

NVIDIA DGX OS SERVER VERSION 2.1.3

DA-08260-213_v01 | March 2018

Release Notes and Update Guide

TABLE OF CONTENTS

NVIDIA DGX OS Server Release Notes for Version 2.1.3	3
Update Advisement	. 3
About Release 2.1	. 4
Highlights and Changes Since Version 2.1.2	. 4
Known Issues	
DGX OS Server Software Content	. 6
Version Reference	. 6
DGX OS Server Release 2.1.3 Update Guide	7
Connecting to the DGX-1 Console	
Performing the Over-the-Network Update	. 8
Performing the Update	
Verifying the Update 1	11
Recovering from an Interrupted Update 1	13
Resolving Version Incompatibilities on Systems with nvidia-docker 1.0.0 1	14
Appendix A. Third Party License Notice1	5

NVIDIA DGX OS SERVER RELEASE NOTES FOR VERSION 2.1.3

This document describes version 2.1.3 of the NVIDIA® DGXTM OS Server Release 2.1 software and update package for performing the over-the-network update.

This software update is provided "over the network", and requires an internet connection and ability to access the NVIDIA public repository.

If your DGX-1 is not connected to a network with internet access, refer to the DGX-1 User Guide for instructions on how to install the software on air-gapped systems.

UPDATE ADVISEMENT

DGX OS Server software

NVIDIA recommends updating to the latest DGX OS Server Release 3.1. See the DGX OS Server 3.1.6 Release Notes for more information.

All DGX-1 systems with DGX OS Server Release 1.x or 2.x should be updated to DGX OS Server version 2.1.3 as it improves system reliability and I/O performance, and includes security fixes.

Customers who want to use NVIDIA Docker Containers Release version 17.09 and later *must* upgrade their DGX-1 systems to version 2.1.1 or later to ensure compatibility with the DGX Optimized Frameworks.

NVIDIA Docker Containers

In conjunction with the DGX OS Server 2.1.3 update, customers should update their NVIDIA Docker containers to the latest container release¹.

Ubuntu Security Updates

Customers are responsible for keeping the DGX-1 up-to-date with the latest Ubuntu security updates using the 'apt-get upgrade' procedure. See the <u>Ubuntu Wiki</u> <u>Upgrades</u> web page for more information.

ABOUT RELEASE 2.1

The following are the primary features of the DGX OS Server Release 2.1:

- Supports DGX-1 using NVIDIA Pascal GPUs
- ▶ Ubuntu Server 14.04 LTS
- ▶ NVIDIA GPU Driver Release 384

HIGHLIGHTS AND CHANGES SINCE VERSION 2.1.2

- NVIDIA GPU Driver Version 384.125
 Includes security updates for driver components.
- Added nvsysinfo utility.
- Upgraded the gcc toolchain with the latest security fixes.

¹ See the *NVIDIA Deep Learning Frameworks* documentation website

^{(&}lt;u>http://docs.nvidia.com/deeplearning/dgx/index.htm</u>) for information on the latest container releases as well as instructions for how to access them.

KNOWN ISSUES

- Network Performance Drop
- ▶ Software Power Cap Not Reported Correctly by nvidia-smi
- ▶ Ipmitool cannot Manipulate FRU Data

Network Performance Drop

Issue

A bug in the Ubuntu kernel 4.4.0-116 results in slower network performance when running server-side UDP workloads.

Details

Depending on when the over-the-network update is performed, the update may incorporate the affected kernel version (4.4.0-116). This may be resolved in a later kernel version, at which point an over-the-network update of the DGX OS software will incorporate the fix.

Software Power Cap Not Reported Correctly by nvidia-smi

Issue

On DGX-1 systems with Pascal GPUs, nvidia-smi does not report Software Power Cap as "Active" when clocks are throttled by power draw.

Workaround

This issue is with nvidia-smi reporting and not with the actual functionality. This will be fixed in a future release.

Ipmitool cannot Manipulate FRU Data

Issue

The installed *ipmitool* cannot manipulate FRU data.

Workaround

If you need to access FRU information, go to the NVIDIA Enterprise Support site (<u>https://nvid.nvidia.com/enterpriselogin</u>) and refer to the Knowledge Base article *DGX-1*

Asset Tag Programming for instructions on using the *ipmi fru* command from the *freeipmi* package.

