

Supported MIG Profiles

This section provides an overview of the supported profiles and possible placements of the MIG profiles on supported GPUs.

A30 MIG Profiles

The following diagram shows the profiles supported on the NVIDIA A30:

Config	GPC Slice #0	GPC Slice #1	GPC Slice #2	GPC Slice #3	OFA	NVDEC	NVJPG	P2P	GPU Direct RDMA
1	4				1	4	1	No	Supported MemBW proportional to size of the instance
2	2		2		0	2+2	0	No	
3	2		1	1	0	2+1+1	0	No	
4	1	1	2		0	1+1+2	0	No	
5	1	1	1	1	0	1+1+1+1	0	No	

Figure 10: Profiles on A30

The following table shows the supported profiles on the A30-24GB product.

Table 5: GPU Instance Profiles on A30

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.6gb	1/4	1/4	0 NVDECs /0 JPEG /0 OFA	1/4	1	4
MIG 1g.6gb+me	1/4	1/4	1 NVDEC /1 JPEG /1 OFA	1/4	1	1 (A single 1g profile can include media extensions)
MIG 2g.12gb	2/4	2/4	2 NVDECs /0 JPEG /0 OFA	2/4	2	2
MIG 2g.12gb+me	2/4	2/4	2 NVDECs /1 JPEG /1 OFA	2/4	2	1 (A single 2g profile can

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
						include media extensions)
MIG 4g.24gb	Full	4/4	4 NVDECs /1 JPEG /1 OFA	Full	4	1

Note

The `1g.6gb+me` profile is only available starting with R470 drivers.

The `2g.12gb+me` profile is only available starting with R525 drivers.

A100 MIG Profiles

The following diagram shows the profiles supported on the NVIDIA A100:

Config	GPC Slice #0	GPC Slice #1	GPC Slice #2	GPC Slice #3	GPC Slice #4	GPC Slice #5	GPC Slice #6	OFA	NVDEC	NVJPG	P2P	GPU Direct RDMA
1	7							1	5	1	No	Supported MemBW proportional to size of the instance
2	4				3			0	2+2	0	No	
3	4				2		1	0	2+1+0	0	No	
4	4				1	1	1	0	2+0+0+0	0	No	
5	3			3				0	2+2	0	No	
6	3			2		1		0	2+1+0	0	No	
7	3			1	1	1		0	2+0+0+0	0	No	
8	2		2		3			0	1+1+2	0	No	
9	2		1	1	3			0	1+0+0+2	0	No	
10	1	1	2		3			0	0+0+1+2	0	No	
11	1	1	1	1	3			0	0+0+0+0+2	0	No	
12	2		2		2		1	0	1+1+1+0	0	No	
13	2		1	1	2		1	0	1+0+0+1+0	0	No	
14	1	1	2		2		1	0	0+0+1+1+0	0	No	
15	2		1	1	1	1	1	0	1+0+0+0+0	0	No	
16	1	1	2		1	1	1	0	0+0+1+0+0+0	0	No	
17	1	1	1	1	2		1	0	0+0+0+0+1+0	0	No	
18	1	1	1	1	1	2		0	0+0+0+0+0+1	0	No	
19	1	1	1	1	1	1	1	0	0+0+0+0+0+0+0	0	No	

Figure 11: Profiles on A100

The following table shows the supported profiles on the A100-SXM4-40GB product. For A100-SXM4-80GB, the profile names will change according to the memory proportion – for example, `1g.10gb`, `1g.10gb+me`, `1g.20gb`, `2g.20gb`, `3g.40gb`, `4g.40gb`, `7g.80gb` respectively.

Table 6: GPU Instance Profiles on A100

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.5gb	1/8	1/7	0 NVDECs /0 JPEG /0 OFA	1/8	1	7
MIG 1g.5gb+me	1/8	1/7	1 NVDEC /1 JPEG /1 OFA	1/8	1	1 (A single 1g profile can include media extensions)
MIG 1g.10gb	1/8	1/7	1 NVDEC /0 JPEG /0 OFA	1/8	1	4
MIG 2g.10gb	2/8	2/7	1 NVDEC /0 JPEG /0 OFA	2/8	2	3
MIG 3g.20gb	4/8	3/7	2 NVDECs /0 JPEG /0 OFA	4/8	3	2
MIG 4g.20gb	4/8	4/7	2 NVDECs /0 JPEG /0 OFA	4/8	4	1
MIG 7g.40gb	Full	7/7	5 NVDECs /1 JPEG /1 OFA	Full	7	1

Note

The `1g.5gb+me` profile is only available starting with R470 drivers.

