How It Works:
Rubrik Integration with Pure Storage FlashArray

Marc Creviere
May 2021
RWP-0581

TABLE OF CONTENTS

CHALLENGES

Rubrik CDM leverages built-in VMware snapshot capabilities to capture point-in-time data from your vSphere environment. When the backup is completed the snapshot is merged and virtual disks are consolidated, resulting in what is typically an imperceptible 'stun' where I/O is paused during the final disk consolidation operation. The Virtual Machines are stunned for the duration of the consolidation. In typical circumstances, this process is completed almost immediately. A VM with considerable amounts of changes gathered in the temporary snapshot can be stunned for a noticeable or disruptive amount of time. This can have adverse effects on guest applications or services. Likewise, the simple existence of a snapshot can negatively impact performance of a guest, largely because of the 'write redirection' required to send new blocks to the delta disk.

Most workloads are tolerant of this and can take advantage of Rubrik's do-no-harm approach to backup operations and native vSphere integration, but some solutions recommend or require additional steps to minimize any impact from the vSphere snapshot process.

## RUBRIK INTEGRATION WITH PURE FLASHARRAY

Using Rubrik's integration with Pure FlashArray, snapshot data can be captured at the storage level and backed up by Rubrik, minimizing impact to production operations. VMs do not incur any additional latency due to the existence of snapshots while backup data is being ingested. Because the backed-up data is captured on a temporarily created volume via a proxy VM there is no VM stun associated with disk consolidation at the conclusion of the backup. Rubrik's forever incremental snapshot approach is still applicable after the first backup as this integration can still make use of change block tracking. Furthermore, this integration can be enabled and disabled without impact to the forever incremental process.

Pure FlashArray's snapshot technology is non-performance impacting. Snapshots are not copy-on-write, nor are they redirect-on-write. They are simply metadata preservations of the point-in-time of the volume. No data is moved or redirected. So no performance is hurt when a FlashArray snapshot is created, in existence, or deleted[1] and globally data-reduced, so it does not significantly impact operation or capacity during backup operations.

Rubrik's Pure Storage array integration is an effective way to protect workloads that are specifically vulnerable to latency caused by vSphere snapshot stun time. For all other workloads Rubrik recommends that standard protection workflows be used.

---

1   Rubrik and FlashArray Integration: Why it matters - Cody Hosterman, Pure Storage

ARCHITECTURE
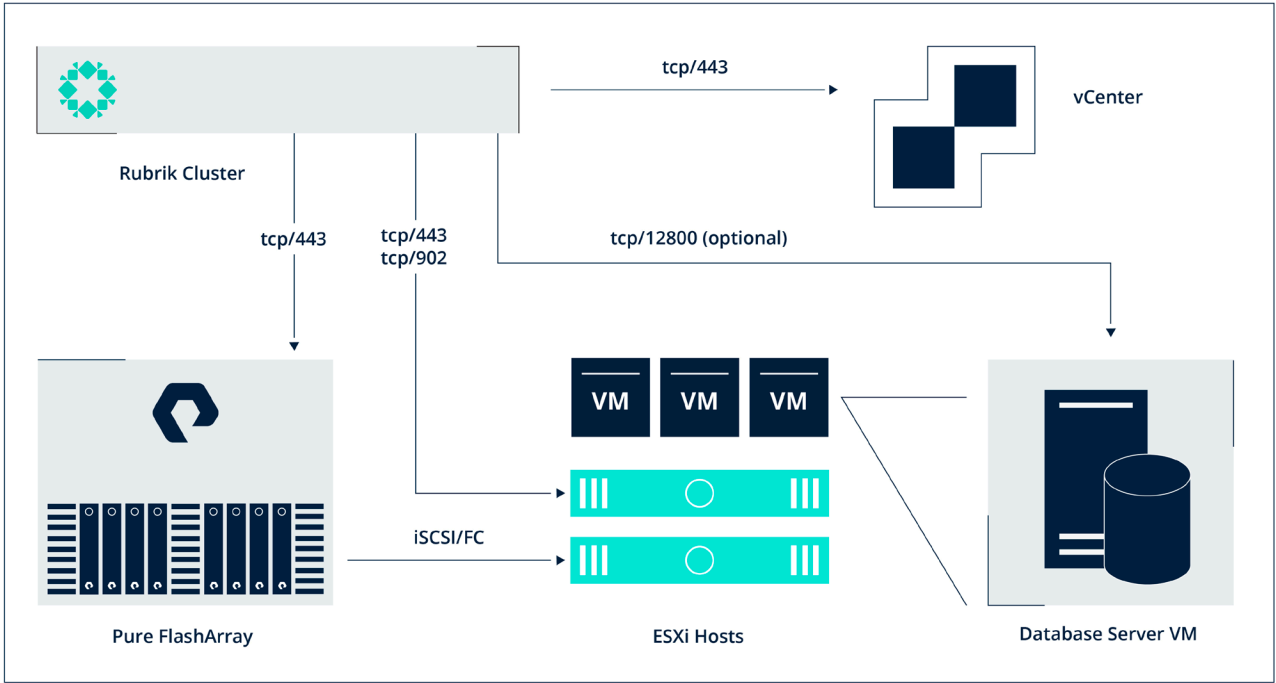
Rubrik CDM will communicate with vCenter and the Pure FlashArray using their respective APIs over HTTPS. Snapshot data will be captured from the ESXi host via NBDSSL, while pre and post scripts will either run via RBS directly to the guest OS or via VMware Tools accessed through the hypervisor host[2].

