Intel® Open Source HD Graphics and Intel Iris™ Plus Graphics

Programmer's Reference Manual

For the 2016 - 2017 Intel Core™ Processors, Celeron™ Processors, and Pentium™ Processors based on the "Kaby Lake" Platform

Volume 4: Configurations

January 2017, Revision 1.0

Creative Commons License

Notices and Disclaimers

Table of Contents

Configurations Overview

The Intel "Gen" Graphics Architecture was first introduced to the market in 2004. Since that time, the architecture and implementation have evolved to add many new features, increase performance, and improve power efficiency.

Each product generation has its own configurations chapter. Each chapter has a section for each project, and each project contains the following subsections:

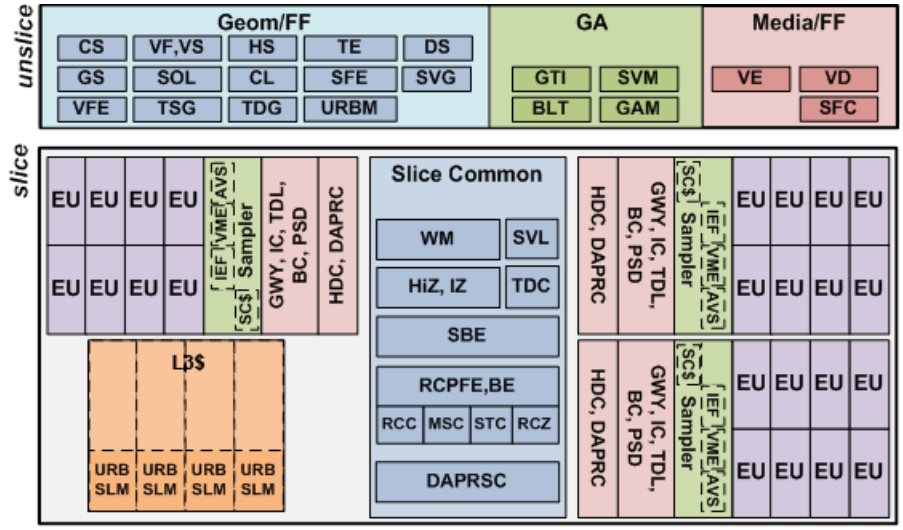
- Top Level Block Diagrams - Show basic feature blocks of the project's graphics architecture, for GT configurations.
- Device Attributes - List details of the graphics configuration options for each project.
- Steppings and Device IDs - Lists all of the current unique GT Die / Packages for a specific project.

Top Level Block Diagrams

The diagrams below show basic feature blocks of the Kaby Lake (KBL) graphics architecture.

