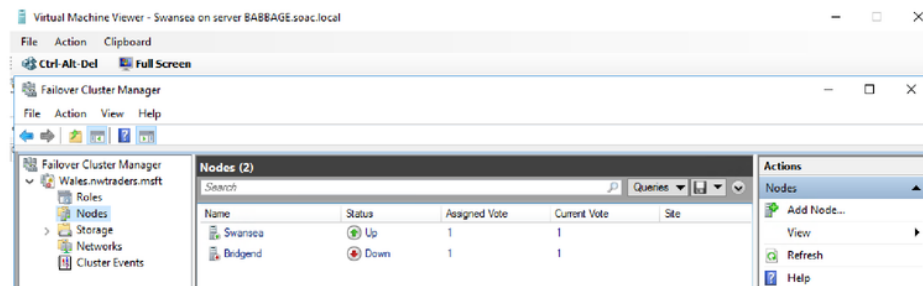


2.4 Test the Failover cluster

‘Power off’ Bridgend service.



Go back to the cluster manager and discover that the owner of the resource has switched to Swansea. Because the service on Bridgend has stopped, it is impossible to recover automatically. Still passing the heartbeat detection mechanism, after losing heartbeat messages (about 6 seconds), the node will be declared to fail.



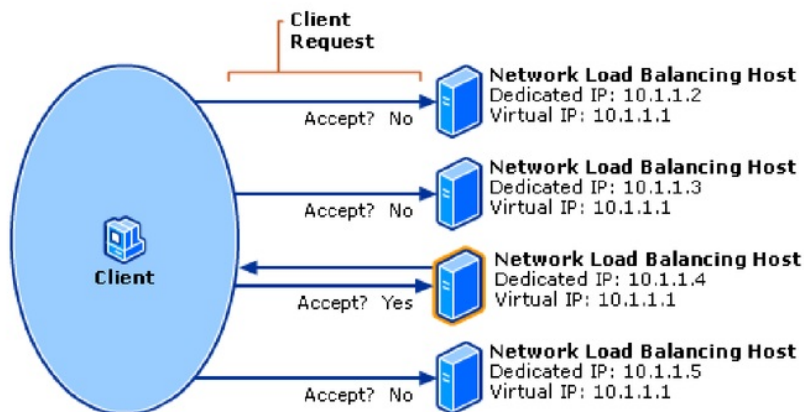
This test shows that when a cluster service on a node is stopped, resources running on the problem node are automatically moved to other normal nodes.

3. Expand Discussion

Network Load Balanced (NLB) Cluster

Network load balancing (NLB) clusters provide scalability and reliability to large/complex applications. In an NLB cluster, connections are routed, using a specified algorithm, to nodes in the cluster in turn. NLB clusters only operate on specific “stateless” applications, which are generally network heavy. Because the applications are stateless, it shouldn’t matter which node receives each request. [1]

The figure below is the work model of NLB cluster.



Service should be designed to be stateless, which is mainly considered from the perspective of scalability. If the server is stateless, then for the client, the request can be sent to any server, and then horizontal expansion can be achieved through load balancing.

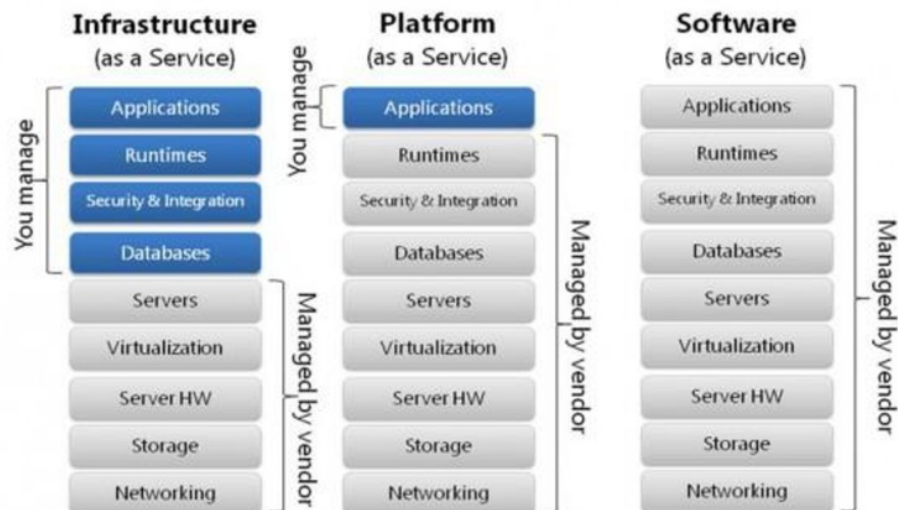
If the server is stateful, then it can't be easily implemented, because the client needs to always send the request to the same server. Failover supports stateful applications, such as Hyper-V and DHCP. It can also run stateless applications.

SPI

Cloud services can be divided into three categories:

- IaaS: Infrastructure-as-a-service
- PaaS: Platform-as-a-service
- SaaS: Software Services, Software-as-a-service

The relationship between different cloud services and software development is shown in the figure [4] .



IaaS

IaaS is the bottom level of cloud services which is mainly providing some basic resources. It is used in three ways: public, private and hybrid clouds [5].

VMware's virtualization technology has been applied to many Windows Server operating systems. Microsoft face the challenge of this powerful competitor. Microsoft's cloud computing strategy involves two different parts of a cloud computing software product and the interoperability that is almost non-existent between them. It has an enterprise private cloud product, which is built on Windows Server 2008, Hyper-V, system centre and its components. Windows Azure is Microsoft's Platform as a Service (PaaS) cloud computing platform.

PaaS

PaaS provides a software deployment platform that abstracts hardware and operating system details and scales seamlessly. Developers only need to focus on their business logic and don't need to focus on the bottom.

SaaS

SaaS means that enterprise put software development, management, deployment to third parties. They do not need to care about technical issues.

References

- [1] "Moodle," [Online]. Available: <https://moodle.uwtsd.ac.uk/course/view.php?id=2812>.
- [2] "IP SAN Fundamentals: An," [Online]. Available: <https://www.oracle.com/technetwork/systems/articles/ip-san-fundamentals-149896.pdf>.
- [3] M. Rouse, "iSCSI (Internet Small Computer System Interface)," [Online]. Available: <https://searchstorage.techtarget.com/definition/iSCSI>.
- [4] D. NG, "SaaS, PaaS and IaaS explained in one graphic," 12 July 2017. [Online]. Available: <https://m.oursky.com/saas-paas-and-iaas-explained-in-one-graphic-d56c3e6f4606>.
- [5] E. Colman, "WHAT'S THE DIFFERENCE BETWEEN SAAS, PAAS, IAAS" 27 August 2013. [Online]. Available: <https://www.computenext.com/blog/when-to-use-saas-paas-and-iaas/>.

FailoverCluster-report

GRADEMARK REPORT

FINAL GRADE

52/100

GENERAL COMMENTS

Instructor

Engagement with Literature Skills (50): 20

Knowledge and Understanding Skills (10): 5

Cognitive and Intellectual Skills (15): 8

Practical Application Skills (15): 12

Transferable Skills for Life and Professional Practice (10): 7

A low number of references of relatively low credibility. Always aim to use journal sources, not web sources. Consider how authoritative a source is on a particular subject.

A little low on detail generally - you're a fair way under the word count and this lack of detail has echoed through to the report overall.

Good practical work, reasonably well-documented.

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19
