

HPCオープンソースソフトウェア 普及部会ワークショップ 多様化するアクセラレータ

東北大学サイバーサイエンスセンター

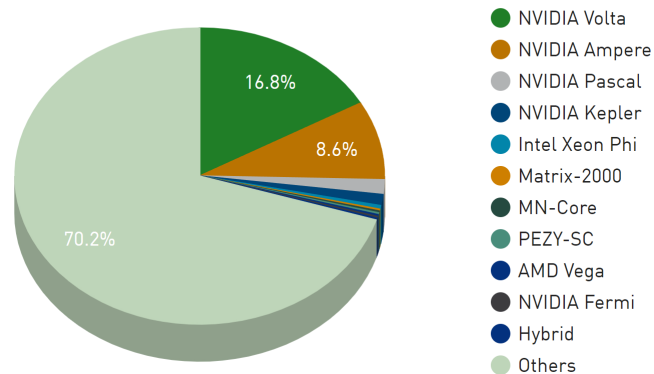
滝沢 寛之

本日の司会進行役

アクセラレータの重要性

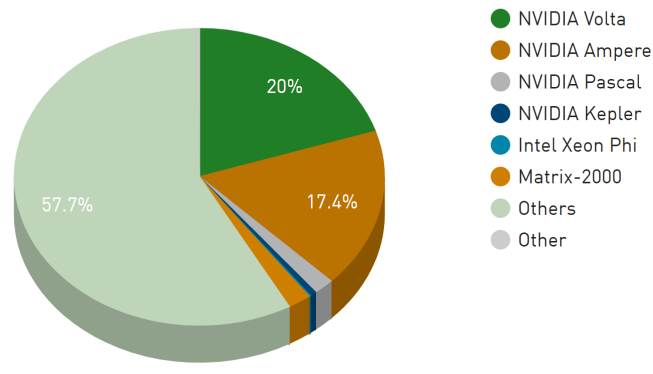
- HPC分野におけるアクセラレータの普及
 - TOP500リスト(Nov. 2021)の上位10システムのうち **8システム** でアクセラレータを使用
 - TOP500リスト全体で **30%のシステム** で採用

Accelerator/CP Family System Share



システム数では約30%

Accelerator/CP Family Performance Share



システム性能では約42%

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPCC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	Perlmutter - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589
6	Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646
7	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
8	JUWELS Booster Module - Bull Sequana XH2000, AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	70,980.0	1,764
9	HPC5 - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, DELL EMC Eni S.p.A. Italy	669,760	35,450.0	51,720.8	2,252
10	Voyager-EUS2 - ND96amsr_A100_v4, AMD EPYC 7V12 48C 2.45GHz, NVIDIA A100 80GB, Mellanox HDR Infiniband, Microsoft Azure Azure East US 2 United States	253,440	30,050.0	39,531.2	

<https://www.top500.org>

主に採用されているアクセラレータ

- ・アクセラレータを使っている149システムのうち**143システム**でNVIDIA社のGPUを利用 (→ **96%**)

