

国立大学法人 大阪大学
サイバーメディアセンター 御中

SX-ACEユーザ向けSQUID説明会

SX-Aurora TSUBASAの概要

2021年01月26日
日本電気株式会社
第一官公ソリューション事業部

Orchestrating a brighter world

NECは、安全・安心・公平・効率という
社会価値を創造し、
誰もが人間性を十分に発揮できる
持続可能な社会の実現を目指します。

目 次

1. はじめに
2. ハードウェアスペック上の違い
3. 次期システム(SQUID)の概要
4. SX-Aurora TSUBASAのご紹介
5. アプリケーション事例

1. はじめに

はじめに

■ 本セッションの内容

- 【2. ハードウェアスペックの違い】

SX-ACE とSX-Aurora TSUBASAのスペックの違いのご説明

SX-Aurora TSUBASAご利用上のポイントをご説明します

- 【3. 次期システム(SQUID)の概要】

次期システム(SQUID)のベクトル型計算機のご紹介です

- 【4. SX-Aurora TSUBASAのご紹介】

SX-Aurora TSUBASA の製品紹介です

- 【5. アプリケーション事例】

アプリケーション実行例のご紹介

アプリケーション移行の具体的な手順は、後半のセッションでご説明します

2. ハードウェアスペックの違い

ハードウェアスペックの比較

ハードウェアスペックの比較

- 共有メモリの単位である、SX-ACE 1台と、SX-Aurora TSUBASA 1VEの比較は以下の通りです。

	SX-ACE	SX-Aurora TSUBASA Type20A	比率
コア数	4	10	2.5倍
演算性能	256 GFlops	3.07 TFlops	11.9倍
メモリ量	64 GB	48GB	0.75倍
メモリ帯域	256 GB/s	1.53 TB/s	5.97倍



※ 共有メモリ量が小さくなるため、1台でメモリを最大近くまで利用していたアプリケーションは、MPI化などのプログラム修正が発生する場合があります。

VEとVH - SX-Aurora TSUBASA アーキテクチャ

ベクトルプロセッサ+x86/Linux アーキテクチャ

- 主にOS処理を行う**ベクトルホスト(VH)**部 + 演算処理を行う**ベクトルエンジン(VE)**部
- アプリケーション全体が実行されるVE 部は、ベクトルプロセッサ及び高速メモリから構成され、VHとPCIe 経由で接続
- OS機能を提供するVHに、x86/Linux マシンを採用することで、ベクトルプロセッサを標準のx86/Linux 環境上で利用可能

VH(ベクトルホスト)

VE(ベクトルエンジン)



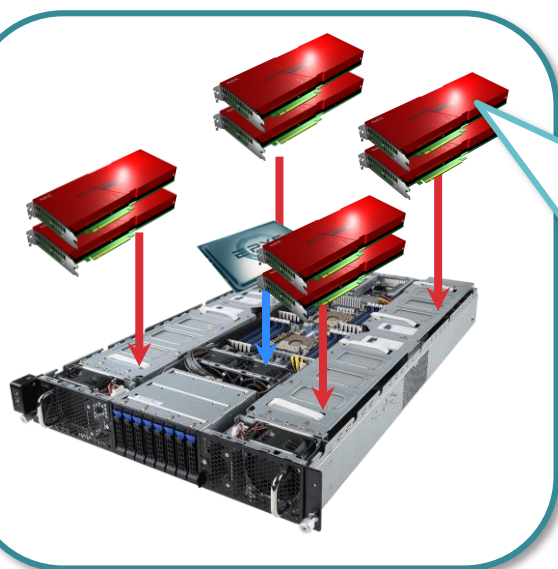
SX-Aurora TSUBASA システム構成

SX-Aurora TSUBASA システム構成

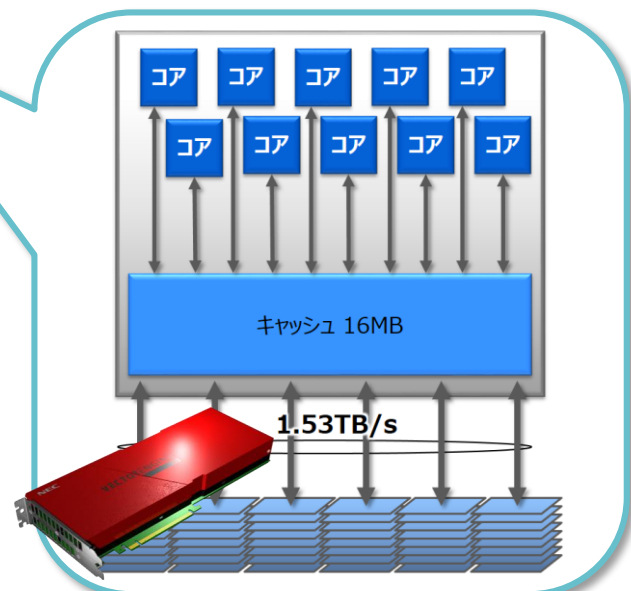
- ベクトルエンジンあたり、10ベクトル演算コア
- ベクトルホストあたり、8ベクトルエンジンを搭載



ベクトルノード群
36VH (2Rack)



VH(ベクトルホスト)
8VE + x86_64 サーバ
(288 VE)



VE(ベクトルエンジン)
10core / VE
(2,880 core)

並列処理

分散メモリ(MPI)並列で利用

共有メモリ(スレッド)並列で利用

SX-Aurora TSUBASA 利用上のポイント

アプリケーションは全てVE上で動作する

- アプリケーションは**全てVE上で動く**ため、実行環境はVEのスペックで決まる
※ NEC SDK for VE でのコンパイルが必要
- 基本は、**VHの CPUとメモリは使用しない**。※VHのCPUを同時利用する高度な使い方あり

OS機能(ファイルアクセスなど)はVH経由で動作する

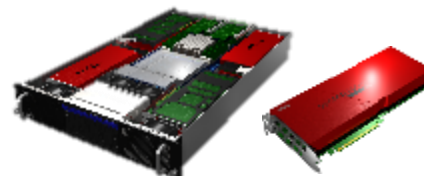
- ファイルアクセスはLinux環境と同じイメージで利用可能
- 前処理等のスクリプトは、VH上で動き**Linux環境の利便性が維持**される
- **glibc サポート**がされ、市販アプリケーションを簡単にコンパイル・実行可能に

VE内は共有メモリ並列、複数VEは分散メモリ並列

- VE内は、10コアを利用した共有メモリ(スレッド)並列が利用可能(**1VE ≡ 1ノード**)
- 複数VEを同時利用する場合は、分散メモリ(MPI)並列を利用可能