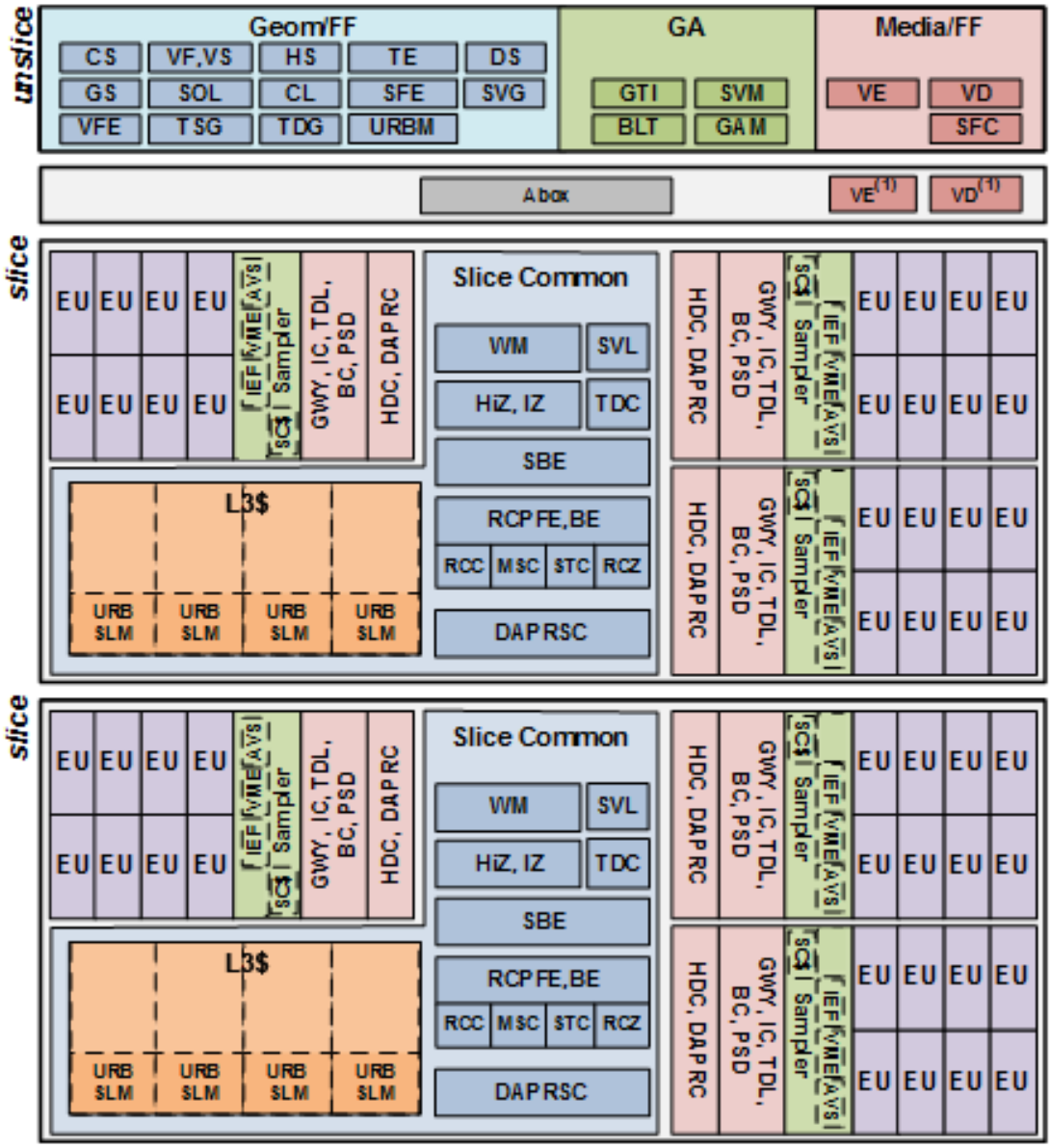
GT2 Configuration

The GT2 configuration contains one Unslice and one Slice with separate power domains for each, although they share a single clock domain.



GT3 Configuration

The GT3 configuration has an identical Unslice to GT2, except that it contains two Slices. Separate clock domains for the Unslice and Slice may be available depending on SKU. The L3 caches of each Slice combine to provide an aggregate L3 cache of twice the size and twice the bandwidth of a single instance. GT3 also has additional media blocks with second instance of VEBox and VDBox each.



This diagram is based on the following functional partitions:

(a) Geometry Fixed Functions (Geom/FF)

(b) Media Fixed Functions (Media/FF)

(c) Global Assets and GT Interface (GA)

(d) One or more Subslices (three shown)

(e) A Slice Common block

(f) An L3 Cache (L3$) block

Note that the combination of (a), (b), and (c) is typically referred to as the “unslice”, while a combination of (d), (e), and (f) is referred to as a compute “slice”.

The functionality in each of these groupings is further broken down as follows:

- Unslice – Fixed function pipelines for 3D, GPGPU, and Media operations, and interface to the outside world.
    - The 3D Geometry / Fixed Function (Geom/FF) block consisting of:
        - 3D fixed function pipeline (CS, VFVS, HS, TE, DS, GS, SOL, SL, SFE, SVG)
        - Video Front-End unit (VFE)
        - Thread Spawner unit (TSG) and the global Thread Dispatcher unit (TDG)
        - Unified Return Buffer Manager (URBM)
        -
    - Media fixed function assets:
        - Video Decode (VD) Box
        - Video Encode (VE) Box
        - Scaler & Format Converter (SFC)
    - The Global Assets (GA) block as the primary interface and memory stream gateway to the outside world, consisting of:
        - GT Interface (GTI)
        - State Variable Manager (SVM)
        - Blitter (BLT)
        - Graphics Arbiter (GAM)
- Subslice (three shown) – A compute unit with supporting fixed- or shared-function assets sufficient for the EU capability.
    - A bank of Execution Units (EUs) – eight per subslice shown
    - Sampler, supporting both media and 3D functions
    - Gateway (GWY)
    - Instruction cache (IC)
    - Local Thread Dispatcher (TDL)
    - Barycentric Calculator (BC)
    - Pixel Shader Dispatcher (PSD)
    - Data Cluster (HDC)
    - Dataport Render Cache (DAPRC) - two per subslice
- Slice Common – Scalable fixed function assets which support the compute horsepower provided two or more subslices.
    - 3D Fixed Function:
        - Windower/Mask unit (WM)
        - Hi-Z (HZ) and Intermediate Z (IZ)
        - Setup Backend (SBE)

        - RCPFE, BE
        - 3D stream caches (RCC, MSC, STC, RCZ)
    - Media Fixed Functions:
        - DAPRSC
        - SVL
        - TDC
- L3 Cache – backing L3 cache for certain memory streams emanating from subslices.
    - L3 Data cache with support for data, URB, and shared local memory (SLM)

Device Attributes

The following table lists detailed GT device attributes for proposed KBL SKUs.

