Accelerating Hyper-Converged Enterprise Virtualization using Proxmox and Ceph

Md. Mahedi Hasan

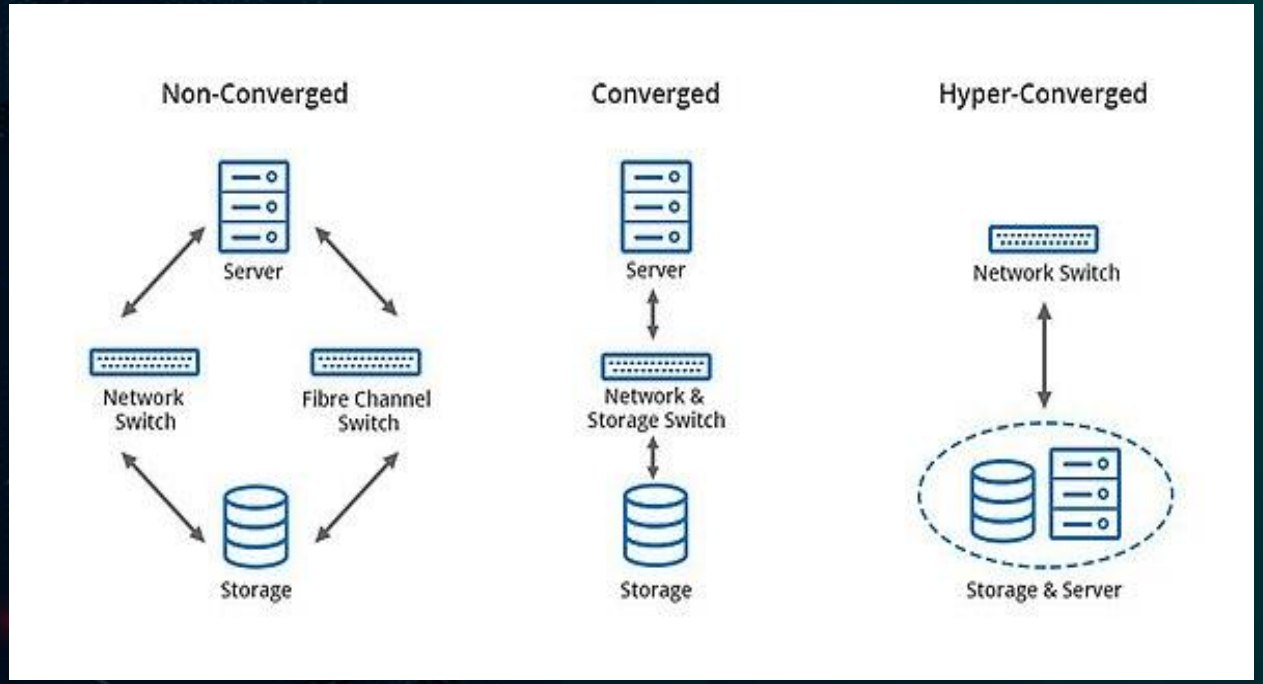
Network Engineer, University of Dhaka

cse.mahedi@gmail.com | www.Mahedi.me

/in/mahedicse/ | /mahedi.cse

Hyper-Converged Infrastructure (HCI)

- ✓ **Hyper-converged infrastructure (HCI)** is a software-defined infrastructure that virtualizes all the elements of conventional "hardware-defined" systems.
- ✓ It's integrating computing (hypervisor), software-defined storage, and virtualized networking (software-defined networking).

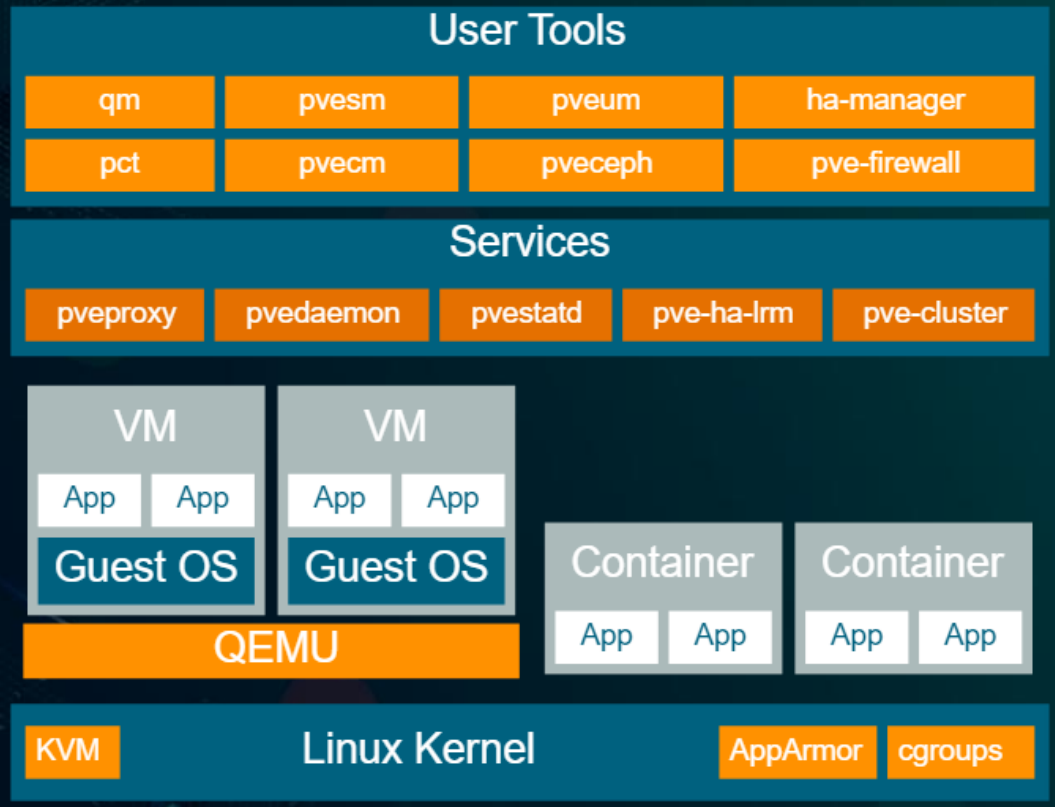


Difference between non-converged, converged and hyper-converged System
Image Source: Wikipedia

Proxmox Virtual Environment (Proxmox VE)

Proxmox Virtual Environment (Proxmox VE):

- ✓ Stands as a leading open-source virtualization platform
- ✓ That revolutionizes the way organizations manage and deploy virtualized workloads.



Proxmox VE Architecture
Source: Proxmox.com

Key Features and Advantages of Ceph

KVM and LXC:
KVM for VM and LXC for containerization

High Availability:
Live Migration and Fault Tolerance minimizing downtime

Backup and Restore:
Built-in backup and restore provide data protection and recovery capabilities.

Web-Based Management:
Web-based interface makes it user-friendly, to manage Proxmox

Storage Integration:
Support ZFS, Ceph, NAS and SAN Storage for flexible and scalable management.

Resource Management:
Precise resource allocation of CPU, Memory, and Storage.

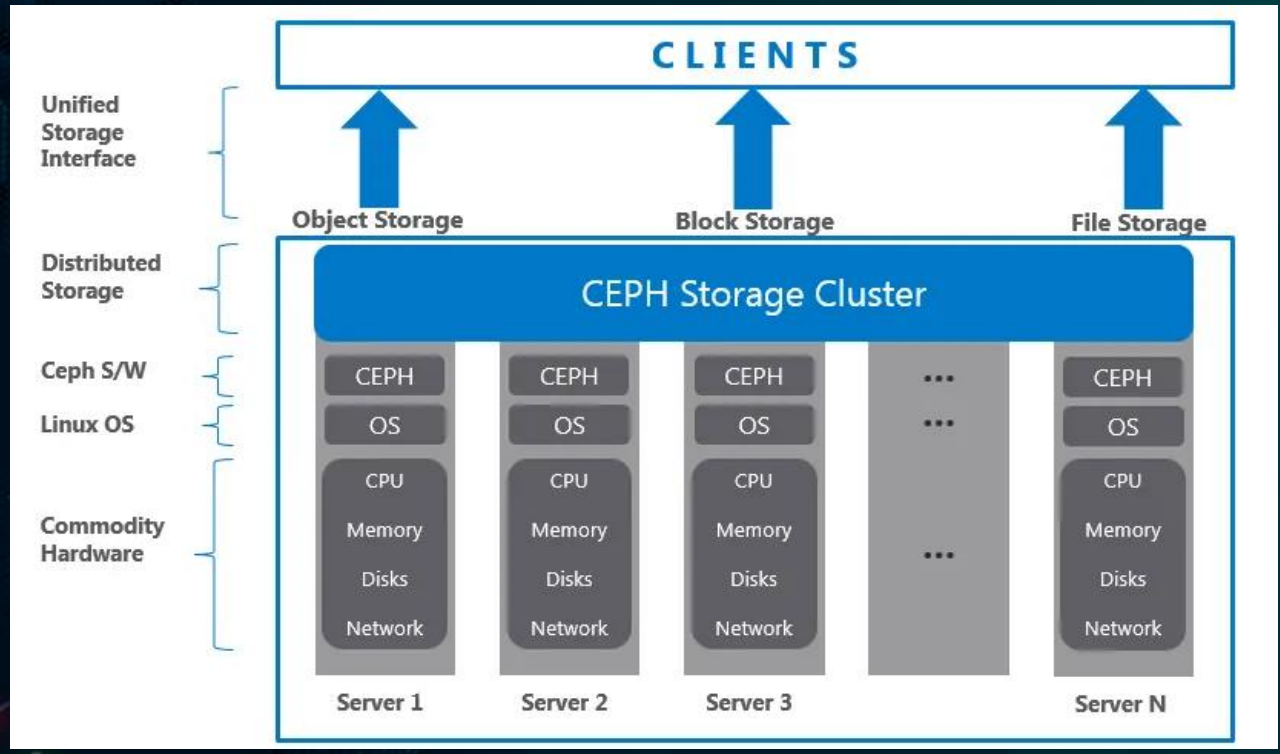
Ceph Storage

Open source, software-defined storage, Support of block, object, and file storage.

Ceph clusters run on any hardware with the help of CRUSH algorithm

Distributes data across the clustered nodes for highly fault-tolerant and scalable storage

Ensures data is both resilient and readily available, even in the hardware failures.



Ceph Cluster Architecture
Image Source: Medium.com

Key Features and Advantages of Ceph

Scalability:
Scales effortlessly by adding more nodes making it suitable for enterprises storage needs.

Data Redundancy:
Redundancy techniques, like replication and erasure coding, to protect data loss and data integrity.

Self-Healing:
Its self-healing capability, automatically detect and repair data inconsistencies, ensuring data consistency and reliability.

Block and File Storage:
Block storage for virtual machines and containers and file storage for shared data access.

Distributed Object Store:
RADOS (Reliable Autonomic Distributed Object Store) forms distributed storage

Ceph Cluster components

OSD (Object Storage Daemon):
Each node runs one or more OSD daemons (one per disk).
It does all data storage, replication and data recovery operations.

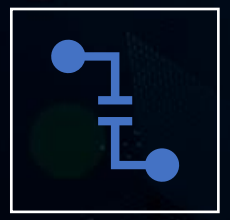
Rados Gateway:
delivers an api service and it connect via S3 or Swift directly with Ceph.

Ceph Manager:
Provide additional monitoring and interfaces to external monitoring and management systems.

Monitor:
Responsible for maintaining a master copy of the cluster map. The Ceph cluster needs a minimum quorum of 3 or more to ensure high availability

Metadata Server: MDS handles all file operations and uses RADOS objects to store data and file system attributes.

Managing Data with Ceph



Ceph Object Storage:

Automatically replicated across different storage devices.

The CRUSH algorithm, a scalable hashing technique, controls how the objects are distributed and accessed, thus avoiding any single point of failure.

Block Storage:

Block Devices, or RADOS Block Devices (RBD), allows Ceph to interact with block storage.

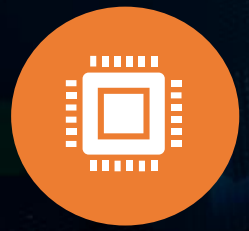
Providing storage solution for virtual machines and support thin provisioning and cache tiering

Ceph File System:

It's a POSIX-compliant file system that uses a Ceph Storage Cluster to store data

Allowing for the usual file operations while adding scalability, reliability, and performance.

