

Scale Computing Networking Guidelines and Recommendations

Reference Guide

Any information listed here is not a substitute for the product's user guide or support and is not covered under the support and warranty for the HC3 system. This is meant only as a reference guide as it may relate to the Scale Computing HC3 system.



Networking Guidelines and Recommendations

You can find the most up-to-date technical documentation on the Scale Computing Portal at:

<http://www.scalecomputing.com/support/login/>

The Scale Computing site also provides the latest product updates and information:

<http://www.scalecomputing.com/>

Provide feedback on this document at:

documentation@scalecomputing.com

Contact support at:

support@scalecomputing.com

<http://www.scalecomputing.com/support/login/>

US: +1-877-SCALE-59 (877-722-5359)

Europe: 0808 234 0699

Document Version 6.0: 12/2017

Corporate Headquarters

The Union 525
525 South Meridian Street
Indianapolis, IN 46225
P. +1 877-722-5359

West Coast Office

360 Ritch Street
Suite 300
San Francisco, CA 94107

EMEA Office

Saunders House
52-53 The Mall
London
W5 3TA
United Kingdom

www.scalecomputing.com

1-877-SCALE-59 (877-722-5359)

Contents

Introduction	5
Minimum 1 GbE Switch Attributes	5
Minimum 10 GbE Switch Attributes	5
Recommended Switches	6
1 GbE Switches	6
10 GbE Switch	6
Known Successful Switches (Untested)	7
1 GbE Switches	7
10 GbE Switch	7
Non-Recommended Switches	7
HC3 10 GbE Cable Requirements	8
HC1100 (10 GbE) HC3 Family Nodes	8
HC2000z and HC4000z (10 GbE) HC3 Family Nodes	8
HC5150 (10 GbE) HC3 Family Nodes	8
Dell Networking	9
Scale Computing Stock	9
Dell N2000 Series: Stacking Cable Product Numbers	9
10 GbE to 1 GbE Converter	9
HP Networking	9
Scale Computing Stock	9
HP 2930F Series: Stacking Module and Cable Product Numbers	9
10 GbE to 1 GbE Converter	9
Mellanox Networking	10
Scale Computing Stock	10
Mellanox SX1012: QSFP to SPF+ Adapter (40 GbE to 10 GbE Adapter)	10

Cisco Networking	10
Cisco Only: 10 GbE to 1 GbE Converter	10
Port Requirements	11
Public and Private System Networks	11
Public Network Provisioning	11
Private Network Provisioning	11
Tips and Examples for Network Provisioning	12
Network Interface Layout	13
HC1100 Series	13
HC2100 Series	13
HC4100 Series	15
HC5100 Series	15
Spanning Tree (STP)	16
Flow Control	16
Switch and Node Redundancy	16
High Availability	17
Feedback and Support	19
Document Feedback	19
Technical Support and Resources	19
Disclaimer	19

Introduction

The content here is meant as a guideline and recommendation for networking equipment and its use with Scale Computing HC3 systems. This guide covers general concepts rather than specific configuration options. If you need information on specific configuration information for a switch or other networking product you can refer to the guides in the Scale Computing Knowledge base or the website of the manufacturers of the switch or product.

NOTE: Scale Computing does offer Professional Service engagements for an additional fee to assist with initial networking and switch configuration for the HC3 system. These services are offered dependent on purchasing the proper recommended hardware products found in this guide. Contact your Scale Computing Sales Representative for full details on Professional Service offerings.

Minimum 1 GbE Switch Attributes

When using 1 GbE switches other than a recommended switch, the following features and attributes should be considered. Optional attributes are strongly encouraged but may not be required depending on your deployment needs.

- Managed switch
- 72/144 mpps for 24/48 port switches respectively
- Allows disabling of spanning tree protocol (STP) at the switch level
- Supports 802.3x flow control
- (Optional*) VLAN support
- (Optional) Allows disabling of spanning tree protocol (STP) on a per port basis
- (Optional) Offers rapid spanning tree protocol (Rapid STP)
- (Optional) Stacking or dedicated interconnect for High Availability (HA)

Minimum 10 GbE Switch Attributes

When using 10 GbE switches other than a recommended switch, the following features and attributes should be considered. Optional attributes are strongly encouraged but may not be required depending on your deployment needs.

