Nutanix Cloud Clusters - Nutanix Cloud Clusters on Google Cloud

[PDF generated February 25 2026. For all recent updates please see the Nutanix Bible releases notes located at https://nutanixbible.com/release_notes.html. Disclaimer: Downloaded PDFs may not always contain the latest information.]

NC2 now extends NCP into the Google Cloud ecosystem, providing a true hybrid multicloud infrastructure stack. NC2 on Google Cloud allows organizations to run directly on Google Compute Engine (GCE) bare-metal instances in their own Google Cloud Virtual Private Cloud (VPC).

Supported Configurations

The solution is applicable to the configurations below (list may be incomplete, refer to documentation for a fully supported list):

Core Use Case(s):

- On-Demand / burst capacity
- Backup / DR
- Cloud Native
- Geo Expansion / DC consolidation
- App migration

Management interfaces(s):

- Nutanix Clusters Portal - Provisioning
- Prism Central (PC) - Nutanix Management
- Google Cloud Console - Google Cloud Management

Supported Environment(s):

- Cloud:
  - AWS
  - Azure
  - Google Cloud
- Bare Metal Instance Types:
  - Z3 High Memory
  - C4 High Memory
  - C4 Standard

Upgrades:

- Part of AOS

Compatible Features:

- AOS Features
- Google Cloud Services

Key terms / Constructs

Nutanix Cloud Clusters (NC2) on Google Cloud integrates Nutanix Cloud Platform (NCP) with the underlying Google Cloud infrastructure. The solution uses the following components:

- NC2 Portal
  - The NC2 Portal is a multicloud control plane used to deploy, manage, expand, and shrink NC2 clusters in supported public clouds, including Google Cloud. The NC2 Portal is responsible for handling cluster provisioning requests and interacting with Google Cloud and the provisioned hosts. It creates cluster specific details and handles the cluster creation and helps to remediate hardware problems.

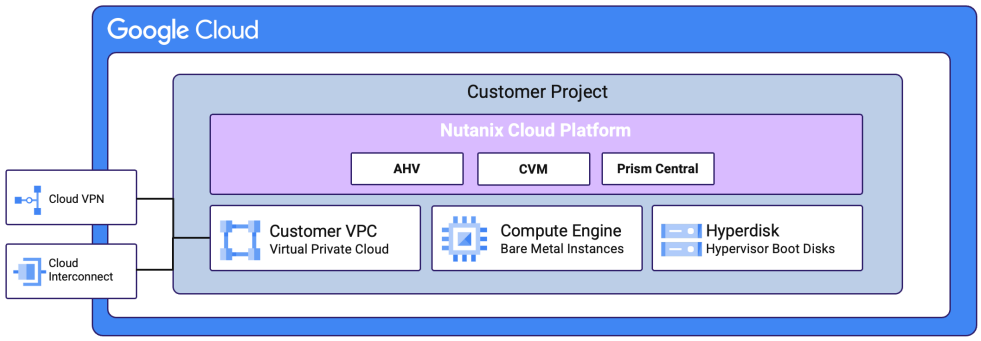
- Region
  - A geographic landmass or area where multiple Availability Zones (sites) are located. A region can have two or more AZs. These can include regions like "us-central1" (Iowa) or us-east1 (South Carolina).

- Availability Zone (AZ)
  - An AZ consists of one or more discrete datacenters interconnected by low latency links. Each site has its own redundant power, cooling, network, etc. Comparing these to a traditional colo or datacenter, these would be considered more resilient as a AZ can consist of multiple independent datacenters.

- Google Compute Engine bare-metal instances
  - Google Compute Engine (GCE) provides the bare-metal instances on which the NCP solution runs. These bare-metal instances appear in the Google Cloud Console like any other GCE instances and are distinct from the Google Bare Metal Solution that is not supported with NC2. Nutanix AHV runs directly on the GCE bare-metal instances, providing the foundation for running VMs.

- Google Cloud Project
  - A Google Cloud project is the fundamental organizing entity for all your resources, APIs, billing, and permissions on Google Cloud.

- Virtual Private Cloud (VPC)
  - A logically isolated segment of the Google Cloud Cloud within a project. Provides a mechanism to secure and isolate environment from others. Can be exposed to the internet or other private network segments (other VPCs, or VPNs).



NC2 on Google Cloud - Overview

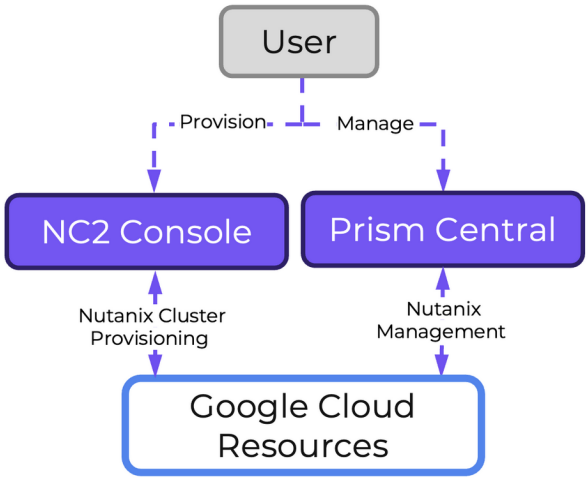
Solution Architecture

From a high-level the NC2 Portal is the main interface for provisioning Nutanix Clusters on Google Cloud and interacting with Google Cloud.

