EMC ISILON SMARTCONNECT

Optimize Scale-out Storage Performance and Availability

ABSTRACT

This white paper covers the built-in client connection balancing functionality found in the EMC Isilon scale-out NAS platform and details how the SmartConnect feature addresses customer challenges in this area. This white paper also compares the two versions of SmartConnect and the differences and benefits the licensed-version adds.

July 2014

Copyright © 2014 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided "as is." EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

VMware are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. EMC^2, EMC, the EMC logo, Isilon, OneFS, FlexProtect, SmartConnect, SmartLock, SmartPools, SnapshotIQ, and SyncIQ are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners.

Part Number H8316.1

TABLE OF CONTENTS

EXECUTIVE SUMMARY 4
AUDIENCE 4

WHAT IS SMARTCONNECT? 4

HOW DOES SMARTCONNECT WORK? 5
Client Connection Management 5
DNS Configuration 6
Client Connection Balancing Policy Options 6
Other Configuration Options 7
Dynamic NFS Failover and Failback with Performance Rebalance 7
Dynamic IP Allocation 8
NFS Failover 9
Rebalance Policy 10
Client Connection Behavior 10

SMARTCONNECT VERSIONS 10

HOW CAN SMARTCONNECT BE USED? 11
Use Case 1: Client Connection Balancing 11
Use Case 2: Connection Management for High-Availability Solution 12
Use Case 3: Connection Management for Compute and Network Service 13
Use Case 4: SmartConnect in maintenance & upgrades scenarios 14
Use Case 5: SmartConnect in mixed protocol environments 15

SUMMARY 15

EXECUTIVE SUMMARY

Today, enterprises face a tremendous increase in the amount of data used to conduct their everyday business. IT managers know that applications are using and storing video, audio, images, research sets and other forms of data, and are pushing the limits of traditional storage—sometimes to the breaking point. In addition, they are faced with the daunting task of managing both an ever-increasing number of clients and demanding enterprise mandates for failure resiliency. Meeting these requirements can seem overwhelming.

It is important that the storage system intelligently manages client connections for the following reasons:

- Client connections to the storage subsystem need to be **transparent** and appear as a single network element to the client.
- Client connections should be **balanced** evenly across all the nodes and prevent bottlenecks.
- Able to provide **varying levels of service** to different clients based on parameters.
- Able to provide a **resilient and reliable** connection. When one of the nodes has been removed from the cluster (either intentionally or otherwise), the clients continue to be able to access the data using the same connection (i.e. the connection fails over to one of the other existing nodes).
- When a new node is added to the cluster, the clients see no difference in access and possibly enhancing performance.

EMC Isilon® SmartConnect™ functionality allows IT Managers to meet the demands of an always-on, 24x7x365 world by ensuring the highest levels of performance and industry leading high-availability. With intelligent client connection load balancing and failover support, SmartConnect simplifies and optimizes scale-out network-attached storage (NAS) performance and availability.

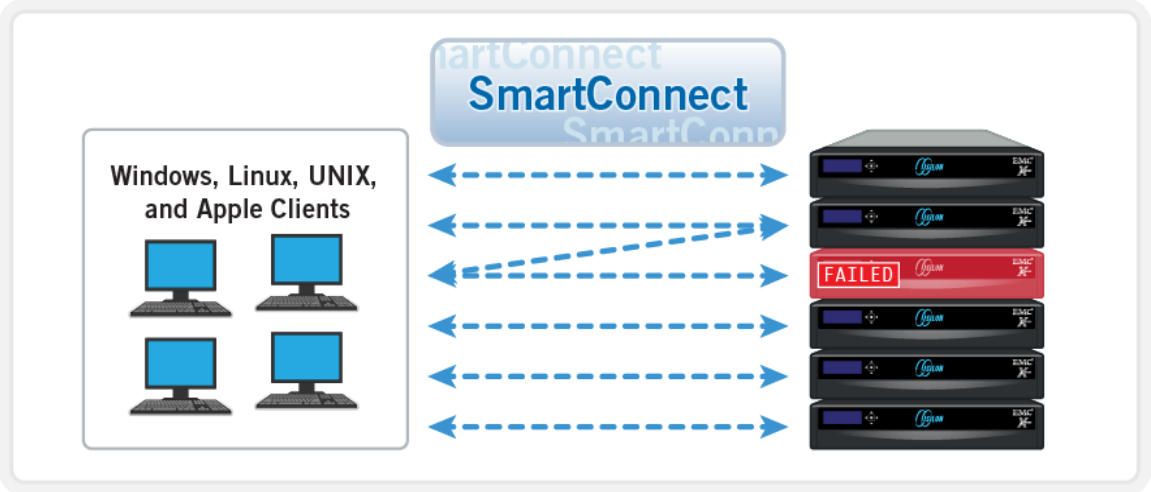
AUDIENCE

This white paper is intended business decision makers, IT managers, storage architects, and implementers.

WHAT IS SMARTCONNECT?