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

Product Configuration Attribute Table				
Product Family	**KBL**			
Architectural Name *	**1x2x6**	**1x3x6**	**1x3x8**	**2x3x8**
SKU Name	**GT1F**	**GT1.5F**	**GT2**	**GT3**
Global Attributes				
Slice count	1	1	1	2
Subslice Count	2	3	3	6
EU/Subslice	6	6	8	8
EU count (total)	12	18	23 / 24 [b]	47 / 48 [b]
Thread Count	7	7	7	7
Thread Count (Total)	84	126	161 / 168	329 / 336
FLOPs/Clk - Half Precision, MAD (peak)	384	576	736 / 768	1504 / 1536
FLOPs/Clk - Single Precision, MAD (peak)	192	288	368 / 384	752 / 768
FLOPs/Clk - Double Precision, MAD (peak)	48	72	92 / 96	188 / 192
Unslice clocking (coupled/decoupled from Cr slice)	coupled	coupled	coupled	coupled
GTI / Ring Interfaces	1	1	1	1
GTI bandwidth (bytes/unslice-clk)	64: R	64: R	64: R	64: R
	64: W	64: W	64: W	64: W
eDRAM Support	N/A	N/A	N/A	0, 64MB
Graphics Virtual Address Range	48 bit	48 bit	48 bit	48 bit
Graphics Physical Address Range	39 bit	39 bit	39 bit	39 bit
Caches & Dedicated Memories				
L3 Cache, total size (bytes)	384K	768K	768K	1536K
L3 Cache, bank count	2	4	4	8
L3 Cache, bandwidth (bytes/clk)	2x 64: R 2x 64: W	4x 64: R 4x 64: W	4x 64: R 4x 64: W	8x 64: R 8x 64: W
L3 Cache, D$ Size (Kbytes)	192K - 256K	512K	512K	1024K
URB Size (kbytes)	128K - 192K	384K	384K	768K
SLM Size (kbytes)	0, 128K	0, 192K	0, 192K	0, 384K
LLC/L4 size (bytes) [1]	~2MB/CPU core	~2MB/CPU core	~2MB/CPU core	~2MB/CPU core
Instruction Cache (IC, bytes)	2x 48K	3x 48K	3x 48K	6x 48K
Color Cache (RCC, bytes)	24K	24K	24K	2x 24K

Product Configuration Attribute Table				
MSC Cache (MSC, bytes)	16K	16K	16K	2x 16K
HiZ Cache (HZC, bytes)	12K	12K	12K	2x 12K
Z Cache (RCZ, bytes)	32K	32K	32K	2x 32K
Stencil Cache (STC, bytes)	8K	8K	8K	2x 8K
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
MIN,MAX, 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
POW, SP (ops/EU/clk)	1	1	1	1
IDIV, SP (ops/EU/clk)	1-6	1-6	1-6	1-6
TRIG, SP (ops/EU/clk)	2	2	2	2
FDIV, SP (ops/EU/clk)	1	1	1	1
Load/Store				
Data Ports (HDC)	2	3	3	6
L3 Load/Store (dwords/clk)	2x 64	3x 64	3x 64	6x 64
SLM Load/Store (dwords/clk)	2x 64	3x 64	3x 64	6x 64
Atomic Inc, 32b - sequential addresses (dwords/clk)	2x 64	3x 64	3x 64	6x 64
Atomic Inc, 32b - same address (dwords/clk)	2x 4	3x 4	3x 4	6x 4
Atomic CmpWr, 32b - sequential addresses (dwords/clk)	2x 32	3x 32	3x 32	6x 32
Atomic CmpWr, 32b - same address (dwords/clk)	2x 4	3x 4	3x 4	6x 4
3D Attributes				
Geometry pipes	1	1	1	1
Samplers (3D)	2	3	3	6
Texel Rate, point, 32b (tex/clk)	8	12	12	24
Texel Rate, point, 64b (tex/clk)	8	12	12	24
Texel Rate, point, 128b (tex/clk)	8	12	12	24
Texel Rate, bilinear, 32b (tex/clk)	8	12	12	24
Texel Rate, bilinear, 64b (tex/clk)	8	12	12	24