---

2 Overview of VSS, Application or Filesystem Consistent Snapshots

PROTECTION WORKFLOW
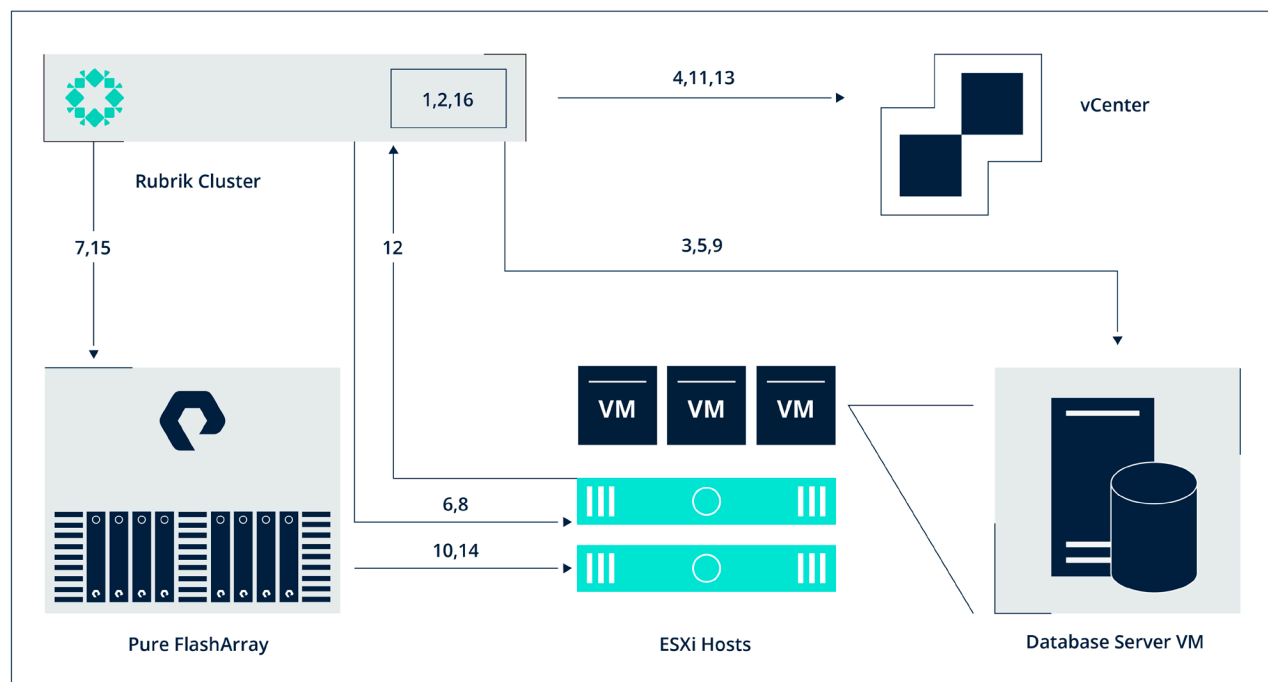
1. Backup is scheduled as per SLA or manual snapshot request.