DGX OS SERVER SOFTWARE CONTENT

The following table provides version information for software included in the DGX OS Server update.

Component	Version	Notes
DGX-OS Server Software	2.1.3	Software versions included in this DGX-1 base OS update package
GPU Driver	384.125	
Ubuntu Server	14.04 LTS	

VERSION REFERENCE

The following table shows the firmware and BIOS versions for the DGX-1 hardware at the time of the update release, and tested versions of Docker and the Docker Engine Utility for NVIDIA GPUs (nvidia-docker).

Component	Version	Notes
DGX-1 BMC	3.12.30	Released versions for DGX-1 hardware (Pascal) at the time of this software release. Information provided for reference purposes.
DGX-1 SBIOS	S2W_3A04	
DGX-1 VBIOS	86.00.41.00.05	
Docker	1.12.6	Tested versions for running NVIDIA Docker containers. Docker software not included. If the software is installed prior to updating the DGX Server OS, then Docker will be updated to the latest stable Docker release.
nvidia-docker	1.0.1	

DGX OS SERVER RELEASE 2.1.3 UPDATE GUIDE

These instructions explain how to update the DGX OS Server software through an internet connection to the NVIDIA public repository, using the DGX-1 over-the-network upgrade script (dgx1-ota-update). The script updates a DGX-1 system image to the latest QA'd versions of the entire DGX-1 software stack, including the drivers.

Perform the updates using commands on the DGX-1 console.

CONNECTING TO THE DGX-1 CONSOLE

Connect to the DGX-1 console using either a direct connection or a remote connection through the BMC.

NOTE: SSH can be used to perform the update. However, if the Ethernet port is configured for DHCP, there is the potential that the IP address can change after the DGX-1 is rebooted during the update, resulting in loss of connection. If this happens, connect using either a direct connection or through the BMC to continue the update process.

Direct Connection

- 1. Connect a display to the VGA connector and a keyboard to any one of the USB ports.
- 2. Power on the DGX-1

Remote Connection through the BMC

This method requires that you have the BMC login credentials.

- 1. Make sure you have connected the IPMI port on the DGX-1 to your LAN.
- 2. Open a Java-enabled browser within your LAN and go to http://<IPMI IP Address>/. Make sure popups are allowed for the BMC address.
- 3. Log in.

If the administrator has not manually created a password, then the username that was created during the initial DGX-1 setup is used for both the BMC username and BMC password.

- 4. From the top menu, click **Remote Control** and then select **Console Redirection**.
- 5. Click **Java Console** to open the popup window.
- 6. If necessary, power on the DGX-1 using the power button icon on the upper right corner of the BMC window.

PERFORMING THE OVER-THE-NETWORK UPDATE

NOTES:

The DGX-1 will automatically reboot during the update process to complete the upgrade of the Linux kernel, CUDA driver, and Mellanox firmware. See the section <u>Recovering from an Interrupted Update</u> in the event the script is interrupted for any reason.

If you have already installed Docker and nvidia-docker, then be sure to perform the steps in the section <u>Resolving Version Incompatibilities on Systems with nvidia-docker</u> <u>1.0.0.</u>

Performing the Update

1. Verify that networking is enabled so that you can access the DGX-1 public repository.

For example:

```
$ ping www.google.com
```

You may require alternate methods of verifying, if your network is not configured for ping.

2. Run the package manager.

```
$ sudo apt-get update
```

3. Install CURL, if not already installed.

\$ sudo apt-get install curl

4. Verify that the DGX-1 public repository is configured.

\$ dpkg -1 dgx1-repo-ubuntu1404

Expected output (last line):

```
ii dgx1-repo-ubuntu1404 1.1-1 amd64 dgx1 repository configuration files
```

If this output appears, then proceed directly to step 5.