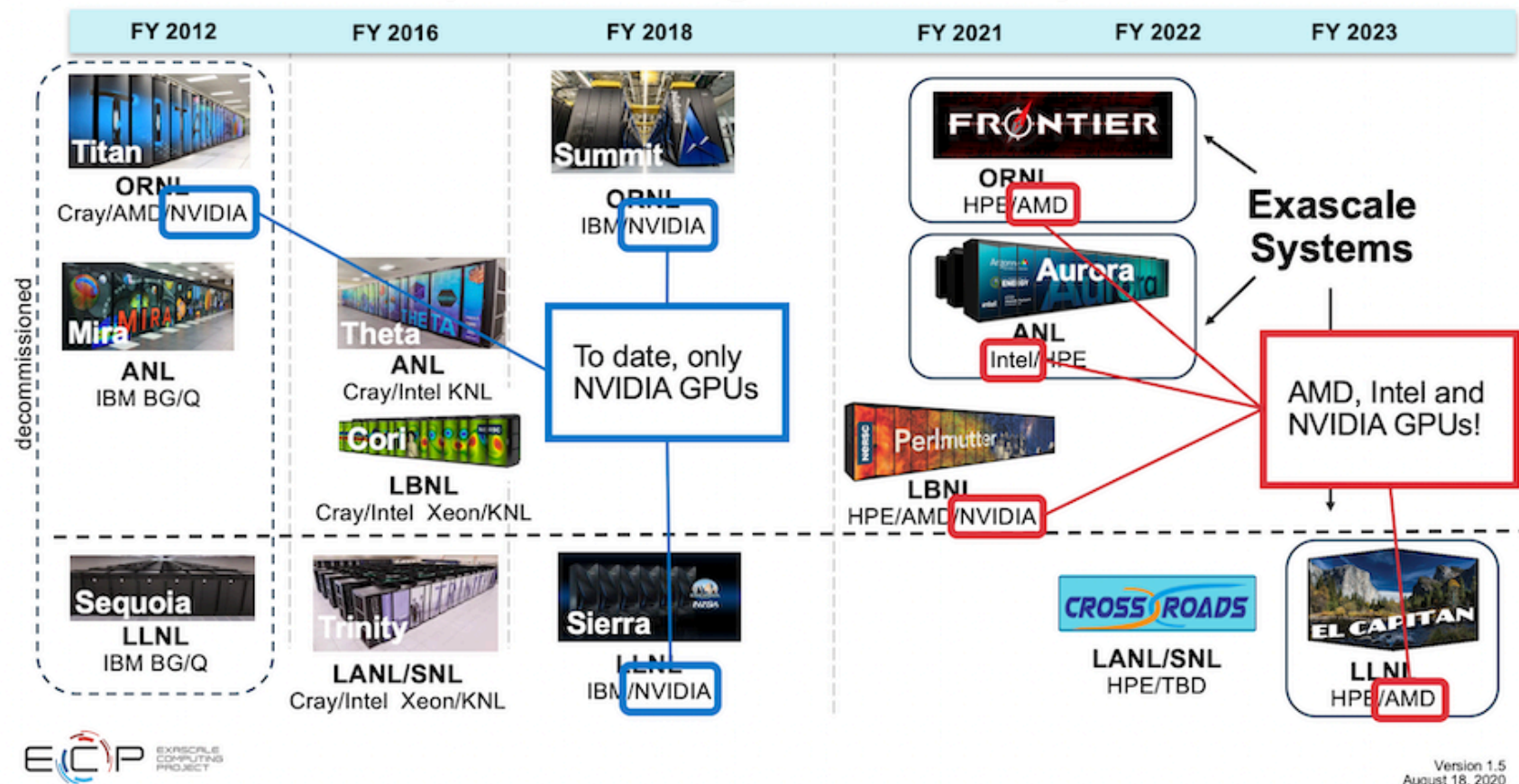


	Accelerator/CP Family	Count	System Share (%)	Rmax (GFlops)	Rpeak (GFlops)	Cores
1	NVIDIA Volta	84	16.8	608,245,890	1,015,712,384	11,475,992
2	NVIDIA Ampere	43	8.6	527,074,700	749,271,060	5,072,700
3	NVIDIA Pascal	8	1.6	54,569,640	80,911,013	1,080,788
4	NVIDIA Kepler	6	1.2	17,082,090	27,962,710	403,712
5	Intel Xeon Phi	2	0.4	4,610,520	6,462,566	369,336
6	Matrix-2000	1	0.2	61,444,500	100,678,664	4,981,760
7	MN-Core	1	0.2	2,181,200	3,389,520	1,664
8	PEZY-SC	1	0.2	1,684,830	2,353,850	822,400
9	AMD Vega	1	0.2	1,661,000	2,472,960	100,800
10	NVIDIA Fermi	1	0.2	2,566,000	4,701,000	186,368
11	Hybrid	1	0.2	3,126,240	5,610,481	152,692
12	Others	351	70.2	1,750,732,014	0	0