Deployment Considerations



Compute nodes: ensure that, compute nodes meet the minimum requirements for proxmox VE. Consider factors like CPU capacity, RAM, and storage.

Storage nodes:

for ceph storage, dedicate nodes with ample storage capacity.

Disk speed and redundancy are critical considerations.

Networking:

high-speed and low-latency networking is essential.

Implement redundant network connections to ensure reliability.

Separation of traffic: for management, storage, and VM/container workloads.

This separation improves security, simplifies troubleshooting, and optimizes performance.

Disadvantages of Proxmox VE

Learning Curve:

Proxmox VE, while user-friendly, may still have a learning curve for those new to virtualization and HCI technologies.

Limited Support:

While Proxmox offers community support, enterprise-level support options are available but at an additional cost.

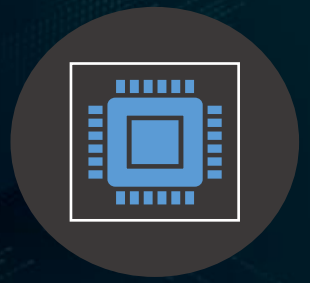
Compatibility:

Proxmox VE primarily supports KVM for virtualization, which may not be compatible with all operating systems or software applications.

Disadvantages of Ceph

Complexity:

its distributed and highly configurable nature, can be complex to set up and manage, especially for those without prior experience.



Resource Intensive:

Ceph can be resource-intensive, particularly in terms of CPU and memory usage.

Data Recovery:

While Ceph is known for its self-healing capabilities, data recovery from failed nodes or storage pools can be time-consuming and may require expertise.

Network Dependencies:

Ceph relies heavily on a high-speed and low-latency network.

Organizations with network constraints may experience performance bottlenecks

Performance Comparison Local Storage vs Ceph

- ✓ **Disk Read/Write Speed in Local Storage:**

```
mahedi@dns1:~$ sudo hdparm -tT /dev/xvda

/dev/xvda:
 Timing cached reads:   17316 MB in  1.99 seconds = 8702.99 MB/sec
 Timing buffered disk reads: 1032 MB in  3.00 seconds = 343.76 MB/sec
mahedi@dns1:~$ sudo dd if=/dev/zero of=test bs=10G count=1 oflag=direct
0+1 records in
0+1 records out
2147479552 bytes (2.1 GB, 2.0 GiB) copied, 8.70109 s, 247 MB/s
```

- ✓ **Disk Read/Write Speed in Ceph Storage:**

```
[root@oia ~]# sudo hdparm -tT /dev/sda

/dev/sda:
 Timing cached reads:   21354 MB in  1.99 seconds = 10733.85 MB/sec
 Timing buffered disk reads: 4042 MB in  3.00 seconds = 1346.62 MB/sec
[root@oia ~]#
[root@oia ~]# sudo dd if=/dev/zero of=test bs=10G count=1 oflag=direct
0+1 records in
0+1 records out
2147479552 bytes (2.1 GB, 2.0 GiB) copied, 2.3077 s, 931 MB/s
```

Case Study of Dhaka University

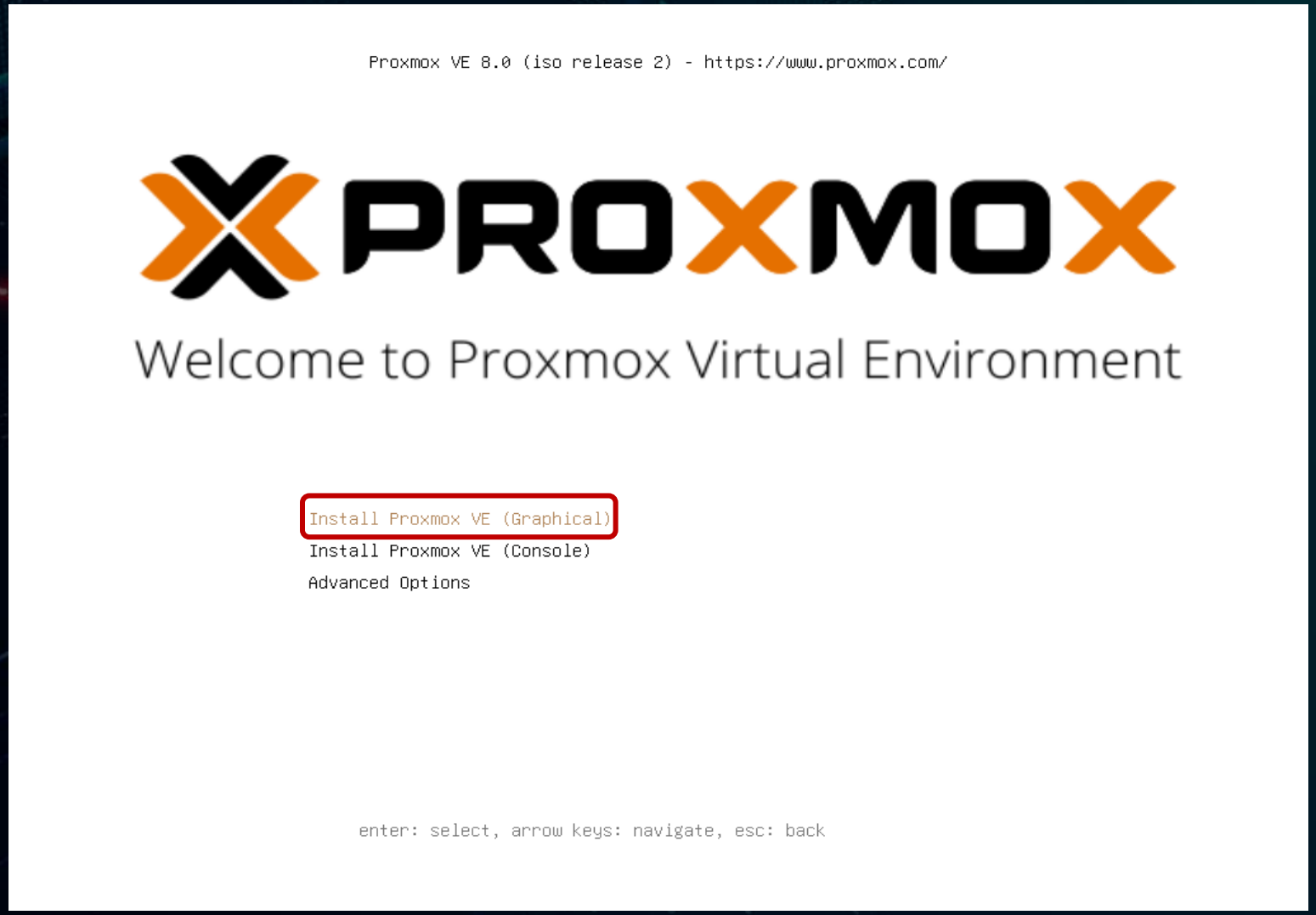
- ✓ **Before Proxmox:**
  - ✓ We are using VMware ESXi virtualization on servers in a rack, but we cannot use clustering due to licensing issues. And licensing costs are higher than our server hardware costs.
  - ✓ When we purchased another rack of servers, we encountered the same issue regarding commercial licensing for virtualization.
  - ✓ We then started using Xenserver, but its free version has limited features.
  - ✓ Meanwhile, ZenServer has stopped providing free version since its new version.
  - ✓ It would not have been possible to ensure high availability as there was no clustering.
  - ✓ We need hyper-converged solutions to ensure high availability and robust performance.

Case Study of Dhaka University

- ✓ **After Proxmox:**
  - ✓ We choose Proxmox and Ceph, which are completely free with no feature restrictions.
  - ✓ Its hyper-converged clustering solutions ensure high availability and automated fault tolerance for our applications.
- ✓ **Benefit:**
  - ✓ This saves roughly $5K/year per server.
  - ✓ Due to the use of Ceph with this, we don't need extra storage devices like SAN.
  - ✓ Overall, the availability of our application is more assured than ever.
  - ✓ Ensure Uptime up to 99.99%

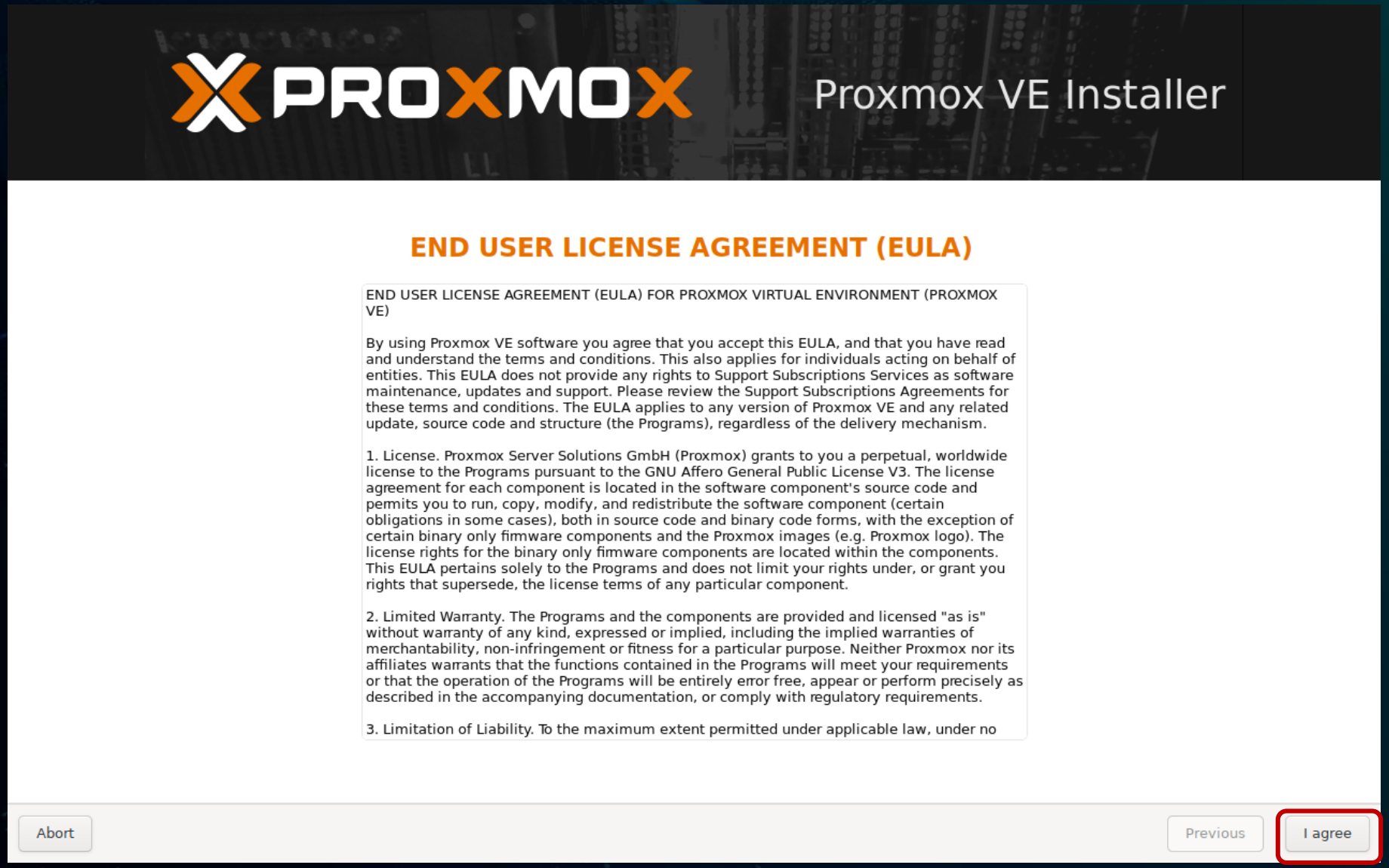
Install Proxmox

✓ Boot Proxmox:



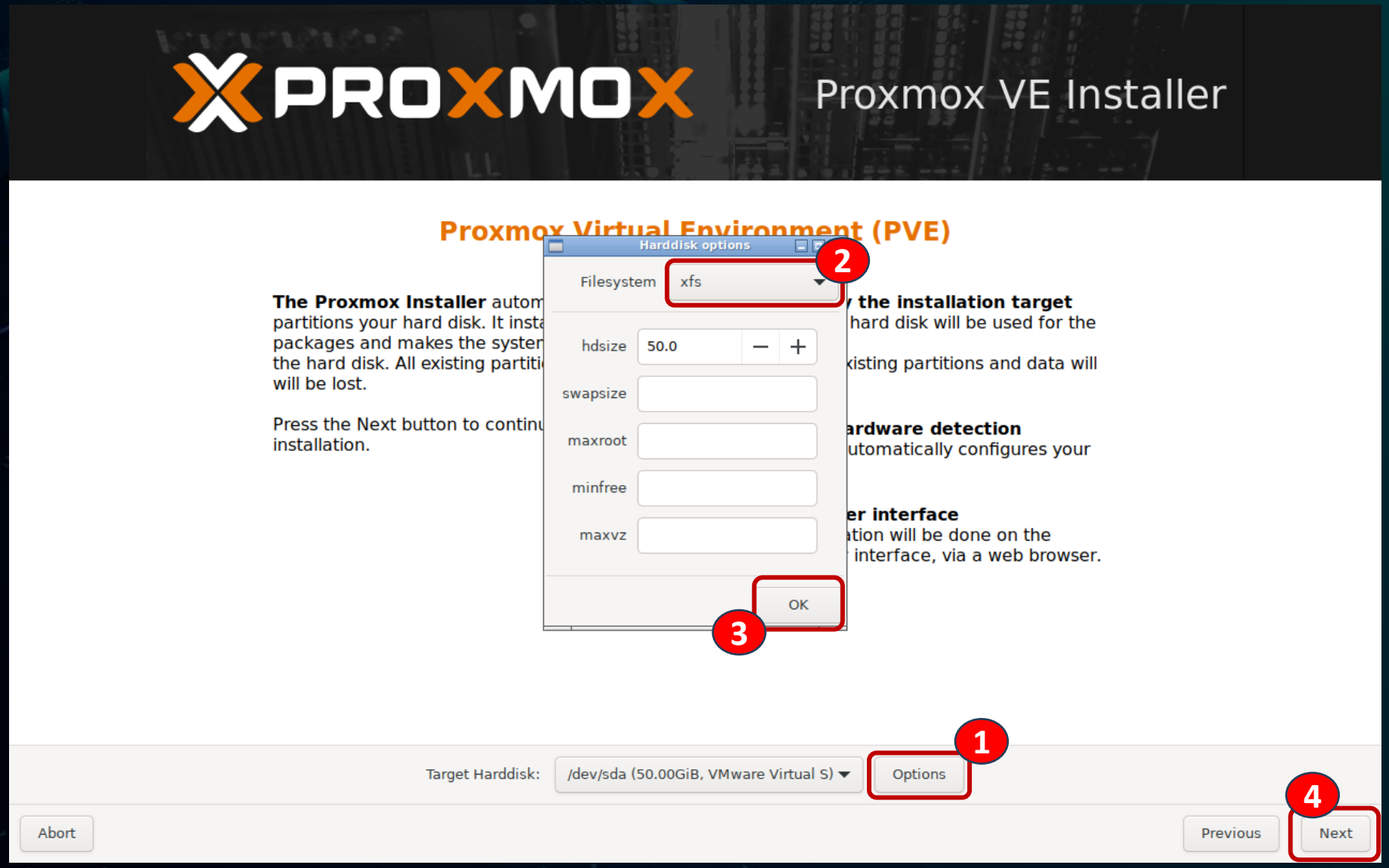
Install Proxmox

✓ Accept license agreement:



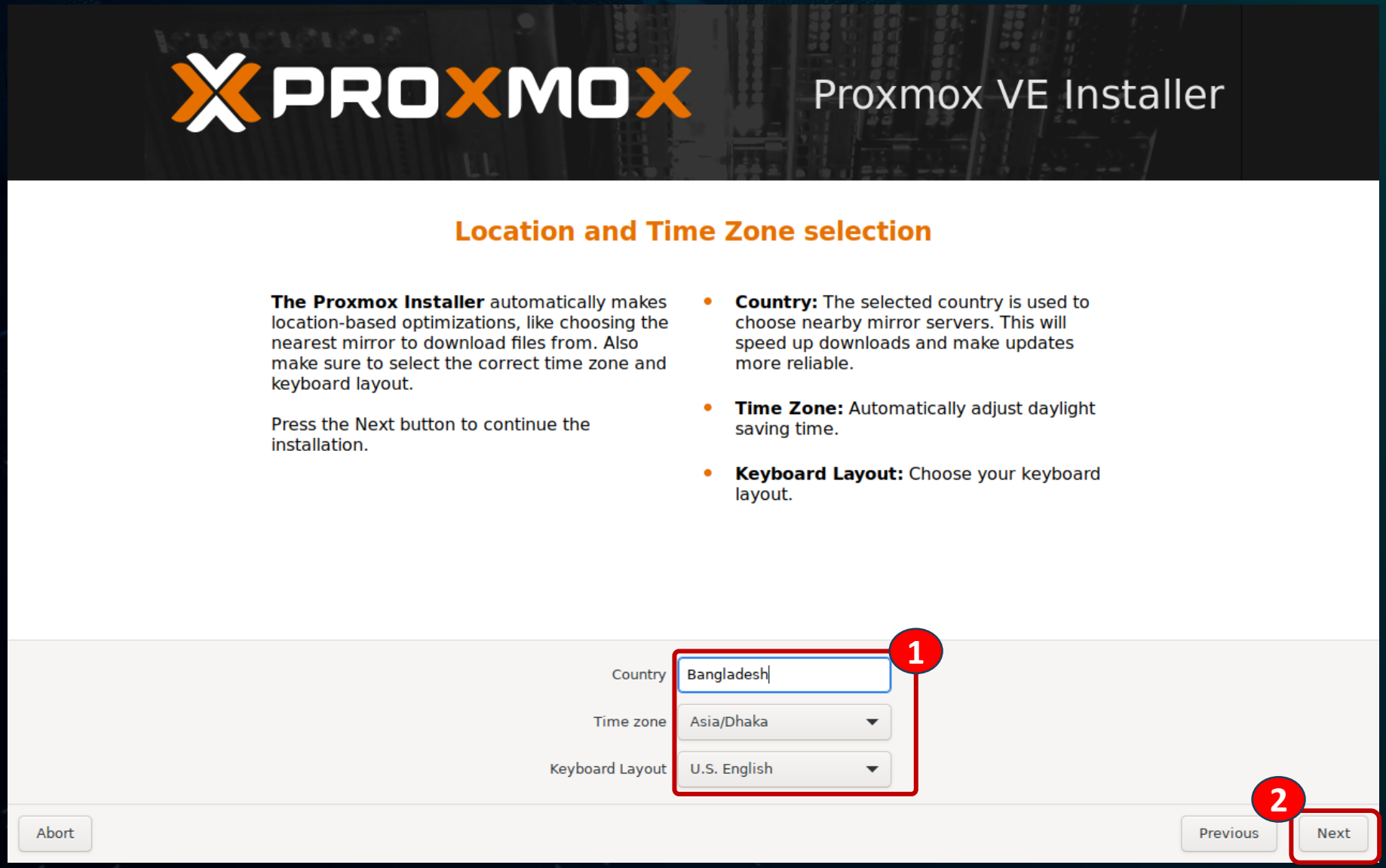
Install Proxmox

✓ Configure Disk:



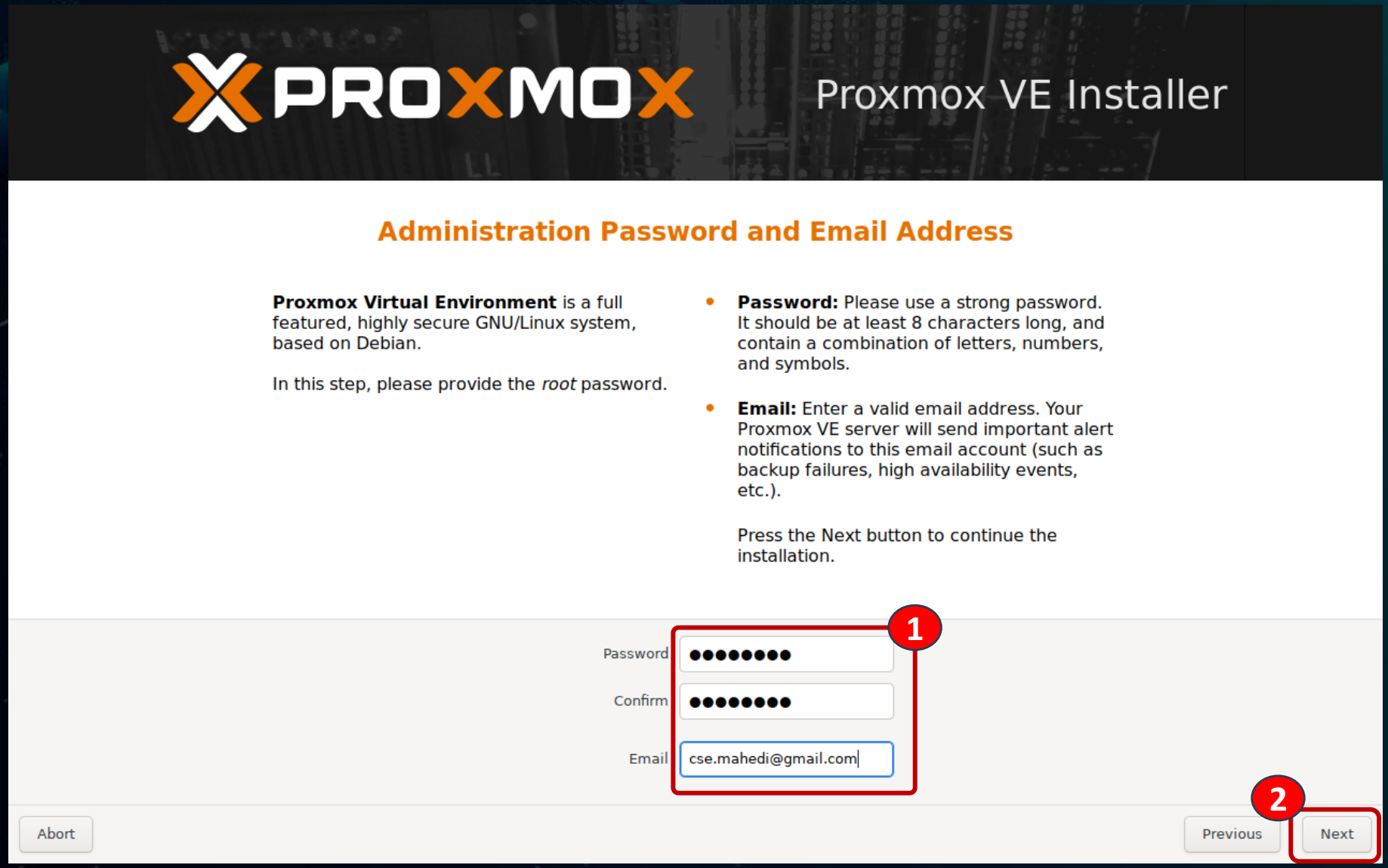
Install Proxmox

✓ Set Location and Timezone:



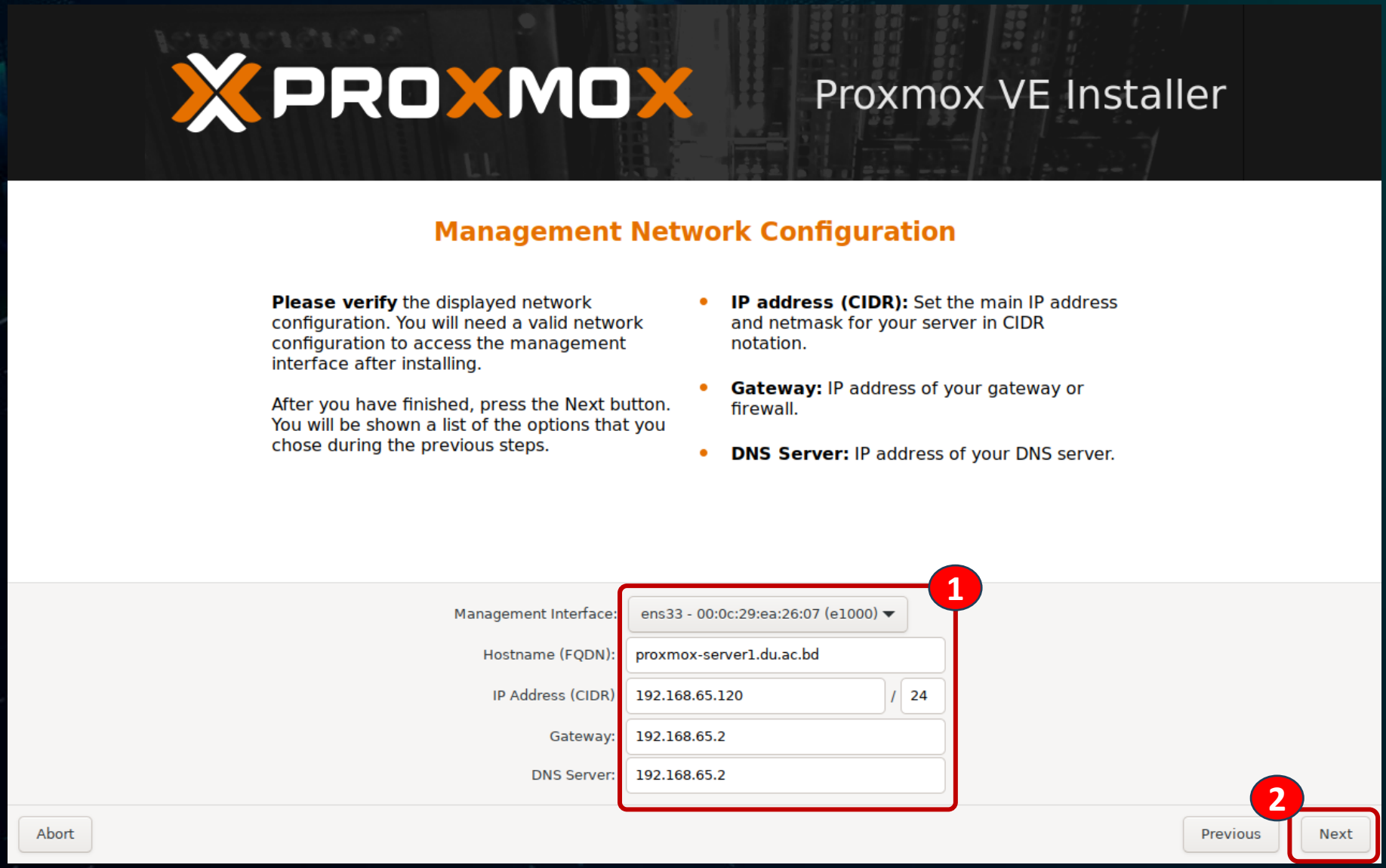
Install Proxmox

✓ Configure Root password and Email:



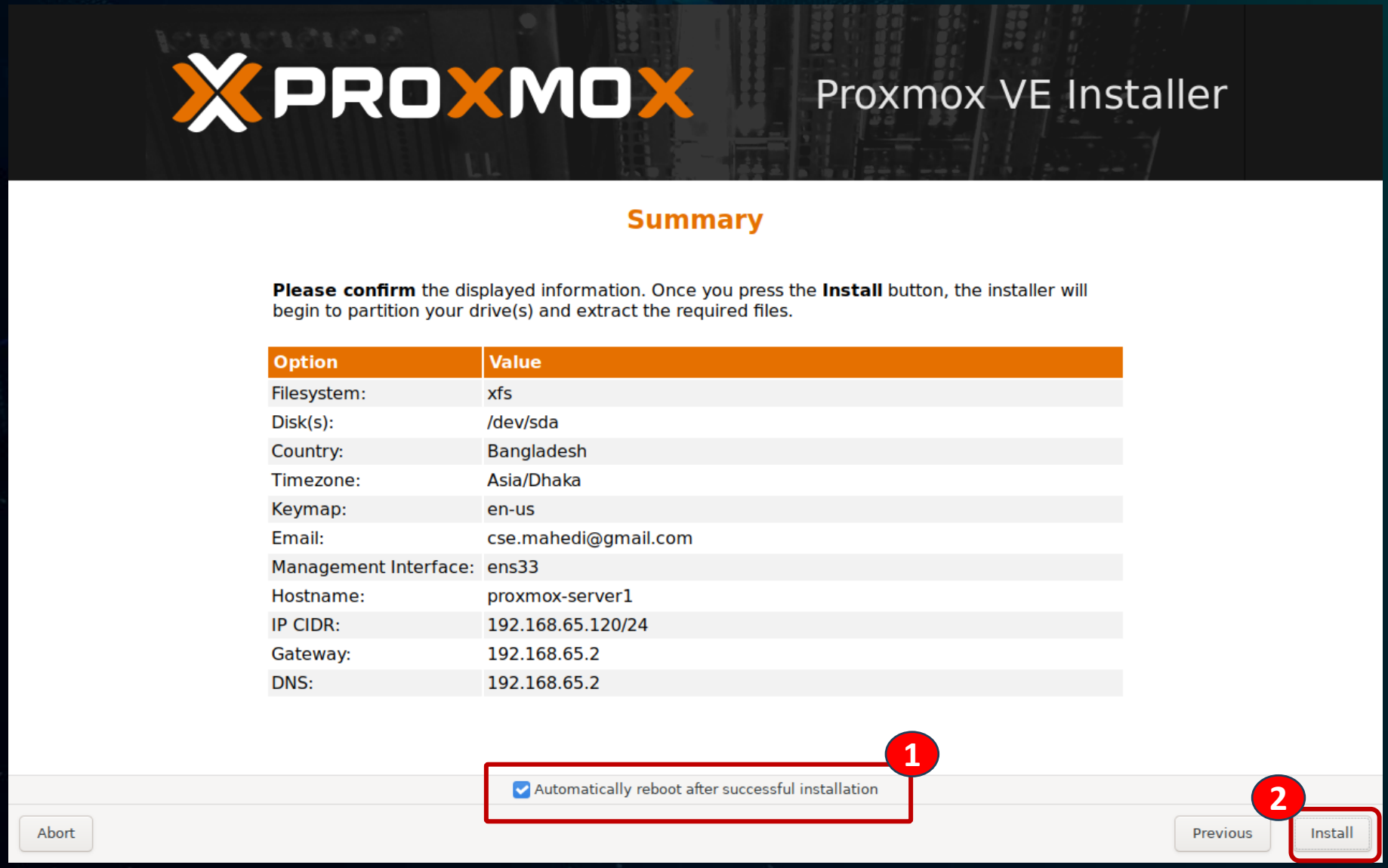
Install Proxmox

✓ Set Hostname and Configure Network:



Install Proxmox

✓ Check Summary and Start Install:



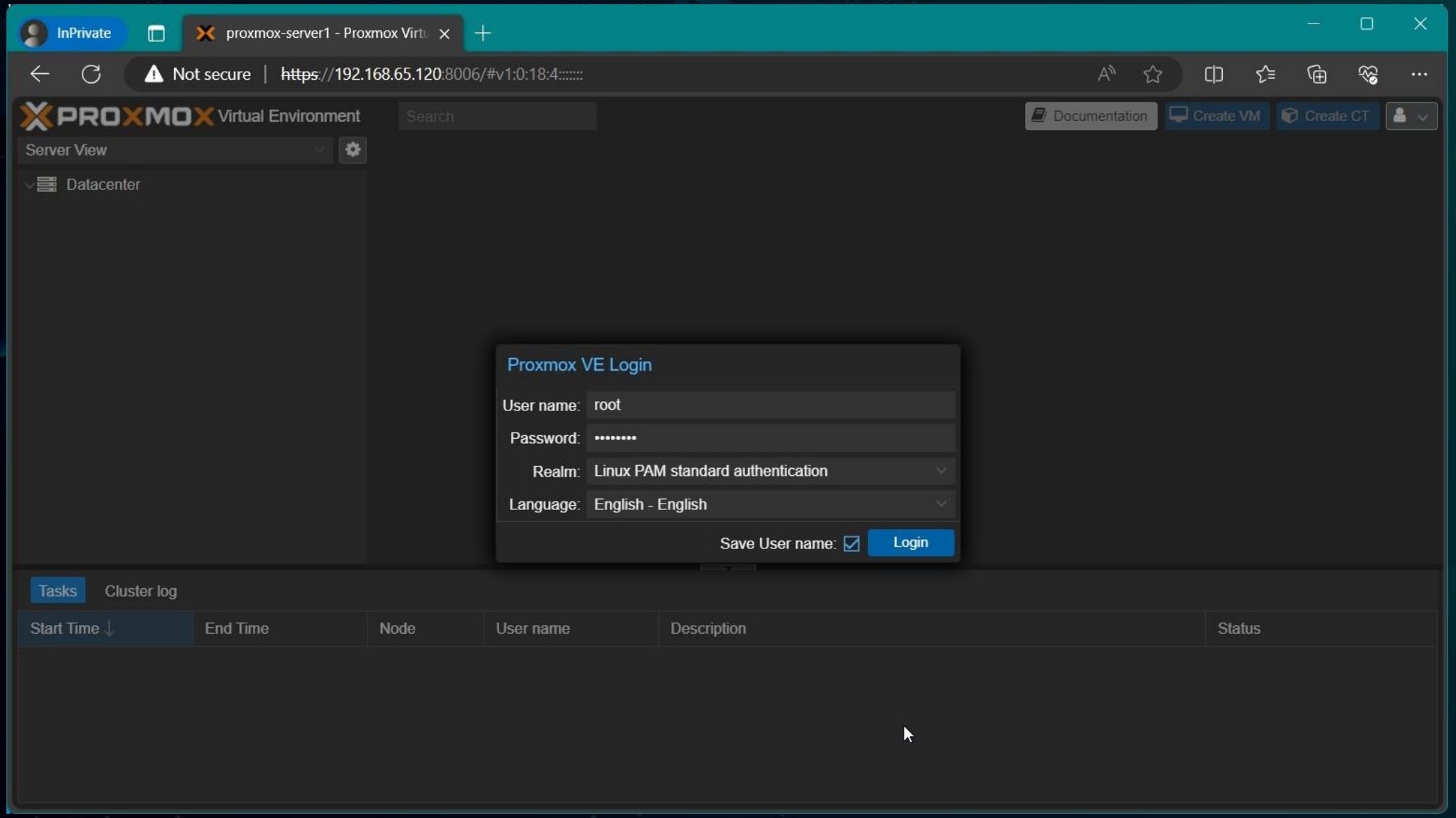
Install Proxmox

✓ Proxmox Login:

```
--------------------------------------------------------------------------------

Welcome to the Proxmox Virtual Environment. Please use your web browser to
configure this server - connect to:

  https://192.168.65.120:8006/

--------------------------------------------------------------------------------

proxmox-server1 login: _
```

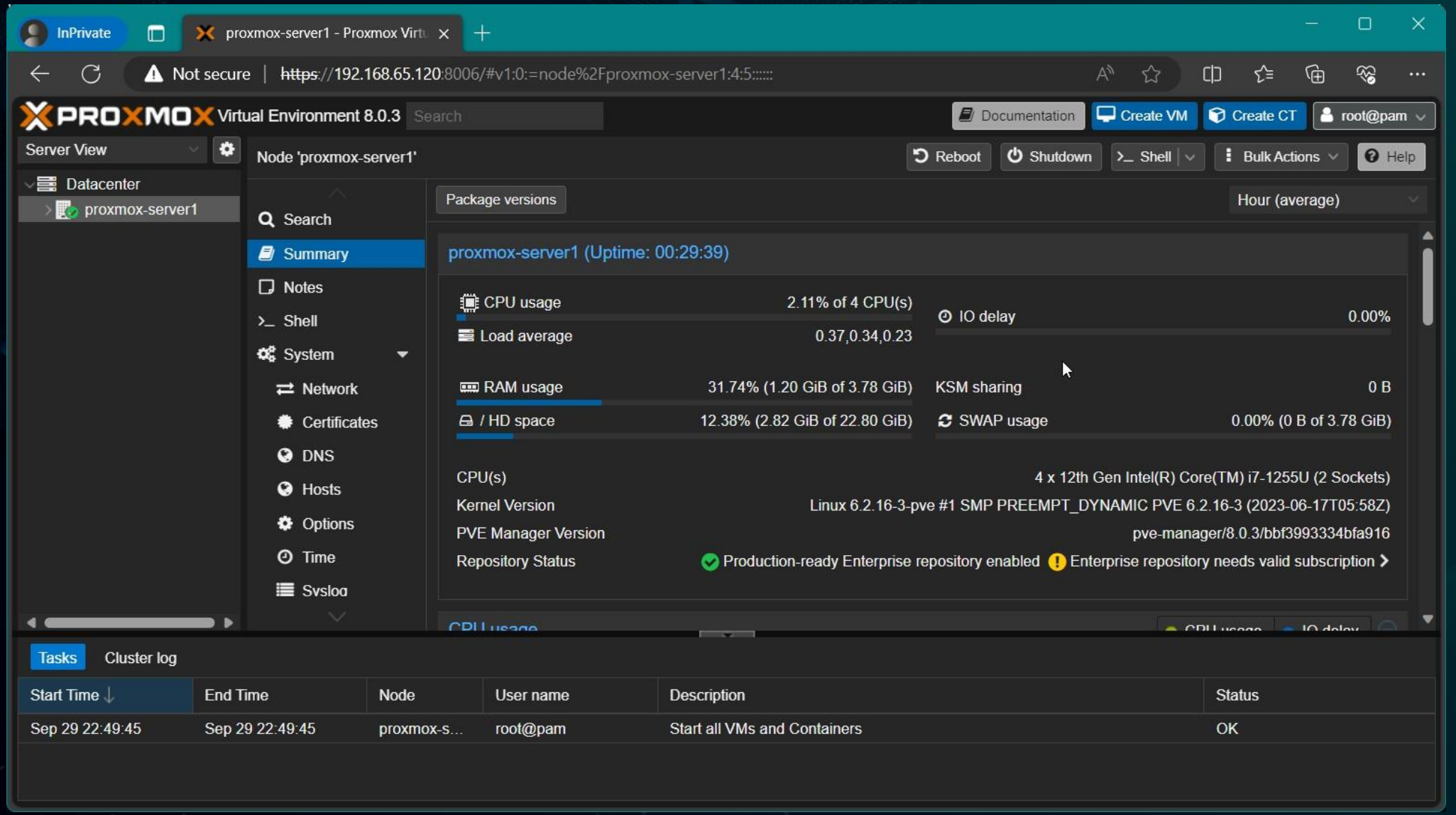
Install Proxmox

✓ Login from browser:

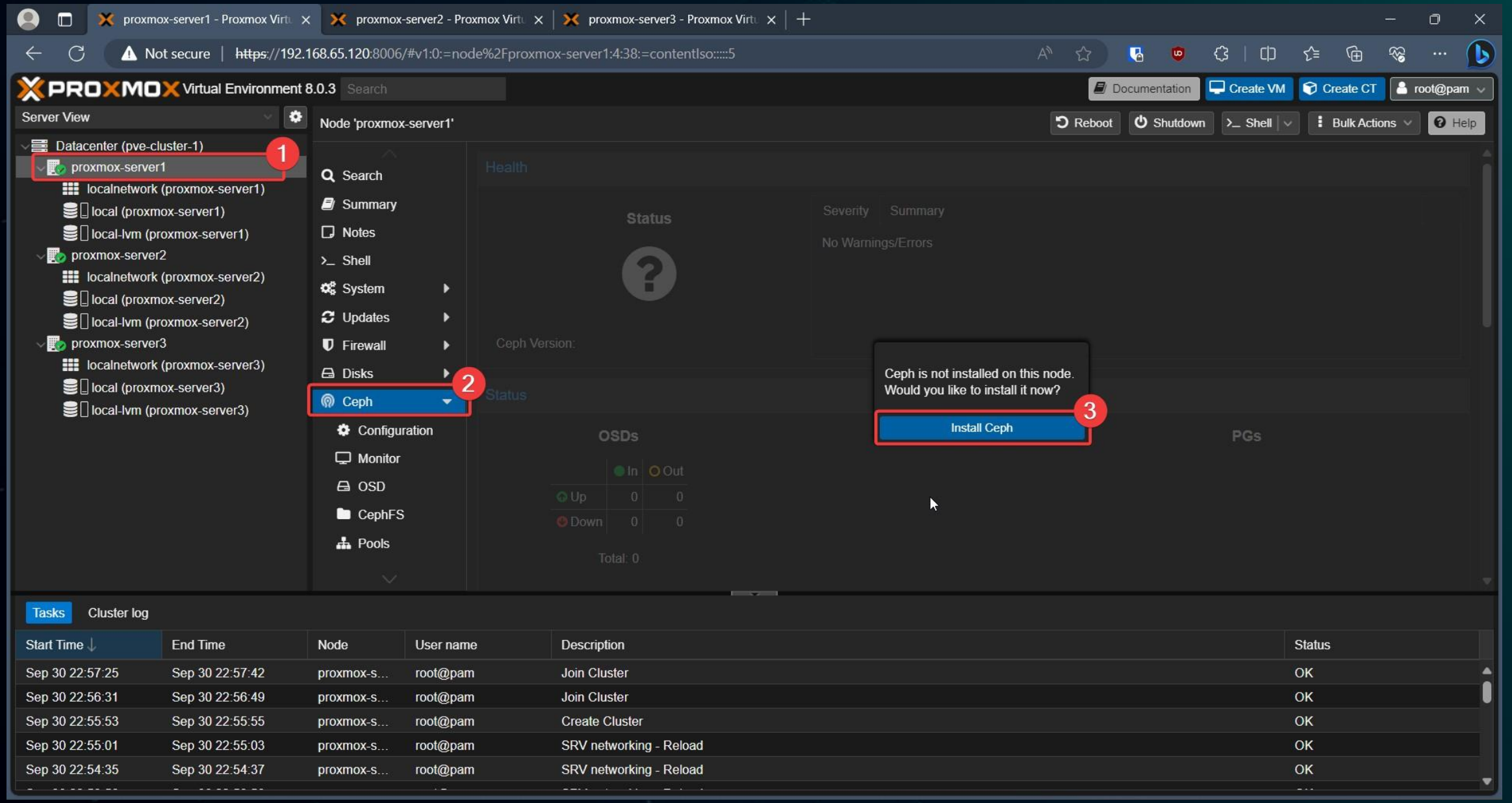


Install Proxmox

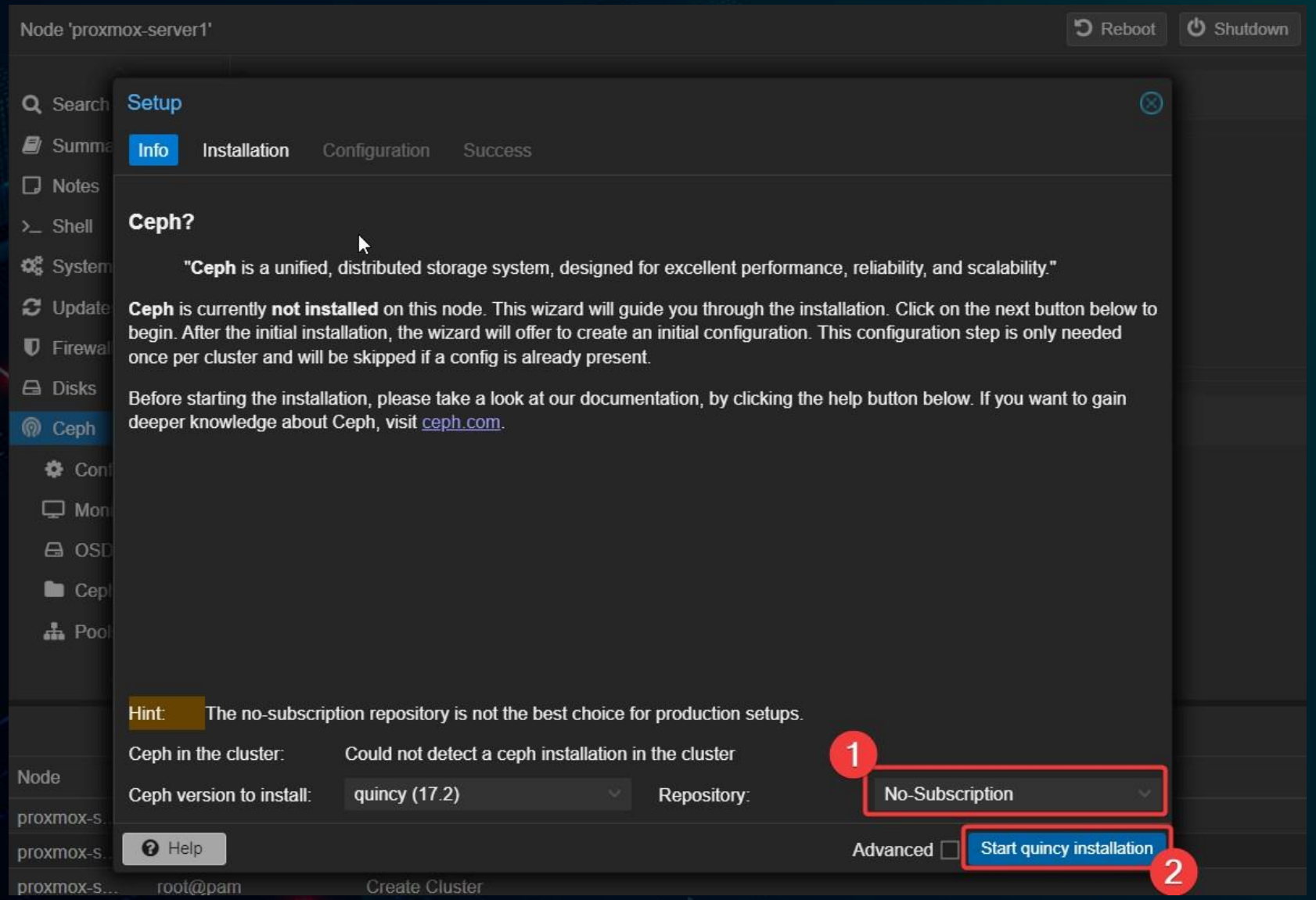
✓ User Dashboard:



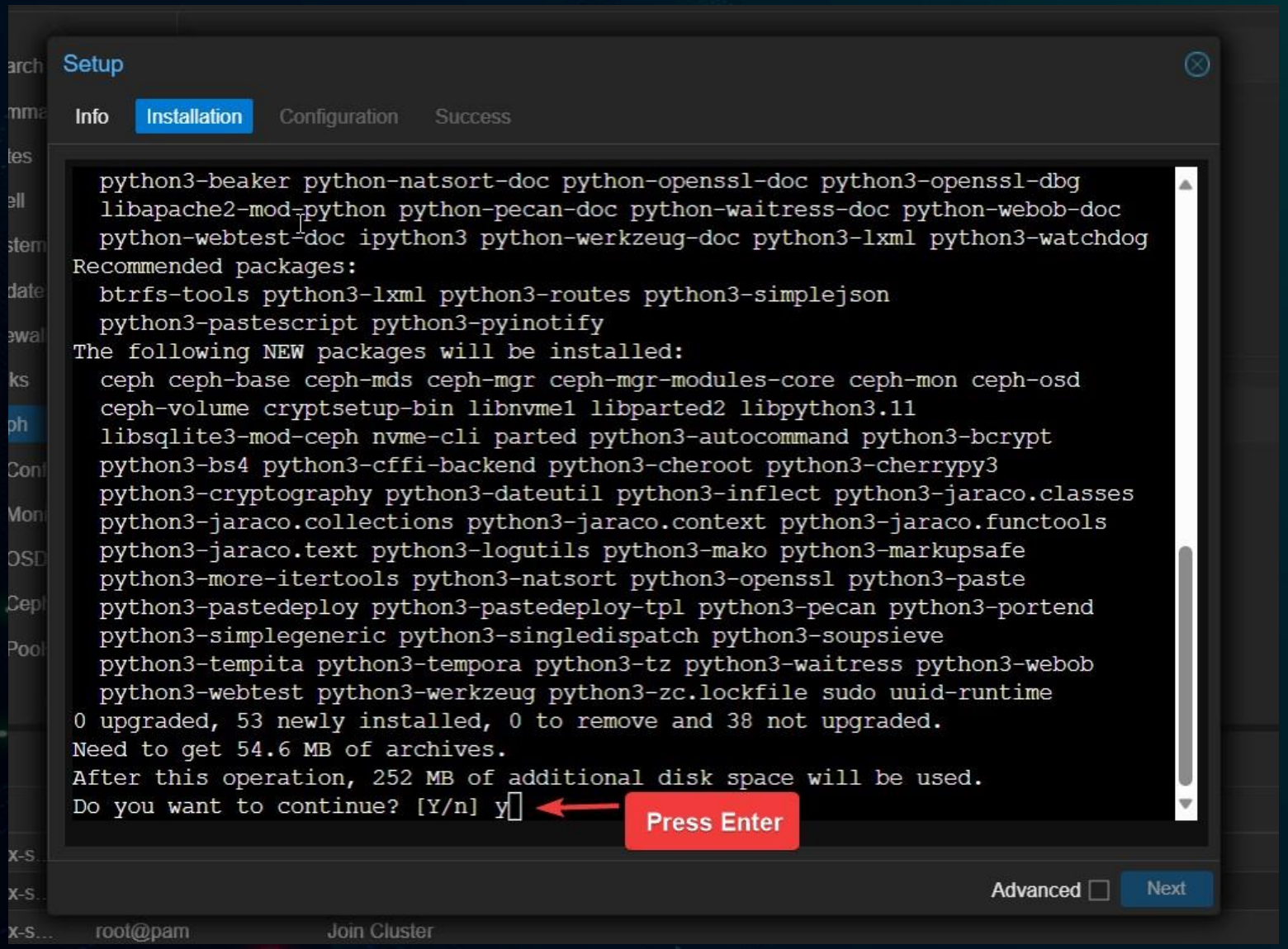
Steps to Install Ceph on Proxmox



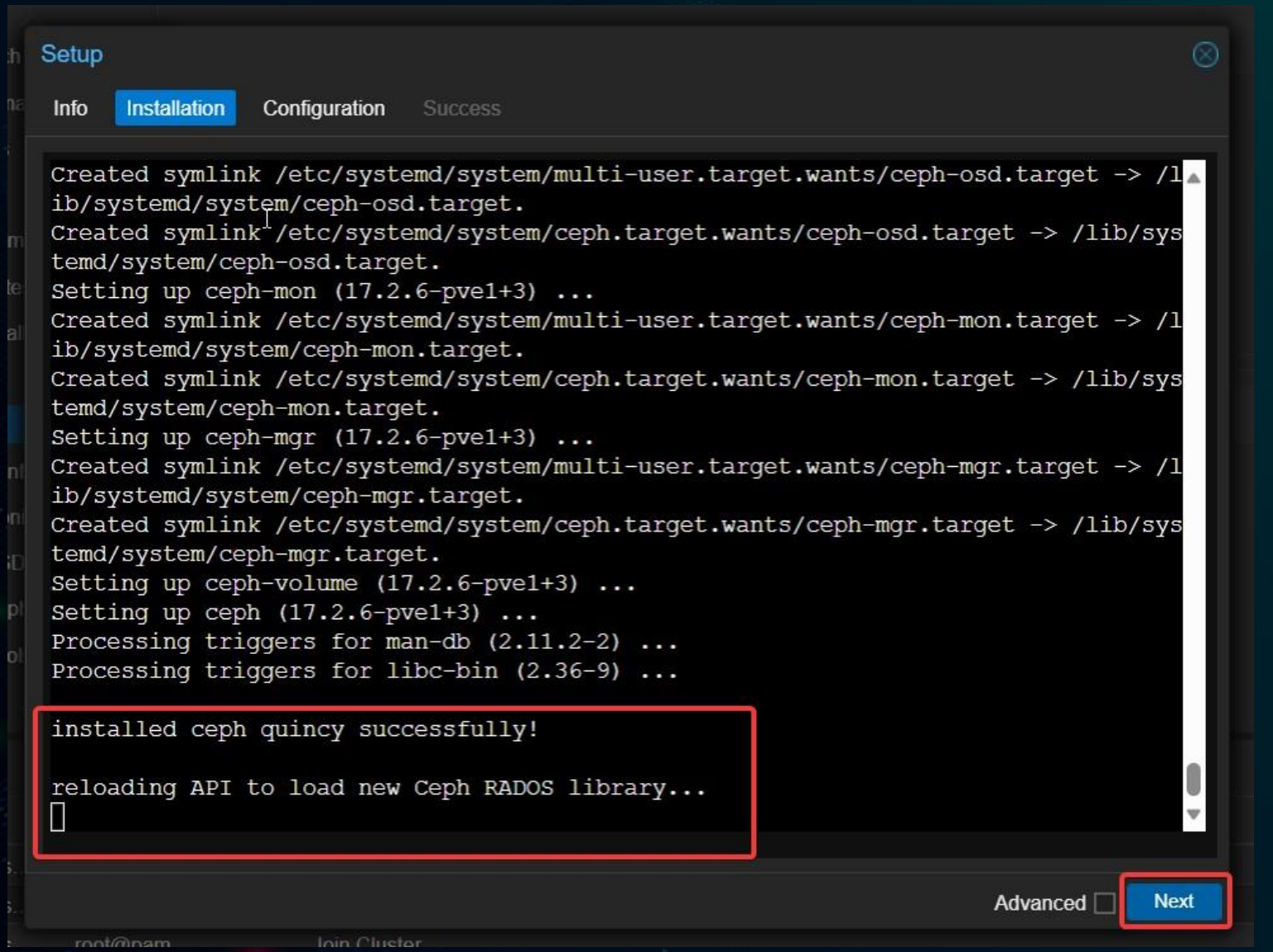
Steps to Install Ceph on Proxmox



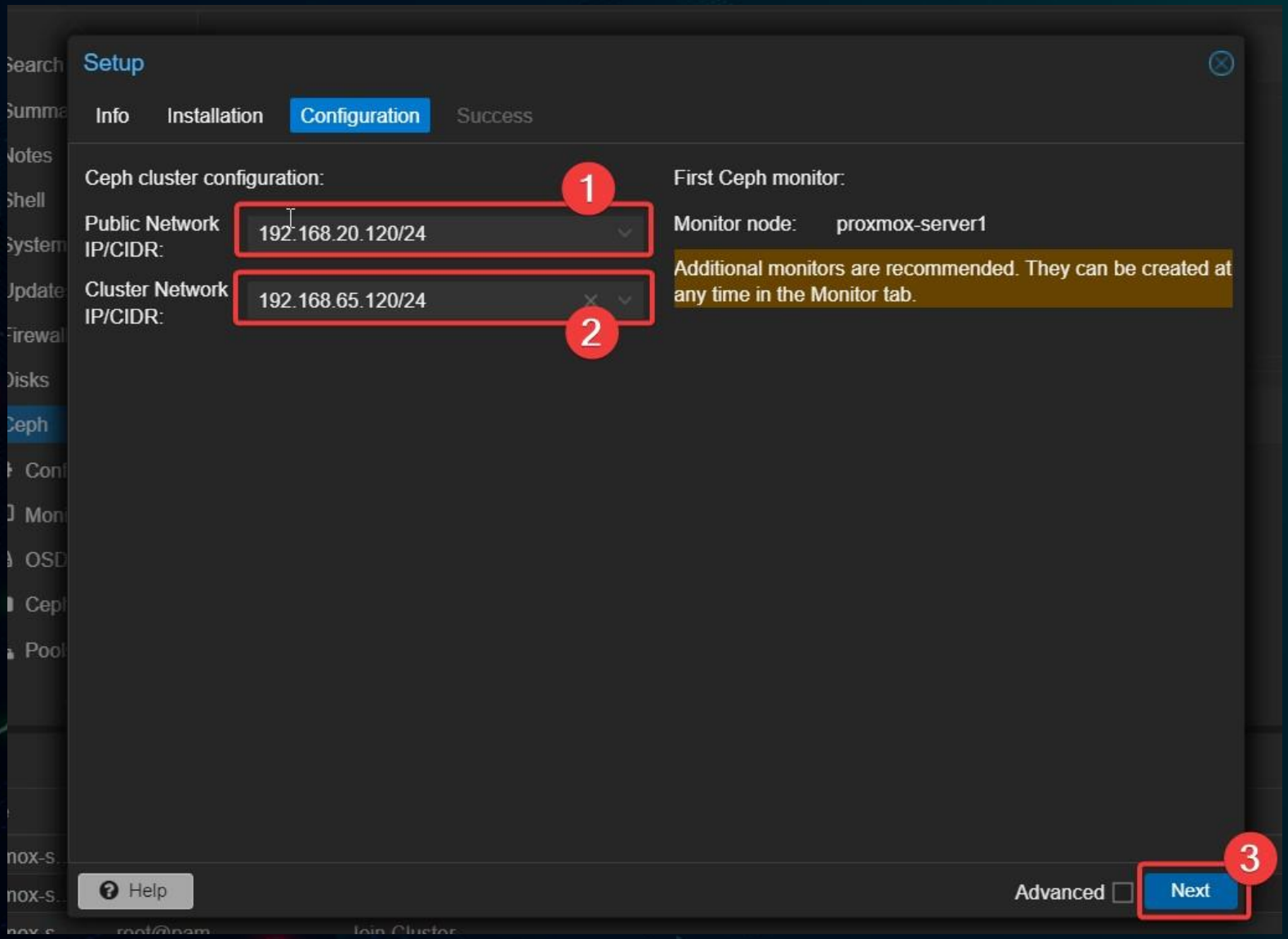
Steps to Install Ceph on Proxmox



Steps to Install Ceph on Proxmox