2. Rubrik node is assigned to the backup job.

3. Guest authentication is verified.

4. Rubrik queries vCenter to identify the vSphere host containing the backup source.

5. (If configured) Pre-script is executed (database 'freeze' as example)

   ◦ If Rubrik Backup Service (RBS) is installed on the guest operating system and connected to Rubrik CDM, this will be used to call scripts within the guest. RBS runs as a service within the guest operating system. Communication between Rubrik CDM is secure and does not require credentials stored in Rubrik CDM to be used at the time of script execution.

   ◦ If RBS is not installed then VMware tools via the ESXi host will be the path for script execution within the guest, and will be authenticated at the time of execution using guest credentials stored in CDM.

6. vSphere snapshot is created.

   ◦ If RBS is installed then it will act as the VSS provider, falling back to vSphere tools as the VSS provider if RBS is not installed. Using one of these tools the operating system is quiesced, or prepared for backup by flushing unwritten data from memory to disk or any application-specific tasks required to ensure that data on disk is in a consistent state for backup. For more information about Rubrik RBS see Rubrik's blog titled Why We Built Our Own VSS Provider.

7. Pure FlashArray creates volume snapshots for any volume containing data for the source VM.

   ◦ In the event that Rubrik CDM is unable to contact the array or unable to create and mount a volume snapshot, the default behavior is to continue the backup without array integration.

8. vSphere snapshot deletion is initiated and virtual disks consolidated.

   ◦ Because the snapshot only existed for a very short amount of time and I/O has optionally been frozen via pre-execution script, stun time and subsequent impact to the running workload is minimized.

9. (If configured) Post-snap script is executed (database 'thaw' as example).

   ◦ Steps 6-8 take place in a matter of seconds, resulting in minimal snapshot delta to be merged and leaving production I/O free to resume while subsequent data protection takes place via a proxy VM (see below) and without interaction with the production guest VM. It is important to note that while snapshot deletion is initiated within seconds, full consolidation is dependent on change data created during steps 6-8. Pre- and post- script duration are dependent on script execution time. **This is where the reduced impact to production vs. normal snapshot operation is fully realized.**

10. Pure FlashArray snapshot is mapped to vSphere cluster hosts and mounted as vSphere datastore(s) if data spans multiple volumes.

    ◦ Cluster storage is re-scanned during datastore create/remove operations.

11. Proxy VM with random naming prefix is created and virtual disks from snapshot volume(s) are attached.

    ◦ The proxy VM is a temporary shell VM that is created with the express purpose of attaching the virtual disks from the snapshotted volume(s) via VMware vStorage APIs for Data Protection (VADP). It remains powered off and thus consumes no CPU or memory resources on the ESXi host/cluster.
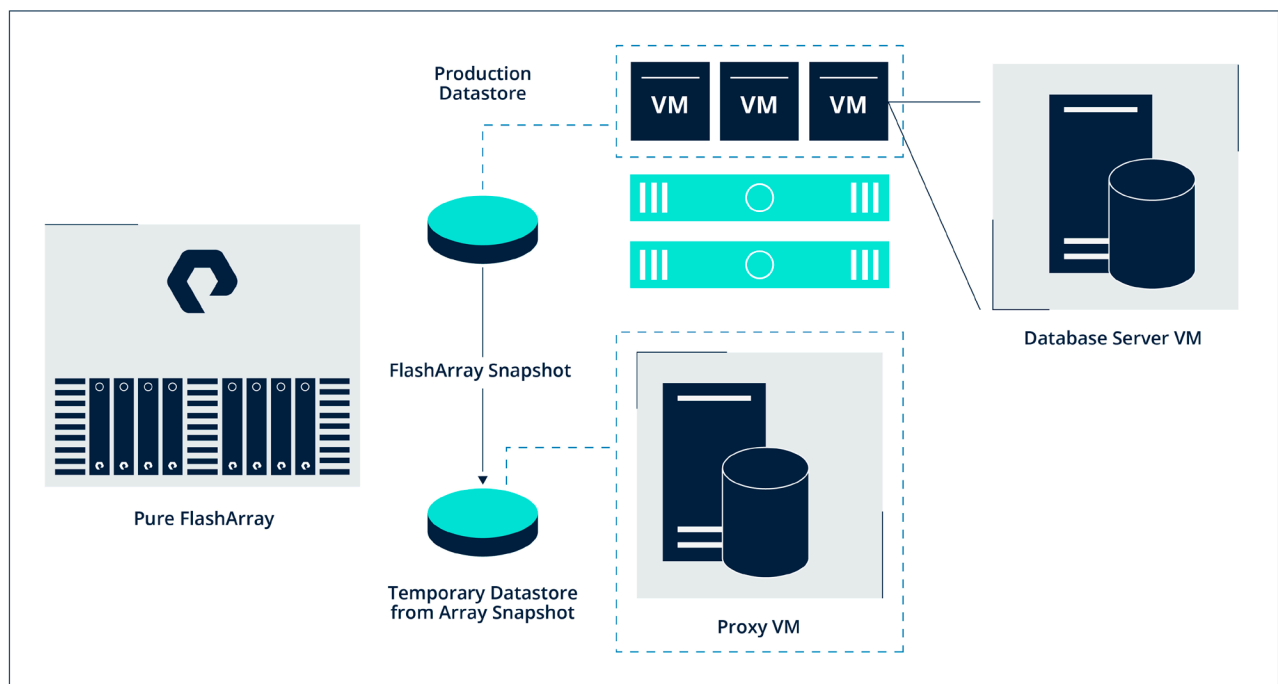
12. Rubrik ingests backup data via proxy VM.