- Managed switch
- SFP+ ports
- Allows disabling of spanning tree protocol (STP) at the switch level
- Supports 802.3x flow control
- (Optional*) VLAN support
- (Optional) Allows disabling of spanning tree protocol (STP) on a per port basis
- (Optional) Offers rapid spanning tree protocol (Rapid STP)
- (Optional) Stacking or dedicated interconnect for High Availability (HA)

NOTE: 10 GbE switches do not have the same limitations as 1 GbE versions. Most small business or enterprise 10 GbE switch models should be compatible with the HC3 family. Contact ScaleCare Support if you have any questions on specific 10 GbE switch models.

Recommended Switches

These switches are expected to work based on published specifications, successful customer deployments, and/or internal testing. Switches not listed here may work as long as they meet the stated requirements above in **Minimum Switch Attributes**.

WARNING

If it is determined that performance or cluster redundancy issues are due to network latency, dropped packets, or other network issues, a recommended switch from Scale Computing may become the required action from ScaleCare Support.

1 GbE Switches

- Lenovo RackSwitch G7028
 - Datasheet: <http://bit.ly/2k0H449>
- Lenovo RackSwitch G8052
 - Datasheet: <http://bit.ly/2k3MRpw>
- Juniper EX Series EX2300-24P
 - Datasheet: <http://juni.pr/2BNZugF>
- Dell Networking N2000 Series
 - Datasheet: <http://dell.to/1UdbYSq>
- HP Aruba Procurve 2920 Series
 - Datasheet: <http://bit.ly/1xcpHkQ>
- HP Aruba Procurve 2930F Series
 - Datasheet: <http://bit.ly/2jLC8T5>

10 GbE Switch

- Lenovo RackSwitch G8124E
 - Datasheet: <http://bit.ly/2AxWPtf>
- Lenovo RackSwitch G8272
 - Datasheet: <http://bit.ly/2AV9zLh>
- Dell Networking N4000 Series
 - Datasheet: <http://bit.ly/2it7tHE>
- Dell X4012
 - Datasheet: <http://bit.ly/1MRTNz6>

NOTE: The X4012 is a low-cost 12-port 10 GbE switch for use with the entire HC3 family. This switch may have limitations for virtual machines with high memory workloads. Contact a support representative if you have questions regarding this limitation.

- Mellanox SX1012
 - Datasheet: <http://bit.ly/12p1Y24>

Known Successful Switches (Untested)

These switches have been seen to be successful in various customer deployments. Switches listed here have not been tested by Scale Computing and Scale Computing does not offer any guidelines for configuration outside what is listed in this guide in **Spanning Tree (STP)** and **Flow Control**. Switches not listed here may work as long as they meet the stated requirements above in **Minimum Switch Attributes**.

WARNING

If it is determined that performance or cluster redundancy issues are due to network latency, dropped packets, or other network issues, a **Recommended Switch** may become the required action from ScaleCare Support.

I GbE Switches

- Cisco 2960XR Series
 - 2960X-48-TD-L
 - 2960X-48-LPS-L
 - Datasheet: <http://bit.ly/Iro00XI>
- HP Procurve 5406ZI Modular Switch (1GbE Module)
 - HP Part Number (Switch): J9866A

10 GbE Switch

- HP Procurve 5406ZI Modular Switch (10GbE Module)
 - HP Part Number (Switch): J9866A
- Netgear XSM7224s Switch (M7300 Series)
 - Netgear Product Number: M7300-24XF
 - Datasheet: <http://bit.ly/IEibIAL>

Non-Recommended Switches

These switches have not performed well or have performed inconsistently in the field in a wide variety of use cases.

- Cisco Catalyst 3650
- Cisco Catalyst 3750 Family
- HP Procurve 1910
- Home, Workgroup, Departmental, or Unmanaged switches do not have sufficient switching performance and will not be functional with the cluster

HC3 10 GbE Cable Requirements

WARNING

There is no industry standard for 10 GbE SFP+ adapters and the protocols they use. It is the customer's responsibility to ensure they are purchasing cables that are compatible with not only their make and model of switch but also the NICs installed in their purchased Scale Computing nodes. Scale Computing is not responsible for improper functionality or performance of the HC3 system due to a cable mismatch and/or malfunction.

