



# **Overview of Hadoop data protection use cases**

NetApp Solutions

NetApp  
October 20, 2023

This PDF was generated from <https://docs.netapp.com/us-en/netapp-solutions/data-analytics/hdcs-sh-overview-of-hadoop-data-protection-use-cases.html> on October 20, 2023. Always check docs.netapp.com for the latest.

# Table of Contents

- Overview of Hadoop data protection use cases. . . . . 1
  - Use case 1: Backing up Hadoop data . . . . . 1
  - Use case 2: Backup and disaster recovery from the cloud to on-premises . . . . . 1
  - Use case 3: Enabling DevTest on existing Hadoop data . . . . . 1
  - Use case 4: Data protection and multicloud connectivity . . . . . 1
  - Use case 5: Accelerate analytic workloads . . . . . 1

# Overview of Hadoop data protection use cases

[Previous: Hadoop data protection and NetApp.](#)

This section provides a high-level description of the data protection use cases, which constitute the focus of this paper. The remaining sections provide more details for each use case, such as the customer problem (scenario), requirements and challenges, and solutions.

## Use case 1: Backing up Hadoop data

For this use case, the In-Place Analytics Module helped a large financial institution reduce the long backup window time from more than 24 hours to just under a few hours.

## Use case 2: Backup and disaster recovery from the cloud to on-premises

By using the data fabric powered by NetApp as building blocks, a large broadcasting company was able to fulfill its requirement of backing up cloud data into its on-premise data center depending on the different modes of data transfers, such as on demand, instantaneous, or based on the Hadoop/Spark cluster load.

## Use case 3: Enabling DevTest on existing Hadoop data

NetApp solutions helped an online music distributor to rapidly build multiple space-efficient Hadoop clusters in different branches to create reports and run daily DevTest tasks by using scheduled policies.

## Use case 4: Data protection and multicloud connectivity

A large service provider used the data fabric powered by NetApp to provide multicloud analytics to its customers from different cloud instances.

## Use case 5: Accelerate analytic workloads

One of the largest financial services and investment banks used the NetApp network-attached storage solution to reduce I/O wait time and accelerate its quantitative financial analytics platform.

[Next: Use case 1 - Backing up Hadoop data.](#)

## Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

## Trademark information

NETAPP, the NETAPP logo, and the marks listed at <http://www.netapp.com/TM> are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.