3. 次期システム(SQUID)の概要

SQUIDではCPU/GPU/ベクトルの3種の計算環境をお使いいただけます。



	CPUノード	GPUノード	ベクトルノード
モデル名	LX 103Bj-8	LX 106Rj-4G	SX-Aurora TSUBASA B401-8
ノード数	1,520	42	288 (VE)
演算器	次期Intel Xeon CPU(Icelake) x2	次期Intel Xeon CPU(Icelake) x2	Type 20A Vector Engine
加速器	-	NVIDIA A100 x8	-
メモリ容量	256GiB	512GiB	48GiB(HBM2) 128GiB (VH)
相互結合網	InfiniBand HDRx1	InfiniBand HDR100x4	Infiniband HDR x2
ローカルディスク	240GB SATA SSD x1	240GB SATA SSD x1	960GB SATA SSD x1(VH)
冷却方式	水冷	水冷	水冷

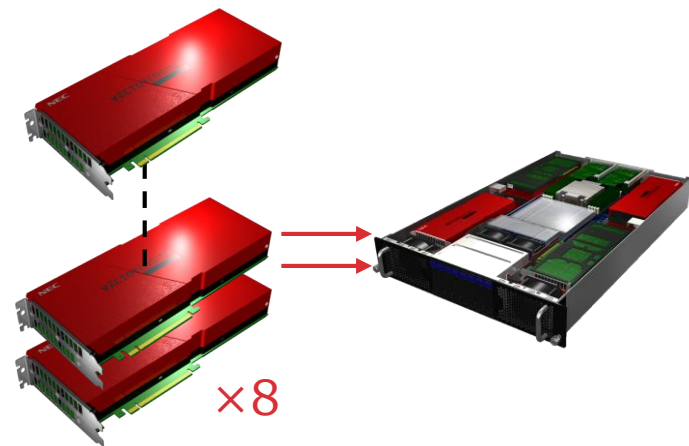
本日のご説明内容

ベクトルノード - HWスペック

「NEC SX-Aurora TSUBASA B401-8」 288VE(36VH)ノードとなります。
VectorEngine (VE) として「Type20A」を有し、VHあたり2Uの筐体に8VEを実装した高密度実装モデルです。

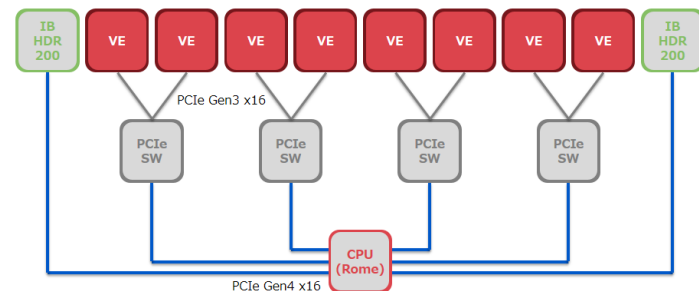
ベクトルエンジン(VE)

項目		構成
VE数		8基 (1VH当たり)
モデル名		Type 20A
VE諸元	演算性能 (倍精度)	307 GFlops / 10core
	メモリ構成	48 GiB (HBM2)
	メモリ帯域	1.53 TB/s
オペレーティングシステム		VEOS



ベクトルホスト(VH)

項目		構成
VH数		36VH
VH諸元	プロセッサ	AMD EPYC 7402P Processor(2.8GHz/24core) x1
	メモリ構成	128GiB (DDR4-3200 16GiB x8)
	ストレージ	960GB SATA SSD x1
	インタフェース	InfiniBand HDR x2, 1000Base-T x1, BMC
オペレーティングシステム		CentOS 8.2 (64bit)



ベクトルエンジン

■ 世界最速クラスのコア

- 307GFlops (DP)
- 614GFlops (SP)