If the **dgx1-repo-ubuntu1404** package is not present, then install the package manually as follows:

a) Download the debian repo using *wget*.

```
$ sudo wget
http://international.download.nvidia.com/dgx1/repos/pool/multiverse/
d/dgx1-repo/dgx1-repo-ubuntu1404_1.1-1_amd64.deb -0 /tmp/dgx1-repo-
ubuntu1404_1.1-1_amd64.deb
```

b) Install the package.

```
$ sudo dpkg -i /tmp/dgx1-repo-ubuntu1404_1.1-1_amd64.deb
```

c) Verify that the package installed.

```
$ dpkg -1 dgx1-repo-ubuntu1404
```

Expected output:

```
ii dgxl-repo-ubuntul404 1.1-1 amd64 dgxl repository configuration files
```

5. Get the new package list.

```
$ sudo apt-get update
```

6. Confirm update script 2.1.3 is available.

\$ apt-cache policy dgx1-ota-update

Expected output (first few lines):

```
dgx1-ota-update:
  Installed: (none) <or existing DGX OS Server version>
  Candidate: 2.1.3
```

7. Install the update script.

\$ sudo apt-get -y --force-yes install dgx1-ota-update

8. Verify that update meta package version 2.1.3 is available.

```
$ apt-cache policy dgx1-ota-update-meta
```

Expected output: dgx1-ota-update-meta: Candidate: 2.1.3 If you do not get this output, repeat the previous step (install the update script).

9. Run the update script.

The script automatically performs several steps:

- Removes packages (nvidia-361, ar-mgr, isert-dkms, etc).
- Installs the update meta package (dgx1-ota-update-meta).
- Reboots the DGX-1 upon completion.

To run the script, enter the following:

```
$ sudo /usr/bin/dgx1-ota-update.sh
```

Enter y at the prompt:

Upgrade NVIDIA DGX1 from 2.0.x to 2.1.3? (y/NO)-:y

• If you are updating from version 1.0, you may be presented with a configuration choice such as the following:

Setting up dgx-limits (1.0-1) ...

Configuration file '/etc/security/limits.d/dgx1- limits.conf'

==> File on system created by you or by a script. ==> File also in package provided by package maintainer. What would you like to do about it ? Your options are: Y or I : install the package maintainer's version N or O : keep your currently-installed version D : show the differences between the versions Z : start a shell to examine the situation The default action is to keep your current version. **** dgx1-limits.conf (Y/I/N/O/D/Z) [default=N] ?

Press **Y** to all such messages to "install the package maintainer's version".

• If you see the following message:

```
A new version of configuration file /etc/default/grub is available, but the version installed currently has been locally modified.
```

What do you want to do about modified configuration file grub? install the package maintainer's version keep the local version currently installed show the differences between the versions show a side-by-side difference between the versions show a 3-way difference between available versions do a 3-way merge between available versions start a new shell to examine the situation

Select install the package maintainer's version.

- You may see the following error message during DKMS configuration for the nv_peer_mem module while running the script: DKMS: install completed. modprobe: ERROR: could not insert 'nv_peer_mem': Exec format error The next section provides instructions for remedying this error.
- 10. When completed, press y at the prompt to reboot the system.Reboot now to finalize 2.1.3? (y/NO) : y
- 11. Wait for the system to come back up, then proceed to the next section.

Verifying the Update

1. Confirm that the Linux kernel version is 4.4.0-116 or later.

\$ uname -a

Expected output:

Linux jws-1 4.4.0-116-generic ...

2. Confirm GPU driver version is 384.125.

\$ nvidia-smi

Expected output (first line)

NVIDIA-SMI 384.125 Driver Version: 384.125

3. Make sure the nvidia-peer-memory module is installed.

\$ lsmod | grep nv_peer_mem

Expected output:

```
nv_peer_mem 16384 0
nvidia 11911168 30 nv_peer_mem,nvidia_modeset,nvidia_uvm
ib_core 143360 13
rdma_cm,ib_cm,ib_sa,iw_cm,nv_peer_mem,mlx4_ib,mlx5_ib,ib_mad,ib_ucm
,ib_umad,ib_uverbs,rdma_ucm,ib_ipoib
```

• If the expected output appears, then skip to step 4.

- If there is no output, then build and install the nvidia-peer-memory module as follows:
- a) Get and install the module.

```
$ sudo apt-get install --reinstall mlnx-ofed-kernel-dkms nvidia-
peer-memory-dkms
```

Expected output:

```
DKMS: install completed.
Processing triggers for initramfs-tools (0.103ubuntu4.2) ...
update-initramfs: Generating /boot/initrd.img-4.4.0-64-generic
```

b) Add the module to the Linux kernel.

```
$ sudo modprobe nv_peer_mem
```

There is no expected output for this command.

- c) Repeat step 3 to verify that the nvidia-peer-memory module is installed.
- 4. If your network is configured for DHCP, then make sure that dynamic DNS updates are enabled.

