Benefits of active, idle, and concurrent HTTP connections

StorageGRID

NetApp
May 03, 2022

This PDF was generated from https://docs.netapp.com/us-en/storagegrid-116/s3/benefits-of-keeping-idle-http-connections-open.html on May 03, 2022. Always check docs.netapp.com for the latest.

Table of Contents

Benefits of active, idle, and concurrent HTTP connections 1
- Benefits of keeping idle HTTP connections open 1
- Benefits of active HTTP connections 1
- Benefits of concurrent HTTP connections 2
- Separation of HTTP connection pools for read and write operations 3

Benefits of active, idle, and concurrent HTTP connections

How you configure HTTP connections can impact the performance of the StorageGRID system. Configurations differ depending on whether the HTTP connection is active or idle or you have concurrent multiple connections.

You can identify the performance benefits for the following types of HTTP connections:

- Idle HTTP connections
- Active HTTP connections
- Concurrent HTTP connections

Benefits of keeping idle HTTP connections open

You should keep HTTP connections open even when client applications are idle to allow client applications to perform subsequent transactions over the open connection. Based on system measurements and integration experience, you should keep an idle HTTP connection open for a maximum of 10 minutes. StorageGRID might automatically close an HTTP connection that is kept open and idle for longer than 10 minutes.

Open and idle HTTP connections provide the following benefits:

- Reduced latency from the time that the StorageGRID system determines it has to perform an HTTP transaction to the time that the StorageGRID system can perform the transaction

  Reduced latency is the main advantage, especially for the amount of time required to establish TCP/IP and TLS connections.

- Increased data transfer rate by priming the TCP/IP slow-start algorithm with previously performed transfers
- Instantaneous notification of several classes of fault conditions that interrupt connectivity between the client application and the StorageGRID system

Determining how long to keep an idle connection open is a trade-off between the benefits of slow start that is associated with the existing connection and the ideal allocation of the connection to internal system resources.

Benefits of active HTTP connections

For connections directly to Storage Nodes or to the CLB service (deprecated) on Gateway Nodes, you should limit the duration of an active HTTP connection to a maximum of 10 minutes, even if the HTTP connection continuously performs transactions.

Determining the maximum duration that a connection should be held open is a trade-off between the benefits of connection persistence and the ideal allocation of the connection to internal system resources.

For client connections to Storage Nodes or to the CLB service, limiting active HTTP connections provides the following benefits:

- Enables optimal load balancing across the StorageGRID system.

  When using the CLB service, you should prevent long-lived TCP/IP connections to optimize load balancing across the StorageGRID system. You should configure client applications to track the duration of each HTTP connection and close the HTTP connection after a set time so that the HTTP connection can be reestablished and rebalanced.

  The CLB service balances load across the StorageGRID system at the time that a client application establishes an HTTP connection. Over time, an HTTP connection might no longer be optimal as load balancing requirements change. The system performs its best load balancing when client applications establish a separate HTTP connection for each transaction, but this negates the much more valuable gains associated with persistent connections.

  > The CLB service is deprecated.

- Allows client applications to direct HTTP transactions to LDR services that have available space.
- Allows maintenance procedures to start.

  Some maintenance procedures start only after all the in-progress HTTP connections are complete.

For client connections to the Load Balancer service, limiting the duration of open connections can be useful for allowing some maintenance procedures to start promptly. If the duration of client connections is not limited, it may take several minutes for active connections to be automatically terminated.

Benefits of concurrent HTTP connections

You should keep multiple TCP/IP connections to the StorageGRID system open to allow parallelism, which increases performance. The optimal number of parallel connections depends on a variety of factors.

Concurrent HTTP connections provide the following benefits:

- Reduced latency

  Transactions can start immediately instead of waiting for other transactions to be completed.

- Increased throughput

  The StorageGRID system can perform parallel transactions and increase aggregate transaction throughput.

Client applications should establish multiple HTTP connections. When a client application has to perform a transaction, it can select and immediately use any established connection that is not currently processing a transaction.

Each StorageGRID system's topology has different peak throughput for concurrent transactions and connections before performance begins to degrade. Peak throughput depends on factors such as computing resources, network resources, storage resources, and WAN links. The number of servers and services and the number of applications that the StorageGRID system supports are also factors.

StorageGRID systems often support multiple client applications. You should keep this in mind when you determine the maximum number of concurrent connections used by a client application. If the client application

consists of multiple software entities that each establish connections to the StorageGRID system, you should add up all the connections across the entities. You might have to adjust the maximum number of concurrent connections in the following situations:

- The StorageGRID system's topology affects the maximum number of concurrent transactions and connections that the system can support.
- Client applications that interact with the StorageGRID system over a network with limited bandwidth might have to reduce the degree of concurrency to ensure that individual transactions are completed in a reasonable time.
- When many client applications share the StorageGRID system, you might have to reduce the degree of concurrency to avoid exceeding the limits of the system.

Separation of HTTP connection pools for read and write operations

You can use separate pools of HTTP connections for read and write operations and control how much of a pool to use for each. Separate pools of HTTP connections enable you to better control transactions and balance loads.

Client applications can create loads that are retrieve-dominant (read) or store-dominant (write). With separate pools of HTTP connections for read and write transactions, you can adjust how much of each pool to dedicate for read or write transactions.

Copyright Information

Copyright © 2022 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.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

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.