The provisioning process can be summarized with the following high-level steps:

1. Create a cluster in NC2 Portal
2. Deployment specific inputs (e.g. Region, AZ, Instance type, VPCs/Subnets, etc.)
3. The NC2 Portal creates associated resources
4. Host agent running on AHV checks-in with Nutanix Clusters on Google Cloud
5. Once all hosts as up, cluster is created

The following shows a high-level overview of the NC2 on Google Cloud interaction:



NC2 on Google Cloud - Management

Node Architecture

Given that the hosts are bare metal, the solution has full control over storage and network resources similar to a typical on-premises Nutanix deployment, witht he exception that Google Cloud's Hyperdisk service is used to provide the AHV Hypervisor boot volumes.

Placement policy

Nutanix Clusters on Google Cloud uses a partition placement policy with 7 partitions. Hosts are striped across these partitions which correspond with racks in an on premises Nutanix deployment. This ensures you can have 1 or 2 (with a 2N/2D configuration) full "rack" failures and still maintain availability.

Storage

Core storage is the exact same as you'd expect on any Nutanix cluster, passing the "local" storage devices to the CVM to be leveraged by Stargate.

Networking

NC2 utilizes Flow Virtual Networking in Google Cloud to create an overlay network to ease administration for Nutanix administrators and reduce networking constraints across Cloud vendors. Flow Virtual Networking is used to abstract the Google Cloud native network by creating overlay virtual networks. On the one hand this abstracts the underlying network in Google Cloud, while at the same time, it allows the network substrate (and its associated features and functionalities) to be consistent with the customer's on-premises Nutanix deployments. You will be able to create new virtual networks (called Virtual Private Clouds or VPCs) within Nutanix, subnets in any address range, including those from the RFC1918 (private) address space and define DHCP, NAT, routing, and security policy right from the familiar Prism Central interface.

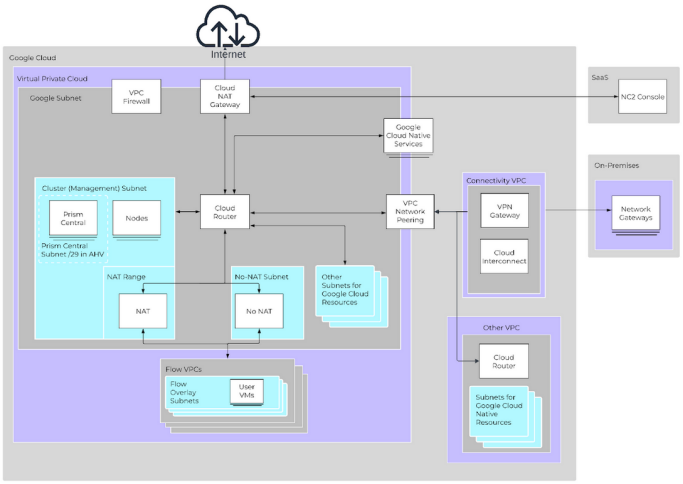
You can deploy Nutanix Cloud Clusters (NC2) on Google Cloud into an existing Google Cloud Virtual Private Cloud (VPC) and its subnets, or you can create a VPC and subnets specifically for NC2 in a project at deployment. VMs and containers running on the cluster can connect to the internet, native Google Cloud services, and on-premises datacenters using WAN networking (virtual private network, Google Cloud Router, Multiprotocol Label Switching).

Flow Virtual Networking for NC2 on Google Cloud has the following capabilities:

- Flow Virtual Networking creates an overlay network for user VMs, allowing you to create virtual networks (VPCs) and subnets in Prism Central independent of the underlying Google Cloud network topology.
- Flow Virtual Networking provides a unified networking experience across on-premises and Google Cloud environments, simplifying hybrid connectivity and policy management.

- For communication between user VMs on different physical hosts, the overlay network uses Geneve encapsulation.
- User VMs can connect to external Google Cloud networks, the internet, and native Google Cloud services, such as Google Kubernetes Engine (GKE) and BigQuery.
- You can provision user VMs to a network regardless of network address translation (NAT).

  > If you provision a user VM to a NAT network, assign a floating IP address to the user VM to facilitate connectivity from outside the Flow Virtual Networking overlay networks.



Prism Central contains the control plane for Flow Virtual Networking. The User VMs using the Flow VPC(s) can communicate to native Google Cloud services and allows the VMs to have parity with native Compute Engine VM instances.

Unlike NC2 in Azure, NC2 in Google Cloud does not require Flow Gateway VMs (FGWs).

Usage and Configuration

At a high level, the deployment process for Nutanix Cloud Clusters (NC2) on Google Cloud follows these steps:

1. Set up or access an active Google Cloud subscription and project.
2. Configure the necessary Google Cloud APIs and create a Service Account with the required IAM roles and permissions.
3. (Optional) Configure networking in Google Cloud, including a Virtual Private Cloud (VPC) and the necessary subnets, or use a VPC created automatically during cluster creation.
4. Create a My Nutanix account and subscribe to the NC2 service.
5. Add your Google Cloud account to an organization in the NC2 console.
6. Create a Nutanix Cluster in a new or existing Google Cloud VPC using the NC2 console.

More to come!