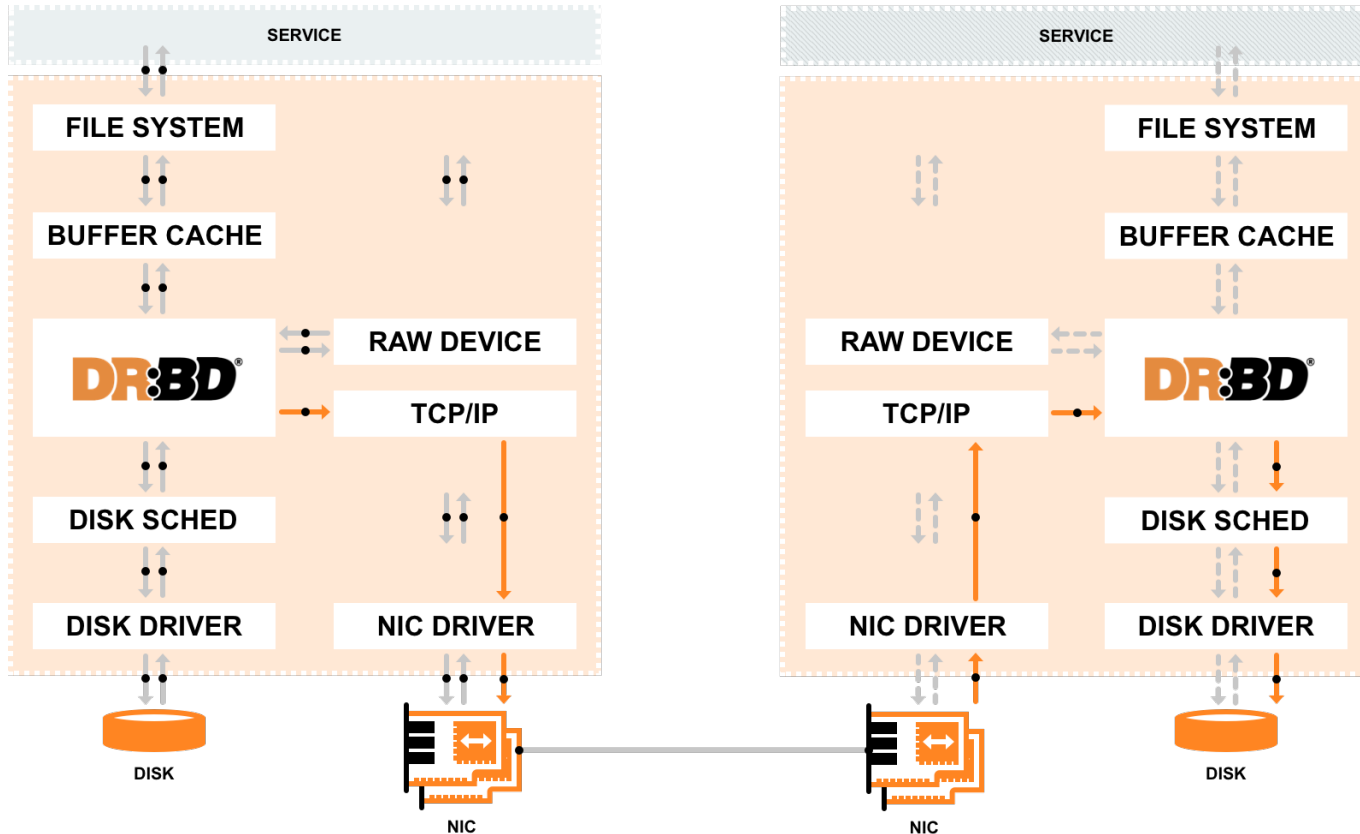




**Majority in DRBD
and the LINBIT CloudStack HCI Appliance**

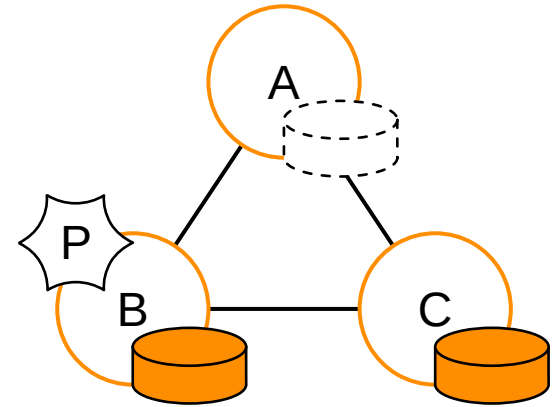
Philipp Reisner, CEO LINBIT

Protecting Data by replication



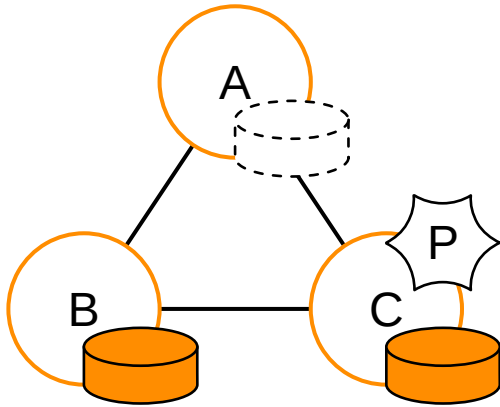
Quorum - Basics

- Three nodes minimum
- Diskless nodes act as tiebreakers
- Node B is in primary role and runs the application



