IBM BladeCenter GPU Expansion Blade and GPU Expansion Blade II

Product Guide (withdrawn product)

The IBM® BladeCenter® GPU Expansion Blade and GPU Expansion Blade II provide the capability to attach next-generation graphics processing unit (GPU) technology on supported server blades. This capability is ideal for many applications written to take advantage of acceleration and visualization performance advantages that are offered in general-purpose computing on GPUs. These products offer the choice of NVIDIA Tesla M2070, M2070Q, M2075, and M2090 GPU. In addition, the IBM BladeCenter GPU Expansion Blade is stackable, allowing clients to stack up to four GPU Expansion Blades on a single-base compute blade, thereby offering a unique density advantage versus the competition. Figure 1 shows the IBM BladeCenter GPU Expansion Blade.

Figure 1. IBM BladeCenter GPU Expansion Blade (attached to a HS22 blade server)

Did you know?

General purpose GPUs are more than just high performance graphics engines. In fact, they do not connect to computer displays and they do not have video out capabilities, Instead, they perform general purpose graphical, scientific, and engineering computing and allow the offloading of compute-intensive tasks across all industries, including life sciences, fluid dynamics, finance, data analytics, atmospheric modeling, as well as large scale graphic rendering. These types of applications operate on data that is broken down easily into small chunks that can be operated on in parallel. GPUs are designed for massive parallelism. Whereas a traditional CPU might have four processing cores, the NVIDIA GPUs have more than 100 times that many cores (up to 512 cores), providing a peak double-precision capacity of 665 Gigaflops (billions of floating point operations per second) instead of approximately 50 Gigaflops. The CPU and GPU work together in this computing model. The main sequential part of the application runs on the CPU, and the computationally-intensive part runs on the GPU.

Part number information

Table 1. Ordering part number and feature code

Description	Part number	Feature code
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2070	46M6740*	5090
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2070Q	46M6772	A10R
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2075	46M6771	A25F
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2070Q	68Y7479	A246
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2075	68Y7478	A245
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2090	00D6881	A2VW

* Special bid or Configure-to-Order (CTO) only

These part numbers include the following items:

- IBM BladeCenter GPU Expansion Blade or GPU Expansion Blade II with a NVIDIA GPU installed
- Documentation CD that contains the *Installation and User's Guide*
- Warranty information and Safety flyer
- Important Notices document
- Technical Update flyer

Figure 2 shows the IBM BladeCenter GPU Expansion Blade attached to an HS22 blade server.

Figure 2. IBM BladeCenter GPU Expansion Blade (attached to a HS22 blade server)

Features

The expansion blades have the following features:

- Support for a high-performance NVIDIA adapter in each expansion blade
  The expansion blade comes with NVIDIA Tesla M2070, NVIDIA Tesla M2070Q, NVIDIA Tesla M2075, or NVIDIA Tesla M2090 adapter card. The use of these expansion blades enables the use of the specialized adapters that would not otherwise be available to BladeCenter customers.
- Ability to stack up to four expansion blades on a single base blade
  You can attach up to four expansion blades, thereby maintaining the BladeCenter density advantage while still having the ability to access the power of these high-performance GPUs as needed without the need for additional server blades and the added complexity and cost that brings. The first expansion blade connects to the server blade using the CFFh expansion slot of the server blade. The second expansion blade attaches to the CFFh connector on the first expansion blade, and so on.

  Note: You cannot attach an IBM BladeCenter PCI Express Gen 2 Expansion Blade or PCI Express Gen 2 Expansion Blade II to the same stack as GPU Expansion Blades.
- CFFh slot still available
  The CFFh expansion connector is accessible on the top-most expansion blade (see Figure 3), even with four expansion blades attached. This lets you maintain the integrated networking capabilities of the blade server when installed in a BladeCenter H or BladeCenter HT chassis.

Figure 3 shows the layout of the expansion blade.