2020年6月現在、NEC調べ

ビックコア

■ 世界最速クラスのデータアクセス性能

- 1.53TB/s

2020年6月現在、NEC調べ

■ テクノロジ

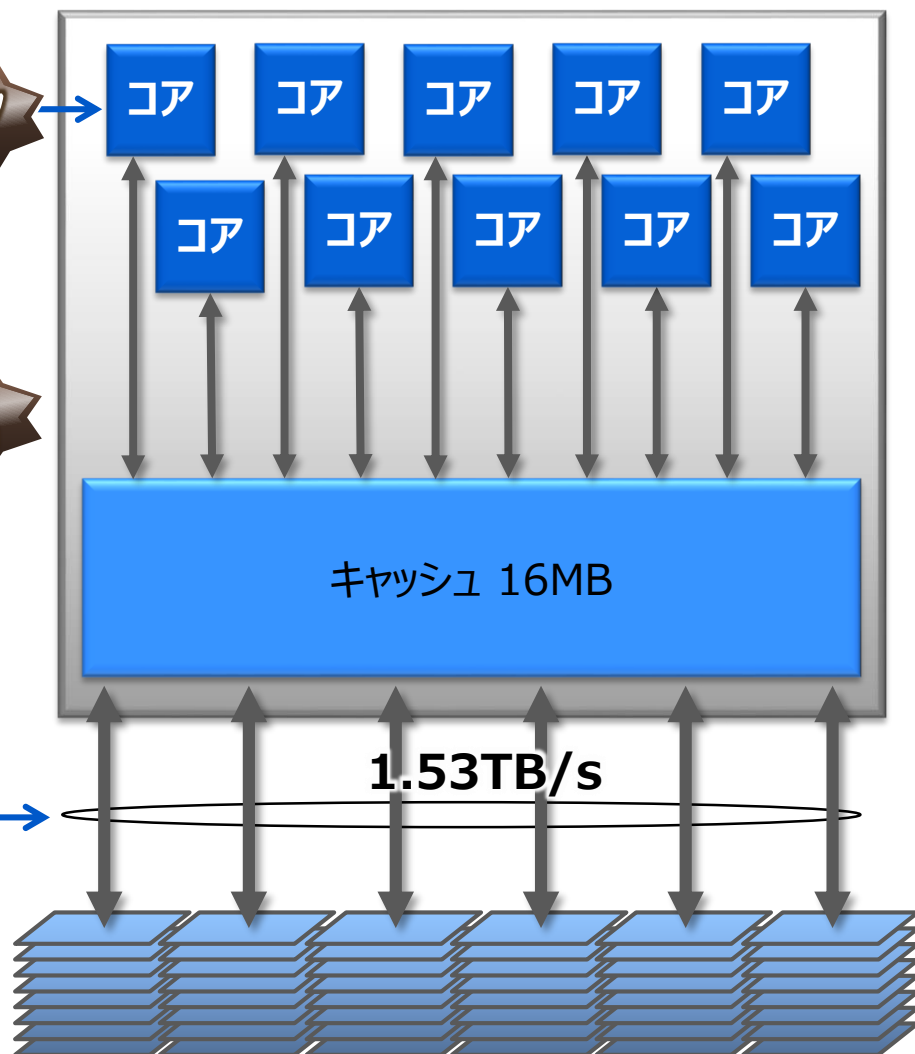
- 世界初HBM2 x6実装

メモリバンド幅重視

世界初



世界初となるCPUと6個の3次元積層メモリHBM2搭載技術をTSMC社と共同開発



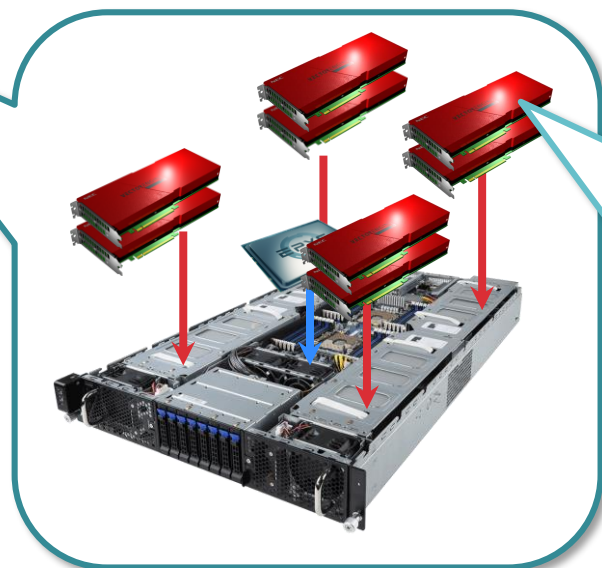
ベクトルノード - システム構成

SX-Aurora TSUBASA システム構成

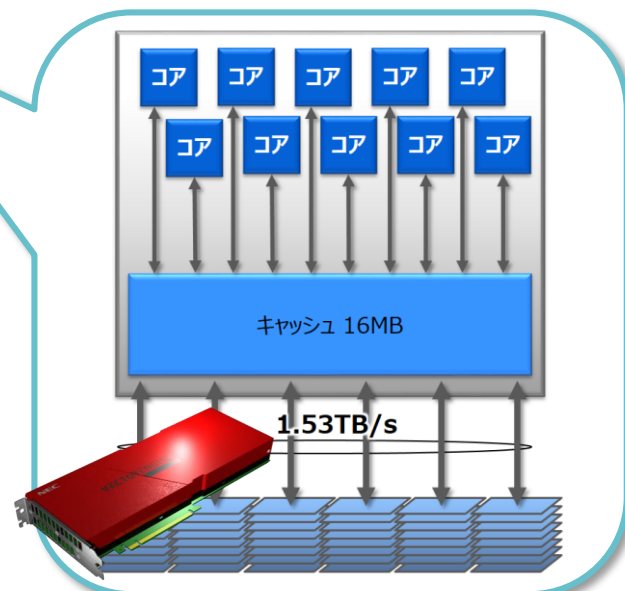
- ベクトルエンジンあたり、10ベクトル演算コア
- ベクトルホストあたり、8ベクトルエンジンを搭載



ベクトルノード群
36VH (2Rack)



VH(ベクトルホスト)
8VE + x86_64 サーバ
(288 VE)



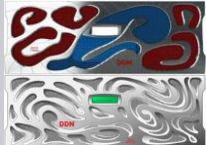
VE(ベクトルエンジン)
10core / VE
(2,880 core)

並列処理

分散メモリ(MPI)並列で利用

共有メモリ(スレッド)並列で利用

■ SQUIDでは2つのファイルシステム、3つの領域がご利用いただけます。

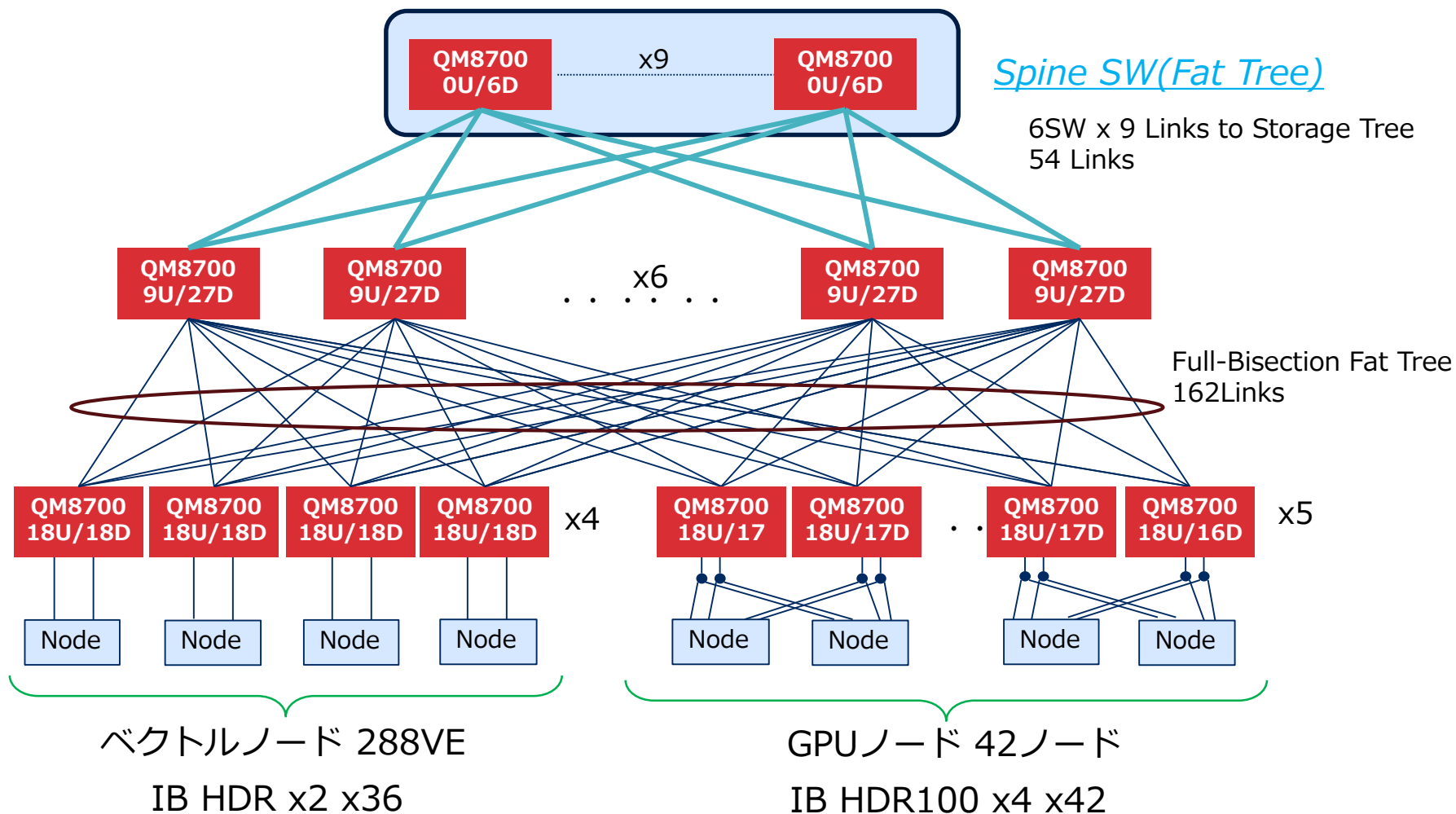
	高速領域	拡張領域	home領域
特徴	ALL Flashの ファイルシステムで 高速なIOが可能	グループ単位で容量の 追加購入が可能	ユーザ単位での最低限 の利用領域
容量	(申請により追加)	5TiB+(申請により追加)	10GiB
ファイル システム	SSD(Lustre)	HDD(Lustre)	
総容量	1.2 PB	20 PB	
ハードウェア	DDN ES400NVX x5 	DDN ES7990X x10 	
OSS数/OST数	20 / 40	20 / 80	
ディスク	15.36TB NVMe SSD	16TB 7,200rpm NL-SAS	