<https://www.top500.org>

アクセラレータ多様化の兆し

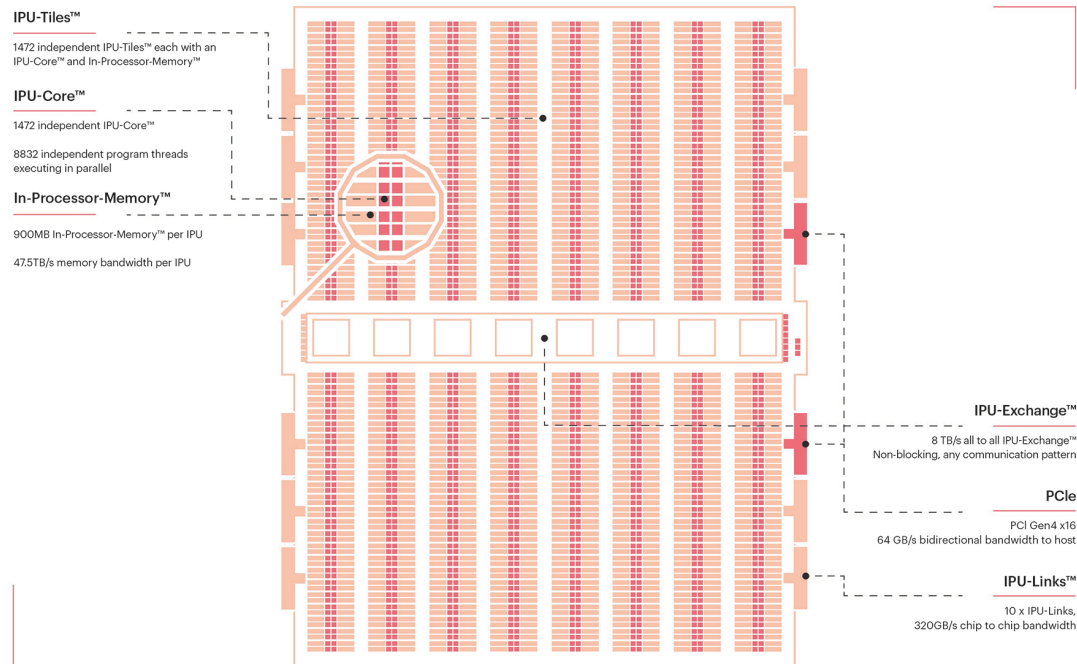
DOE HPC Roadmap to ECP Target Exascale Systems



<https://www.hpcwire.com/2021/03/11/rice-oil-and-gas-keynote-offers-exascale-lessons-and-thoughts-on-technology-uptake/>

