■ NetApp

Conclusion

NetApp Solutions

NetApp October 19, 2021

This PDF was generated from https://docs.netapp.com/us-en/netapp-solutions/ai/ai-edge-conclusion.html on October 21, 2021. Always check docs.netapp.com for the latest.

Table of Contents

Conclusion	 	
Acknowledgments	 	
Where to find additional information	 	
Version history	 	

Conclusion

Previous: Architecture sizing options.

Al-driven automation and edge computing is a leading approach to help business organizations achieve digital transformation and maximize operational efficiency and safety. With edge computing, data is processed much faster because it does not have to travel to and from a data center. Therefore, the cost associated with sending data back and forth to data centers or the cloud is diminished. Lower latency and increased speed can be beneficial when businesses must make decisions in near-real time using Al inferencing models deployed at the edge.

NetApp storage systems deliver the same or better performance as local SSD storage and offer the following benefits to data scientists, data engineers, AI/ML developers, and business or IT decision makers:

- Effortless sharing of data between AI systems, analytics, and other critical business systems. This data sharing reduces infrastructure overhead, improves performance, and streamlines data management across the enterprise.
- Independently scalable compute and storage to minimize costs and improve resource usage.
- Streamlined development and deployment workflows using integrated Snapshot copies and clones for instantaneous and space-efficient user workspaces, integrated version control, and automated deployment.
- Enterprise-grade data protection for disaster recovery and business continuity. The NetApp and Lenovo solution presented in this document is a flexible, scale-out architecture that is ideal for enterprise-grade Al inference deployments at the edge.

Acknowledgments

- J.J. Falkanger, Sr. Manager, HPC & Al Solutions, Lenovo
- Dave Arnette, Technical Marketing Engineer, NetApp
- Joey Parnell, Tech Lead E-Series Al Solutions, NetApp
- Cody Harryman, QA Engineer, NetApp

Where to find additional information

To learn more about the information described in this document, refer to the following documents and/or websites:

NetApp AFF A-Series arrays product page

https://www.netapp.com/data-storage/aff-a-series/

NetApp ONTAP data management software—ONTAP 9 information library

http://mysupport.netapp.com/documentation/productlibrary/index.html?productID=62286

• TR-4727: NetApp EF-Series Introduction

https://www.netapp.com/pdf.html?item=/media/17179-tr4727pdf.pdf

NetApp E-Series SANtricity Software Datasheet

https://www.netapp.com/pdf.html?item=/media/19775-ds-3171-66862.pdf

• NetApp Persistent Storage for Containers—NetApp Trident

https://netapp.io/persistent-storage-provisioner-for-kubernetes/

- MLPerf
 - https://mlcommons.org/en/
 - http://www.image-net.org/
 - https://mlcommons.org/en/news/mlperf-inference-v11/
- NetApp Cloud Sync

https://docs.netapp.com/us-en/occm/concept_cloud_sync.html#how-cloud-sync-works

TensorFlow benchmark

https://github.com/tensorflow/benchmarks

Lenovo ThinkSystem SE350 Edge Server

https://lenovopress.com/lp1168

Lenovo ThinkSystem DM5100F Unified Flash Storage Array

 $https://lenovopress.com/lp1365-thinksystem-dm5100f-unified-flash-storage-array \cite{Continuous} \ci$

Version history

Version	Date	Document version history
Version 1.0	March 2021	Initial release
Version 2.0	October 2021	Updated with EF and MLPerf Inference v1.1

Copyright Information

Copyright © 2021 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.