※ 内容は今後変更になる可能性があります。

ネットワーク環境

SQUIDでは、InfiniBand HDR (200Gbps) の相互接続網がいただけます

- 各計算環境内は、ベクトル/GPUノード内は全て、CPUノード内は513ノード毎にフルバイセクションバンド幅が保証されます。



4. SX-Aurora TSUBASA のご紹介

SX-2

1983



Technology: Bipolar
CPU Frequency: 166 MHz
CPU Performance: 1.3 GFlops
CPU Memory Bandwidth: 10.7 GB/sec

SX-3

1989



Technology: Bipolar
CPU Frequency: 340 MHz
CPU Performance: 5.5 GFlops
CPU Memory Bandwidth: 12.8 GB/sec

SX-4

1994



Technology: 350 nm
CPU Frequency: 125 MHz
CPU Performance: 2.0 GFlops
CPU Memory Bandwidth: 16.0 GB/sec

SX-5

1998



Technology: 250 nm
CPU Frequency: 250 MHz
CPU Performance: 8.0 GFlops
CPU Memory Bandwidth: 64.0 GB/sec

SX-6

2001



Technology: 150 nm
CPU Frequency: 500 MHz
CPU Performance: 8.0 GFlops
CPU Memory Bandwidth: 32.0 GB/sec

SX-7

2002



Technology: 150 nm
CPU Frequency: 552 MHz
CPU Performance: 8.8 GFlops
CPU Memory Bandwidth: 35.3 GB/sec

SX-8

2004



Technology: 90 nm
CPU Frequency: 1.0 GHz
CPU Performance: 16.0 GFlops
CPU Memory Bandwidth: 64.0 GB/sec

SX-9

2007



Technology: 65 nm
CPU Frequency: 3.2 GHz
CPU Performance: 102.4 GFlops
CPU Memory Bandwidth: 256.0 GB/sec

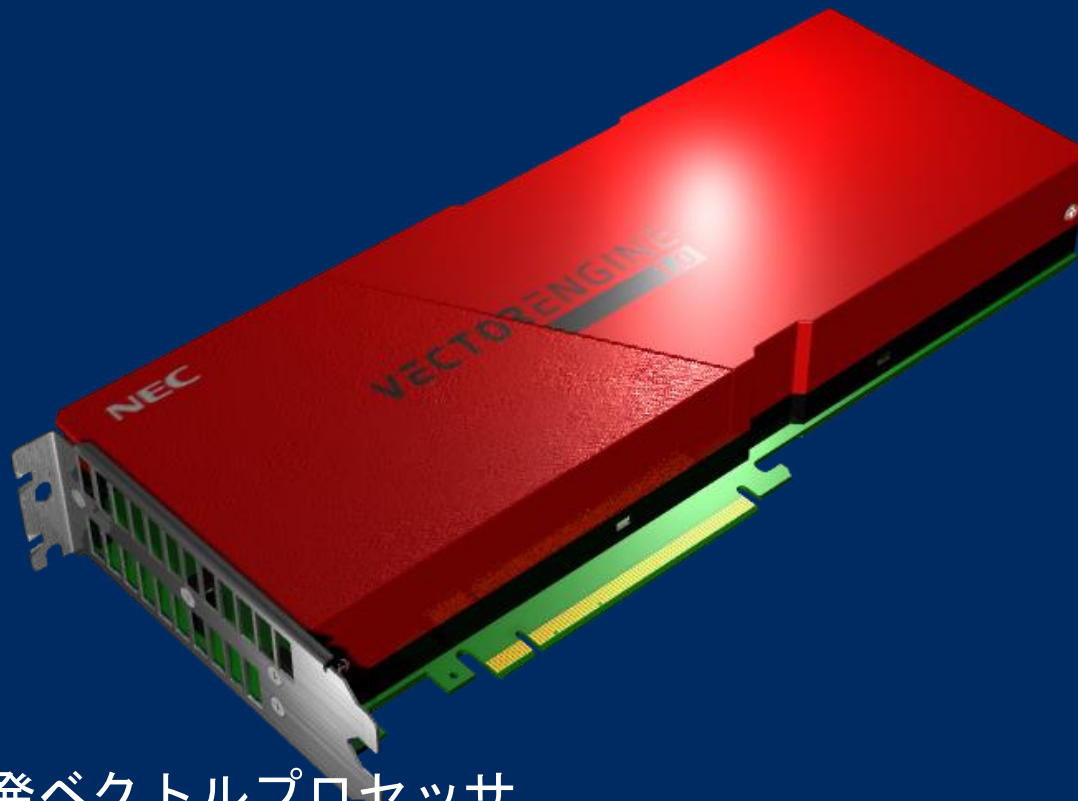
SX-ACE®

2013



Technology: 28 nm
CPU Frequency: 1.0 GHz
CPU Performance: 256.0 GFlops
CPU Memory Bandwidth: 256.0 GB/sec

**Over 30 years
Experience
For
High Sustained
Performance**



- 新規開発ベクトルプロセッサ
- アクセラレータ型とは一線を画すPCIeカード実装
- プロセッサあたり10コア
- 3.07TF 性能
- 1.53TB/s メモリバンド幅
- Fortran/C/C++ による標準的なプログラミングをサポート