    ◦ Since the snapshot was taken on the production VM, CBT (change block tracking) is preserved and all backups after the first are appropriately incremental.

13. Once backup data is collected, proxy VM is deleted from vSphere.

    ◦ Because a proxy VM is being used, there is no VM stun due to disk consolidation affecting the production workload at completion of backup.

14. Pure FlashArray snapshot volume datastore is unmounted from vSphere.

15. Pure FlashArray snapshot volume is removed from the array.

16. Backup workflow is completed.

SNAPSHOT STORAGE AND RECOVERY

Once data has been protected with this method it functions as any other Rubrik-protected virtual machine snapshot. Data will be indexed and can be replicated, archived, and restored using any of Rubrik's standard processes.

LIMITATIONS

- Continuous Data Protection (CDP) is not compatible with this workflow.
- Guests stored on Virtual Volumes (vVols) are not compatible with this workflow.

CONFIGURATION

Prerequisites:

- SLA Domain created with desired configuration.
- Solution-specific pre/post backup scripts configured on guest (Epic database freeze/thaw scripts, as an example).
- VM guest authentication is configured.
- VMs to be protected are stored on Pure FlashArray.
- Hostname/IP of the vCenter containing VMs to be included.
- Username and password of vCenter account with appropriate permissions for Rubrik backup operations[3]
- Hostname/IP of the Pure FlashArray
- Username and password of Pure FlashArray account with 'storage admin' access privileges

1. Add vCenter server to Rubrik cluster. Note - The most up to date installation instructions can be found in the *Adding vCenter Server connection information* section in the Rubrik CDM User Guide.

    a. Log in to the Rubrik CDM web UI.

    b. Click the gear icon.

    c. Click **vCenter Servers**.
       The vCenter Servers page appears.

    d. Click the blue **+** icon.
       The Add vCenter dialog box appears.

    e. In **vCenter IP**, type the resolvable hostname or IP address of the vCenter Server.
       For an IPv6 address, enclose the address in square brackets. For example:
       `[fd9d:22d3:cd28:7257::1]`

    f. In **vCenter Username**, type the username assigned to the Rubrik cluster.

    g. In **vCenter Password**, type the password assigned to the Rubrik cluster.

---

3 Minimum vCenter Server privileges for Rubrik CDM

h. Optional: Turn on the automatic linking feature by clicking the **Automatically link discovered virtual machines** checkbox.

i. Click **Advanced Setting** to add a Certificate Authority (CA) certificate for TLS validation.
The dialog box expands to show the Trusted Root Certificate field.

j. Paste the text of the trusted CA root certificate for the vCenter into the Trusted Root Certificate field.

k. Click **Add**.

2. Add Pure FlashArray to Rubrik cluster. Note - The most up to date installation instructions can be found in the Manage storage arrays section in the Rubrik CDM User Guide.

   a. Log in to the Rubrik CDM web UI.

   b. Click the gear icon on the top bar of the web UI.

   c. Click **Storage Arrays**.

   d. Click the blue **+** icon.

   e. In **Array Type**, select **Pure Storage**.

   f. In **Hostname**, type the IPv4 address or resolvable hostname of the storage array.

   g. In **Username**, type the username for an account with 'storage admin' privileges on the storage array.

   h. In **Password**, type the password for the account.

   i. Click **Add**.