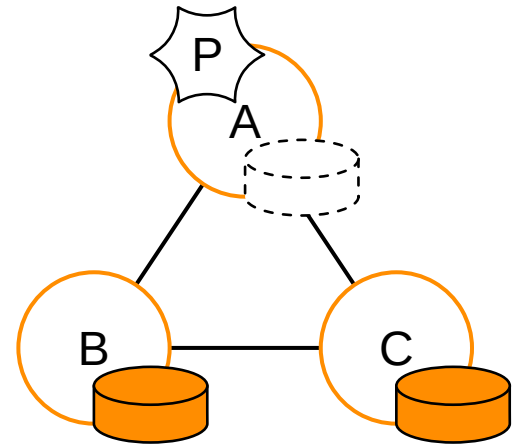
Quorum - Primary Movable

- Primary role and application are freely movable



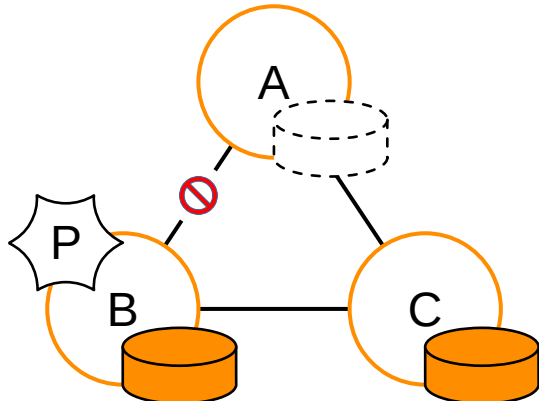
Quorum - Diskless nodes

- When it runs on a diskless node, it reads in a load balancing fashion from the nodes with backing storage
- The diskless nodes is also called the tiebreaker or the witness node