Product Configuration Attribute Table				
Texel Rate, bilinear, 128b (tex/clk)	2	3	3	6
Texel Rate, trilinear, 32b (tex/clk)	8	12	12	24
Texel Rate, trilinear, 64b (tex/clk)	4	6	6	12
Texel Rate, trilinear, 128b (tex/clk)	1	1.5	1.5	3
Texel Rate, aniso 2x, MIP Linear,, 32b (tex/clk)	2	3	3	6
Texel Rate, aniso 4x, MIP Linear,, 32b (tex/clk)	1	1.5	1.5	3
Texel Rate, ansio 8x, MIP Linear,, 32b (tex/clk)	0.5	0.75	0.75	1.5
Texel Rate, ansio 16x, MIP Linear,, 32b (tex/clk)	0.25	0.375	0.375	0.75
HiZ Rate, (ppc)	64	64	64	2x 64
IZ Rate, (ppc)	16	16	16	2x 16
Stencil Rate (ppc)	64	64	64	2x 64
(500 MHz, DDR-2400 or eDRAM; Range depends on dynamic compression ratio)				
Pixel Rate, fill, 32bpp (pix/clk, RCC hit)	8	8	8	16
Pixel Rate, fill, 32bpp (pix/clk, LLC hit @ 1.0x unslice clk) [2]				
Pixel Rate, fill, 32bpp (pix/clk, LLC hit, @ 1.5x unslice clk) [2]	N/A	N/A	N/A	
Pixel Rate, fill, 32bpp (pix/clk, memory, @ 1.0x unslice clk) [2]				
Pixel Rate, fill, 32bpp (pix/clk, memory, @ 1.5x unslice clk) [2]	N/A	N/A	N/A	
(500 MHz, DDR-2400 or eDRAM; Range depends on dynamic compression ratio)				
Pixel Rate, blend, 32bpp (p/clk, RCC hit)	8	8	8	16
Pixel Rate, blend, 32bpp (p/clk, LLC hit, @ 1.0x unslice clk) [2]				
Pixel Rate, blend, 32bpp (p/clk, LLC hit, @ 1.5x unslice clk) [2]	N/A	N/A	N/A	
Pixel Rate, blend, 32bpp (pix/clk, memory, @ 1.0x unslice clk) [2]				
Pixel Rate, blend, 32bpp (pix/clk, memory, @ 1.5x unslice clk) [2]	N/A	N/A	N/A	
Media Attributes				
Samplers (media)	2	3	3	6
VDBox Instances	1	1	1	2
VEBox Instances	1	1	1	2
SFC Instances	1	1	1	1

Product Configuration Attribute Table				
Display Attributes				
Display Pipes	3	3	3	3
Display Planes per Pipe	3	3	3	3
DDI ports	2	2	2	2
eDP ports	1	1	1	1

Footnotes:

* Architectural Name = Slice Count x Subslice Count x EUs per Subslice

[a] SKU naming & details has not yet been decided.

[b] One EU reserved for die recovery purposes.

Steppings and Device IDs

The following table lists currently proposed variations of GT Die / Packages for Gen9 (KBL). Prior to manufacturing, this information is subject to change at any time.

CPU SKU	GT SKU	CPU Stepping	GT/Disp Stepping	Device0 ID	Native Device2 ID	GT Device2 Revision ID	Comments
KBL-U 2+2	GT2	H0	C0/B0	0x5904	0x5916	0x2	MB
KBL-Y 2+2	GT2	H0	C0/B0	0x590C	0x591E	0x2	MB
KBL-H 4+2	GT2	B0	F0/C0	0x5910	0x591B	0x4	DT, uP Server
KBL-S 4+2	GT2	B0	F0/C0	0x591F	0x5912	0x4	DT, uP Server
KBL-U 2+3	GT3	J1	D1/B1	0x5904	0x5926	0x6	
KBL-S 2+2	GT2	S0	F0/C0	0x590F	0x5912	0x4	

Kaby Lake SKUs and Device IDs

The following table details all currently planned SKUs for KBL. This list is subject to change at any time based on roadmap plans.

SKU Type	CPU SKU	GT SKU	Device 2 ID
U	KBL U - ULT 2+1F	GT1	0x5906
U	KBL U - ULT 2+3E, 28W	GT3	0x5927
U	KBL U - ULT 2+3E, 15W	GT3	0x5926
Y	KBL Y - ULX 2+2	GT2	0x591E
S	KBL DT 4+2	GT2	0x5912
S	KBL DT 2+2	GT2	0x5912
S	KBL DT 2+1F	GT1	0x5902
H	KBL Halo 4+2	GT2	0x591B
SRV	KBL SRV 4+2	GT2	0x591A
WS	KBL WKS 4+2	GT2	0x591D