SmartConnect is a licensable software module of the EMC Isilon OneFS® operating system that optimizes performance and availability by enabling intelligent client connection load balancing and failover support. Through a single host name, SmartConnect enables client connection load balancing and dynamic NFS failover and failback of client connections across storage nodes to provide optimal utilization of the cluster resources. SmartConnect eliminates the need to install client side drivers, enabling the IT administrator to easily manage large numbers of clients with confidence. And in the event of a system failure, file system stability and availability are maintained.

To a client system, the cluster appears as a single network element. Both cluster and client performance can be enhanced when connections are more evenly distributed among nodes. SmartConnect provides intelligent connection balancing that does not require extensive configuration by users. Even in its minimum implementation, it can remove nodes that have gone offline from the request queue, and prevent new clients from mounting a node that is no longer available. In addition, SmartConnect can be configured so new nodes are automatically added to the connection balancing pool. By adding this intelligence, customers will be able to tune their Isilon cluster to achieve the resource utilization and performance goals for their business. SmartConnect simplifies management, enhances availability, and boosts performance of the cluster.



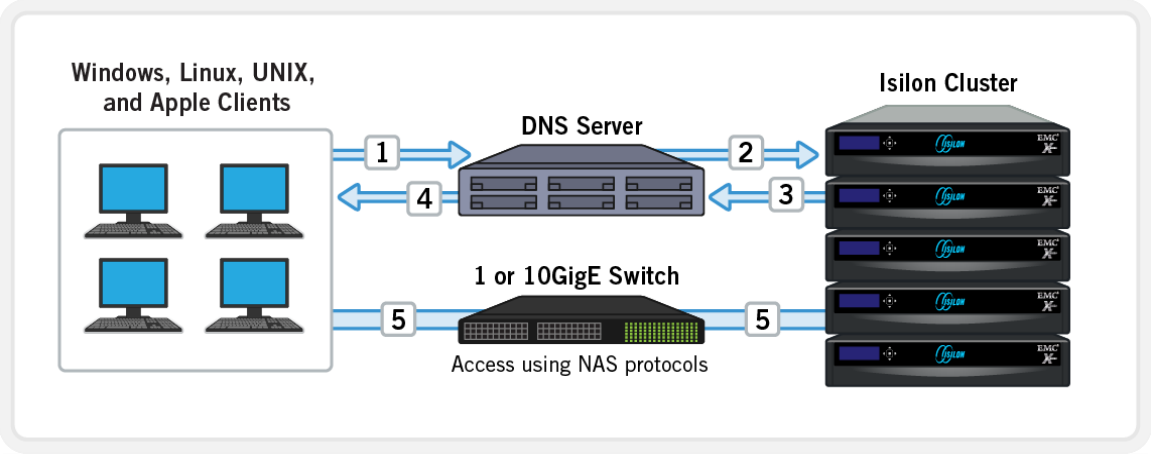
HOW DOES SMARTCONNECT WORK?

Client Connection Management

SmartConnect leverages the customer's existing DNS server by providing a layer of intelligence within the OneFS software application. Specifically, all clients are configured to make requests from the resident DNS server using a single DNS host name. Since all clients point to a single host name, it makes it easy to manage large numbers of clients. The resident DNS server will forward the lookup request for the delegated zone to the delegated zone's server of authority, in this case the SmartConnect Service IP (SIP) address on the cluster. If the node providing the SmartConnect service becomes unavailable, the SIP will move to a different node in the pool automatically.

The steps followed during client connection to Isilon cluster are:

1. The client attempts to connect to the Isilon cluster using a SmartConnect name which appears to the client as the hostname of the cluster. It does so by requesting a lookup for that host name from the environment's DNS server.
2. The environment's DNS Server determines that the SmartConnect name should be resolved by the SmartConnect Service IP (SIP) based on the delegation entry in the DNS. The DNS server queries SmartConnect using this SIP.
3. SmartConnect will act as the authority for this DNS lookup and will provide an IP address of a node based on the load balancing policy that has been selected for that zone.
4. The environment's DNS server responds back to the client's lookup with the IP address that SmartConnect provided.
5. Clients can then initiate connection to the node by mounting/mapping to the appropriate node based on the IP address returned from the DNS lookup.



All of this is transparent to the end user who can be confident that client connections are being balanced across the cluster ensuring optimal resource utilization and performance. If a node goes down, SmartConnect automatically removes the node IP from the available list of nodes, thereby ensuring that a client does not attempt to mount the unavailable node. Again, all of this is transparent to the customer/administrator.

The delegated server of authority is always the node with the lowest ID, unless it has either voluntarily or involuntarily surrendered its authority status. This node should always be available and if the status of the node changes and becomes unavailable, it will voluntarily surrender its role as server of authority.

DNS Configuration

Most customers with an Active Directory Service (ADS) domain have a resident DNS server and leveraging this server is a typical configuration for SmartConnect. However, use of the resident DNS server is not required for SmartConnect to work. For example, if an environment is a closed, secure, private network, a customer may be looking to limit the number of devices on the network and have no need or not want to add additional devices such as a DNS server. In those instances, customers can configure their clients to use SmartConnect as their DNS server and will be able to use the cluster with the connection balancing features. However, SmartConnect will not act as a general DNS server, and will not answer requests for anything other than the Isilon cluster.

A delegation (NS) entry needs to be added for the SmartConnect name, pointing to the SmartConnect Service IP as the nameserver.