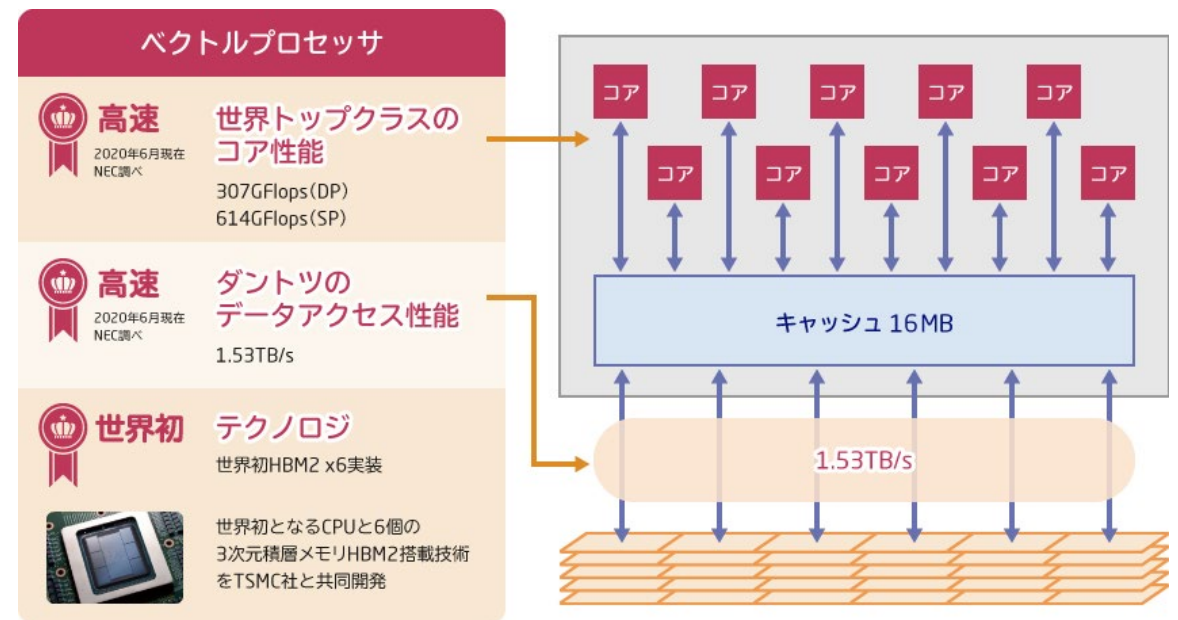
GPU以外の演算アクセラレータの例

GraphCore Colossus MK2 GC200 IPU



<https://www.graphcore.ai/>

NEC SX-Aurora TSUBASA Vector Engine Type 20A



<https://jpn.nec.com/hpc/sxauroratsubasa/features/index.html>

2028年のメニーコア型システムの予測性能 (WP1.0.0より)

- 最も積極的な予測でも最大1.8 EFLOPS (富岳の性能の3.37倍)

	30MW			40MW			50MW		
	60%	70%	80%	60%	70%	80%	60%	70%	80%
ソケット数	46620	54390	62160	62160	72520	82880	77700	90650	103600
総コア数	3.3×10^6	3.8×10^6	4.4×10^6	4.4×10^6	5.1×10^6	5.8×10^6	5.4×10^6	6.3×10^6	7.3×10^6
PFLOPS	815	950	1086	1086	1267	1448	1358	1584	1810
DDR 総 BW (PB/s)	102	120	137	137	160	182	171	200	228
HBM 総 BW (PB/s)	307	358	410	410	478	547	512	598	683
DDR 総容量 (PB)	17	20	23	23	27	31	29	34	39
HBM 総容量 (PB)	4	5	5	5	6	7	7	8	9
インジェク ション BW (Tb/s)	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
総 I/O 性能 (TB/s)	34	34	34	34	34	34	34	34	34
総ストレ ージ容量 (EB)	3.45	3.45	3.45	3.45	3.45	3.45	3.45	3.45	3.45

← システムの電力制約の仮定
← CPUの電力バジェット

158,976
 8.3×10^6
537
—
163
—
4.85
0.33

参考) 富岳の諸元

2028年のGPU混載型システムの予測性能 (WP1.0.0より)

- 最も積極的な予測で最大18.0 EFLOPS (富岳の性能の33.5倍)

保守的な予測の場合
(NVIDIA GPUの性能
トレンドから外挿)

	30MW			40MW			50MW		
	60%	70%	80%	60%	70%	80%	60%	70%	80%
GPU 数	50661	59104	67548	67548	78806	90064	84435	98508	112580
総コア数	5.3×10^8	6.2×10^8	7.1×10^8	7.1×10^8	8.3×10^8	9.4×10^8	8.8×10^8	1.0×10^9	1.2×10^9
PFLOPS	1279	1492	1706	1706	1940	2474	2132	2487	2843
HBM 総BW (PB/s)	91	107	122	122	143	163	153	178	204
HBM 総容量 (PB)	1	1	2	2	2	2	2	3	3

積極的な予測の場合
(CPUとの性能比の
トレンドから外挿)

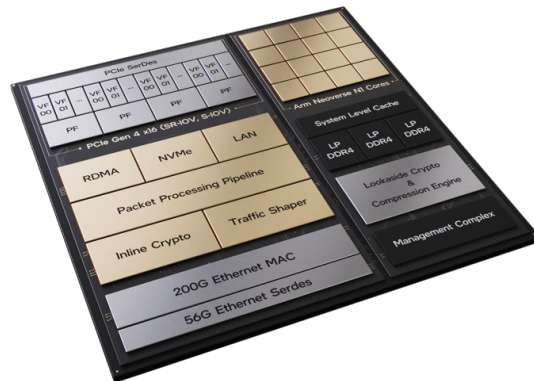
	30MW			40MW			50MW		
	60%	70%	80%	60%	70%	80%	60%	70%	80%
GPU 数	50661	59104	67548	67548	78806	90064	84435	98508	112580
総コア数	3.4×10^9	3.9×10^9	4.5×10^9	4.5×10^9	5.2×10^9	6.0×10^9	5.6×10^9	6.5×10^9	7.5×10^9
PFLOPS	8083	9431	10778	10778	12574	14371	13472	15718	17963
HBM 総BW (PB/s)	334	390	445	445	520	594	557	650	743
HBM 総容量 (PB)	4	5	6	6	7	8	8	9	10