HC1100 (10 GbE) HC3 Family Nodes

- Intel X710 NICs
- Any SFP passive or active limiting direct attach copper cable that complies with the SFF-8431 v4.1 and SFF-8472 v10.4 specifications
- SFF-8472 identifier must have value 03h for SFP or SFP+ (verify the value with the cable manufacturer)
- Maximum cable length for passive cables is 7 meters
- Support for active cables requires Intel® Network Connections software version 15.3 or later
- Intel® Ethernet SFP SR Optics and Intel® Ethernet SFP LR Optics are the only 10-Gbps optical modules supported; you cannot use other brands of SFP (10 GbE) modules with these adapters
- Only the modules listed in the table under “What SFP+ and SFP modules are supported?” on the Intel FAQ page will work with the X710. Find the latest FAQ [here](#).

HC2000z and HC4000z (10 GbE) HC3 Family Nodes

- Intel X520 Dual Port NICs
- Direct attach passive copper cables compliant with the SFF-8431 v4.1 and SFF-8472 v10.4 specifications required
- **Supports SFP Direct Attached PASSIVE cables (Twin Axial) ONLY**
- SFP optical modules **cannot** be used with this adapter; if you plug an optical module into this adapter, the module will not be recognized
- Maximum cable length is 7 meters
- For more information visit Intel's website for the X520 NIC specifications [here](#)

HC5150 (10 GbE) HC3 Family Nodes

- Intel X710 NICs
- Any SFP passive or active limiting direct attach copper cable that complies with the SFF-8431 v4.1 and SFF-8472 v10.4 specifications
- SFF-8472 identifier must have value 03h for SFP or SFP+ (verify the value with the cable manufacturer)
- Maximum cable length for passive cables is 7 meters
- Support for active cables requires Intel® Network Connections software version 15.3 or later
- Intel® Ethernet SFP SR Optics and Intel® Ethernet SFP LR Optics are the only 10-Gbps optical modules supported; you cannot use other brands of SFP (10 GbE) modules with these adapters
- Only the modules listed in the table under “What SFP+ and SFP modules are supported?” on the Intel FAQ page will work with the X710. Find the latest FAQ [here](#).

Dell Networking

Customers are required to provide all of their own network cables. If a “Scale Computing Stock” subsection is shown, Scale Computing resells the listed products in that section. Any Dell switches purchased through Scale Computing **do not** include cables—cables are a separate purchase.

Scale Computing Stock

- 10 GbE
 - Dell x4012 10 GbE Backplane Switch: B12-DL-10
 - Passive Direct Attach Copper Twinax 3 meter cable: BSFP3m-Cable
- 1 GbE
 - Dell Networking N2024 1 GbE Backplane Switch: B24-DL

Dell N2000 Series: Stacking Cable Product Numbers

- 3 meter cable: 470-AAPX
- 1 meter cable: 470-AAPW
- .25 meter cable: 470-AAPV

10 GbE to 1 GbE Converter

Though not recommended, it may be necessary to convert from a 10 GbE model to a 1 GbE model. Customers are required to provide their own network cables and converters for this purpose. These are some recommended converters that have performed well in a variety of use cases, but be aware that these may not work for all environments or configurations and should not be seen as the only option for conversion. The links provided are for product reference only and are not necessarily meant as the avenue for procurement.

- PC Connection StarTech Gigabit SFP (GLCTST) Converter: <http://bit.ly/2fRdqNP> *
- CDW Proline Cisco GLC-T SFP Converter: <http://bit.ly/2g6XQvq>

HP Networking

Customers are required to provide all of their own network cables. If a “Scale Computing Stock” subsection is shown, Scale Computing resells the listed products in that section. Any HP switches purchased through Scale Computing **do not** include cables—cables are a separate purchase.