BIND server:

In BIND, a new name server (NS) record needs to be added to the existing authoritative DNS zone specifying the server of authority for the new sub-zone. For that, an A record must be added, specified in the NS record that points to the SIP address of the cluster. For example, if the SmartConnect zone name is *cluster.example.com*, the DNS entries would looks like:

```
>>  cluster.example.com        IN     NS     sip.example.com
>>  sip.example.com            IN     A      {IP address}
```

Windows DNS Server:

In the Microsoft DNS wizard, a "New Delegation" record will be added in the forward lookup zone for the parent domain, which is equivalent to the NS record mentioned above.

Client Connection Balancing Policy Options

SmartConnect Advanced offers CPU utilization, connection counting, and aggregate throughput client connection policies in addition to the simple round robin policy. In the event that a node is unavailable it is removed from rotation until the node is available again. Each SmartConnect zone has its own load balancing policy.

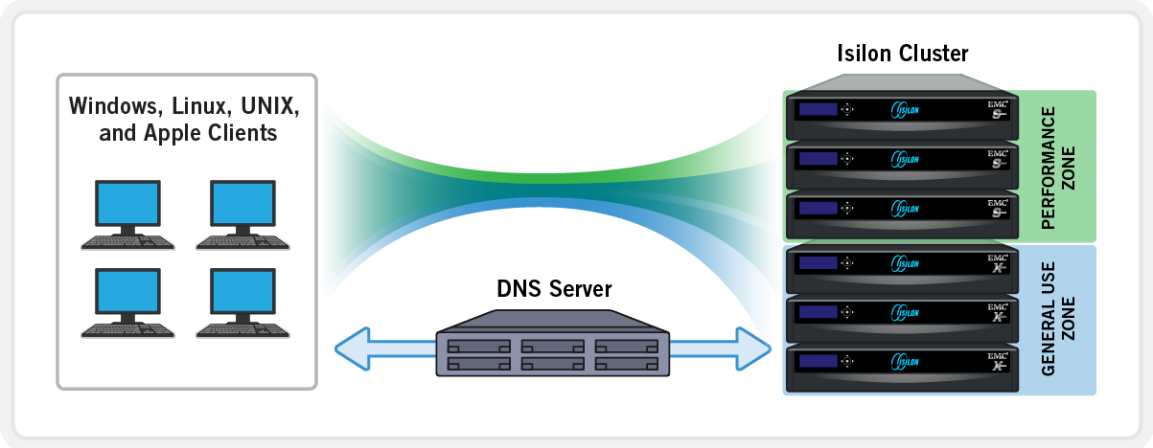
- **Round Robin** – This connection method works on a rotating basis, so that, as one node IP address is handed out, it moves to the back of the list; the next node IP address is handed out, and then it moves to the end of the list; and so on. This follows an orderly sequence to distribute client connections. This is the default state (once SmartConnect is activated) if no other policy is selected.
- **CPU Utilization** – This connection method examines CPU load on each node, and then attempts to distribute the connections to balance the workload evenly across all nodes in the cluster.
- **Connection Count** – In this algorithm, the number of established TCP connections is determined, and an attempt is made to balance connections evenly per node.
- **Network Throughput** – This method relies on an evaluation of the overall file system throughput per node, and then client connection balancing policies are used to distribute throughput consumption.

Each node will collect these statistics regularly (CPU Utilization – every 5-seconds, Connection Count and Network Throughput – every 10-seconds) and send to the delegated server of authority. This information is maintained in the delegated server of authority for the one minute (sliding window) and will be used to determine where a new connection request will be sent. These status messages also double up as the heart-beat from the nodes.

Other Configuration Options

SmartConnect provides the flexibility to implement different connection balancing strategies to meet various needs of the enterprise. The two enabling options include:

- **Zoning Strategies:** SmartConnect enables the ability to create zones by node and interface groupings. Zoning based on specific node interfaces gives a customer the flexibility to apply connection balancing policy across different work departments, subnets, or other groupings of clients. For example, the cluster could be made up different types of Isilon nodes (with differing performance characteristics). With SmartConnect, the IT manager will now be able to specify which clients can connect to which of these specific nodes. The diagram below illustrates two zones - one Performance and other General Use. SmartConnect policies can be defined to target client connection to these two zones.



- **Inclusion/Exclusion:** Another feature of SmartConnect is the ability to decide which Isilon nodes should participate in a specific connection balancing configuration strategy. In other words, any specific node(s) can be selected to be excluded or included from any or all balancing schemes for each Isilon cluster.