計算以外にも活躍の場

nVIDIA BlueField-2

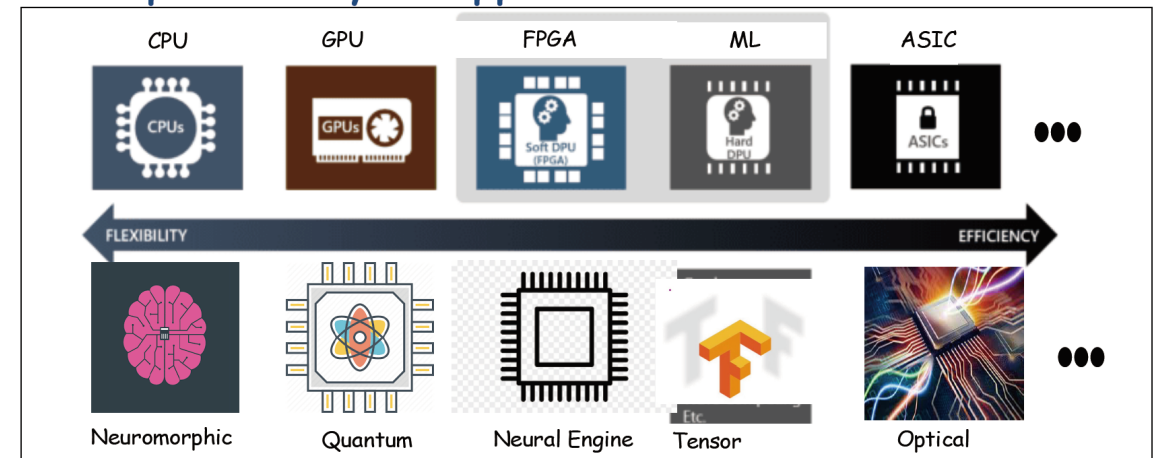


Intel Mount Evans



Future HPC Systems Will be Customized...

◆ You will be able to dial up what you need in your computer for your application mix ...