※性能値はType 20Aの値です。

SX-Aurora TSUBASAの特長

科学技術の進展と産業の高度化を加速する
HPC（シミュレーション）／AI／ビッグデータ解析向け新プラットフォーム

SX-Aurora TSUBASA



POINT
1 高性能

POINT
2 使いやすさ

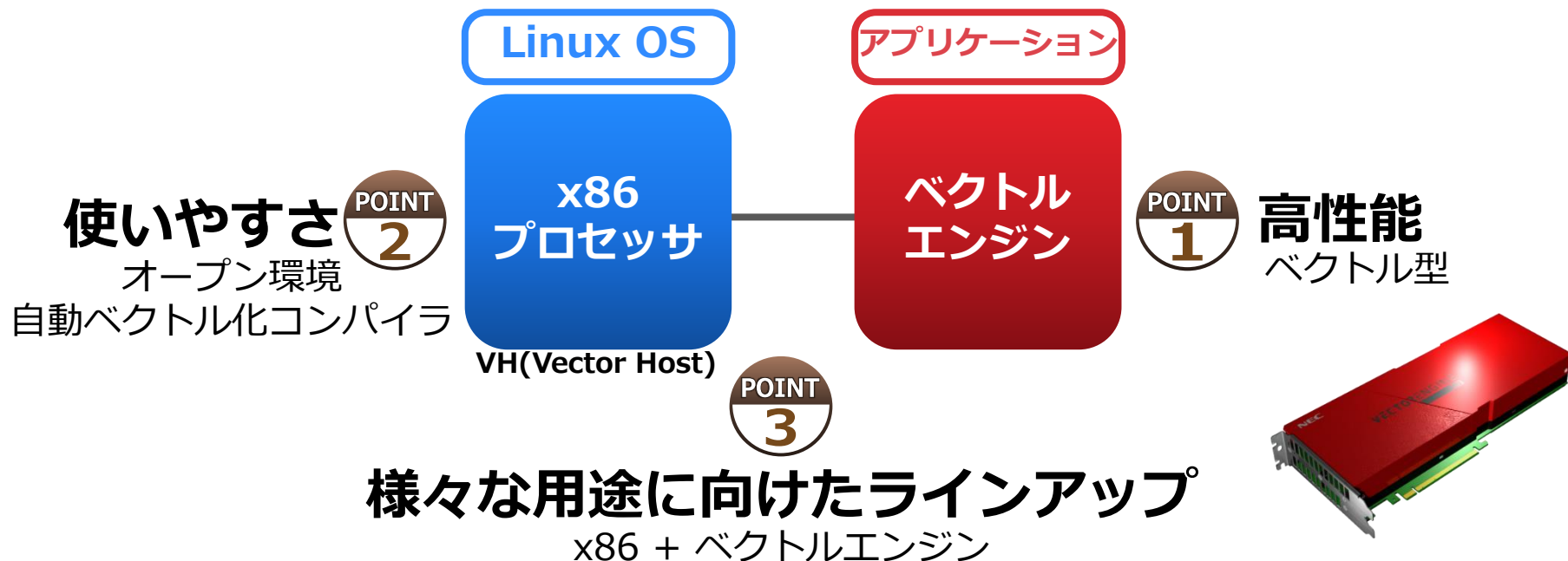
POINT
3 様々な用途に向けた
ラインアップ

新アーキテクチャ

ベクトルプロセッサとx86プロセッサを融合したハイブリッドアーキテクチャ採用

- ① ベクトルプロセッサの高性能を従来SXシリーズから継承
- ② Linux OSによるオープン化、自動ベクトル化コンパイラによる容易なプログラミング
- ③ x86 + ベクトルエンジン構成による豊富な製品ラインアップ

SX-Aurora TSUBASA

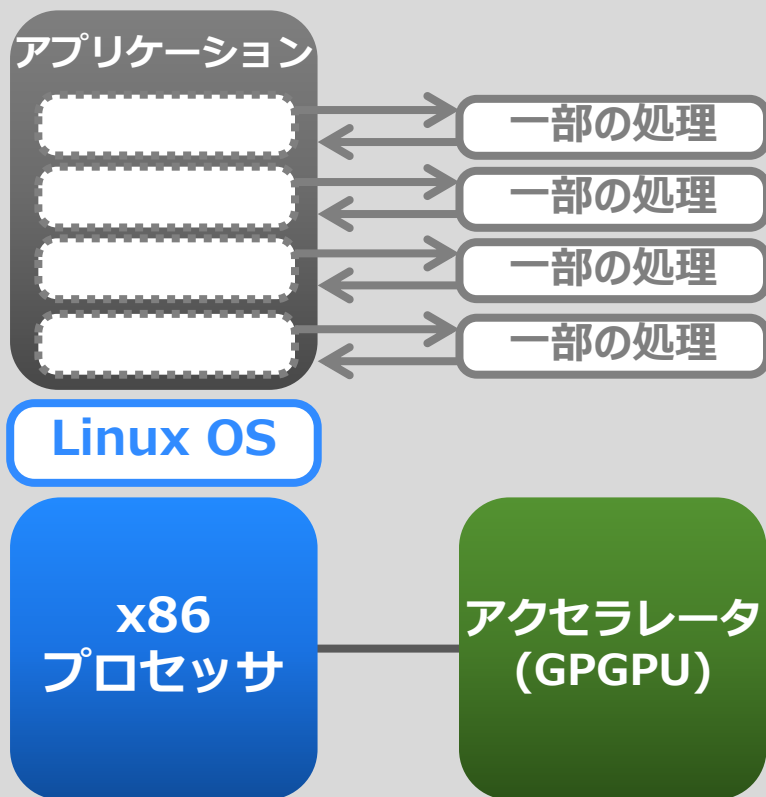


①高性能：新アーキテクチャによる高速化

■ GPGPUの持つ性能ボトルネックの解消

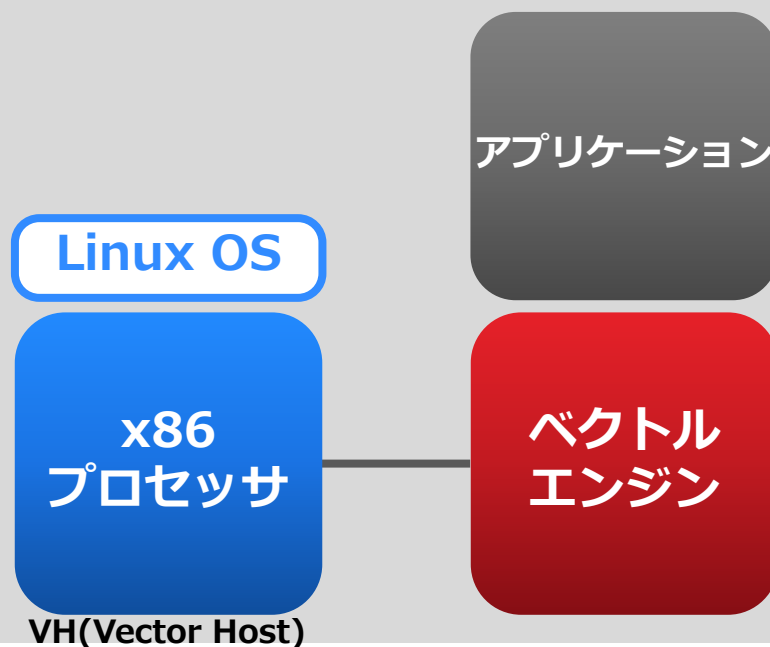
- アクセラレータ：プログラムの一部のみを処理。x86/アクセラレータ間の通信がネック
- SX-Aurora TSUBASA：プログラムを丸ごと実行