Check whether /etc/resolv.conf is a link to /run/resolvconf/resolv.conf.

```
$ ls -l /etc/resolv.conf
```

Expected output:

```
lrwxrwxrwx 1 root root 29 Dec 1 21:19 /etc/resolv.conf ->
../run/resolvconf/resolv.conf
```

If the expected output appears, then you have completed the over-the-network update and can stop here, otherwise enable dynamic DNS updates as follows:

a) Launch the *Resolvconf Reconfigure* package.

\$ sudo dpkg-reconfigure resolvconf

The Configuring resolvconf screen appears.

- b) Select **<Yes>** when asked whether to prepare /etc/resolv.conf for dynamic updates.
- c) Select **<No>** when asked whether to append original file to dynamic file.
- d) Select **<OK>** at the *Reboot recommended* screen.

You do not need to reboot.

You are returned to the command line.

If you are connected to the DGX-1 using SSH, then the next two steps will usually cause you to lose your connection. Therefore, you should perform the

steps using a direct display and keyboard connection to the DGX-1. Alternatively, you can skip the following steps and reboot the system instead.

e) Bring down the interface, where <network interface> is em1 or em2, whichever you have set up as your primary network interface:

```
$ sudo ifdown <network interface>
```

Expected output:

ifdown: interface <network interface> not configured

f) Bring up the interface, where <network interface> is em1 or em2, whichever you have set up as your primary network interface:

```
$ sudo ifup <network interface>
```

Expected output (last line):

bound to <IP address> -- renewal in ...

g) Repeat step 4 to confirm that /etc/resolv.conf is a link to /run/resolvconf/resolv.conf.

RECOVERING FROM AN INTERRUPTED UPDATE

If the script is interrupted during the update, such as from a loss of power or loss of network connection, you can attempt to recover as follows:

- 1. Restore power or restore the network connection, whichever caused the interruption.
 - If the system encounters a kernel panic after you restore power and reboot the DGX-1, you will not be able to perform the over-the-network update. You will need to re-image the DGX-1 with the latest 2.x image and then update the software.
 - If you are successfully returned to the Linux command line, continue the following steps.
- 2. Reset the script.

```
$ sudo dpkg --configure -a
```

3. Continue following the instructions from step 9 in the <u>Performing the Over-the-</u> <u>Network Update</u> instructions.

RESOLVING VERSION INCOMPATIBILITIES ON SYSTEMS WITH NVIDIA-DOCKER 1.0.0

As part of the upgrade of the DGX Software to v2.1, a new version of Docker (dockerengine) is automatically installed. If your system was already installed with the nvidiadocker package 1.0.0, it will not be compatible with the updated docker-engine.

To resolve the incompatibility, re-install docker-engine 1.12.6 by entering the following, including entering **y** at the prompt to continue.

```
$ sudo apt-get install --reinstall docker-engine=1.12.6-0~ubuntu-trusty
Reading package lists... Done
Building dependency tree
Reading state information... Done
Recommended packages:
  aufs-tools cgroupfs-mount cgroup-lite git apparmor
The following packages will be DOWNGRADED:
  docker-engine
0 upgraded, 0 newly installed, 1 downgraded, 0 to remove and 0 not
upgraded.
Need to get 19.3 MB of archives.
After this operation, 9,383 kB of additional disk space will be used.
Do you want to continue? [Y/n] y
Fetched 19.3 MB in 1s (17.7 MB/s)
dpkg: warning: downgrading docker-engine from 17.05.0~ce-0~ubuntu-
trusty to 1.12.6-0~ubuntu-trusty
(Reading database ... 102380 files and directories currently
installed.)
Preparing to unpack .../docker-engine_1.12.6-0~ubuntu-trusty_amd64.deb
. . .
docker stop/waiting
Unpacking docker-engine (1.12.6-0~ubuntu-trusty) over (17.05.0~ce-
0~ubuntu-trusty) ...
Processing triggers for man-db (2.6.7.1-lubuntul) ...
Processing triggers for ureadahead (0.100.0-16) ...
Setting up docker-engine (1.12.6-0~ubuntu-trusty) ...
Installing new version of config file /etc/init.d/docker ...
Installing new version of config file /etc/default/docker ...
Installing new version of config file /etc/bash_completion.d/docker ...
Installing new version of config file /etc/init/docker.conf ...
docker start/running, process 5138
Processing triggers for ureadahead (0.100.0-16) ...
$
```