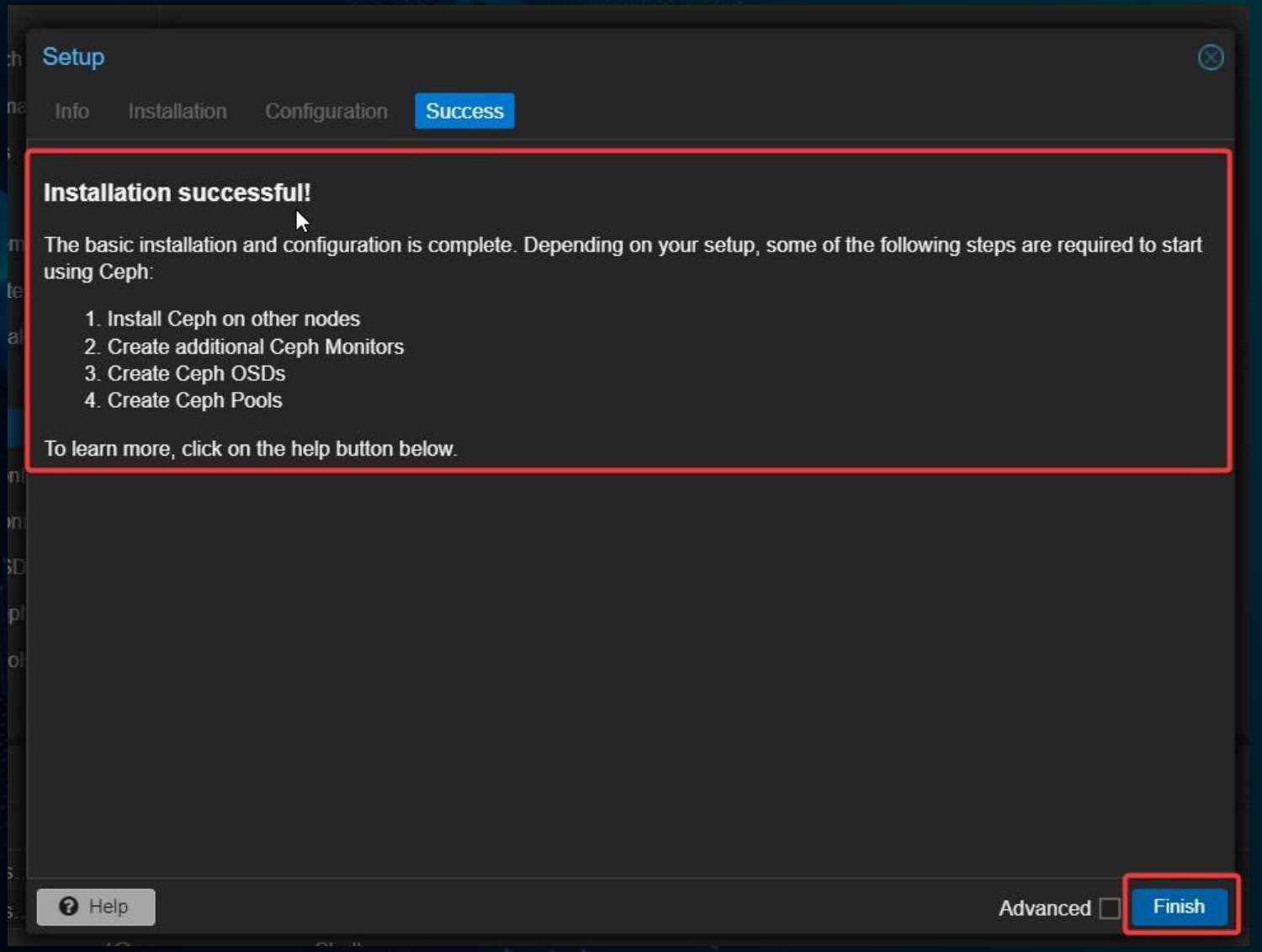


Steps to Install Ceph on Proxmox

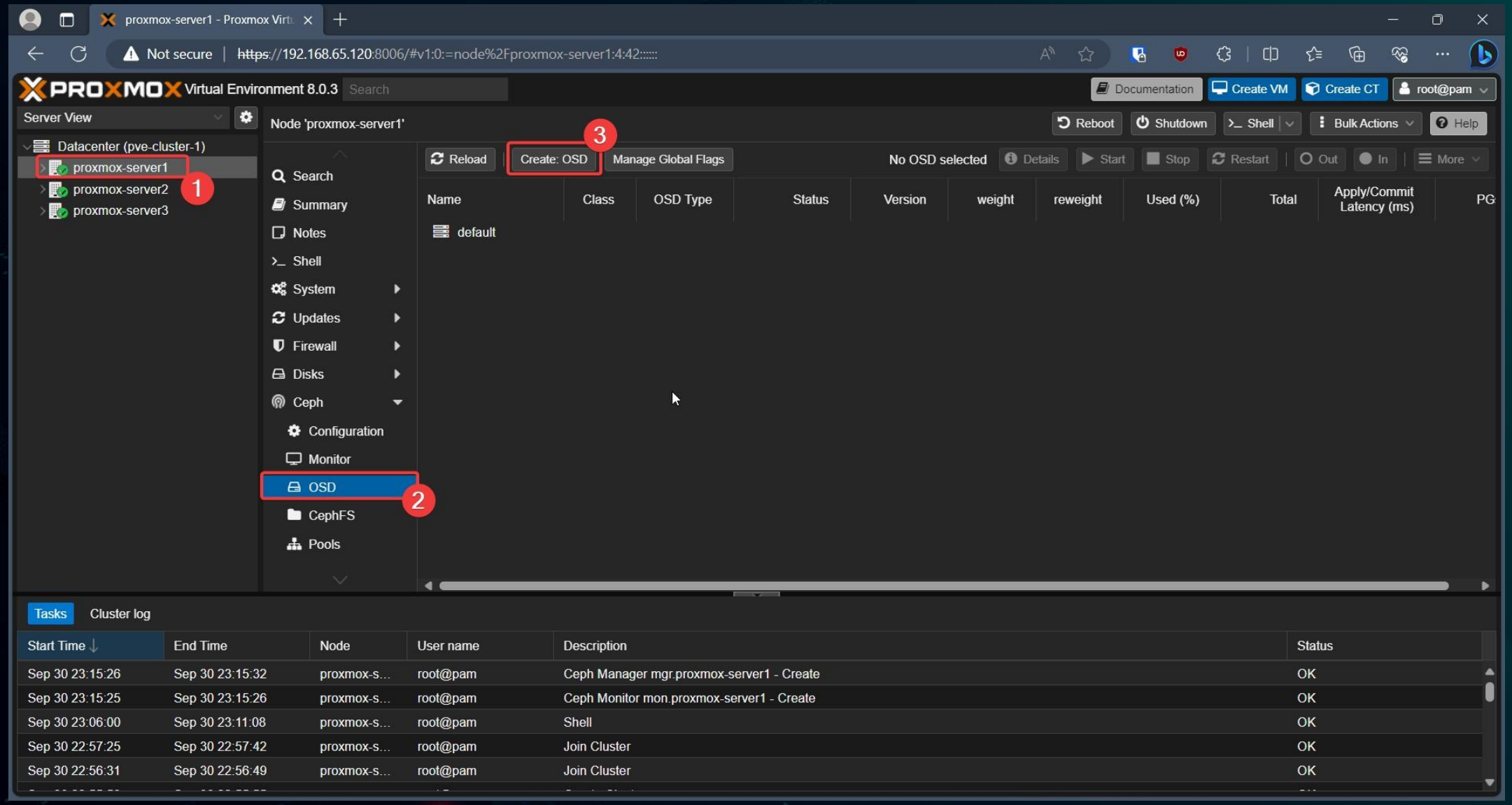


Steps to Install Ceph on Proxmox

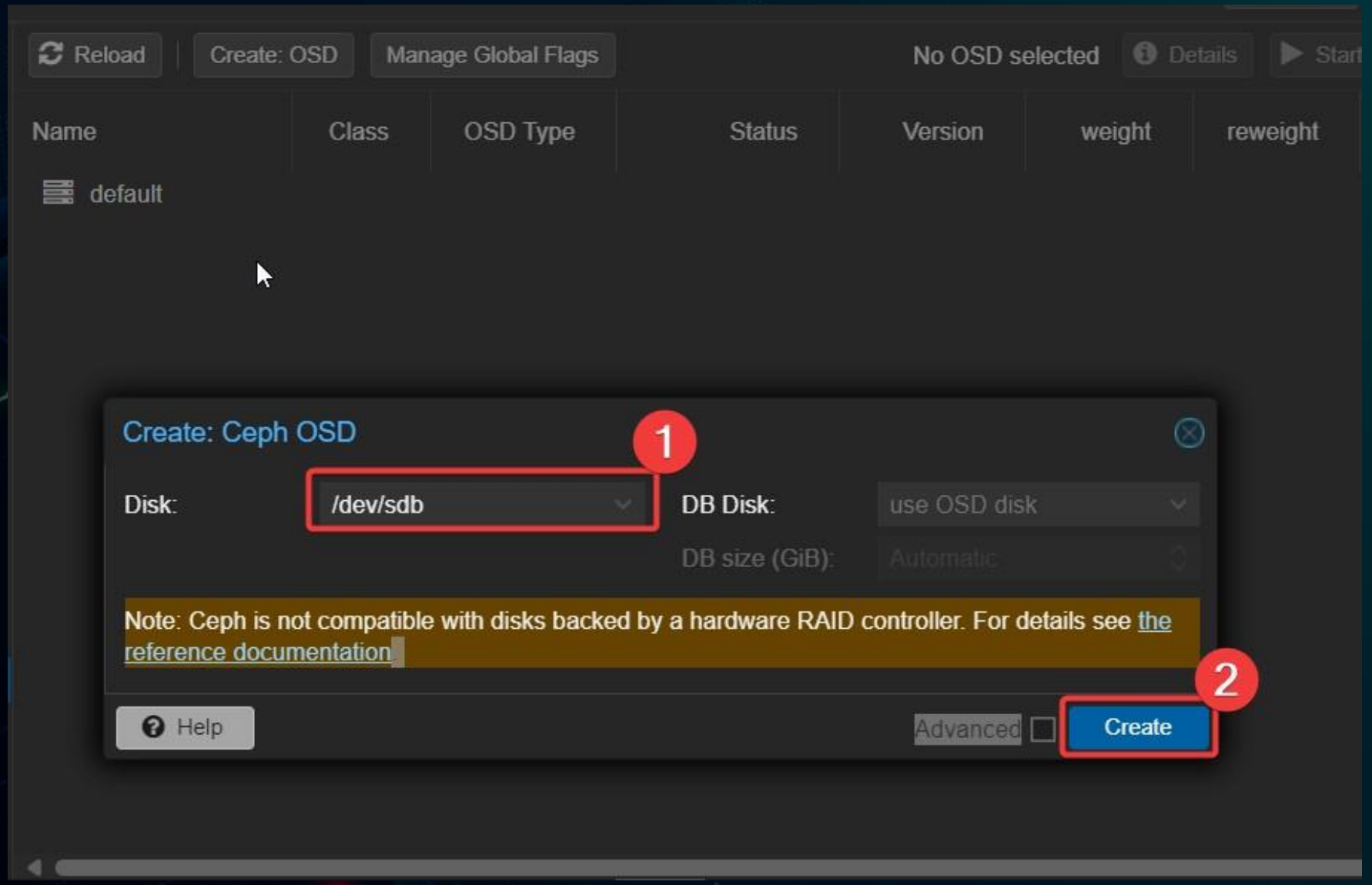
- Install successfully, repeat these steps on other nodes:



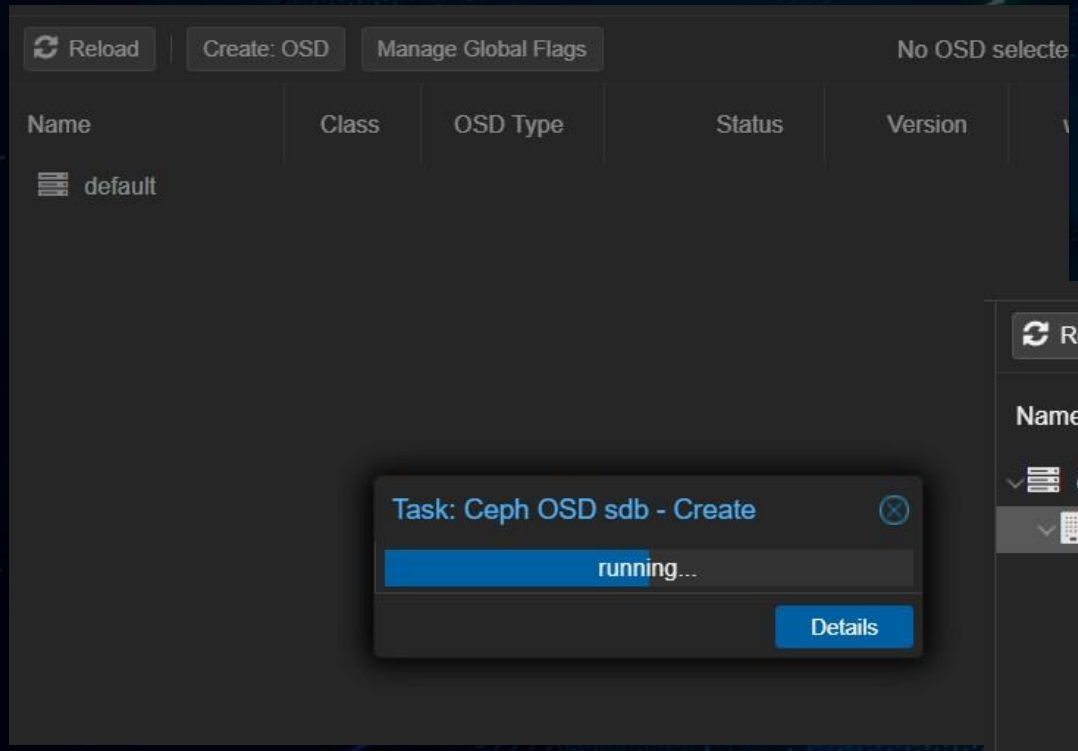
Create Ceph OSD Daemons

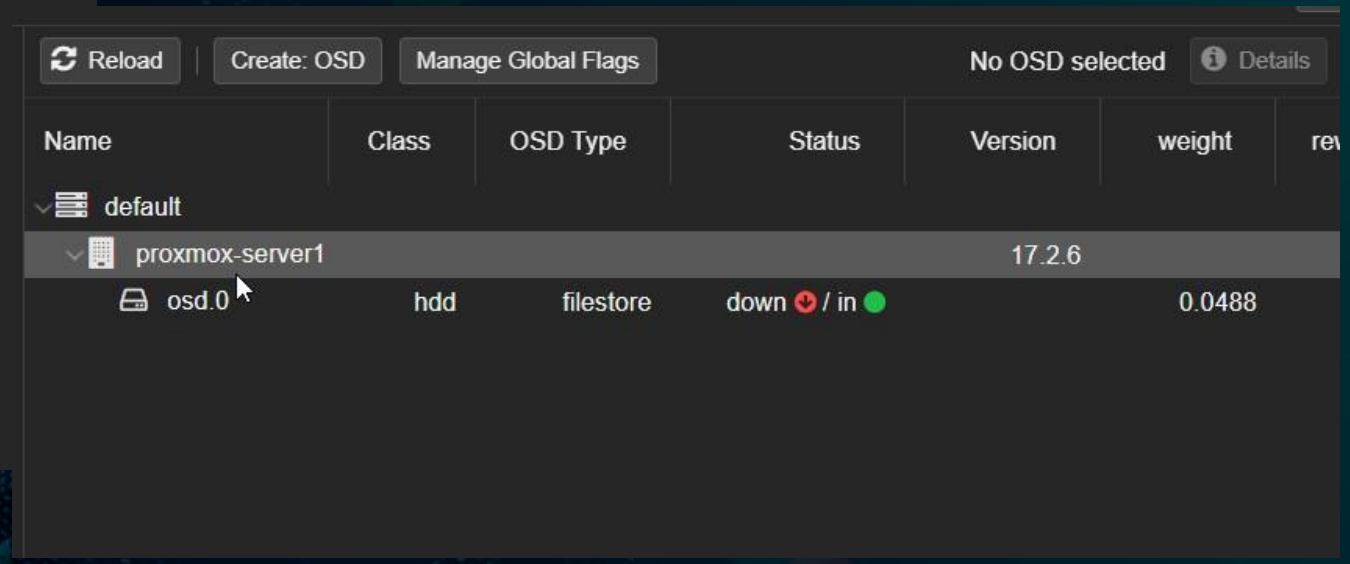


Create Ceph OSD Daemons

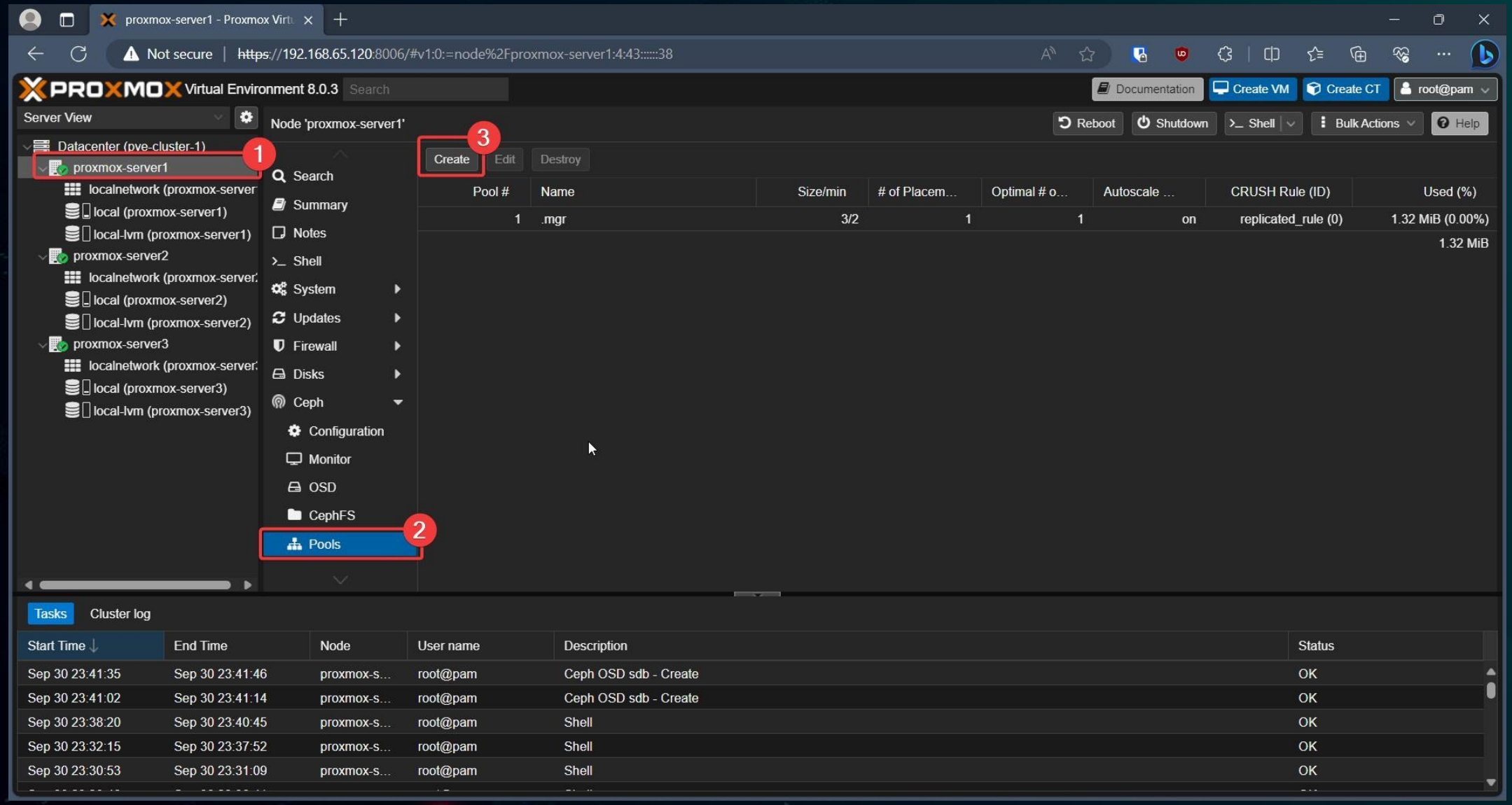


Create Ceph OSD Daemons

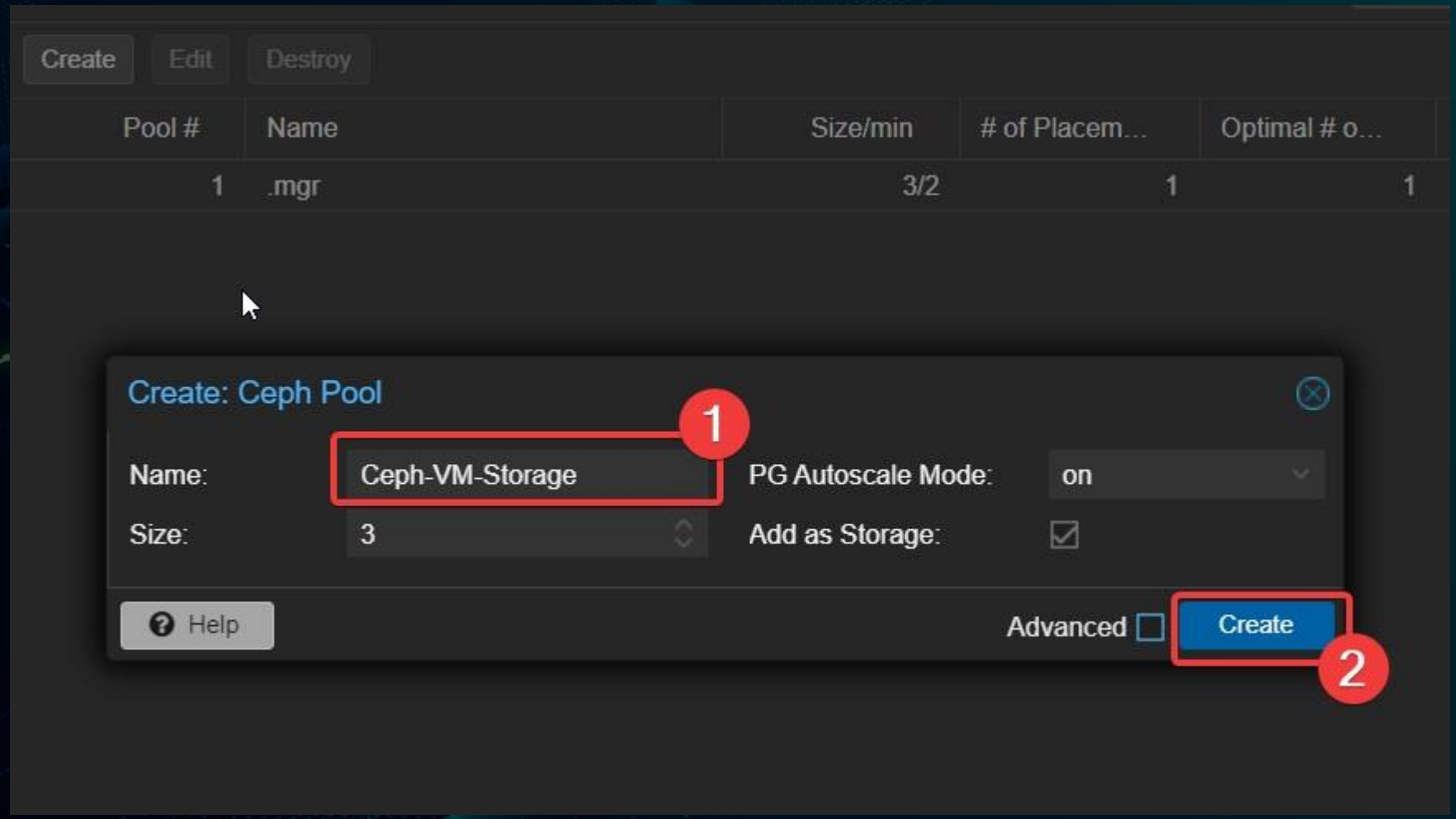




Create Ceph Pool

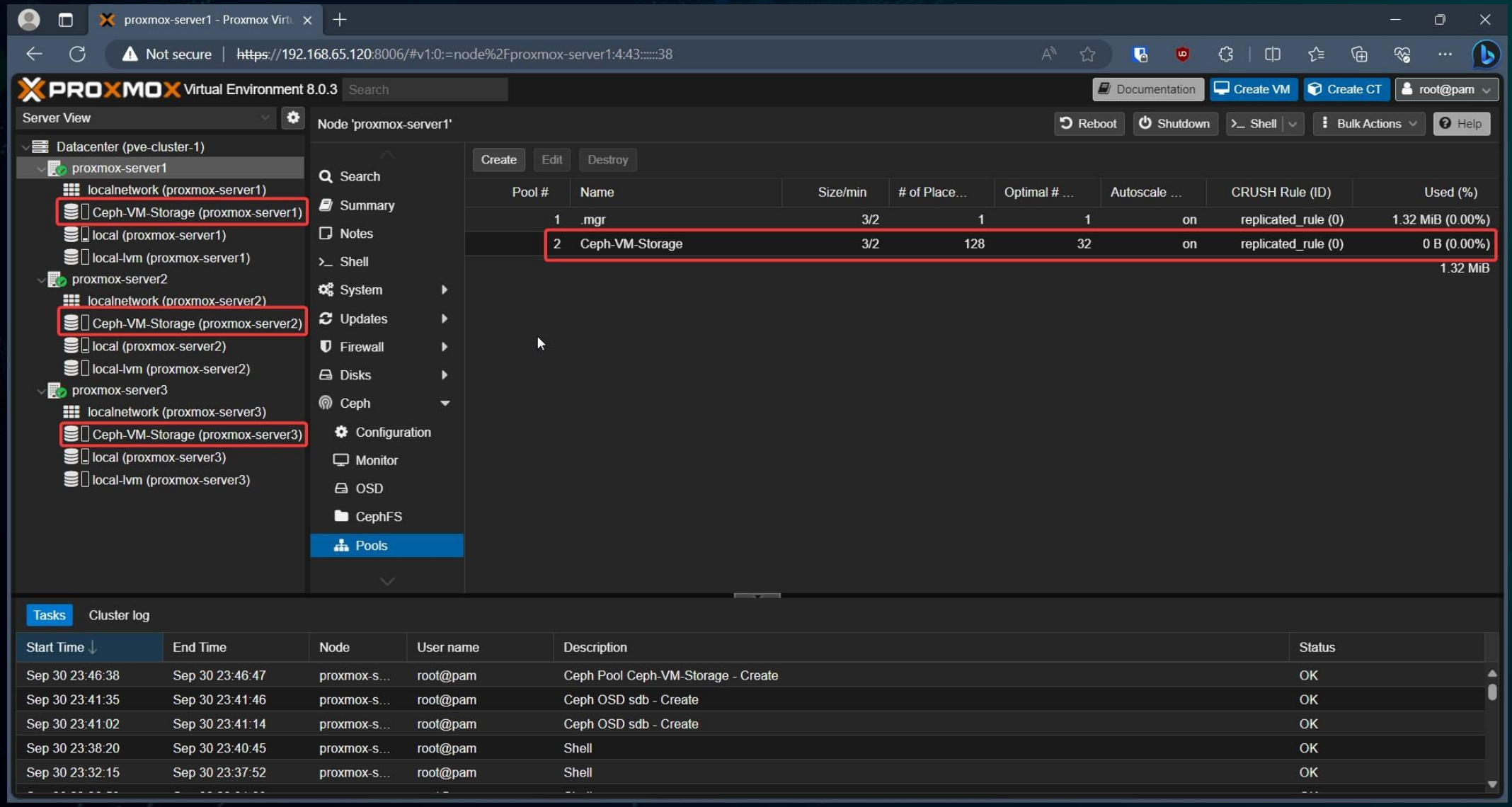


Create Ceph Pool



Create Ceph Pool

- Pool is ready and available in all servers to store VMs:



Live demo

- ✓ **Live Demo from test lab (Max 5mins) / Recorded Video**

Thank You

?