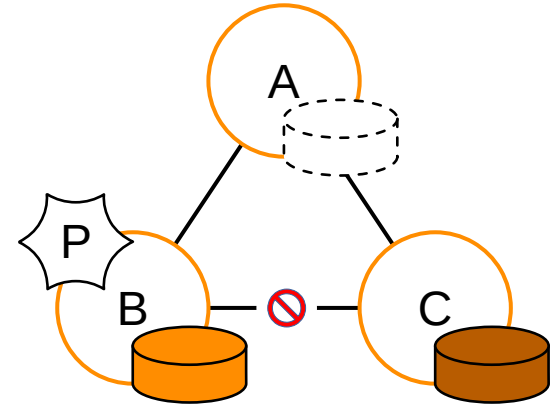
Network partitioning

- Network partitioning between A and B has no impact



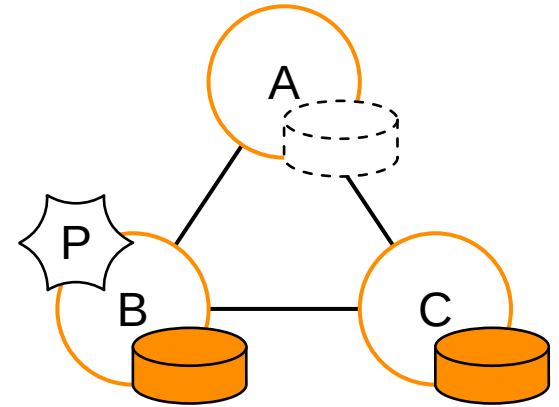
Network partitioning

- Network partitioning between A and C leads to C's replica becoming Outdated



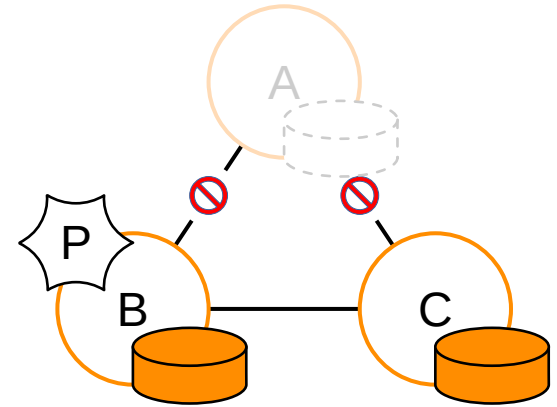
Network partitioning

- DRBD resyncs C upon reconnect
- Only blocks touched in the meantime on the primary



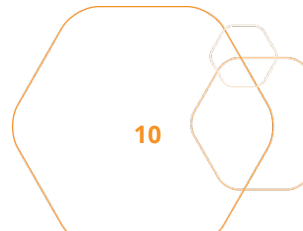
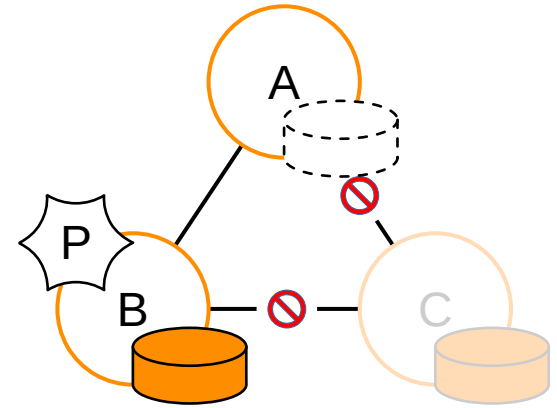
Network partitioning

- Isolate A
- A loses quorum – can no longer promote – it also has no access to data
- B & C keep quorum



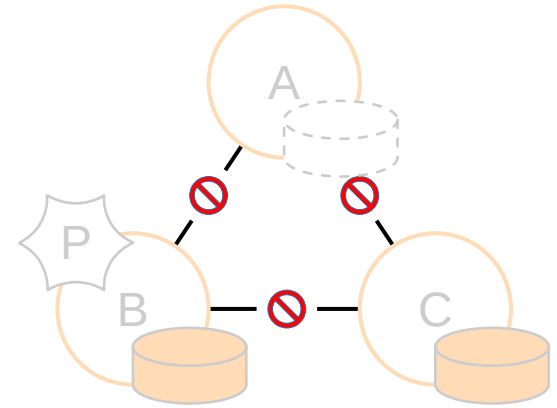
Network partitioning

- Isolate C
- C loses quorum – can no longer promote
- B & A keep quorum



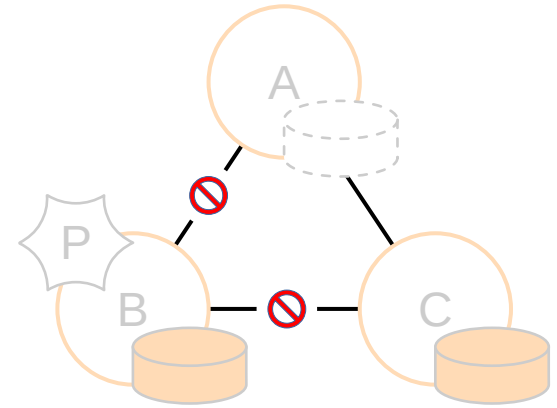
Diskless switching partition

- A switches to C, step 1
- All nodes without quorum
- Primary B, freezes I/O or completes I/O requests with errors, depends on config



