Intel® Arc™ A-Series Graphics and Intel Data Center GPU Flex Series

Open-Source Programmer's Reference Manual

For the discrete GPUs code named "Alchemist" and "Arctic Sound-M"

Volume 4: Configurations

March 2023, Revision 1.0

Notices and Disclaimers

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks

Customer is responsible for safety of the overall system, including compliance with applicable safety-related requirements or standards.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document, with the sole exceptions that a) you may publish an unmodified copy and b) code included in this document is licensed subject to Zero-Clause BSD open source license (0BSD). You may create software implementations based on this document and in compliance with the foregoing that are intended to execute on the Intel product(s) referenced in this document. No rights are granted to create modifications or derivatives of this document.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

Table of Contents

Configurations

This chapter contains configurations details as described in the following sections:

- Top Level Block Diagrams
- Device Attributes
- Steppings and Device IDs

Product Mapping Table

Product Mapping Table		
Product Family	**Alchemist**	
Base Die	**512**	**128**
Status	**POR**	**POR**
Render Engine	XeHPG 8x4x16	XeHPG 2x4x16
Media Engine	XeHPM	XeHPM
Display Engine	XeHPD	XeHPD
In-Package Memory	N/A	N/A
Main Memory	Up to 256bit GDDR6 @ 18GTs	Up to 96bit GDDR6 @ 16GTs

Top Level Block Diagrams

512 [8x4x16] Block Diagram:

Device Attributes Xe_{HPG}

NOTE: This information is preliminary, and subject to change.

Product Configuration Attribution Table 512

Base Die	**512**			
Geometry Config	**4x4**	**6x4**	**7x4**	**8x4**
Compute Config	**2x8**	**2x12**	**1x16+1x12**	**4x8 || 2x16**
Status	**POR**	**POR**	**POR**	**POR**
Global Attributes				
Slice count	4	6	7	8
Dual-Subslice (DSS) Count	16	24	28	32
EU/DSS	16	16	16	16
EU count (total)	256	384	448	512
Threads / EU	8	8	8	8
Thread Count (Total)	2048	3072	3584	4096
FLOPs/Clk - Half Precision, MAD (peak)	8192	12288	14336	16384
FLOPs/Clk - Single Precision, MAD (peak)	4096	6144	7168	8192
FLOPs/Clk - Double Precision, MAD (peak)	N/A	N/A	N/A	N/A
IOPs/Clk - Int8 DP4AS (peak)	65536	98304	114688	131072
FLOPs/Clk - FP16 DP4AS (peak)	32768	49152	57344	65536
GTI	4	8	8	8
GTI bandwidth (bytes/unslice-clk)	r: 256	r: 512	r: 512	r: 512
	w: 256	w: 512		w: 512
Graphics Virtual Address Range	48 bit	48 bit	48 bit	48 bit
Graphics Physical Address Range (DM) Max Per Device Memory = 64 GB	36 bit	36 bit	36 bit	36 bit
Graphics Physical Address Range (System Memory)	46 bit	46 bit	46 bit	46 bit
Caches & Dedicated Memories				
Device Cache, total size (bytes)[1]	8192k	16384k	16384k	16384k
Device Cache, bank count[1]	16	32	32	32
Device Cache, bandwidth (bytes/clk)	16x 64 R|W	32x 64 R|W	32x 64 R|W	32x 64 R|W

Base Die	512			
URB Size (bytes)	4x 480k	8x 480k	8x 480k	8x 480k
L1 Cache, total size (bytes)	16x 192k	24x 192k	28x 192k	32x 192k
L1 Cache, bandwidth (bytes/clk)	16x (256R|128W)	24x (256R|128W)	28x (256R|128W)	32x (256R|128W)
SLM Size (bytes)[1]	16x 128k	24x 128k	28x 128k	32x 128k
SLM, bandwidth (bytes/clk)	16x (256R|128W)	24x (256R|128W)	28x (256R|128W)	32x (256R|128W)
Instruction Cache (instances, bytes ea.)	16x 96k	24x 96k	28x 96k	32x 96k
Color Cache (RCC, bytes)	4x 16k	8x 16k	8x 16k	8x 16k
MSC Cache (MSC, bytes)	8x 16k	12x 16k	14x 16k	16x 16k
HiZ Cache (HZC, bytes)	8x 4k	12x 4k	14x 4k	16x 4k
Z Cache (RCZ, bytes)	8x 16k	12x 16k	14x 16k	16x 16k
Stencil Cache (STC, bytes)	8x 6k	12x 6k	14x 6k	16x 6k
Instruction Issue Rates				
FMAD, SP (ops/EU/clk)	8	8	8	8
FMUL, SP (ops/EU/clk)	8	8	8	8
FADD, SP (ops/EU/clk)	8	8	8	8
CMP, SP (ops/EU/clk)	8	8	8	8
INV, SP (ops/EU/clk)	2	2	2	2
SQRT, SP (ops/EU/clk)	2	2	2	2
RSQRT, SP (ops/EU/clk)	2	2	2	2
LOG, SP (ops/EU/clk)	2	2	2	2
EXP, SP (ops/EU/clk)	2	2	2	2
TRIG, SP (ops/EU/clk)	2	2	2	2
Load/Store				
Data Ports (HDC)	16	24	28	32
Device Cache Load/Store (bytes/clk)	1024	1536	1792	2048
SLM Load/Store (bytes/clk)	4096/2048	6144/3072	7168/3584	8192/4096
Atomic, Local 32b (bytes/clk)	16	24	28	32
Atomic, Global 32b (bytes/clk)	256	384	448	512
3D Attributes				
Geometry pipes	4	6	7	8
Samplers (3D)	16	24	28	32
2D Texel Rate, point, 32b (tex/clk)	128	192	224	256
2D Texel Rate, point, 64b (tex/clk)	128	192	224	256
2D Texel Rate, point, 128b (tex/clk)	128	192	224	256