APPENDIX A. THIRD PARTY LICENSE NOTICE

This NVIDIA product contains third party software that is being made available to you under their respective open source software licenses. Some of those licenses also require specific legal information to be included in the product. This section provides such information.

libressl (OpenSSL)

LibreSSL <u>http://www.libressl.org/</u> version 2.4.1 is provided under the following terms:

```
* Copyright (c) 1998-2011 The OpenSSL Project. All rights reserved.
*
* Redistribution and use in source and binary forms, with or without
* modification, are permitted provided that the following conditions *
are met:
*
* 1. Redistributions of source code must retain the above copyright
* notice, this list of conditions and the following disclaimer.
*
* 2. Redistributions in binary form must reproduce the above copyright
* notice, this list of conditions and the following disclaimer in
* the documentation and/or other materials provided with the
* distribution.
*
* 3. All advertising materials mentioning features or use of this
* software must display the following acknowledgment:
* "This product includes software developed by the OpenSSL Project
* for use in the OpenSSL Toolkit. (http:// www.openssl.org/)"
*
```

* 4. The names "OpenSSL Toolkit" and "OpenSSL Project" must not be used to * endorse or promote products derived from this software without * prior written permission. For written permission, please contact * openssl-core@openssl.org. * 5. Products derived from this software may not be called "OpenSSL" * nor may "OpenSSL" appear in their names without prior written * permission of the OpenSSL Project. * 6. Redistributions of any form whatsoever must retain the following * acknowledgment: * "This product includes software developed by the OpenSSL Project * for use in the OpenSSL Toolkit (http:// www.openssl.org/)" * THIS SOFTWARE IS PROVIDED BY THE OpenSSL PROJECT ``AS IS'' AND ANY * EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE * IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR * PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE OpenSSL PROJECT OR * ITS CONTRIBUTORS 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.

Mellanox (OFED)

MLNX OFED (<u>http://www.mellanox.com/</u>) version 3.3-1.0.4.0 is provided under the following terms:

Copyright (c) 2006 Mellanox Technologies. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED.

IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS 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.

Notice

THE INFORMATION IN THIS DOCUMENT AND ALL OTHER INFORMATION CONTAINED IN NVIDIA DOCUMENTATION REFERENCED IN THIS DOCUMENT IS PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE INFORMATION FOR THE PRODUCT, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the product described in this document shall be limited in accordance with the NVIDIA terms and conditions of sale for the product.

THE NVIDIA PRODUCT DESCRIBED IN THIS DOCUMENT IS NOT FAULT TOLERANT AND IS NOT DESIGNED, MANUFACTURED OR INTENDED FOR USE IN CONNECTION WITH THE DESIGN, CONSTRUCTION, MAINTENANCE, AND/OR OPERATION OF ANY SYSTEM WHERE THE USE OR A FAILURE OF SUCH SYSTEM COULD RESULT IN A SITUATION THAT THREATENS THE SAFETY OF HUMAN LIFE OR SEVERE PHYSICAL HARM OR PROPERTY DAMAGE (INCLUDING, FOR EXAMPLE, USE IN CONNECTION WITH ANY NUCLEAR, AVIONICS, LIFE SUPPORT OR OTHER LIFE CRITICAL APPLICATION). NVIDIA EXPRESSLY DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY OF FITNESS FOR SUCH HIGH RISK USES. NVIDIA SHALL NOT BE LIABLE TO CUSTOMER OR ANY THIRD PARTY, IN WHOLE OR IN PART, FOR ANY CLAIMS OR DAMAGES ARISING FROM SUCH HIGH RISK USES.

NVIDIA makes no representation or warranty that the product described in this document will be suitable for any specified use without further testing or modification. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to ensure the product is suitable and fit for the application planned by customer and to do the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA does not accept any liability related to any default, damage, costs or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document, or (ii) customer product designs.

Other than the right for customer to use the information in this document with the product, no other license, either expressed or implied, is hereby granted by NVIDIA under this document. Reproduction of information in this document is permissible only if reproduction is approved by NVIDIA in writing, is reproduced without alteration, and is accompanied by all associated conditions, limitations, and notices.

Trademarks

NVIDIA, the NVIDIA logo, and DGX-1 are trademarks and/or registered trademarks of NVIDIA Corporation in the Unites States and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2018 NVIDIA Corporation. All rights reserved.