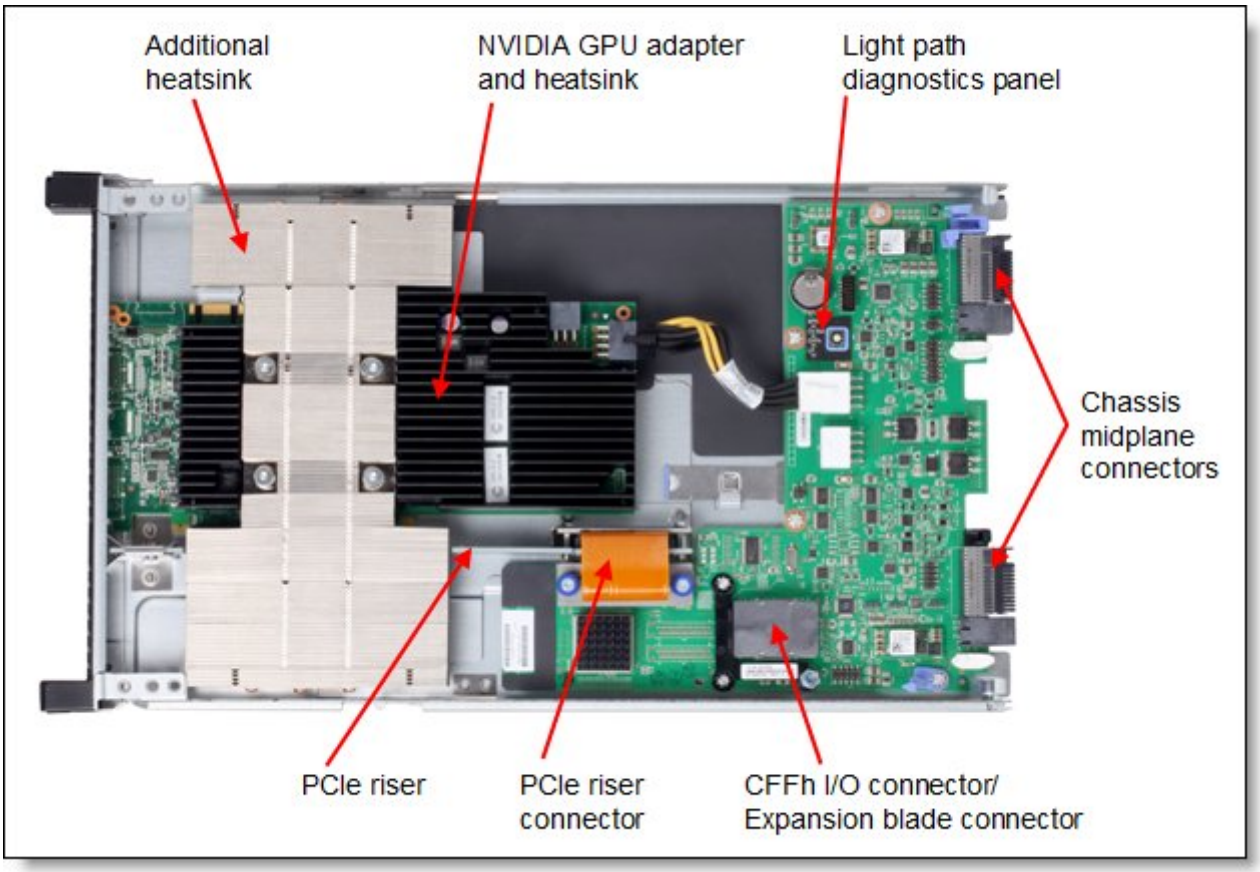

Figure 3. Layout of the expansion blade

Specifications

The expansion blade has the following specifications:

- 30 mm expansion blade, attaches to a supported server blade through the PCIe x16 CFFh connector
- Standard NVIDIA GPU computing module
- Up to four expansion blades can be attached to the base server. With four expansion blades attached, the total width of the server is 150 mm.
- The CFFh slot remains usable on the top-most expansion blade, for use in BladeCenter H and HT configurations. Supported CFFh cards are listed in Table 4.