Dynamic NFS Failover and Failback with Performance Rebalance

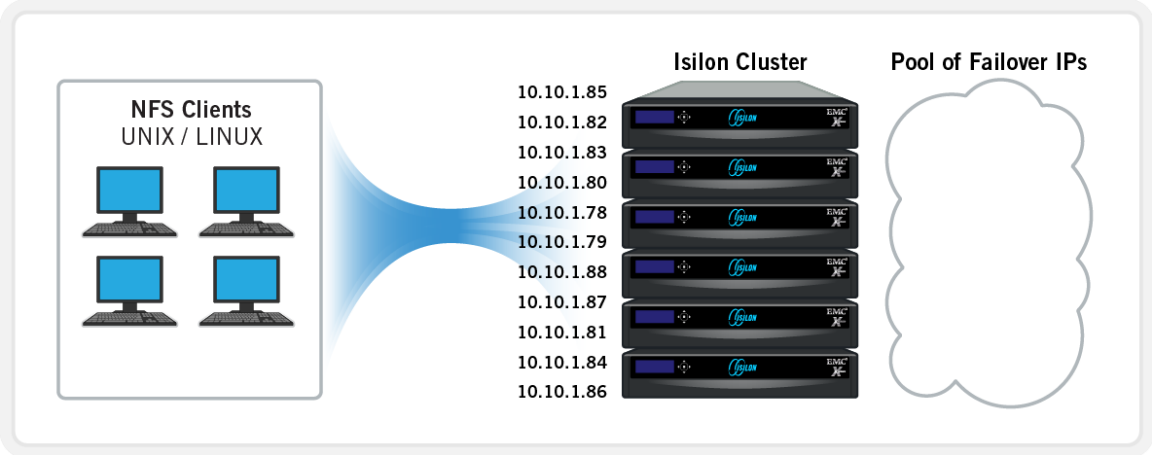
Traditional storage systems with two-way failover typically sustain at least 50% degradation in performance when a storage controller fails as all clients must fail over to the remaining controller. With Isilon, during failover, clients are evenly distributed across all remaining nodes in the cluster minimizing performance impact. This section describes this functionality in detail.

Dynamic IP Allocation

SmartConnect uses a virtual IP failover scheme that is specifically designed for Isilon scale-out NAS and does not require any client side drivers. Dynamic IP ("NFS failover IP") allocation is a feature of SmartConnect Advanced that distributes all IP addresses to the nodes participating in the IP address pool. IP address allocation controls how OneFS assigns IP addresses to the node interfaces in an IP address pool. If a node or an interface becomes unavailable, its IP addresses are automatically moved to other available node interfaces in the pool, which preserves NFS connections gracefully. When the offline node is brought back online, SmartConnect can be used to rebalance the NFS clients across the entire cluster maximizing storage and performance utilization.

Dynamic IP allocation has the following advantages:

- It enables NFS failover, which provides continuous NFS service on a cluster.
- It provides high-availability because dynamic IP addresses are available to clients at all times.



It is also important to understand the difference between "Static node IPs" vs. "NFS failover IPs". There is one Static IP address for each front-side network interface per node. Static node IPs do not participate in failover activity. Proper configuration will typically have all non-NFS clients connecting to Static node IP's.

The number of NFS failover IPs to create per zone is determined by the number of NFS clients that will be participating in a failover. Generally, optimal balancing behavior will be achieved when there is one NFS failover IP per client with a minimum of at least one failover IP per node in the cluster (i.e. if a 6-node cluster that has 20 NFS clients, a minimum of 6 NFS failover IP's should be

configured. But an optimal configuration will involve 20 NFS failover IP's. When it is not possible to reasonably configure a one to one relationship between number of NFS clients and NFS failover IPs, the decision should be based on the average client workload.

IP reallocation is triggered when:

- Node status changes
  - Node crashes
  - Node is power cycled
  - Node is rebooted
  - Node's back-end network fails or switch is rebooted
  - Spontaneous cluster split
- SmartConnect Zone change occurs – For example, when a node is added or removed
- Network status of external interface changes
- Manual triggered by admin
- Cronjob is setup to trigger it periodically (variation of the manual trigger case)

NFS Failover

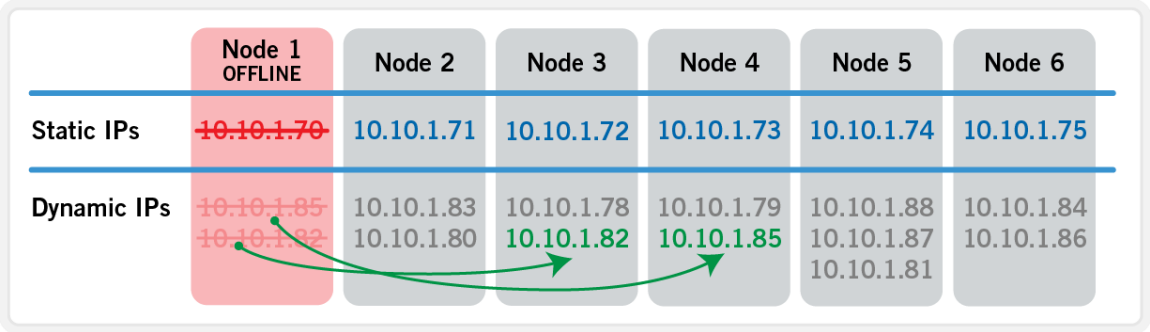
NFS failover provides high-availability by redistributing IP addresses among node interfaces in an IP address pool when one or more interfaces are unavailable. With NFS failover, SmartConnect Advanced ensures that all of the IP addresses in the pool are assigned to an available node. When a node goes down, the dynamic access IP addresses of the node will be redistributed among the remaining available nodes. Subsequent NFS client connections to the dynamically assigned IPs will be directed to the node newly assigned the address.