The `1g.10gb` profile is only available starting with R525 drivers.

H100 MIG Profiles

The following diagram shows the profiles supported on the NVIDIA H100:

Config	GPC Slice #0	GPC Slice #1	GPC Slice #2	GPC Slice #3	GPC Slice #4	GPC Slice #5	GPC Slice #6	OFA	NVDEC	NVJPG	P2P	GPU Direct RDMA
1	7							1	7	7	No	Supported MemBW proportional to size of the instance
2	4				3			0	4+3	4+3	No	
3	4				2		1	0	4+2+1	4+2+1	No	
4	4				1	1	1	0	4+1+1+1	4+1+1+1	No	
5	3			3				0	3+3	3+3	No	
6	3			2		1		0	3+2+1	3+2+1	No	
7	3			1	1	1		0	3+1+1+1	3+1+1+1	No	
8	2		2		3			0	2+2+3	2+2+3	No	
9	2		1	1	3			0	2+1+1+3	2+1+1+3	No	
10	1	1	2		3			0	1+1+2+3	1+1+2+3	No	
11	1	1	1	1	3			0	1+1+1+1+3	1+1+1+1+3	No	
12	2		2		2		1	0	2+2+2+2+1	2+2+2+2+1	No	
13	2		1	1	2		1	0	2+1+1+2+1	2+1+1+2+1	No	
14	1	1	2		2		1	0	1+1+2+2+1	1+1+2+2+1	No	
15	2		1	1	1	1	1	0	2+1+1+1+1+1	2+1+1+1+1	No	
16	1	1	2		1	1	1	0	1+1+2+1+1+1	1+1+2+1+1+1	No	
17	1	1	1	1	2		1	0	1+1+1+1+2+1	1+1+1+1+2+1	No	
18	1	1	1	1	1	2		0	1+1+1+1+1+2	1+1+1+1+1+2	No	
19	1	1	1	1	1	1	1	0	1+1+1+1+1+1+1	1+1+1+1+1+1+1	No	

Figure 12: Profiles on H100

The following table shows the supported profiles on the H100 80GB product (PCIe and SXM5).

Table 7: GPU Instance Profiles on H100

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.10gb	1/8	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	7
MIG 1g.10gb+me	1/8	1/7	1 NVDEC /1 JPEG /1 OFA	1/8	1	1 (A single 1g profile can include media extensions)
MIG 1g.20gb	1/4	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	4
MIG 2g.20gb	2/8	2/7	2 NVDECs /2 JPEG /0 OFA	2/8	2	3

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 3g.40gb	4/8	3/7	3 NVDECs /3 JPEG /0 OFA	4/8	3	2
MIG 4g.40gb	4/8	4/7	4 NVDECs /4 JPEG /0 OFA	4/8	4	1
MIG 7g.80gb	Full	7/7	7 NVDECs /7 JPEG /1 OFA	Full	8	1

The following table shows the supported profiles on the H100 94GB product (PCIe and SXM5).

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.12gb	1/8	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	7
MIG 1g.12gb+me	1/8	1/7	1 NVDEC /1 JPEG /1 OFA	1/8	1	1 (A single 1g profile can include media extensions)
MIG 1g.24gb	1/4	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	4
MIG 2g.24gb	2/8	2/7	2 NVDECs /2 JPEG /0 OFA	2/8	2	3
MIG 3g.47gb	4/8	3/7	3 NVDECs /3 JPEG /0	4/8	3	2

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
			OFA			
MIG 4g.47gb	4/8	4/7	4 NVDECs /4 JPEG /0 OFA	4/8	4	1
MIG 7g.94gb	Full	7/7	7 NVDECs /7 JPEG /1 OFA	Full	8	1

The following table shows the supported profiles on the H100 96GB product (H100 on GH200).