Scale Computing Stock

- 1 GbE
 - HP ProCurve 2930F 24 Port Backplane Switch: B24-HP

HP 2930F Series: Stacking Module and Cable Product Numbers

- Stacking module: JL325A
- 1 meter cable: J9735A
- .5 meter cable: J9734A

10 GbE to 1 GbE Converter

Though not recommended, it may be necessary to convert from a 10 GbE model to a 1 GbE model. Customers are required to provide their own network cables and converters for this purpose. These are some recommended converters that have performed well in a variety of use cases, but be aware

that these may not work for all environments or configurations and should not be seen as the only option for conversion. The links provided are for product reference only and are not necessarily meant as the avenue for procurement.

- PC Connection StarTech Gigabit SFP (GLCTST) Converter: <http://bit.ly/2fRdqNP> *
- CDW Proline Cisco GLC-T SFP Converter: <http://bit.ly/2g6XQvq>

Mellanox Networking

Customers are required to provide all of their own network cables. If a “Scale Computing Stock” subsection is shown, Scale Computing resells the listed products in that section. Any Mellanox switches purchased through Scale Computing **do not** include cables—cables are a separate purchase.

Scale Computing Stock

- 10 GbE
 - Mellanox MSX1012 Backplane Switch: B48-MX
 - Mellanox stacking cable: BMX-Stacking (MC2210128-003)
 - Mellanox 3 meter breakout cable: BMX-Breakout (MC2609130-003)
- Racking
 - Short-depth 1U rack installation kit for 1 or 2 MSX1012 backplane switches: FRK-BMX

Mellanox SX1012: QSFP to SPF+ Adapter (40 GbE to 10 GbE Adapter)

- Network adapter: MAMIQ00A-QSA

Cisco Networking

Customers are required to provide all of their own network cables. If a “Scale Computing Stock” subsection is shown, Scale Computing resells the listed products in that section.

Cisco Only: 10 GbE to 1 GbE Converter

Though not recommended, it may be necessary to convert from a 10 GbE model to a 1 GbE model. Customers are required to provide their own network cables and converters for this purpose. These are some recommended converters that have performed well in a variety of use cases, but be aware that these may not work for all environments or configurations and should not be seen as the only option for conversion. The links provided are for product reference only and are not necessarily meant as the avenue for procurement.

- PC Connection SFP-GBIC Converter: <http://bit.ly/2fDdZwy>
- CDW SFP-GBIC Converter: <http://bit.ly/1AV1drl>

Port Requirements

The following are the network ports the HC3 system may use for various tasks and should be enabled in the HC3 system network environment.

- **Network Time Protocol (NTP)** — User Datagram Protocol (UDP), Port 123
- **Outgoing Simple Mail Transfer Protocol (SMTP) Server** — Transmission Control Protocol (TCP), Port 25, 587, where 25 is the non-secure default for many environments and 587 is the default for many secure SMTP servers; this is dependent on environment configurations
- **Native HC3 VM Replication** — TCP, Port 10022
- **Outbound Secure Shell (SSH)** — TCP, Port 22, LAN network outbound only for use with remote-support.scalecomputing.com
- **System Software Updates** — TCP, Port 80, 443, LAN network outbound only for use with updates.scalecomputing.com

Public and Private System Networks

The HC3 system has two distinct physical networks in which it participates. A public network provides a path to allow access to the HC3 web interface as well as access to VMs running on the system and is known as the LAN network. A private network, known as the Backplane network, is used for inter-node communication. This includes critical system operations such as the mirroring of data blocks for redundancy between the nodes. It is critical that the Backplane network is isolated to a single HC3 system only (physically or through VLANs) to ensure system stability and performance.

Public Network Provisioning

The LAN ports on any Scale Computing node are in an active / passive bond used for failover. This means that only one LAN port is ever active at a time. There is a default primary and secondary port; which physical port this correlates to in the HC3 system is different between 1 GbE and 10 GbE nodes, but will always be designated LAN0. For redundancy, both the LAN0 and LAN1 ports on a node should always be cabled in to your switch(es).

LAN IP addresses should be assigned from your primary data network and are used by your system's nodes to communicate data between the system and network. These addresses are used **only** for data communication and HC3 web interface access. An internal firewall blocks all other traffic outside HC3 system features.

Private Network Provisioning

The backplane ports on any Scale Computing node are in an active / passive bond used for failover. This means that only one backplane port is ever active at a time. There is a default primary and secondary port; which physical port this correlates to in the HC3 system is different between 1 GbE and 10 GbE nodes but will always be designated Backplane0. For redundancy, both the Backplane0 and Backplane1 ports on a node should always be cabled in to your switch(es).