アクセラレータ型



SX-Aurora TSUBASA

プログラムを丸ごと
ベクトルエンジンで実行

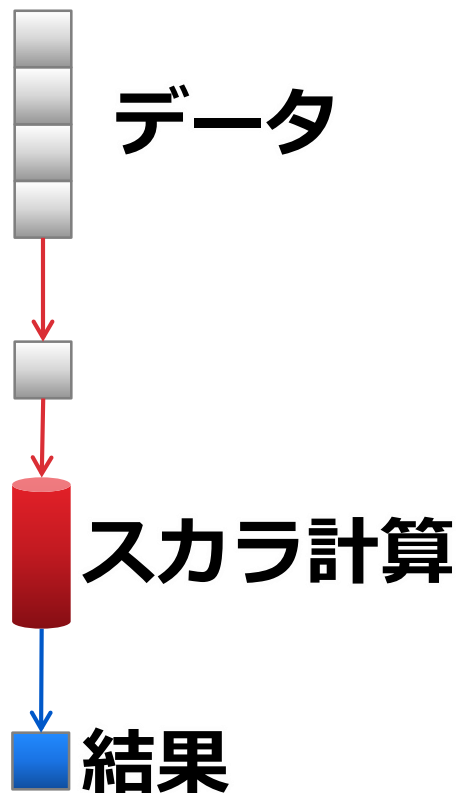


①高性能：ベクトルアーキテクチャ

ベクトル型プロセッサにより、一度に膨大な処理を実行。
大規模データの高速処理を実現

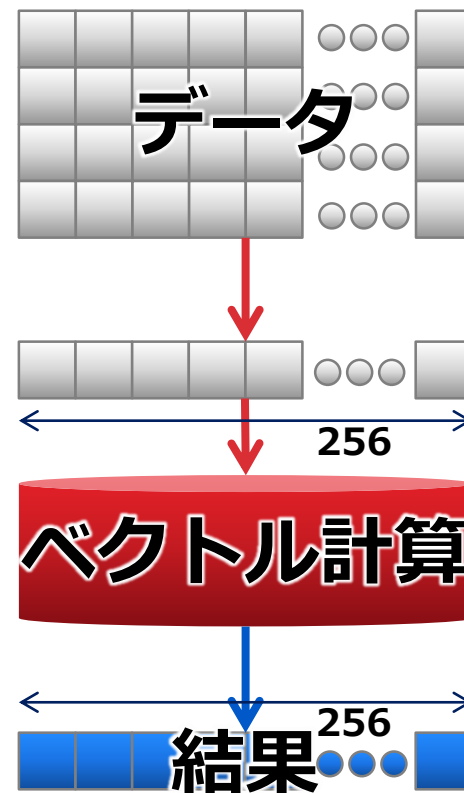
汎用プロセッサ

小さい単位でデータ処理を実行
業務処理・webサーバなど向け



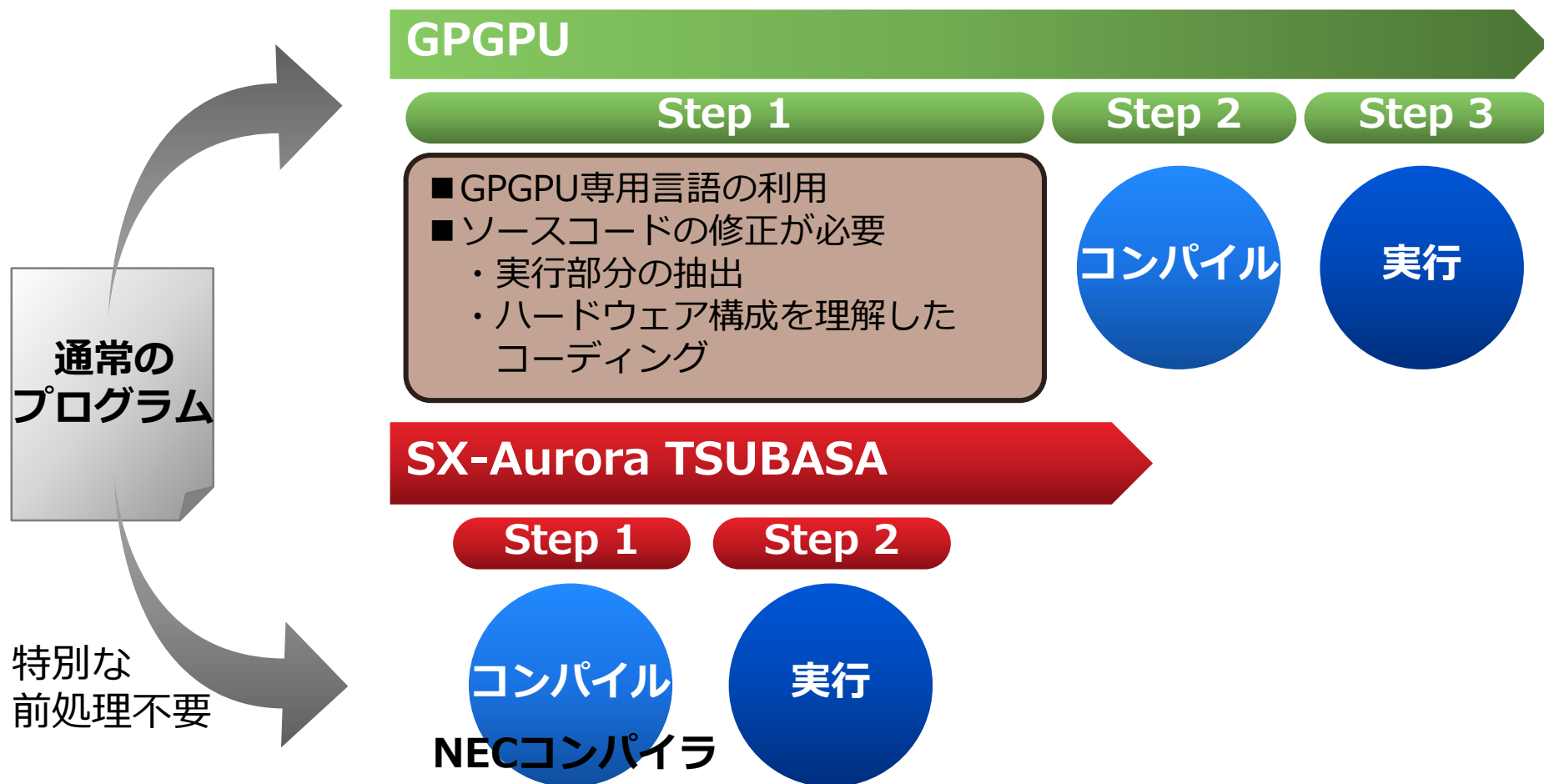
ベクトルプロセッサ

一度に膨大なデータ処理を実行
シミュレーション/AI/ビッグデータなど向け



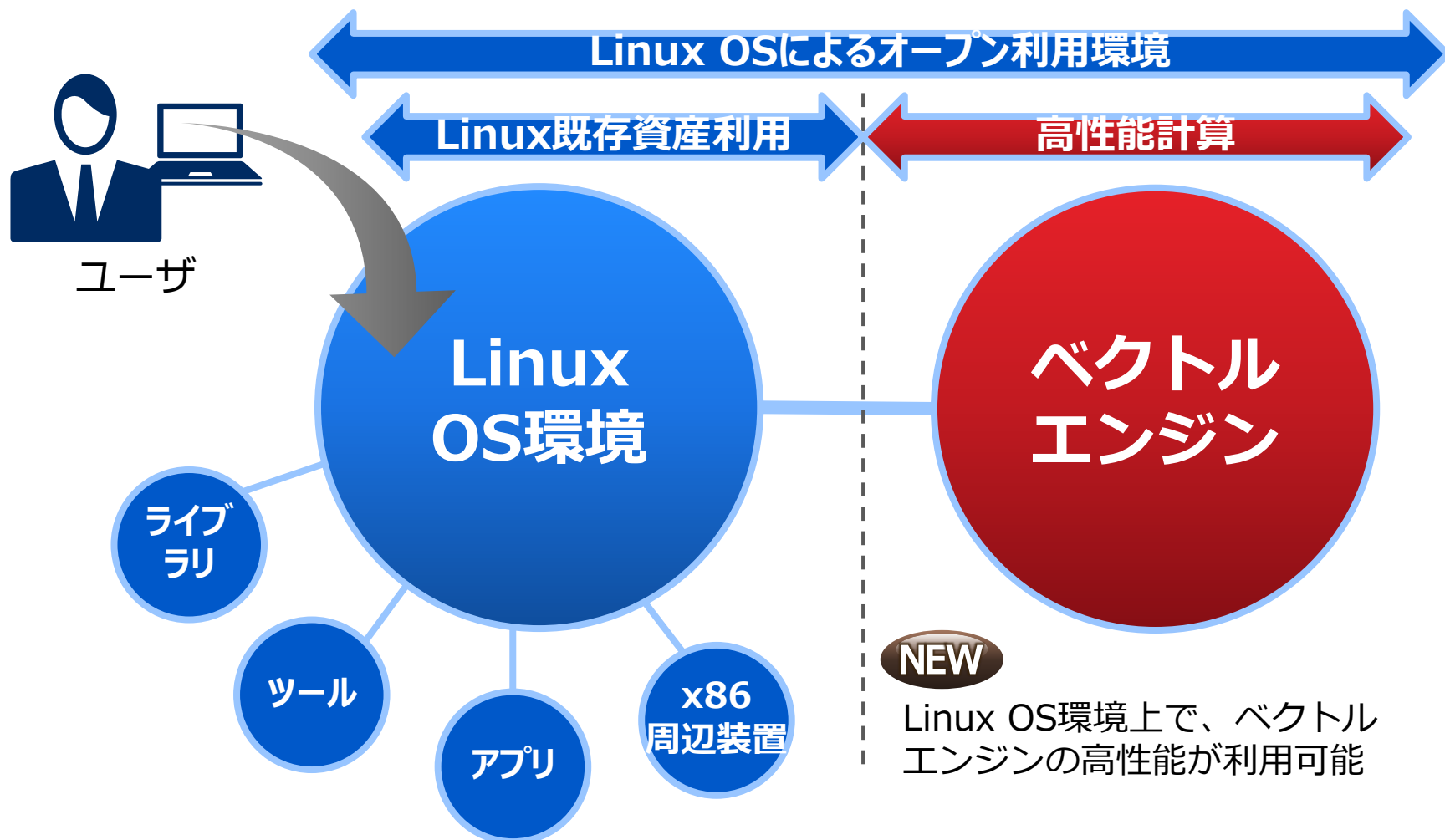
②使いやすさ：プログラミング

- GPGPUは、専用言語の利用、GPGPU実行部分の抽出、及びハードウェア構成を理解したソースコードの修正が必要
- SX-Aurora TSUBASAは自動ベクトル化コンパイラで使いやすさと高性能を両立



②使いやすさ：オープン環境への対応

従来SXシリーズの専用OSから、Linux OSに移行。オープン環境での豊富な資産を利用可能



③様々な用途に向けた豊富なラインアップ

ベクトルエンジンの超高性能を、デスクトップタワーから、大規模データセンター対応モデルまで幅広くご提供し、利用・適用範囲を大幅に拡大

データセンタモデル 高密度ラックモデル

データセンタ・計算センタでの巨大処理

例：大規模AI・ビッグデータ、
大規模シミュレーションなど

- 高性能・高スケーラビリティ・
低導入/低運用コスト

オンサイトモデル

製造業などのシミュレーション、
AI・ビッグデータ利用

例：大規模需要予測、
製造シミュレーションなど

エッジモデル

AI・ビッグデータシステム組込用途

例：店舗毎の需要予測、医療画像、
レーダ解析等

- プログラム開発者向け

A511シリーズ

64VE～搭載



A412シリーズ

2U8VE

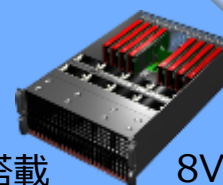
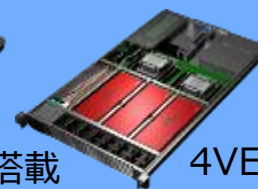


A311シリーズ

2VE搭載

4VE搭載

8VE搭載



A111シリーズ

1VE搭載



SX-Aurora TSUBASAのHW性能を最大限に引き出すための充実したベクトルプログラム開発環境(NEC SDK for VE)をご提供