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.12gb	1/8	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	7
MIG 1g.12gb+me	1/8	1/7	1 NVDEC /1 JPEG /1 OFA	1/8	1	1 (A single 1g profile can include media extensions)
MIG 1g.24gb	1/4	1/7	1 NVDEC /1 JPEG /0 OFA	1/8	1	4
MIG 2g.24gb	2/8	2/7	2 NVDECs /2 JPEG /0 OFA	2/8	2	3
MIG 3g.48gb	4/8	3/7	3 NVDECs /3 JPEG /0 OFA	4/8	3	2

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 4g.48gb	4/8	4/7	4 NVDECs /4 JPEG /0 OFA	4/8	4	1
MIG 7g.96gb	Full	7/7	7 NVDECs /7 JPEG /1 OFA	Full	8	1

H200 MIG Profiles

The following diagram shows the profiles supported on the NVIDIA H200:

Config	GPC Slice #0	GPC Slice #1	GPC Slice #2	GPC Slice #3	GPC Slice #4	GPC Slice #5	GPC Slice #6	OFA	NVDEC	NVJPG	P2P	GPU Direct RDMA
1	7							1	7	7	No	Supported MemBW proportional to size of the instance
2	4				3			0	4+3	4+3	No	
3	4				2		1	0	4+2+1	4+2+1	No	
4	4				1	1	1	0	4+1+1+1	4+1+1+1	No	
5	3			3				0	3+3	3+3	No	
6	3			2		1		0	3+2+1	3+2+1	No	
7	3			1	1	1		0	3+1+1+1	3+1+1+1	No	
8	2		2		3			0	2+2+3	2+2+3	No	
9	2		1	1	3			0	2+1+1+3	2+1+1+3	No	
10	1	1	2		3			0	1+1+2+3	1+1+2+3	No	
11	1	1	1	1	3			0	1+1+1+1+3	1+1+1+1+3	No	
12	2		2		2		1	0	2+2+2+2+1	2+2+2+2+1	No	
13	2		1	1	2		1	0	2+1+1+2+1	2+1+1+2+1	No	
14	1	1	2		2		1	0	1+1+2+2+1	1+1+2+2+1	No	
15	2		1	1	1	1	1	0	2+1+1+1+1+1	2+1+1+1+1	No	
16	1	1	2		1	1	1	0	1+1+2+1+1+1	1+1+2+1+1+1	No	
17	1	1	1	1	2		1	0	1+1+1+1+2+1	1+1+1+1+2+1	No	
18	1	1	1	1	1	2		0	1+1+1+1+1+2	1+1+1+1+1+2	No	
19	1	1	1	1	1	1	1	0	1+1+1+1+1+1+1	1+1+1+1+1+1+1	No	

Figure 13: Profiles on H200

The following table shows the supported profiles on the H200 141GB product.

Table 8: GPU Instance Profiles on H200

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.18gb	1/8	1/7	1 NVDECs /1 JPEG /0 OFA	1/8	1	7

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units	L2 Cache Size	Copy Engines	Number of Instances Available
MIG 1g.18gb+me	1/8	1/7	1 NVDEC /1 JPEG /1 OFA	1/8	1	1 (A single 1g profile can include media extensions)
MIG 1g.35gb	1/4	1/7	1 NVDECs /1 JPEG /0 OFA	1/8	1	4
MIG 2g.35gb	2/8	2/7	2 NVDECs /2 JPEG /0 OFA	2/8	2	3
MIG 3g.71gb	4/8	3/7	3 NVDECs /3 JPEG /0 OFA	4/8	3	2
MIG 4g.71gb	4/8	4/7	4 NVDECs /4 JPEG /0 OFA	4/8	4	1
MIG 7g.141gb	Full	7/7	7 NVDECs /7 JPEG /1 OFA	Full	8	1

B200 MIG Profiles

The following diagram shows the profiles supported on the NVIDIA B200:

Config	Slice #0	Slice #1	Slice #2	Slice #3	Slice #4	Slice #5	Slice #6	OFA	NVDEC	NVJPEG	P2P	GPU Direct RDMA
1	7							1	7	7	No	Supported MemBW proportional to size of the instance
2	4				3			0	4+3	4+3	No	
3	4				2		1	0	4+2+1	4+2+1	No	
4	4				1	1	1	0	4+1+1+1	4+1+1+1	No	
5	3			3				0	3+3	3+3	No	
6	3			2		1		0	3+2+1	3+2+1	No	
7	3			1	1	1		0	3+1+1+1	3+1+1+1	No	
8	2		2		3			0	2+2+3	2+2+3	No	
9	2		1	1	3			0	2+1+1+3	2+1+1+3	No	
10	1	1	2		3			0	1+1+2+3	1+1+2+3	No	
11	1	1	1	1	3			0	1+1+1+1+3	1+1+1+1+3	No	
12	2		2		2		1	0	2+2+2+2+1	2+2+2+2+1	No	
13	2		1	1	2		1	0	2+1+1+2+1	2+1+1+2+1	No	
14	1	1	2		2		1	0	1+1+2+2+1	1+1+2+2+1	No	
15	2		1	1	1	1	1	0	2+1+1+1+1+1	2+1+1+1+1+1	No	
16	1	1	2		1	1	1	0	1+1+2+1+1+1	1+1+2+1+1+1	No	
17	1	1	1	1	2		1	0	1+1+1+1+2+1	1+1+1+1+2+1	No	
18	1	1	1	1	1	2		0	1+1+1+1+1+2	1+1+1+1+1+2	No	
19	1	1	1	1	1	1	1	0	1+1+1+1+1+1+1	1+1+1+1+1+1+1	No	

Figure 14: Profiles on B200

The following table shows the supported profiles on the B200 180 GB product.

