



How historical data is retained in Data Warehouse

OnCommand Insight

NetApp
September 13, 2024

Table of Contents

- How historical data is retained in Data Warehouse 1
- Data retention, ETL, and time periods 2

How historical data is retained in Data Warehouse

Data is maintained in Data Warehouse according to a schedule. As data gets older, the data record retention is reduced.

Data Warehouse retains historical data based on the data marts and granularity of the data, as shown in the following table.

Data mart	Measured object	Granularity	Retention period
Performance marts	Volumes and internal volumes	Hourly	14 days
Performance marts	Volumes and internal volumes	Daily	13 months
Performance marts	Application	Hourly	13 months
Performance marts	Host	Hourly	13 months
Performance marts	Switch performance for port	Hourly	5 weeks
Performance marts	Switch performance for host, storage, and tape	Hourly	13 months
Performance marts	Storage node	Hourly	14 days
Performance marts	Storage node	Daily	13 months
Performance marts	VM performance	Hourly	14 days
Performance marts	VM performance	Daily	13 months
Performance marts	Hypervisor performance	Hourly	14 days
Performance marts	Hypervisor performance	Daily	13 months
Performance marts	VMDK performance	Hourly	14 days
Performance marts	VMDK performance	Daily	13 months
Performance marts	Disk performance	Hourly	14 days
Performance marts	Disk performance	Daily	13 months

Capacity marts	All (except individual volumes)	Daily	13 months
Capacity marts	All (except individual volumes)	Monthly representative	14 months and beyond
Inventory marts	Individual volumes	Current state	1 day (or until next ETL)

After 13 months (which is configurable), Data Warehouse retains only one record per month instead of one record per day for capacity, performance, and resource data in the following fact tables:

- Chargeback fact table (dwh_capacity.chargeback_fact)
- File System Utilization fact table (dwh_fs_util.fs_util_fact)
- Host fact table (dwh_sa.sa_host_fact)
- Internal Volume Capacity fact table (dwh_capacity.internal_volume_capacity_fact)
- Ports fact table (dwh_ports.ports_fact)
- Qtree Capacity fact table (dwh_capacity.qtree_capacity_fact)
- Storage and Storage Pool Capacity fact table (dwh_capacity.storage_and_storage_pool_capacity_fact)
- Volume Capacity fact table (dwh_capacity.vm_capacity_fact)
- Storage Node Hourly Performance (storage_node_hourly_performance_fact) and Storage Node Daily Performance (storage_node_daily_performance_fact) fact tables

Data retention, ETL, and time periods

OnCommand Insight Data Warehouse retains data obtained from the Extract, Transform, and Load (ETL) process for different time periods based on the different data marts and time granularity of the data.

Performance Marts and hourly granularity for volumes and internal volumes

The OnCommand Insight Data Warehouse records the hourly averages, hourly maximums, and access bit for each hour of the day (24 data points) for 14 days. The access bit is a Boolean value that is true if the volume is accessed or false if the volume is not accessed during the hourly interval. All 24 data points for the preceding day are obtained during the first ETL process of the day.

You do not need to run one ETL process per hour to gather the hourly data points. Running additional ETL processes during the day does not obtain any performance information from the OnCommand Insight Servers.

Performance Marts and daily granularity for volumes and internal volumes

Each day when the ETL is processed, the daily averages for the preceding day are calculated and populated within Data Warehouse. The daily average is a summary of the 24 data points for the previous day. The performance data marts retain daily summaries for volumes and internal volumes for 13 months.

Capacity marts and daily granularity

The Capacity marts provide daily measurements for various capacity facts on a daily basis for a period of 13 months. The capacity facts in Data Warehouse are current as of the last data source acquisition for the device prior to the ETL.

Capacity marts and monthly granularity

Data Warehouse retains daily capacity data for 13 months. After the 13-month threshold is reached, the capacity data is summarized on a monthly basis. The monthly data is based on the values reflected by the date that is the month representative date.

The following table shows which monthly data is included in the monthly summary:

Date	Is Month Representative value	Allocated capacity
Jan 1	1 (True)	50 TB
Jan 2	0 (False)	52 TB
...
Jan 31	0 (False)	65 TB
Feb 1	1 (True)	65 TB

Based on the table, a monthly report would show 50 TB allocated for January and 65 TB allocated for February. All of the other capacity values for January would not be included in the monthly summary.

Inventory mart

The Inventory data mart is not historical. Each time an ETL process is run, the Inventory mart is erased and rebuilt. Therefore, any reports generated out of the Inventory mart do not reflect historical inventory configuration.

Copyright information

Copyright © 2024 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.