The NVIDIA Tesla M2070, M2070Q, and M2075 cards have the following specifications:

- NVIDIA Fermi GPU engine
- 448 processor cores operating at 1.15 GHz
- 6 GB GDDR5 (graphics DDR) memory operating at 1.566 GHz
- PCIe x16 host interface
- 225W (TDP) power consumption
- Double Precision floating point performance (peak): 515 Gflops
- Single Precision floating point performance (peak): 1.03 Tflops

The NVIDIA Tesla M2070Q combines the Fermi GPU with a NVIDIA Quadro advanced visualization engine in the same GPU. NVIDIA Tesla M2075 supports Dynamic Power Scaling compared to M2070.

The NVIDIA Tesla M2090 card has the following specifications:

- NVIDIA Fermi GPU engine
- 512 processor cores operating at 1.3 GHz
- 6 GB GDDR5 (graphics DDR) memory operating at 1.85 GHz
- PCIe x16 host interface
- 225W (TDP) power consumption
- Double Precision floating point performance (peak): 665 Gflops
- Single Precision floating point performance (peak): 1.33 Tflops

Operating environment

This card is supported in the following environment:

- Temperature: 10° C to 35° C (50° F to 95° F)
- Operating power: 270 W
- Dimensions: 24.5 cm x 5.9 cm x 45.7 cm (9.7 in. x 2.3 in. x 18.0 in.)

Supported servers

The IBM BladeCenter GPU Expansion Blade and GPU Expansion Blade II are supported attached to the IBM BladeCenter servers listed in Table 2.

TTable 2. Supported servers

		HS12 (8028)	HS21 (8853)	HS21 XM (7995)	HS22 (7870)	HS22V (7871)	HS23 (7875)	HX5 (7872)	HX5 (7873)	JS12 (7998-60X)	JS23/43 (7778)	LS22 (7901)	LS42 (7902)	PS700/1/2 (8406)	PS703/4 (7891)
GPU Expansion Blade with NVIDIA Tesla M2070	46M6740*	N	N	N	Y	N	N	N	N	N	N	N	N	N	N
GPU Expansion Blade with NVIDIA Tesla M2070Q	46M6772	N	N	N	Y	N	N	N	N	N	N	N	N	N	N
GPU Expansion Blade with NVIDIA Tesla M2075	46M6771	N	N	N	Y	N	N	N	N	N	N	N	N	N	N
GPU Expansion Blade II with NVIDIA Tesla M2070Q	68Y7479	N	N	N	Y	N	Y	N	N	N	N	N	N	N	N
GPU Expansion Blade II with NVIDIA Tesla M2075	68Y7478	N	N	N	Y	N	Y	N	N	N	N	N	N	N	N
GPU Expansion Blade II with NVIDIA Tesla M2090	00D6881	N	N	N	Y	N	Y	N	N	N	N	N	N	N	N

* Special bid or Configure-to-Order (CTO) only

The expansion blade attaches to the blade server unit through the CFFh connector on the system board of the server blade. The top cover of the blade server needs to be removed and the expansion blade attaches to it, forming a double-width blade server unit. Additional expansion blades can be attached, for a total of four. A supported CFFh adapter can be attached to the top-most expansion blade if the server unit is installed in a BladeCenter H or BladeCenter HT chassis.

Supported BladeCenter chassis

Table 3 lists the supported chassis in which the IBM BladeCenter GPU Expansion Blades can be installed.