Backplane IP addresses should be assigned from a private network used solely by your system's nodes to communicate with other nodes of the same system and should be non-routable from any other part of the network. These addresses are used **only** for inter-node communication, system data striping, and system data mirroring. The system backplane is critical to system operation as the backbone of the system, and it is what makes the separate nodes a single HC3 system once they are initialized. An internal firewall blocks all backplane traffic aside from the HC3 system traffic.

WARNING

Due to their importance, backplane IP addresses are permanent. Once set, they cannot be changed without an entire system reset, which would wipe all data and configurations on the system.

Tips and Examples for Network Provisioning

Here you can find some tips to keep in mind when you are assigning LAN and backplane IP addresses:

- Backplane IP addresses should match the last octet of LAN IP addresses when possible; for example, LAN IP 192.168.100.10 would have a matching backplane IP of 1.1.1.10.
- LAN IP addresses should be provisioned in the same block, with room for node additions later; for example, a four node system would have a LAN IP scheme of 10.100.1.11, 10.100.1.12, 10.100.1.13, and 10.100.1.14 and then 10.100.1.15-18 would be blocked out for subsequent nodes down the road as necessary.
- Always keep your network environment in mind when assigning IPs; not every recommendation is going to be applicable or possible for every network environment.

WARNING

It is against Scale Computing best practices to use a publicly routable IP address scheme for the LAN IP addresses on your nodes. Choose a non-publicly routable IP addressing scheme for the LAN ports. If it is necessary to have public access to VMs on the system, utilize individual VM VLAN capabilities to control access to the VMs only, not to the system nodes.

Below is an example IP scheme for a four node system. This is **not** the required IP scheme, but an example of proper assignment using best practices. You should use an IP scheme that is compatible with your network environment.

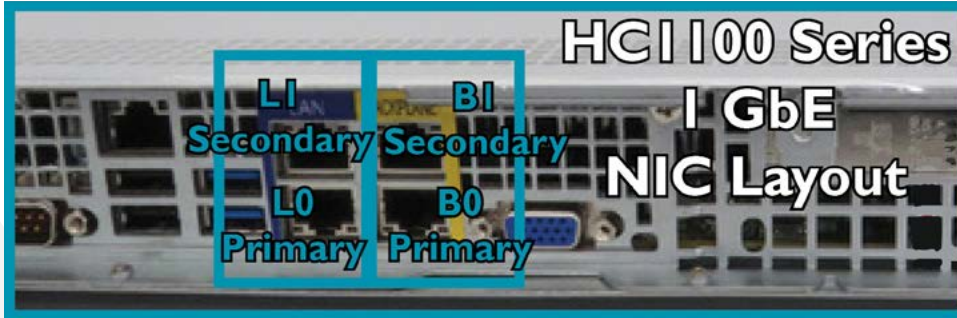
- LAN IP - 192.168.155.20
- LAN IP - 192.168.155.21
- LAN IP - 192.168.155.22
- LAN IP - 192.168.155.23
- LAN Subnet Mask - 255.255.255.0
- LAN Gateway 192.168.155.1
- Backplane IP - 10.10.10.20
- Backplane IP - 10.10.10.21
- Backplane IP - 10.10.10.22
- Backplane IP - 10.10.10.23
- Backplane Subnet Mask (Immutable) - 255.255.255.0

Network Interface Layout

Physical NIC port configurations for 1 GbE and 10 GbE node options.