Table 9: GPU Instance Profiles on B200

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units			L2 Cache Size	Copy Engines	Number of Instances
			NVDEC	JPEG	OFA			
MIG 1g.23gb	1/8	1/7	1	1	0	1/8	2	7
MIG 1g.23gb+me	1/8	1/7	1	1	1	1/8	2	1 (A single 1g profile can include media extensions)
MIG 1g.45gb	2/8	1/7	1	1	0	2/8	2	4
MIG 2g.45gb	2/8	2/7	1	1	0	2/8	3	3
MIG 3g.90gb	4/8	3/7	1	1	0	4/8	6	2

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units			L2 Cache Size	Copy Engines	Number of Instances
			NVDEC	JPEG	OFA			
MIG 4g.90gb	4/8	4/7	1	1	0	4/8	8	1
MIG 7g.180gb	Full	Full	1	1	1	Full	16	1

RTX PRO 5000 Blackwell MIG Profiles

The following diagram shows the profiles supported on the NVIDIA RTX PRO 5000 Blackwell:

Config	Slice #0 Slice #1		NVENC	NVDEC	NVJPEG	OFA	P2P	GPU Direct RDMA
1	2		3	3	1	0	No	Supported MemBW proportional to size of the instance
2	1	1	3+0	3+0	1+0	1	No	

Figure 15: Profiles on NVIDIA RTX PRO 5000 Blackwell

The following table shows the supported profiles on the RTX PRO 5000 Blackwell 48GB products:

Table 10: GPU Instance Profiles on RTX PRO 5000

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units				L2 Cache Size	Copy Engines	Number of Instances
			NVDEC	NVENC	JPEG	OFA			
MIG 1g.12gb	1/4	1/4	1	1	1	0	1/4	1	1
MIG 1g.12gb+me	1/4	1/4	1	1	1	1	1/4	1	1
MIG 1g.12gb+gfx	1/4	1/4	1	1	1	0	1/4	1	1

Profile Name	Fraction of Memory	Fraction of SMs	Hardware Units				L2 Cache Size	Copy Engines	Number of Instances
			NVDEC	NVENC	JPEG	OFA			
MIG 1g.12gb-me	1/4	1/4	0	0	0	0	1/4	1	3
MIG 2g.24gb-me	1/2	1/2	0	0	0	0	1/2	2	1
MIG 4g.48gb	Full	Full	3	3	1	1	Full	4	1
MIG 4g.48gb+gfx	Full	Full	3	3	1	1	Full	4	1

RTX PRO 6000 Blackwell MIG Profiles

The following diagram shows the profiles supported on the NVIDIA RTX PRO 6000 Blackwell Workstation Edition, Max-Q Workstation Edition, and RTX PRO 6000 Blackwell Server Edition:

Config	Slice #0	Slice #1	Slice #2	Slice #3	NVENC	NVDEC	NVJPEG	OFA	P2P	GPU Direct RDMA
1	4				4	4	4	1	No	Supported MemBW proportional to size of the instance
2	2		2		2+2	2+2	2+2	0	No	
3	2		1	1	2+1+1	2+1+1	2+1+1	0	No	
4	1	1	2		1+1+2	1+1+2	1+1+2	0	No	
5	1	1	1	1	1+1+1+1	1+1+1+1	1+1+1+1	0	No	
6	2		1	1	4+0+0	4+0+0	4+0+0	1	No	
7	1	1	1	1	4+0+0+0	4+0+0+0	4+0+0+0	1	No	

Figure 16: Profiles on NVIDIA RTX PRO 6000 Blackwell Workstation Edition and RTX PRO 6000 Blackwell Max-Q Workstation Edition

The following table shows the supported profiles on the RTX PRO 6000 Blackwell Workstation Edition, Max-Q Workstation Edition, and RTX PRO 6000 Blackwell Server Edition 96GB products:

Table 11: GPU Instance Profiles on RTX PRO 6000