It is only the dynamic IPs (also called "NFS failover IPs") that SmartConnect manages in the implementation of failover for NFS clients. NFS failover IPs are configured when setting up SmartConnect zone when failover is "enabled". An administrator will create a pool of these NFS failover IPs that the cluster will hand out to NFS clients. When a node is brought down, the NFS failover IPs (and NFS clients connected to them) are dynamically moved, based on a configured "NFS failover load balancing" policy. When the disabled node is brought back on-line, SmartConnect rebalances all of the NFS failover IPs back across the cluster including the rejoined node based on the configured load balancing algorithm.

Here's an example of how NFS failover and failback works (based on the illustrations shown below). This is a six node Isilon cluster. Each node has a single Static node IP (10.10.1.70 - 75). A pool of NFS failover IP's have been created and distributed across the cluster (10.10.1.78 – 88).

	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
Static IPs	10.10.1.70	10.10.1.71	10.10.1.72	10.10.1.73	10.10.1.74	10.10.1.75
Dynamic IPs	10.10.1.85 10.10.1.82	10.10.1.83 10.10.1.80	10.10.1.78	10.10.1.79	10.10.1.88 10.10.1.87 10.10.1.81	10.10.1.84 10.10.1.86

"Node 1" in the Isilon cluster goes offline. The NFS failover IPs (and connected clients) associated with Node 1 failover to the remaining nodes based on a rebalancing policy/algorithm. The Static node IP for Node 1 is no longer available.

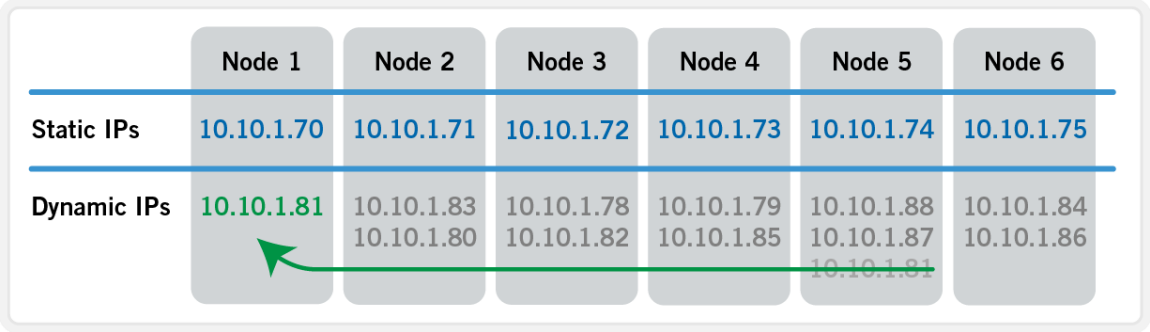


Rebalance Policy

IP rebalancing is a feature of SmartConnect Advanced that controls how IP addresses are redistributed when node interface members for a given IP address pool become available again after a period of unavailability. The rebalance policy could be:

- **Automatic Failback:** The policy automatically redistributes the IP addresses. This is triggered by a change to either the cluster membership, external network configuration or a member network interface.
- **Manual Failback:** IP rebalancing done manually from either Command line (isi smartconnect) or WebUI. This will cause all NFS failover IPs to rebalance within their respective Flexnet subnet. This will not affect the static node IPs that non-NFS clients should be mounted to (i.e. CIFS) if SmartConnect is properly configured.
- **Scheduled:** Implemented via cronjob on the cluster. By editing the crontab for "isi SmartConnect nfs rebalance" command the admin could set various policies such as hourly, daily, weekly.

In the above example, when the Node 1 is brought back online, SmartConnect rebalances the NFS failover IPs (and connected clients) back across the entire cluster based on the chose rebalancing algorithm.



Client Connection Behavior

If a node goes offline that has client connections established, the behavior is protocol-specific. If the IP address gets moved off an interface because that interface went down, the TCP connection is reset. NFS will re-establish the connection with the IP on the new interface and retry the last NFS operation. For SMB though, the SMBv1 and v2 protocols are stateful and so when an IP is moved to an interface on a different node, the connection is broken because the state is lost. HTTP protocol may recover gracefully as well.

SmartConnect Versions

SmartConnect is available in two versions – Basic and Advanced. The SmartConnect Basic version of the application manages client connections using a simple round robin client connection balancing policy within a single management zone. The Basic version is

included with Isilon's OneFS operating system as a standard feature. The SmartConnect Advanced version, a licensable software module, offers advanced options such as CPU utilization, connection counting, and aggregate throughput client connection balancing policies in addition to the simple round robin policy. This version allows multiple management zones to be defined to support multiple subnets and supports dynamic NFS failover and Performance Rebalance across Isilon clusters. In addition, a license key must be obtained for SmartConnect Advanced activation.