コンパイラ

- Fortranコンパイラ
- C/C++コンパイラ
- HPF コンパイラ
- LLVM-IR Vectorizer

デバッグ支援

- GDB
- NEC Parallel Debugger
- strace

NEC MPI

- 最新MPI仕様(MPI3.1)準拠
- InfiniBand(HDR), Ethernet通信
- 最適なプロセス間通信手段の自動選択 & ゼロコピー転送
- Scalar-Vector Hybrid MPI

NEC Numeric Library Collection (数値計算ライブラリ集)

- BLAS/LAPACK/ScaLAPACK
- ASL/ASL FFTW3インタフェース
- Stencil Code Accelerator
- SBLAS/HeteroSolver

性能解析支援ツール

- PROGINF/FTRACE
- FtraceViewer
- gprof

NLCPy (Python対応ライブラリ)

- 基本演算/数学関数(Universal関数)
- 配列操作関数
- 乱数生成
- ソート

1) コンパイラ - Fortran Compiler, C/C++ Compiler

NEC Fortran Compiler, C/C++ Compiler

自動ベクトル化機能

- コンパイラがベクトル処理できるループ文を自動検出しベクトル化
- IF文のTHEN、ELSEブロック、リダクション演算などを含むループもベクトル化
- ベクトル処理の効率が増すように多重ループを変形

自動並列化機能

- 共有メモリ型並列処理できるループ・ブロックを自動検出し並列化
- 並列処理コード・逐次処理コードを実行時に選択する条件並列化など多彩な並列化手法を装備

対応する言語仕様

- Fortran
 - ISO/IEC 1539-1:2004 Fortran 2003 準拠
 - ISO/IEC 1539-1:2010 Fortran 2008 部分サポート (順次対応範囲を拡大予定)
 - CRAYポインタなど他社拡張言語仕様にも対応
- C/C++
 - ISO/IEC 9899:2011 C11 準拠
 - ISO/IEC 14882:2014 C++14 準拠
 - ISO/IEC 14882:2017 C++17 部分サポート (順次対応範囲を拡大予定)
 - GCC拡張言語仕様にも対応
- OpenMP
 - OpenMP 4.5 準拠業
 - 業界標準の共有メモリ型並列化機能をサポート