Dongarra@RSD2020

アプリ開発における関心事

- どのアプリケーション分野や用途で使うべきか？効果は？
- アクセラレータがプログラマからどのように見えるか？
- 投資した労力が無駄にならないか？

などなど



研究者にとっては **活躍のチャンス！**



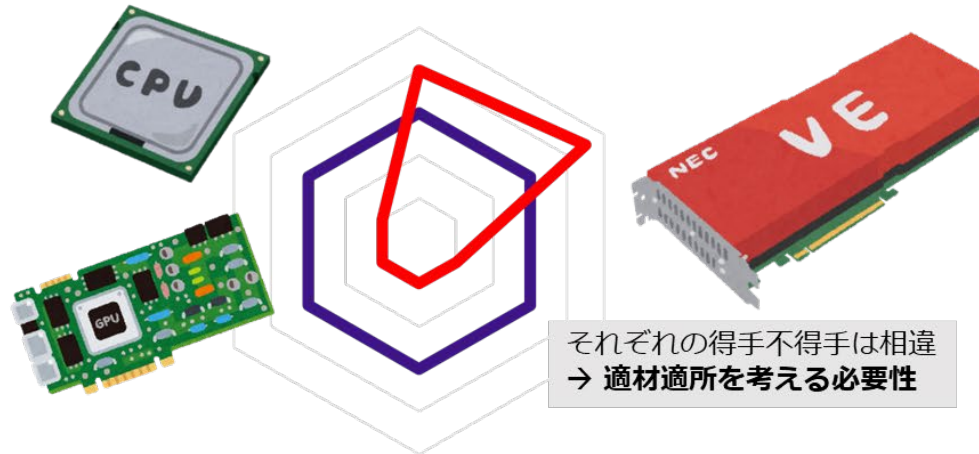
東北大学での取り組み

スーパーコンピュータの運用

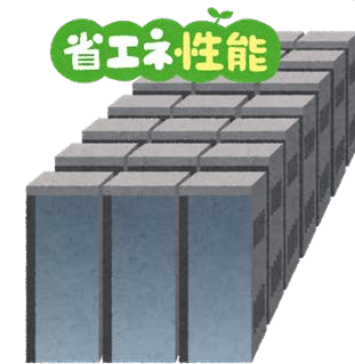


複雑なシステムの性能を引き出すプログラミング

異種複数のプロセッサを搭載する現在のスパコン



アプリケーション最適化技術



消費電力低減は最重要技術



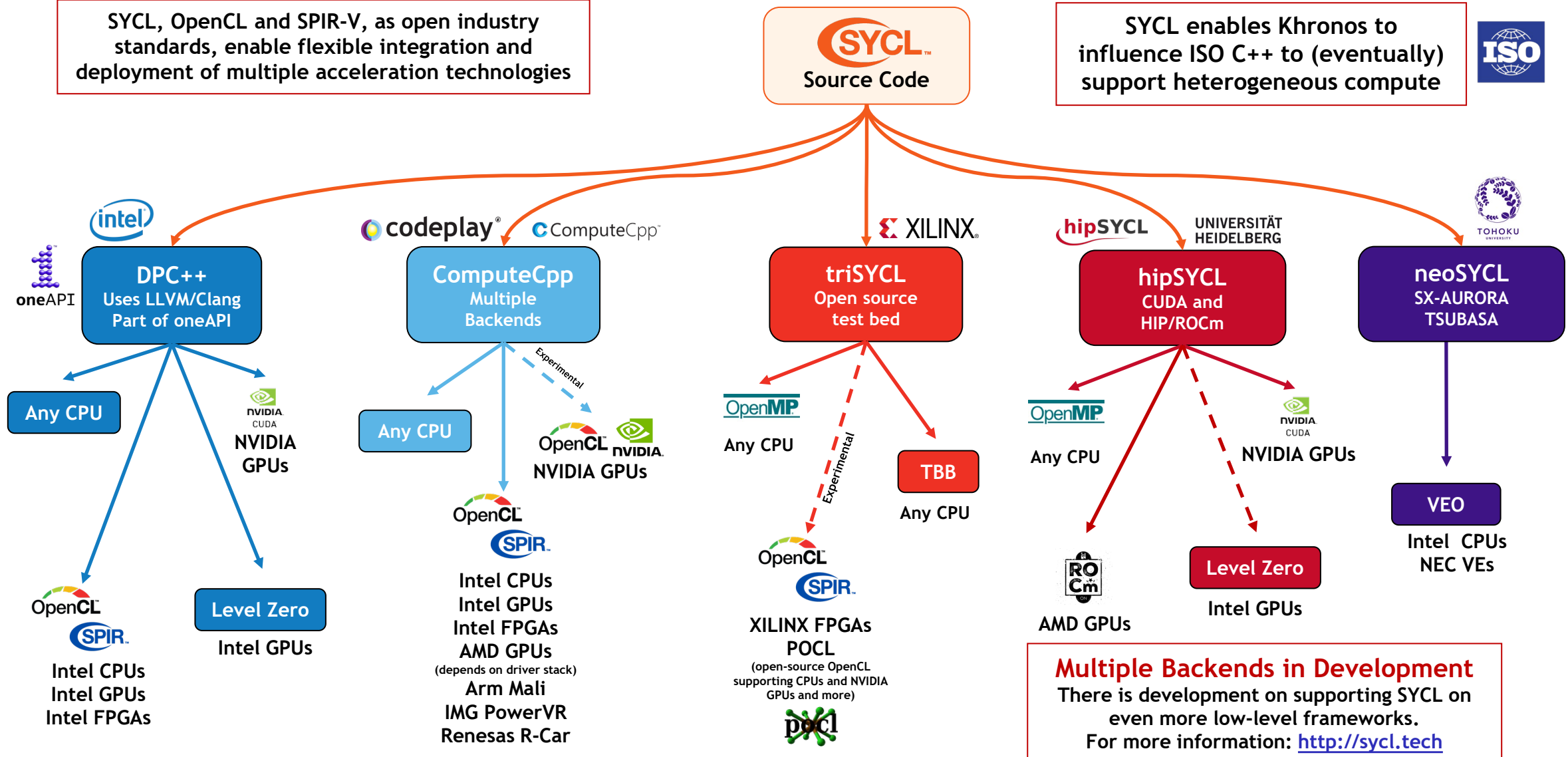
センサーやIoTデバイスと連携する社会インフラとしてのスパコン

次世代スーパーコンピューティング技術の開拓

SYCL Implementations in Development (2021/04/27)

SYCL, OpenCL and SPIR-V, as open industry standards, enable flexible integration and deployment of multiple acceleration technologies

SYCL enables Khronos to influence ISO C++ to (eventually) support heterogeneous compute



本日の予定

13:00-13:20 開会挨拶・趣旨説明

13:20-13:45 佐々木 邦暢（エヌビディア合同会社）
「GPUとDPUによるアクセラレーテッド コンピューティング」

13:45-14:10 Greg, Oakes (AMD)
「AMD Instinct GPU and ROCm technical overview」

14:10-14:35 中野 守（Graphcore Japan KK）
「AI革新プロセッサIPU」

14:45-15:10 大内山 浩、高藤良史（インテル株式会社）
「インテルXPU戦略」

15:10-15:35 政岡 靖久（日本電気株式会社）
「ハイブリッドアクセラレータ SX-Aurora TSUBASA」

15:35-16:00 鯨岡 俊則（SambaNova）
「SambaNova Systemsのご紹介」

16:10-17:30 パネルディスカッション