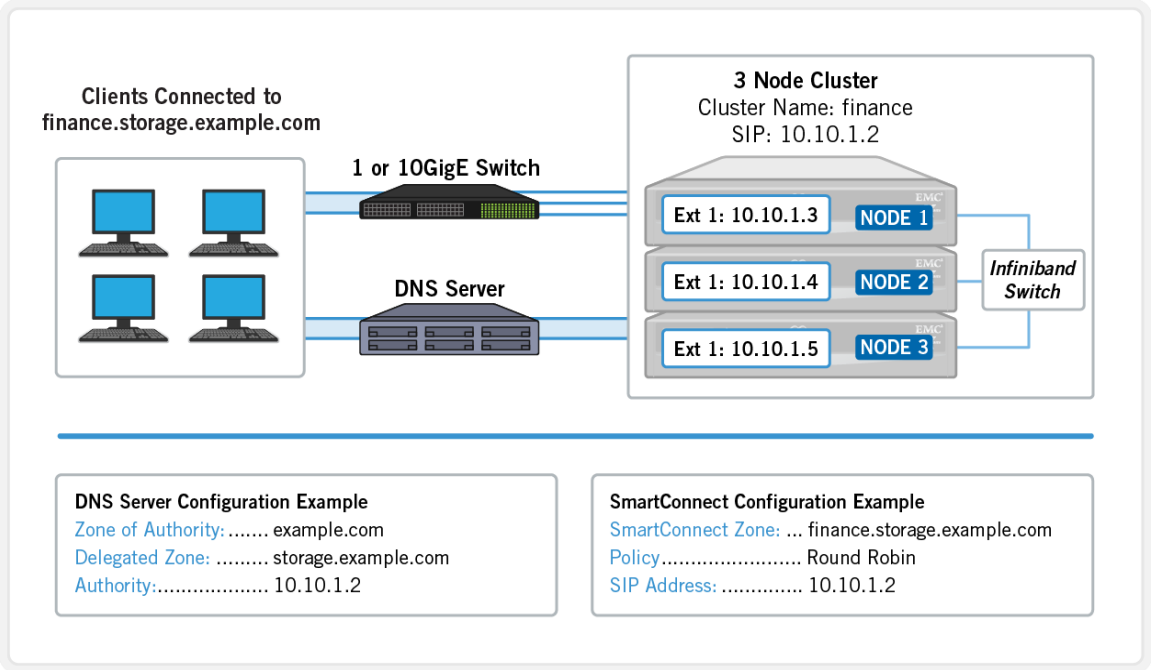
	SmartConnect Basic	SmartConnect Advanced
Load Balancing - Round Robin	Yes	Yes
Load Balancing - CPU Utilization	No	Yes
Load Balancing - Connection Count	No	Yes
Load Balancing - Network Throughput	No	Yes
NFS Failover	No	Yes
IP Allocation	Static	Dynamic
SmartConnect Zones	Single Zone per Subnet	Multiple Zones
Rebalance Policy	No	Yes
IP Failover Policy	No	Yes (configurable)

How can SmartConnect be used?

This section reviews a number of configuration examples that demonstrate configuration options for SmartConnect. The examples assume the reader has working understanding of configuring the Isilon clustered storage solution, SmartConnect Zones, Flexnet Subnet, and DNS Zones. The examples covered include the following use cases:

Use Case 1: Client Connection Balancing

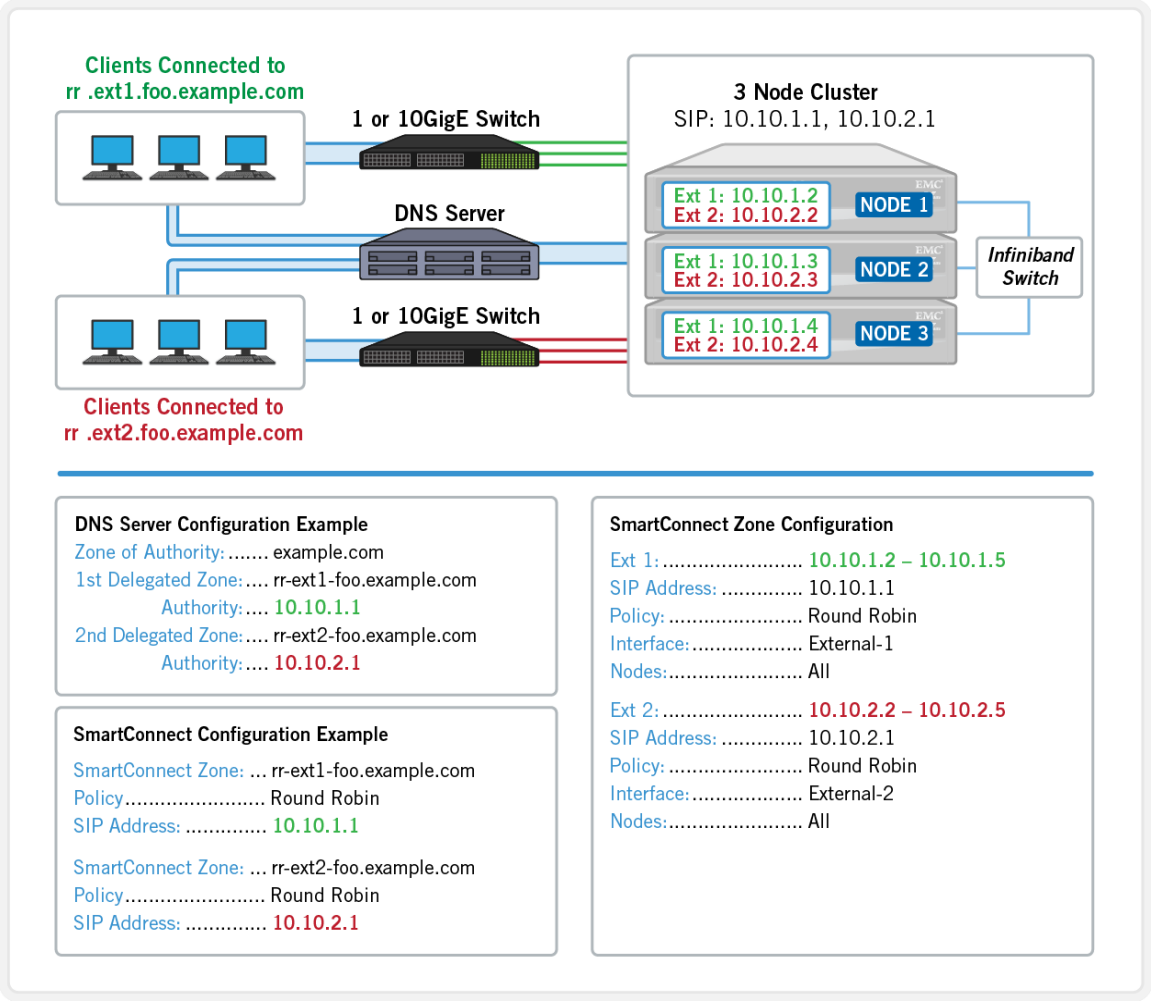
This first configuration example is a basic setup that performs connection management for client connections at time of mount request. Only one of the front side NIC ports are used on each node of the cluster (single-subnet) and the single SmartConnect zone includes all three nodes in the cluster. For this three node cluster, the Static node IP's are 10.10.1.3 – 5. NFS failover is not enabled in this configuration.