Table 3. Supported BladeCenter chassis

	BladeCenter S	BladeCenter E	BladeCenter H	BladeCenter T	BladeCenter HT
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2070	N	Y*	Y	N	Y
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2070Q	N	Y*	Y	N	Y
IBM BladeCenter GPU Expansion Blade with NVIDIA Tesla M2075	N	Y*	Y	N	Y
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2070Q	N	Y*	Y	N	Y
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2075	N	Y*	Y	N	Y
IBM BladeCenter GPU Expansion Blade II with NVIDIA Tesla M2090	N	Y*	Y	N	Y

* Supported in BladeCenter E models 8677-4xx and 3xx only; Chassis requires 2,320 W AC Power Supply options, 46M0508

Supported CFFh expansion cards

Table 4 lists the CFFh expansion cards that are supported or planned to be supported in the IBM BladeCenter GPU Expansion Blades. The use of these CFFh expansion cards requires that the expansion blade be installed in a BladeCenter H or BladeCenter HT chassis.

Table 4. Supported CFFh expansion cards

Description	Part number	Feature code	GPU Expansion Blade	GPU Expansion Blade II
2/4 Port Ethernet Expansion Card (CFFh)	44W4479	5476	Yes	Yes
2-Port 40 Gb InfiniBand Expansion Card (CFFh)	46M6001	0056	Yes	Yes
Broadcom 10 Gb Gen 2 2-port Ethernet Expansion Card (CFFh)	46M6168	0099	Yes	Yes
Broadcom 10 Gb Gen 2 4-port Ethernet Expansion Card (CFFh)	46M6164	0098	Yes	Yes
Brocade 2-port 10GbE Converged Network Adapter	81Y1650	5437	No	Yes
Emulex Virtual Fabric Adapter (CFFh)	49Y4235	5755	Yes	Yes
Emulex Virtual Fabric Adapter Advanced (CFFh)	49Y4275	2435	Yes	Yes
Emulex 10GbE Virtual Fabric Adapter II (CFFh)	90Y3550	A1XG	Yes	Yes
Emulex 10GbE Virtual Fabric Adapter Advanced II (CFFh)	90Y3566	A1XH	Yes	Yes
InfiniBand 4X DDR Expansion Card (CFFh)	43W4423	2991	No	No
Mellanox 2-port 10Gb E'net Expansion Card (CFFh)	90Y3570	A1NW	Yes	Yes
QLogic 2-port 10 Gb Converged Network Adapter (CFFh)	42C1830	3592	Yes	Yes
QLogic Ethernet and 8 Gb Fibre Channel Expansion Card (CFFh)	44X1940	5485	Yes	Yes

Consult IBM ServerProven to see the latest list of CFFh expansion cards that are supported in the Expansion Blades:
http://ibm.com/servers/eserver/serverproven/compat/us/

Operating system support

The following operating systems have been tested for compatibility with the IBM BladeCenter GPU Expansion Blade.

- Windows Server 2008 HPC Edition (64-bit)
- Windows HPC Server 2008 (64-bit)
- Windows 2008 R2 (64-bit)
- Red Hat Enterprise Linux 5 (64-bit)
- SUSE Linux Enterprise Server Edition 11 for x86_64 (64-bit)

Related publications

For more information, refer to these documents:

- IBM U.S. Announcement Letter for GPU Expansion Blade
  http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-231
- IBM U.S. Announcement Letter for GPU Expansion Blade II
  http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS112-044
- *IBM BladeCenter GPU Expansion Blade Installation and User's Guide*
  http://www.ibm.com/support/
- *IBM BladeCenter Interoperability Guide*
  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5073016
- IBM Redbooks publication *IBM BladeCenter Products and Technology*, SG24-7523
  http://www.redbooks.ibm.com/abstracts/sg247523.html

Related product families

Product families related to this document are the following:

- Blade Servers

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
8001 Development Drive
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION ”AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document, TIPS0798, was created or updated on March 6, 2012.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/TIPS0798
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/TIPS0798.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at https://www.lenovo.com/us/en/legal/copytrade/.

The following terms are trademarks of Lenovo in the United States, other countries, or both:
Lenovo®
BladeCenter®
ServerProven®

The following terms are trademarks of other companies:

Linux® is the trademark of Linus Torvalds in the U.S. and other countries.

Windows Server® and Windows® are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.