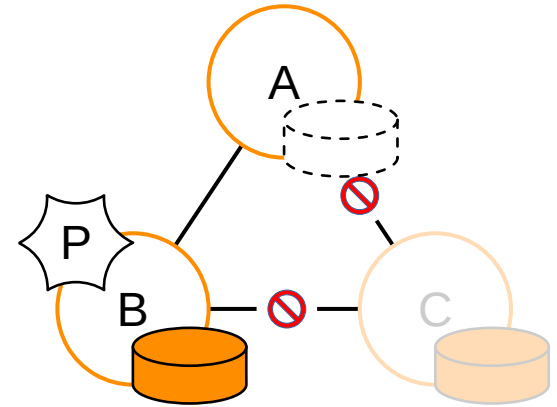
Diskless switching partition

- A switches to C, step 2
- C can not regain quorum by connecting to the tiebreaker node(s)



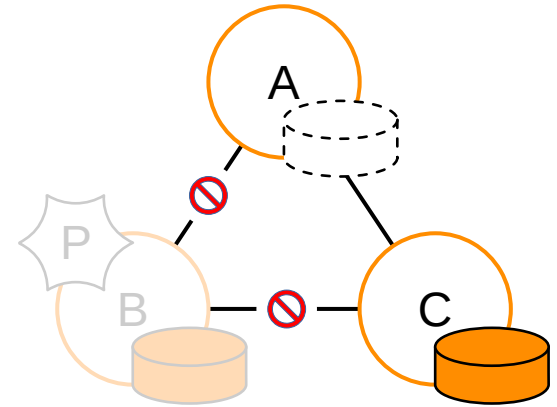
Diskless not switching partition

- A restores quorum in the partition A & B
- I/O on B resumes
- Application on B thaws
- The partition preserves quorum over a reboot of B



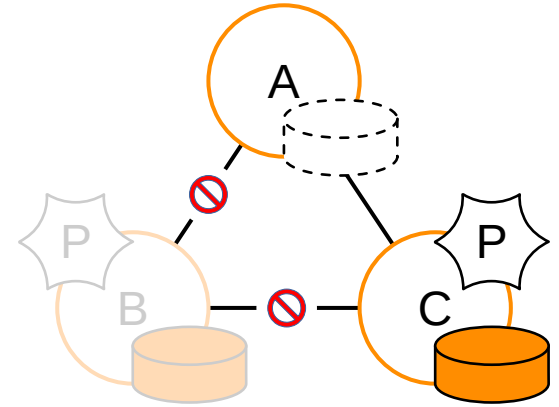
Network partitioning

- B gets isolated
- A & C are quorate
- B freezes I/O and the application
- A & C advertise a positive promotion_score



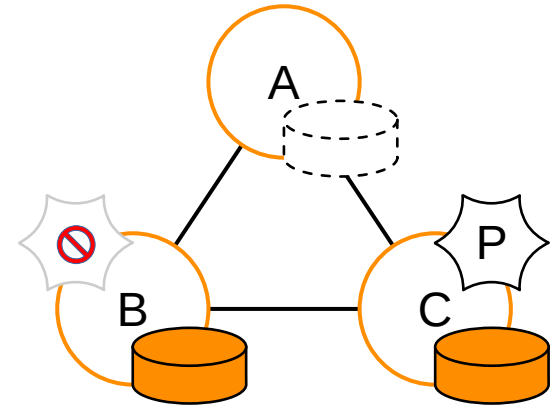
Network partitioning

- A & C are quorate
- A clustermanager promotes C to primary and starts the application
- B does not know, and stays frozen



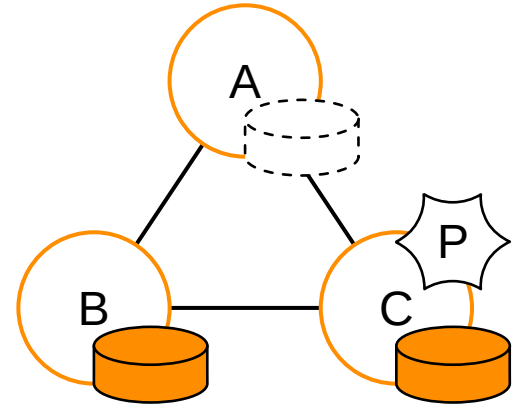
Network partitioning

- B can recover by reconnecting with the other partition
- Or by `drbdadm -force secondary`
- The application gets I/O errors and needs to close the DRBD device (terminate)