The DNS points to the SIP (10.10.1.2) as the authority when a lookup is performed for isilon.com. OneFS will then respond with a node's IP address in round-robin order, distributing client connections across the cluster.

Use Case 2: Connection Management for High-Availability Solution

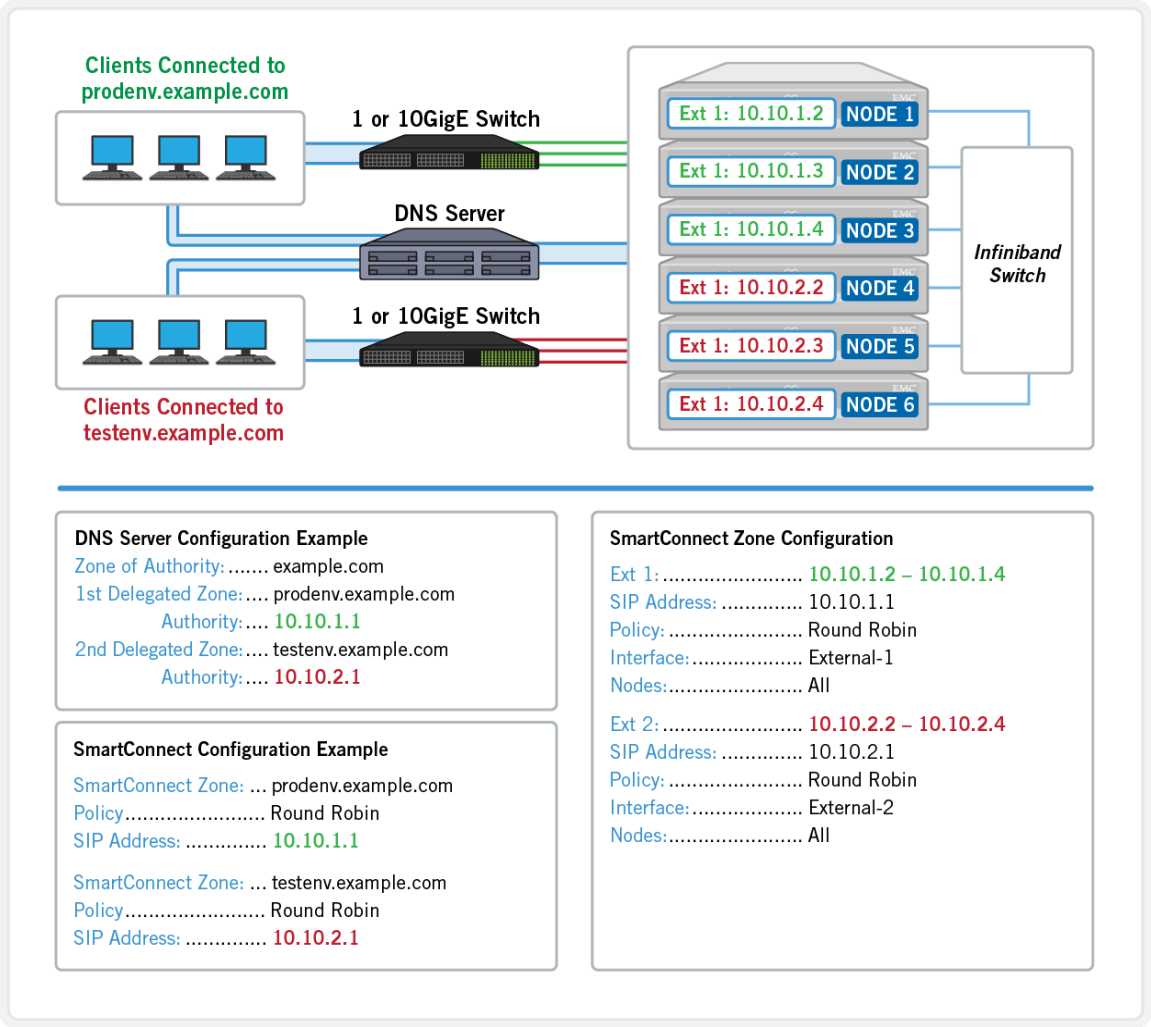
In this example, there are two front-side switches for a high availability solution. Multiple-subnets are configured on the Isilon cluster to use both front-side switches. Two available NIC ports on each node in the cluster are configured to separate subnets. The DNS configuration has zone delegation defined per subnet. The SmartConnect zones and Flexnet subnet are configured to allow for leveraging both external front-side NIC ports (Flexnet subnet Ext 1 and Ext 2). The Static node IP's are for Ext 1: 10.10.1.2 – 5 for Ext 2: 10.10.2.2 – 5. NFS failover is not enabled in this configuration. This configuration allows for a more optimal set-up for larger number of clients sitting on two separate networks.