Base Die	512			
2D Texel Rate, bilinear, 32b (tex/clk)	128	192	224	256
2D Texel Rate, bilinear, 64b (tex/clk)	128	192	224	256
2D Texel Rate, bilinear, 128b (tex/clk)	32	48	56	64
2D Texel Rate, trilinear, 32b (tex/clk)	64	96	112	128
2D Texel Rate, trilinear, 64b (tex/clk)	64	96	112	128
2D Texel Rate, trilinear, 128b (tex/clk)	16	24	28	32
2D Texel Sample Rate, aniso 2x (MIP nearest), 32b (tex/clk)	128	192	224	256
2D Texel Sample Rate, aniso 4x (MIP nearest), 32b (tex/clk)	64	96	112	128
2D Texel Sample Rate, ansio 8x (MIP nearest), 32b (tex/clk)	32	48	56	64
2D Texel Sample Rate, ansio 16x (MIP nearest), 32b (tex/clk)	16	24	28	32
3D Texel Sample Rate, point, 32b (tex/clk)	128	192	224	256
3D Texel Sample Rate, point, 64b (tex/clk)	128	192	224	256
3D Texel Sample Rate, point, 128b (tex/clk)	64	96	112	128
3D Texel Sample Rate, bilinear, 32b (tex/clk)	64	96	112	128
3D Texel Sample Rate, bilinear, 64b (tex/clk)	64	96	112	128
3D Texel Sample Rate, bilinear, 128b (tex/clk)	16	24	28	32
HiZ Rate, (ppc)	8x 16x8	12x 16x8	14x 16x8	16x 16x8
IZ Rate, (ppc)	8x 4x4	12x 4x4	14x 4x4	16x 4x4
Stencil Rate (ppc)	8x 16x8	12x 16x8	14x 16x8	16x 16x8
Pixel Rate, fill/blend, 32bpp (pix/clk)	64	96	112	128

Notes:

For all SKUs, engine reset timeout is 5ms

Product Configuration Attribution Table – 128

Base Die	**128**	
Geometry Config	**2x3**	**2x4**
Compute Config	**1x6**	**1x8**
Status	**POR**	**POR**
Global Attributes		
Slice count	2	2
Dual-Subslice (DSS) Count	6	8
EU/DSS	16	16
EU count (total)	96	128
Threads / EU	8	8
Thread Count (Total)	768	1024
FLOPs/Clk - Half Precision, MAD (peak)	3072	4096
FLOPs/Clk - Single Precision, MAD (peak)	1536	2048
FLOPs/Clk - Double Precision, MAD (peak)	N/A	N/A
IOPs/Clk - Int8 DP4AS (peak)	24576	32768
FLOPs/Clk - FP16 DP4AS (peak)	12288	16384
GTI	2	2
GTI bandwidth (bytes/unslice-clk)	r: 128	r: 128
	w: 128	w: 128
Graphics Virtual Address Range	48 bit	48 bit
Graphics Physical Address Range (DM) Max Per Device Memory = 64 GB	36 bit	36 bit
Graphics Physical Address Range (System Memory)	46 bit	46 bit
Caches & Dedicated Memories		
Device Cache, total size (bytes)[1]	4096k	4096k
Device Cache, bank count[1]	8	8
Device Cache, bandwidth (bytes/clk)	8x 64 R|W	8x 64 R|W
URB Size (bytes)	2x 480k	2x 480k
L1 Cache, total size (bytes)	8x 192k	8x 192k
L1 Cache, bandwidth (bytes/clk)	6x (256R|128W)	8x (256R|128W)
SLM Size (bytes)[1]	6x 128k	8x 128k
SLM, bandwidth (bytes/clk)	6x (256R|128W)	8x (256R|128W)
Instruction Cache (instances, bytes ea.)	6x 96k	8x 96k
Color Cache (RCC, bytes)	2x 16k	2x 16k
MSC Cache (MSC, bytes)	4x 16k	4x 16k