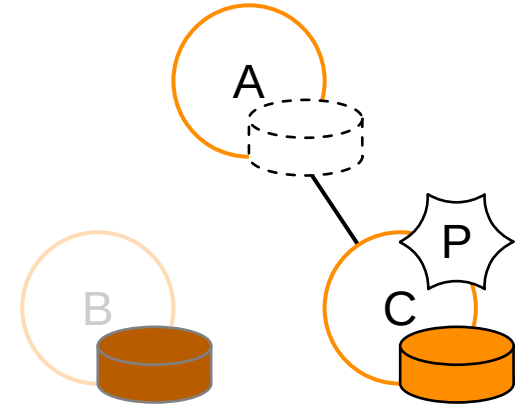
Network partitioning

- After the application terminated
DRBD clears the force-io-errors flag
- Resync from C to B
- Recovery completed



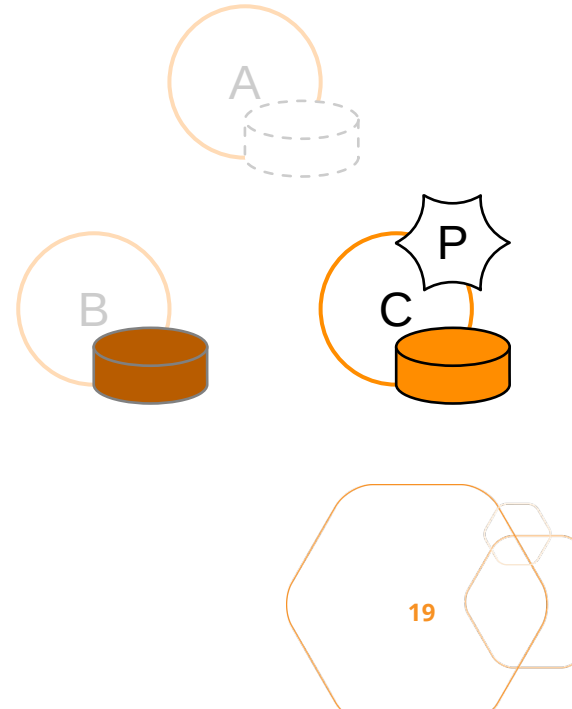
Gracefull disconnect

- B gracefully disconnects
- B outdates its disk when it leaves a partition with a primary node
- A graceful shutdown behaves in the same way



Gracefull disconnect

- C knows that B is Outdated and A is diskless
- It keeps quorum although it is just one out of three



LINBIT CloudStack HCI Appliance

LINBIT CloudStack HCI Appliance



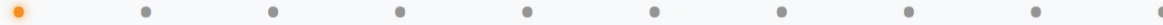
- Bootable ISO File
- About 20 minutes setup time
- Primary & Secondary storage configured
- Everything highly available
- 3 nodes minimum
- Tech preview at the moment
- Enterprise options will come in January 2025



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Welcome to LINBIT CloudStack HCI

Press the button below to start the setup.

Start Setup



Version 0.0.0.f8a395d55ddb8d3af6fe9cf76bbf8fa97206bc32

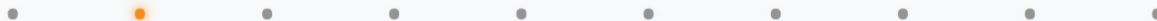
LINBIT is a registered trademark of [LINBIT HA-Solutions GmbH](#).
CloudStack is a registered trademark of the [Apache Software Foundation](#).



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Please select the nodes for your cluster

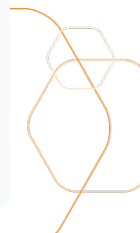
LINBIT VSAN will automatically discover all of its instances on your local network. Please choose which nodes to include in your new cluster. For quorum reasons, at least three nodes are required to form a cluster.

<input checked="" type="checkbox"/>	uninitialized-a768f4305ce26b5dab9f3ad7	192.168.122.249	▼
<input checked="" type="checkbox"/>	uninitialized-d500f2d9f9387ff695a12011	192.168.122.221	▼
<input checked="" type="checkbox"/>	uninitialized-feb454908729f50689511bcb	192.168.122.75	▼

[Are your Nodes missing?](#)

Back

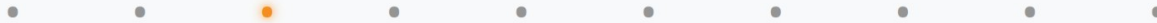
Continue



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Enter a Name for Your Cluster

Nodes

hci3	192.168.122.75
hci1	192.168.122.221
hci2	192.168.122.249

Back

Continue

Version 0.0.0.f8a395d55ddb8d3af6fe9cf76bbf8fa97206bc32

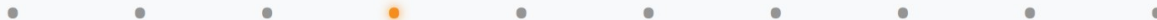
LINBIT is a registered trademark of [LINBIT HA-Solutions GmbH](#).
CloudStack is a registered trademark of the [Apache Software Foundation](#).