Use Case 3: Connection Management for Compute and Network Service

In some customer sites, there are shared servers that host multiple shares and/or exports for different needs. Connection and access to the different shares might have different requirements depending on the client. With SmartConnect, we can manage which nodes the clients connect to and is based on the SmartConnect zone to which it connects.

For example, in the following illustration, Nodes 1, 2, and 3 are high performing nodes (S- or X-Series nodes) and Nodes 4, 5 and 6 are more for general use (NL-Series node). The company has two separate workflows for production and test environments each of which have to access the same network share/NFS import on the Isilon node. With the SmartConnect setting as shown in the diagram below, all traffic for the production environment (prodenv.isilon.com) will go to Node 1, 2, and 3, and all traffic for the test environment (testenv.isilon.com) will go to Nodes 4, 5 and 6.



This provides some separation and different levels of service for the test and production environments. The production environment is allowed the use of the high performance nodes and the test environment has access to the same data through the other set of nodes.

It is important to note that an interface can be made to be a member of multiple pools, SmartConnect zones as well as multiple subnets (as shown in the example above). There are effectively no limits that the system places on how a customer can configure their network environment.

Another variation of this use case is when the IT administrator can create a zone for a GigE and another for the 10GigE interfaces and provide different levels of network connections to clients.

Use Case 4: SmartConnect in maintenance & upgrades scenarios

SmartConnect can be used effectively using two important maintenance scenarios

- **Pull out Nodes for Maintenance**: IT administrators can use the "suspend" and "resume" commands when they have to pull out nodes out of the SmartConnect rotation in order to perform maintenance
- **Rolling Upgrades:** For minor releases, Isilon cluster nodes are upgraded node-by-node (also called the rolling upgrade). During the period the nodes have to be rebooted, it is bounced and will not receive any new connections (similar to the suspect command mentioned above). The failover IPs are drained to the other nodes, which will then serve the client requests.

Use Case 5: SmartConnect in mixed protocol environments

Isilon customers have deployed mixed workloads with both SMB and NFS connections to the same cluster import/share. In these cases, a different SmartConnect zone is created for each of these workloads. The NFS clients can be put in a dedicated SmartConnect zone that will facilitate failover while the SMB clients are put into another SmartConnect zone that will not participate in failover. This will ensure all SMB clients mount to the "Static node IPs" which do not failover. If an SMB client is put into a zone where NFS failover is enabled, the clients will experience more frequent lost connection on those SMB mounts requiring re-establishment of connections.

Summary

SmartConnect provides intelligence and automation that improve availability, increase productivity, and help IT managers to avoid unnecessary complexity in their NAS environment. Whether there are a few clients or thousands or whether there is a 3-node or 144-node Isilon scale-out NAS cluster, SmartConnect makes client management simple—requiring few management resources. For NFS environments, SmartConnect provides automated N-way failover/failback and automated rebalancing across the cluster, handling both failure scenarios and load distribution. SmartConnect delivers industry-leading levels of high-availability and optimized performance for Isilon's scale-out NAS solutions.

Through a single virtual host name, IT managers can use SmartConnect as the cornerstone for providing seamless enterprise-wide client access to an Isilon storage cluster. Clients and storage nodes can be added and the cluster scaled on the fly while a system is in production, making Isilon's scale-out NAS platform with SmartConnect the industry's most flexible, powerful, and easy-to-manage clustered storage solution. With its intelligent client connection load balancing and NFS failover support, SmartConnect achieves breakthrough levels of performance and availability, enabling IT Managers to meet the ever-increasing demands being placed on them to ensure always-on, always-available performance with Isilon scale-out NAS solutions.