3. (Optional) Install the Rubrik Backup Service (RBS) on the Application Guest

   a. Refer to the **Rubrik Backup Service** section of the Rubrik CDM User Guide for installation instructions for your operating system.

4. Enable Array Integration for virtual machines to be backed up using this methodology.

   a. Log in to the Rubrik CDM web UI.

   b. Browse to the VM to be backed up.

   c. Click the ellipsis icon ⋯ near the top right of the VM overview UI.

   d. Click **Enable Array Integration**. Note - this option will only be available for guests on Pure FlashArray storage.

5. Configure Pre/Post backup scripts.

   a. Log in to the Rubrik CDM web UI.

   b. Browse to the VM to be backed up.

   c. Click the ellipsis icon ⋯ near the top right of the VM overview UI.

   d. Click **Configure Pre/Post Scripts**.

   e. (Optional) In **Pre-Backup Script Path**, type the full path for the Pre-Backup Script.
   The full path is relative to the root of the guest OS file system. Please note that for Windows virtual machines the script must be in batch/cmd format and can call other scripts.

   f. (Optional) Select **Cancel Backup if Pre-Backup Scripts Fails**.
   Any script exit status other than 0 is considered a script failure. When this box is selected, the Rubrik CDM displays a notification of the script failure and the value of the exit status in the Activity Log.

   g. (Required when available) In **Timeout**, type an integer value.
   The value represents the number of seconds before the Rubrik cluster terminates the Pre-Backup Script because the script cannot be completed.

h. (Optional) In **Post-Snap Script Path**, type the full path for the Post-Snap Script.
    The full path is relative to the root of the guest OS file system. Please note that for Windows virtual machines the script must be in batch/cmd format and can call other scripts.

i. (Required when available) In **Timeout**, type an integer value.
    The value represents the number of seconds before the Rubrik cluster terminates the Post-Snap Script because the script cannot be completed.

j. (Optional) In **Post-Backup Script Path**, type the full path for the Post-Backup Script.
    The full path is relative to the root of the guest OS file system.

k. (Required when available) In **Timeout**, type an integer value.
    The value represents the number of seconds before the Rubrik cluster terminates the Post-Backup Script because the script cannot be completed.

l. Click **Apply**.

6. Assign VM to desired SLA.

CONCLUSION

By utilizing Rubrik's integration with Pure FlashArray, specific workloads with especially high sensitivity to latency can be protected without the impact that can result from the existence or removal of a vSphere snapshot. By leveraging Rubrik-orchestrated guest, host, and array operations the data capture will be offloaded to a proxy VM and the production server can continue to run without performance impact. This allows applications such as the Caché database used by Epic Systems Electronic Health Record system to be backed up as a full VM in a way that ensures maximum performance and availability.

While this functionality serves to reduce impact to the virtual machine it is configured for, it is **not recommended as a default methodology for protection at scale**. Array integrated snapshots should be used selectively where specifically required as it does nominally increase operational overhead of protecting the vSphere environment.

Additionally, VMware made improvements to VMFS snapshot consolidation in VMFS 6, so if the environment being protected still has VMFS 5 datastores efficiencies can be realized for both backup methodologies by migrating to a VMFS 6 datastore.

SOURCES AND NOTES

1   Rubrik and FlashArray Integration: Why it matters - Cody Hosterman, Pure Storage
    https://blog.purestorage.com/purely-technical/rubrik-and-flasharray-integration-why-it-matters/

2   Overview of VSS, Application or Filesystem Consistent Snapshots.
    https://support.rubrik.com/s/article/000001619

3   Minimum vCenter Server privileges for Rubrik CDM
    https://rubrik-docs.s3-us-west-1.amazonaws.com/en-us/5.3/ug/cdm_user_guide/vcenter_privs.html

ADDITIONAL INFORMATION

Overview of virtual machine snapshots in vSphere
https://kb.vmware.com/s/article/1015180

Best practices for using VMware snapshots in the vSphere environment
https://kb.vmware.com/s/article/1025279

VERSION HISTORY

Version	Date	Summary of Changes
1.0	May 2021	Initial Release