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



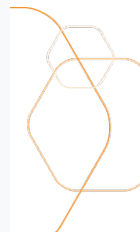
Setting up nodes

Please wait while your nodes are being registered with LINBIT. After registering, you will be able to see your new nodes on your Active Contracts page at my.linbit.com and LINBIT's software stack will be installed and configured.

- ✓ Configuring System
- ✓ Generating LINSTOR Controller Configuration
- ✓ Waiting for LINSTOR Controller to become available
- ✓ Adding Satellites to LINSTOR

Your new nodes have been successfully registered with your account and all required LINBIT packages have been installed.

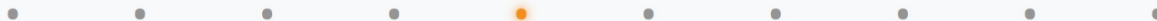
Continue



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



CloudStack Node Setup

Select the nodes to use as CloudStack Hypervisors. These nodes will be reconfigured to host a bridge interface for the given subnet. This subnet is used to expose CloudStack VMs to the outside world.

VM Network

192.168.122.0/24

CloudStack Hypervisor Nodes



hci2



hci1



hci3

Configure the default CloudStack Network. CloudStack assigned IP Addresses from the configured ranges to system and instance VMs. The reserved ranges need to be unassigned.

Gateway: ?

192.168.122.

1

DNS: ?

192.168.122.1

System IP Start: ?

192.168.12

10

System IP End: ?

192.168.12

20

Instance IP Start: ?

192.168.12

100

Instance IP End: ?

192.168.12

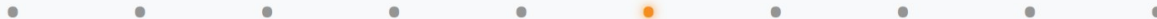
200



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Physical storage

Below is a representation of which storage devices are present on which of your nodes. Check the respective boxes to combine multiple storage devices into a storage pool. LINSTOR will then view this storage pool as a single "chunk" of storage and use it to create virtual storage volumes.

Hover over the entries to see the path for each device on a particular node.

Important: To appear on this list, a storage device:

- Must be greater than 1 GiB.
- Must be **completely empty**. This includes file systems, LVM signatures, and others. If there is existing data on the device, wipe it first (for example, using `wipefs -a`).

	All	hci1	hci2	hci3
20 GiB HDD	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>



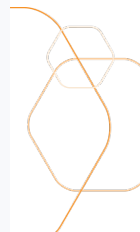
Storage Pool Name



Add to existing pool



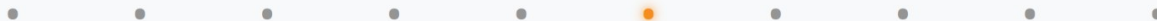
LVM Thin



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Physical storage

Below is a representation of which storage devices are present on which of your nodes. Check the respective boxes to combine multiple storage devices into a storage pool. LINSTOR will then view this storage pool as a single "chunk" of storage and use it to create virtual storage volumes.

Hover over the entries to see the path for each device on a particular node.

Important: To appear on this list, a storage device:

- Must be greater than 1GiB.
- Must be **completely empty**. This includes file systems, LVM signatures, and others. If there is existing data on the device, wipe it first (for example, using `wipefs -a`).

> Storage Pool **pool1**



	All	hci1	hci2	hci3
--	-----	------	------	------



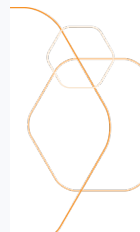
Storage Pool Name



Add to existing pool



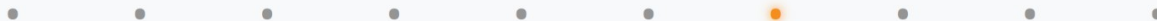
LVM Thin



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Resource groups

Here you can create resource groups from your storage pools.

A resource group is a LINSTOR concept. It can be thought of as a template for resource creation. It defines a storage pool to use and how many replicas of the data to create.

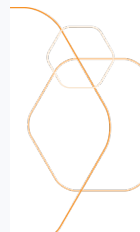
When the resource group is instantiated, for example, when it is used to spawn a resource, LINSTOR will automatically take storage from the specified storage pool and place the resource on the appropriate number of nodes.