HC1100 Series

1 GbE port configuration for the **HC1100**, **HC1150**, **HC1150D**, and **HC1150DF** models.



10 GbE port configurations for the **HC1100**, **HC1150**, **HC1150D**, and **HC1150DF** models.



HC2100 Series

1 GbE port configuration for the **HC2100** model.

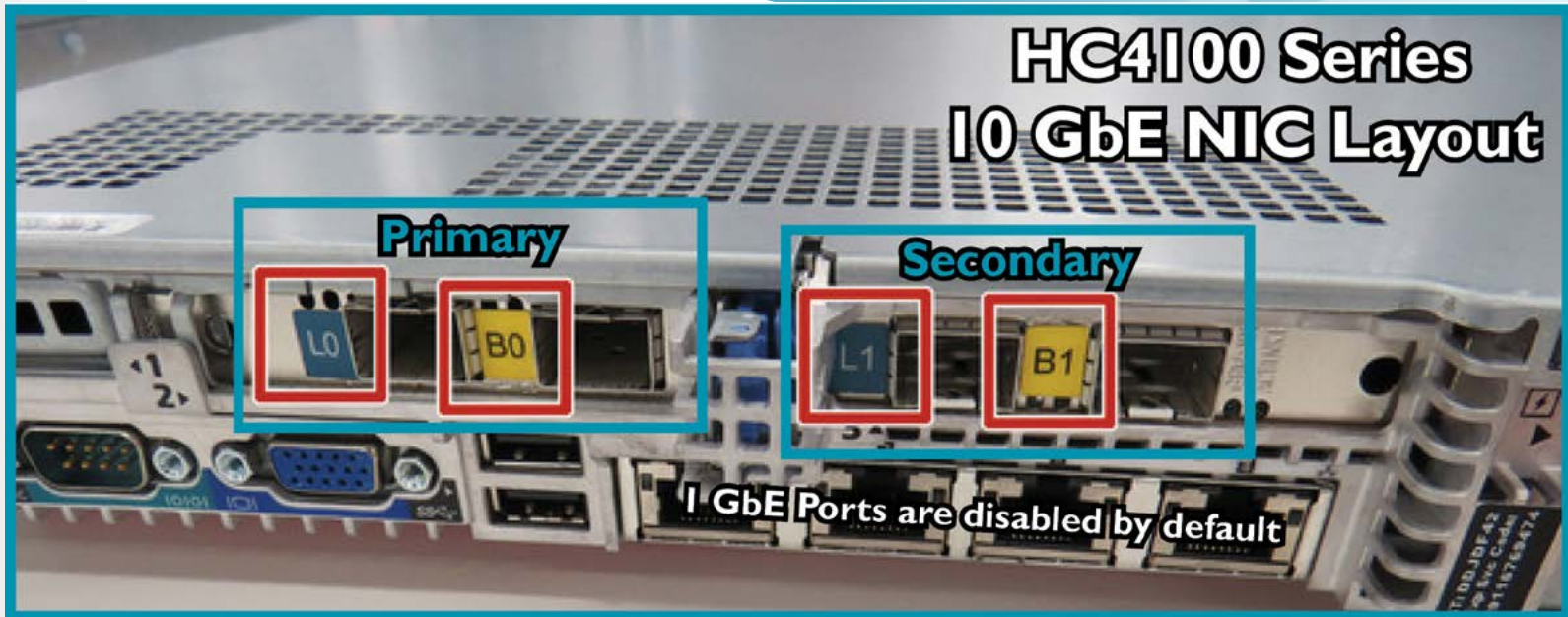


HC2100 Series 10 GbE NIC Layout



HC4100 Series

10 GbE port configuration for the **HC4100** and **HC4150** models. The **HC4100 Series** is 10 GbE only.



HC5100 Series

10 GbE port configuration for the **HC5150D** model. This **HC5100 Series** is currently 10 GbE only.



Spanning Tree (STP)

The Spanning Tree Protocol (STP) is a network protocol that ensures a loop-free topology for bridged local area networks (LANs). STP allows a network design to include spare (redundant) links to provide automatic backup paths via STP without the need for manual intervention.

When STP is enabled, the protocol monitors the participating ports/VLANs. Should there be a change in topology (a port goes active or a port goes down) STP blocks traffic on participating ports until the network topology is determined. When a topology change is discovered, the ports participating in STP are in a Blocking state; they will then move through a Listening, Learning, and then finally a Forwarding state (when traffic is then forwarded).

For this reason, STP should be disabled on the ports that Scale nodes are connected to. Typically, this is accomplished by setting portfast or edge-port mode on the switch port.

If STP is required for the customer LAN, then please enable Rapid STP on the Scale node ports. Rapid STP allows a switch port to rapidly transition into the forwarding state during these situations.