「HPCを加速するアクセラレータの挑戦と将来展望」



パネルディスカッション

「HPCを加速するアクセラレータの挑戦と将来展望」

コーディネーター 滝沢寛之 (東北大学)

趣旨説明

- 今後、ワークロードの多様化に伴いHPCの分野においても様々なアクセラレータの利用が期待されている。現状で広く使われているアクセラレータに対して、各社はどのような戦略でビジネスを展開していこうとしているのか。アクセラレータを利用した変革に挑む各社が、現在のアクセラレータとの差別化や将来展望を討議する。



パネリスト

- ・井崎 武士 (エヌビディア合同会社)
- ・大内山 浩 (インテル株式会社)
- ・鯨岡 俊則 (SambaNova)
- ・中野 守 (Graphcore Japan KK)
- ・政岡 靖久 (日本電気株式会社)
- ・吉田 佳司 (日本AMD株式会社)



議題

1. 技術面およびビジネス面での今後の戦略

- 現在、HPCシステムに求められる役割が多様化しており、従来の科学技術計算だけではなく、AI・機械学習のワークロード実行はもちろんのことで、クラウドやデータセンターとしての役割も期待されるようになっていきます。多様なワークロードを限られた電力でより効率的に実行するために、多様なアクセラレータの活用が今後ますます広がっていくと思われます。そのようなHPC分野の広がりに対してどのように貢献しようとしているのか、アクセラレータに関する技術的およびビジネス的な今後の展開の戦略を教えてください。

2. アクセラレータ向けプログラミングの理想形

- アクセラレータの存在をプログラマに意識させるべきでしょうか？ つまり、明示的にオフロードすべきか、将来的にはプログラマから見えなくなっていくべきでしょうか？ 見える場合には、プログラマは何をどの程度まで指示すべきでしょうか？ プログラマへの見せ方という観点から、アクセラレータのあるべき姿を教えてください。

3. アクセラレータという存在の長期的な将来像

- GPUコンピューティングが登場して間もないころは、アクセラレータはいずれCPUの機能として統合されていく存在であり、暫定的な存在だという意見も多かったと記憶しています。物理的に独立したアクセラレータではなくて、ほんとうはCPUに統合されていくほうが使いやすい面も多々あるかと思います。「アクセラレータ」という存在そのものについて、やや長期的な視点で御社の考える将来像を教えてください。

4. 最後にひとこと

- HPC分野のオープンソースソフトウェア開発者に向けて、他にもアピールするポイントがあれば教えてください。