

## ■ AMD Instinct MI シリーズ

モデル	Instinct MI300X	Instinct MI300A	Instinct™ MI250X	Instinct™ MI250	Instinct™ MI210	Instinct™ MI100
ライン	Instinct MI300 Series	Instinct MI300 Series	Instinct MI200 Series	Instinct MI200 Series	Instinct MI200 Series	Instinct MI100 Series
プラットフォーム	サーバー	サーバー	サーバー	サーバー	サーバー	サーバー
発売日	12/06/2023	12/06/2023	11/08/2021	11/08/2021	3/22/2022	11/16/2020
GPUアーキテクチャ	CDNA3	CDNA3	CDNA2	CDNA2	CDNA2	CDNA
リソグラフィ	TSMC 5nm   6nm FinFET	TSMC 5nm   6nm FinFET	TSMC 6nm FinFET	TSMC 6nm FinFET		TSMC 7nm FinFET
ストリーミングプロセッサ	19,456	14,592	14,080	13,312	6,656	7,680
Matrix Cores	1,216	912	-	-	-	-
演算ユニット	304	228	220	208	104	120
Peak Engine Clock	2100 MHz	2100 MHz	1700 MHz	1700 MHz	1700 MHz	1502 MHz
Peak Eight-bit Precision (FP8) Performance (E5M2, E4M3)	2.61 PFLOPs	1.96 PFLOPs	-	-	-	-
Peak Eight-bit Precision (FP8) Performance with Structured Sparsity (E5M2, E4M3)	5.22 PFLOPs	3.92 PFLOPs	-	-	-	-
ピーク時半精度浮動小数点 (FP16)	1.3 PFLOPs	980.6 TFLOPs	383 TFLOPs	362.1 TFLOPs	181 TFLOPs	184.6 TFLOPs
Peak Half Precision (FP16) Performance with Structured Sparsity	2.61 PFLOPs	1.96 PFLOPs	-	-	-	-
Peak Single Precision (TF32 Matrix) Performance	653.7 TFLOPs	490.3 TFLOPs	-	-	-	-
Peak Single Precision (TF32) Performance with Structured Sparsity	1.3 PFLOPs	980.6 TFLOPs	-	-	-	-
Peak Single Precision Matrix (FP32) Performance	163.4 TFLOPs	122.6 TFLOPs	95.7 TFLOPs	90.5 TFLOPs	45.3 TFLOPs	46.1 TFLOPs
Peak Double Precision Matrix (FP64) Performance	163.4 TFLOPs	122.6 TFLOPs	95.7 TFLOPs	90.5 TFLOPs	45.3 TFLOPs	-
ピーク時単精度浮動小数点 (FP32)	163.4 TFLOPs	122.6 TFLOPs	47.9 TFLOPs	45.3 TFLOPs	22.6 TFLOPs	23.1 TFLOPs
ピーク時倍精度浮動小数点 (FP64)	81.7 TFLOPs	61.3 TFLOPs	47.9 TFLOPs	45.3 TFLOPs	22.6 TFLOPs	11.5 TFLOPs
Peak INT4 Performance			383 TOPs	362.1 TOPs	181 TOPs	184.6 TOPs
Peak INT8 Performance	2.6 POPs	1.96 POPs	383 TOPs	362.1 TOPs	181 TOPs	184.6 TOPs
Peak INT8 Performance with Structured Sparsity	5.22 POPs	3.92 POPs	-	-	-	-
Peak bfloat16	1.3 PFLOPs	980.6 TFLOPs	383.0 TFLOPs	362.1 TFLOPs	181.0 TFLOPs	92.3 TFLOPs
Peak bfloat16 with Structured Sparsity	2.61 PFLOPs	1.96 PFLOPs	-	-	-	-
最大メモリーサイズ	192 GB	128GB	128GB	128GB	64GB	32GB
メモリータイプ	HBM3	HBM3	HBM2e	HBM2e	HBM2e	HBM2
メモリー・インターフェイス	8192 bit	8192 bit	8192-bit	8192-bit	4096-bit	4096-bit
Memory Clock	5.2 GHz	5.2 GHz	1.6GHz	1.6GHz	1.6GHz	1.2GHz
メモリー帯域幅	5.3 TB/s	5.3 TB/s	3276.8 GB/s	3276.8 GB/s	1638.4 GB/s	1228.8 GB/s
メモリー-ECC対応	○	○	○	○	○	○
RAS Support	○	○	○	○	○	○
Page Retirement	○	○	○	○	○	○
形状	OAM Module	APU SH5 Socket	OAM Module	OAM Module	PCIeアドインカード	PCIeアドインカード
占有PCIeスロット数	N/A (OAM module)	N/A (APU)	N/A (OAM module)	N/A (OAM module)	2	2
Busタイプ	PCIe® 5.0 x16	PCIe® 5.0 x16	PCIe® 4.0 x16	PCIe® 4.0 x16	PCIe® 4.0 x16	PCIe® 4.0 x16, PCIe® 3.0 x16
Infinity Fabric™ Links	8	8	8	8	3	3
Peak Infinity Fabric™ Link Bandwidth	128 GB/s	128 GB/s	100 GB/s	100 GB/s	100 GB/s	92 GB/s
Total Board Power (TBP)	750 W Peak	550W   760W Peak	500W   560W Peak	500W   560W Peak	300W Peak	300W Peak
冷却	Passive OAM	Passive & Active	Passive OAM	Passive OAM	Passive	Passive
カード補助電源コネクタ	不要	不要	不要	不要	8 pin 12V EPS x1	PCIe 8-pin x2