Flow Control

Flow control is useful for managing the data rates between two links. It helps prevent a fast sending connection from overwhelming a slower receiving connection and causing retransmits. Scale recommends enabling flow control on ports where a Scale cluster connects to your LAN network.

Switch and Node Redundancy

Scale Computing nodes are designed for network and data redundancy. This section offers tips and best practices for fully utilizing these features in the HC3 system. Not all high availability features will be available in Single Node Systems (SNS).

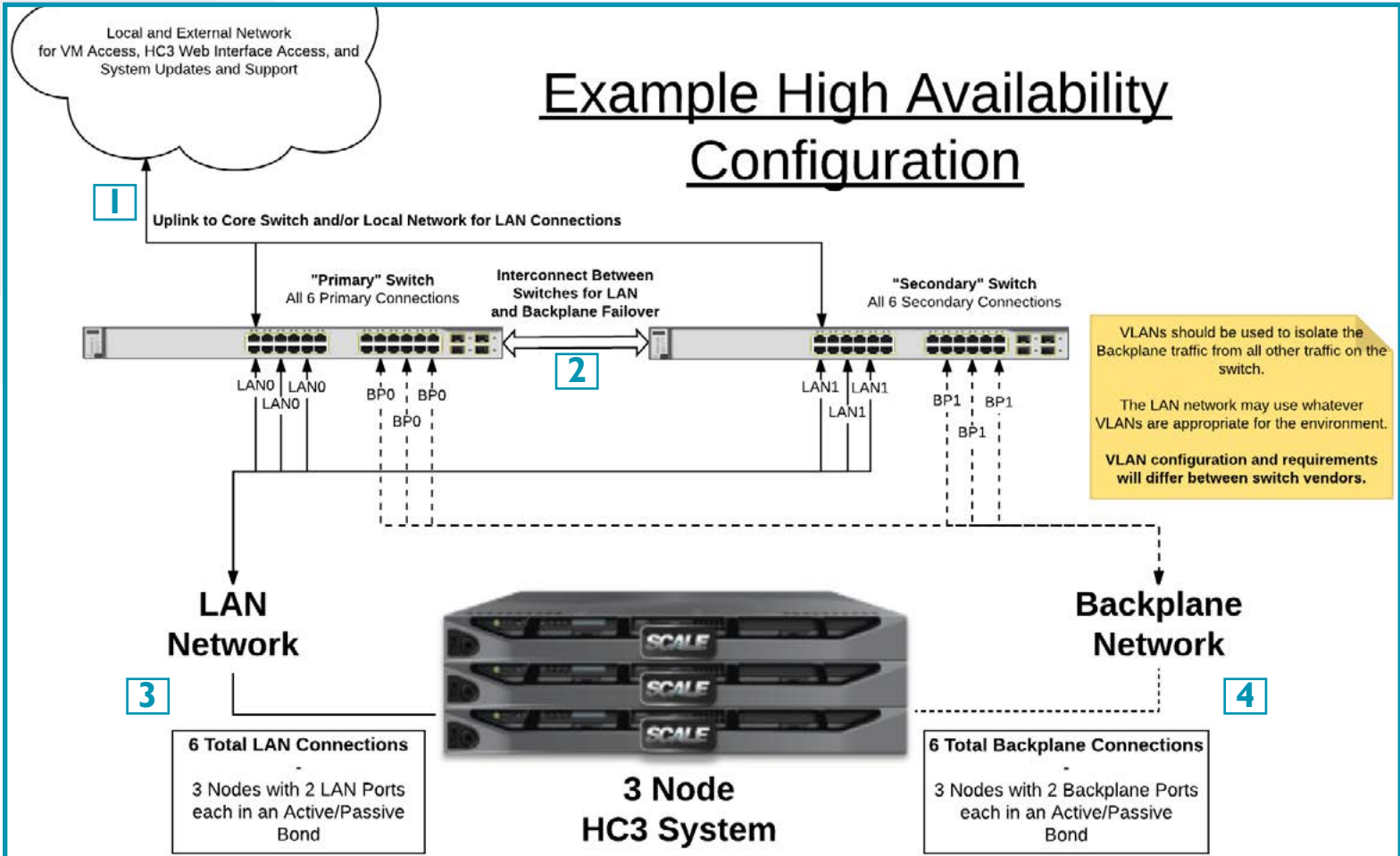
- (*) Use two switches for each HC3 system and balance one LAN and one backplane connection on each switch, separated by VLANs with interconnects between the switches; this configuration ensures not only port failover on the nodes, but also port failover on the switches and full switch redundancy if a switch were to fail.
- Use two different UPS for each of the system nodes' power supplies and switches, and even different power circuits; this configuration ensures the best possible node and switch power redundancy.
- Always use all four network ports on the nodes, both LAN and both backplane.
- Always use both power supplies on the nodes.
- Nodes should not be geographically dispersed amongst different physical locations; when you set up nodes for your system, select one physical location for each system.
- (*) Scale Computing provides non-disruptive rolling software updates, reliant on enough free RAM being available to operate with one node unavailable.
- (*) Automated VM failover is configured in the event of a node failure scenario; this ensures running VMs will resume operation on the running nodes if their current node fails for any reason.