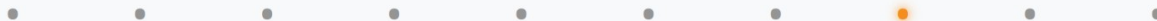
> Resource Group rg2	
> Resource Group rg3	

Resource Group:

Storage pool:

Replica count:





LINSTOR Controller high-availability

Right now the LINSTOR Controller is just running on a single node. When that node goes down, you lose the control plane and you will not be able to create, delete, or modify resources.

To make the system more robust, we will now make the LINSTOR Controller highly available using LINSTOR itself. This makes it possible to host the LINSTOR Controller on any node in the cluster.

Please select a resource group to use for the LINSTOR Controller database:

Service IP: [?](#)

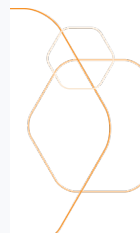
192.168.122.



3

rg3 (19.96 GiB available)

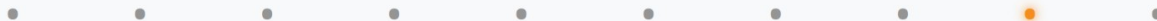
Back



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



CloudStack Setup

Select the Resource Group used to store the CloudStack Management database, along with the service IP that the CloudStack interface will be reachable at.

Resource Group: [?](#)

rg3 (19.89 GiB available)

Service IP: [?](#)

192.168.122.

5

Configure a NFS export as secondary storage for CloudStack.

Resource Group: [?](#)

rg2 (19.9 GiB available)

Service IP: [?](#)

192.168.122.

4

Size:

15

GiB

Use all available

Back

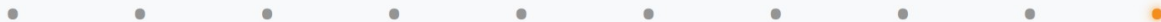
Continue



LINBIT CloudStack HCI Appliance



LINBIT® CloudStack® HCI



Setting up CloudStack

Please wait while CloudStack is being set up. After all steps are complete, you will be forwarded to the CloudStack interface.

- ✓ Starting CloudStack Management Server
- ✓ Starting CloudStack Secondary Storage on NFS
- ✓ Wait for CloudStack API to come online
- ✓ Setting up CloudStack Hypervisor Nodes

Setup is now complete. Log in using the following credentials:

Username

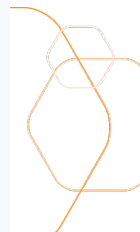
admin

Password

password

Back

Go to CloudStack



LINBIT CloudStack HCI Appliance



☰ Default view ▾

Create ▾ AC admin cloud

- Dashboard
- Compute
- Storage
- Network
- Images
- Events
- Projects
- Roles
- Accounts
- Domains
- Infrastructure
- Service offerings
- Configuration
- Tools

🌐 All zones C Fetch latest

🏠 Infrastructure

Pods	Clusters
1	1
Hosts	Hosts in alert state
3	0
Primary storage	System VMs
2	2
Virtual routers	Instances
0	0

☁ Compute

Memory
 19.51%
1.50 GiB Allocated | 7.69 GiB Total

CPU
 4.39%
1.00 GHz Allocated | 22.80 GHz Total

CPU cores
 33.33%
2 Allocated | 6 Total

GPU
 0%
0 Allocated | 0 Total

📁 Storage

Primary storage used
 4.29%
5.14 GiB Used | 119.74 GiB Total

Primary storage allocated
 0.00%
0.00 GiB Allocated | 239.48 GiB Total

Secondary storage
 0%
0.00 GiB Allocated | 0.00 GiB Total

🌐 Network

Shared Network IPs
 1.98%
2 Allocated | 101 Total

Management IP addresses
 18.18%
2 Allocated | 11 Total

🚨 Alerts

- 19 Nov 2024 13:04:20 [ALERT.MANAGEMENT](#)
Management server node 192.168.122.5 is up
- 19 Nov 2024 13:04:17 [ALERT.MANAGEMENT](#)
Management network CIDR is not configured originally. Set it default to 192.168.122.0/24

[View Alerts](#)

📅 Events

- 19 Nov 2024 13:06:08 [SG.AUTH.INGRESS](#)
(admin) Successfully completed Adding Ingress Rule
[SecurityGroup](#) [default](#)
- 19 Nov 2024 13:06:10 [TRAFFIC.TYPE.CREATE](#)
(admin) Successfully created entity for Creating Physical Network TrafficType. Zone Id: 1



LINBIT CloudStack HCI Appliance



<https://linbit.com/linbit-cloudstack-hci-appliance>

Get it Today!





Thank you

<https://www.linbit.com>