Base Die	128	
HiZ Cache (HZC, bytes)	4x 4k	4x 4k
Z Cache (RCZ, bytes)	4x 16k	4x 16k
Stencil Cache (STC, bytes)	4x 6k	4x 6k
Instruction Issue Rates		
FMAD, SP (ops/EU/clk)	8	8
FMUL, SP (ops/EU/clk)	8	8
FADD, SP (ops/EU/clk)	8	8
CMP, SP (ops/EU/clk)	8	8
INV, SP (ops/EU/clk)	2	2
SQRT, SP (ops/EU/clk)	2	2
RSQRT, SP (ops/EU/clk)	2	2
LOG, SP (ops/EU/clk)	2	2
EXP, SP (ops/EU/clk)	2	2
TRIG, SP (ops/EU/clk)	2	2
Load/Store		
Data Ports (HDC)	6	8
Device Cache Load/Store (bytes/clk)	384	512
SLM Load/Store (bytes/clk)	1536/768	2048/1024
Atomic, Local 32b (bytes/clk)	6	8
Atomic, Global 32b (bytes/clk)	96	128
3D Attributes		
Geometry pipes	2	2
Samplers (3D)	6	8
2D Texel Rate, point, 32b (tex/clk)	48	64
2D Texel Rate, point, 64b (tex/clk)	48	64
2D Texel Rate, point, 128b (tex/clk)	48	64
2D Texel Rate, bilinear, 32b (tex/clk)	48	64
2D Texel Rate, bilinear, 64b (tex/clk)	48	64
2D Texel Rate, bilinear, 128b (tex/clk)	12	16
2D Texel Rate, trilinear, 32b (tex/clk)	24	32
2D Texel Rate, trilinear, 64b (tex/clk)	24	32
2D Texel Rate, trilinear, 128b (tex/clk)	6	8
2D Texel Sample Rate, aniso 2x (MIP nearest), 32b (tex/clk)	48	64
2D Texel Sample Rate, aniso 4x (MIP nearest), 32b (tex/clk)	24	32
2D Texel Sample Rate, ansio 8x (MIP nearest), 32b (tex/clk)	12	16
2D Texel Sample Rate, ansio 16x (MIP nearest), 32b (tex/clk)	6	8
3D Texel Sample Rate, point, 32b (tex/clk)	48	64

Base Die	128	
3D Texel Sample Rate, point, 64b (tex/clk)	48	64
3D Texel Sample Rate, point, 128b (tex/clk)	24	32
3D Texel Sample Rate, bilinear, 32b (tex/clk)	24	32
3D Texel Sample Rate, bilinear, 64b (tex/clk)	24	32
3D Texel Sample Rate, bilinear, 128b (tex/clk)	6	8
HiZ Rate, (ppc)	4x 16x8	4x 16x8
IZ Rate, (ppc)	4x 4x4	4x 4x4
Stencil Rate (ppc)	4x 16x8	4x 16x8
Pixel Rate, fill, blend 32bpp (pix/clk)	24	32

Notes:

For all SKUs, engine reset timeout is 5ms

Device Attributes Xe_{HPM}

NOTE: This information is preliminary, and subject to change.

Product Configuration Attribute Table		
Product Family	**Alchemist**	
SKU Name	**512**	**128**
Media Attributes		
Sampler (VME)	0	0
Sampler (AVS)	0	0
VDBox Instances	2	2
VEBox Instances	2	2
SFC Instances	2	2
WGBox Instances	0	0

Device Attributes Xe_{HPD}

Refer to Display Overview.

Steppings and Device IDs

The following table details all currently planned SKUs. This information is subject to change at any time based on roadmap plans. In general:

Index	SoC	Dev2 ID	Xe-cores	EU	Config	VDBoxes	Gfx String
1	ACM/DG2-G10	0x56A0	32	512	8x4x16	2	Intel® Arc ™ A770 Graphics
2	ACM/DG2-G10	0x56A1	28	448	7x4x16	2	Intel® Arc ™ A750 Graphics
4	ACM/DG2-G10	0x5690	32	512	8x4x16	2	Intel® Arc ™ A770M Graphics
5	ACM/DG2-G10	0x5691	24	384	6x4x16	2	Intel® Arc ™ A730M Graphics
6	ACM/DG2-G10	0x5692	16	256	4x4x16	2	Intel® Arc ™ A550M Graphics
7	ATS-M150	0x56C0	32	512	8x4x16	2	Intel® Data Center GPU Flex 170
13	ACM/DG2-G11	0x56A5	8	128	2x4x16	2	Intel® Arc ™ A380 Graphics
14	ACM/DG2-G11	0x56A6	6	96	2x3x16	2	Intel® Arc ™ A310 Graphics
15	ACM/DG2-G11	0x5693	8	128	2x4x16	2	Intel® Arc ™ A370M Graphics
16	ACM/DG2-G11	0x5694	6	96	2x3x16	2	Intel® Arc ™ A350M Graphics
19	ACM/DG2-G11	0x56B0	8	128	2x4x16	2	Intel® Arc ™ Pro A30M Graphics
20	ACM/DG2-G11	0x56B1	8	128	2x4x16	2	Intel® Arc ™ Pro A40/A50 Graphics
21	ATS-M75 (x2)	0x56C1	8 (x2)	128 (x2)	2x4x16 (x2)	2 (x2)	Intel® Data Center GPU Flex 140

*******DevID today is represented by the SGUnit PCIE ID. Final DevID/RevID definitions may change.