High Availability

HC3 nodes come equipped with two Network Interface Cards (NICs) for High Availability (HA). These NIC ports work in an **Active/Passive** mode, meaning one NIC will be active and the other NIC will be inactive; the inactive NIC will only become active in the event that the primary NIC becomes unavailable.

NOTE: Active/Passive is the only available mode for Scale Computing NICs at this time.

The image below shows an example HA configuration for a three node cluster. This is **NOT** the only configuration possible nor is it the recommended configuration for all nodes and switches—but the example does take into account failover and high availability at the NIC, hardware, and switch level.



- 1. Uplink** — Each switch containing a LAN connection should have an uplink to the local network and/or core switch for VM and HC3 web interface access. The LAN connection should ideally have internet access for system updates and remote support as well.
- 2. Interconnect** — The interconnect shown in the image is generic for any physical and/or virtual bridge between the two switches that will allow the LAN and Backplane connections to continue communication in the event that one or more primary NICs may become unavailable. For the duration of the period the secondary port(s) may be active (failover from the primary to the secondary NIC is automated and seamless) it is important that the NIC on the secondary switch can continue to communicate with those on the primary switch—particularly on the Backplane network so that the

node in question is not isolated from the system.

3. **LAN Network** — The LAN network is for VM data access, HC3 web interface access, system updates, and remote support access.
4. **Backplane Network** — The Backplane network is for inter-system communication **ONLY**. The Backplane IPs should be non-public and non-routable in the local network. No outside traffic should touch the system Backplane network and two or more HC3 systems should **NEVER** share a backplane switch or VLAN.

Feedback and Support

Document Feedback

Scale Computing welcomes your suggestions for improving our documentation. If you have any comments or suggestions please send your feedback to documentation@scalecomputing.com.

Technical Support and Resources

There are many technical support resources available for use. Access this document, and many others, at <http://www.scalecomputing.com/support/login/>.

Online Community Scale Computing has an online forum and community! This is a great medium with which to solicit the advice of your peers, benefit from their experience, find and discuss documentation, and participate in ongoing conversations. Please note this community is not intended to provide ScaleCare Support assistance or replace other Scale Computing communication channels. Find the **Scale Legion HC3 Discussion Forum** at <https://scalelegion.community>

Online Support You can submit support cases and view account information online through the Portal at <http://www.scalecomputing.com/support/login/>. You can also Live Chat with support through www.scalecomputing.com during standard hours Monday-Friday 8-8.

Telephone Support Support is available for critical issues 24/7 by phone at 1-877-SCALE-59 (877-722-5359) in the US and at 0808 234 0699 in Europe. Telephone support is recommended for the fastest response on priority issues, and the only response after standard Support hours.

Professional Resources Scale Computing offers many professional service options for both remote and on-site assistance in getting your new system off the ground quickly and knowledgeably. Contact your Scale Computing sales representative today to discuss our offerings that include options like the HC3 Premium Installation.

Disclaimer

Any information listed here is subject to change at any time. Shifting hardware and software requirements and industry changes mean that not all current recommendations or requirements can or will be sustained. This information is not a substitute for the listed product's user guide or support and is not covered under the support and warranty for the HC3 cluster. Scale Computing is not responsible for any issues or damages arising out of the use of this reference guide.