1) コンパイラ - Fortran Compiler, C/C++ Compiler

glibc 対応

■ glibc に対応したことで、広く使用されているOSSが簡単にコンパイルできるようになります。

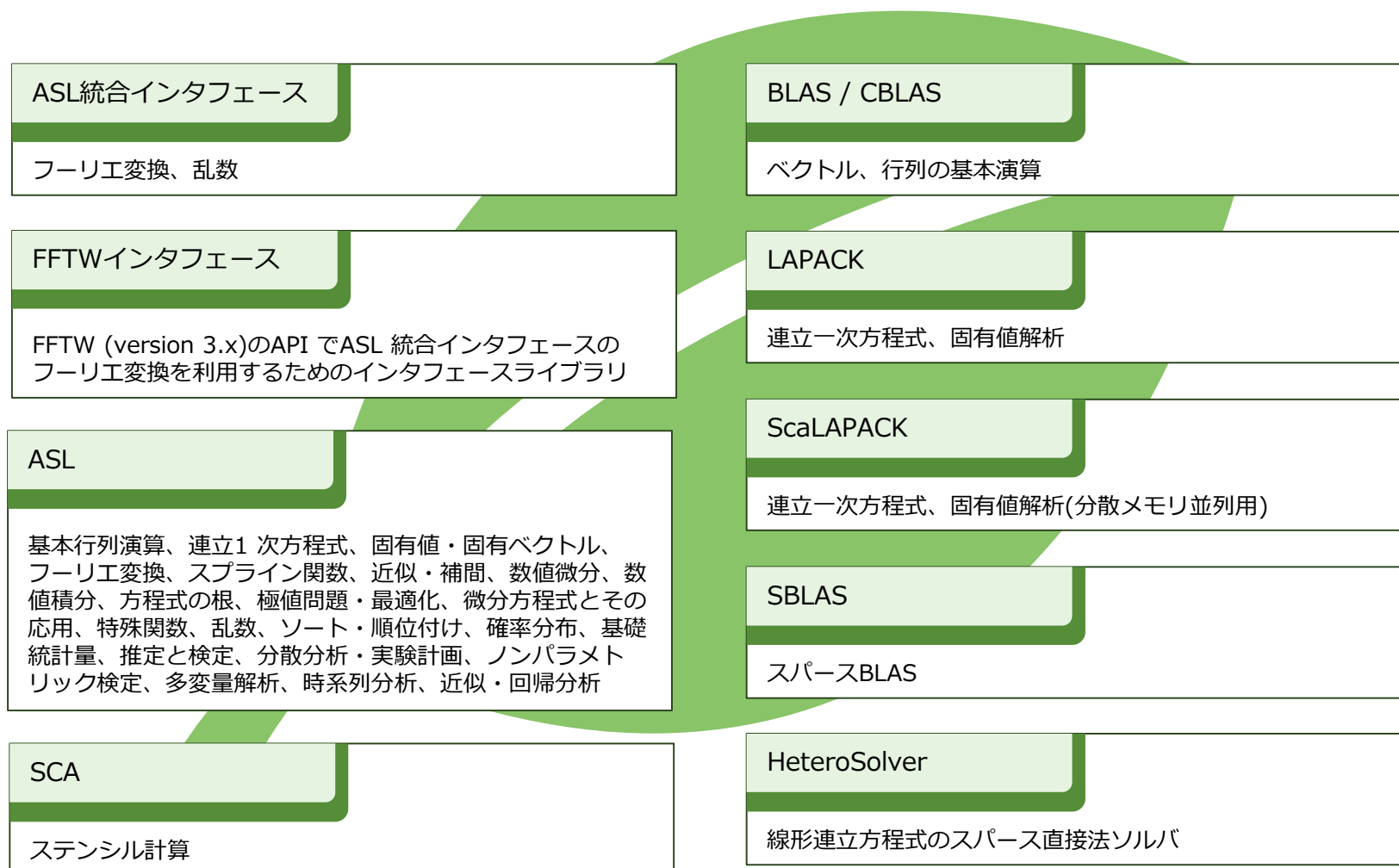
- NetCDF (Data I/O Library)
 - HDF5 (Data I/O Library)
 - WRF (Weather calculation)
 - VASP (first-principles calculation)
 - QuantumEspresso (first-principles calculation)
 - etc
- Makefileの簡単な修正 (NEC SDK for VE のコンパイラ、及びオプションの指定) をし、“configure && make” を実行することでコンパイルできます。
- SX-Aurora TSUBASAでの実行には、既存コードの再コンパイルが必要です。

The GNU C Library



2) NEC Numeric Library Collection

NEC Numeric Library Collectionは、以下のライブラリから構成されています。



2) NEC Numeric Library Collection

HW性能を最大化する科学技術計算ライブラリ

数値解析で頻出する計算を高速化

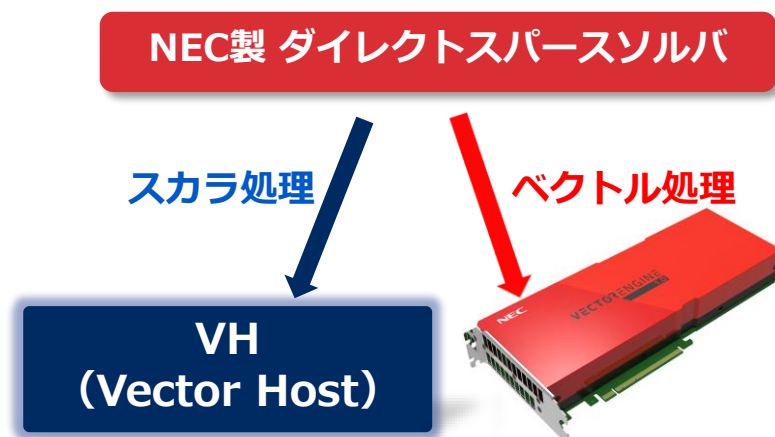
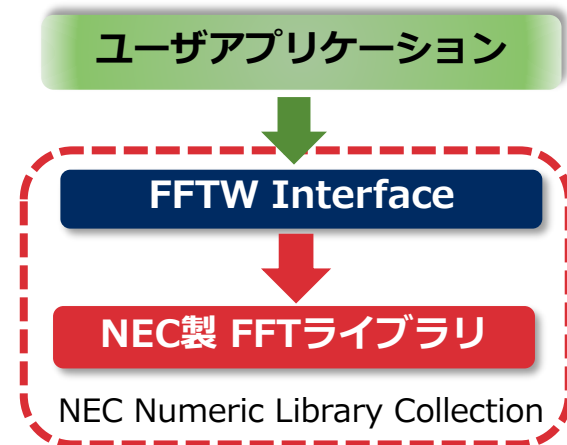
- BLAS, LAPACK, ScaLAPACK
- FFTW インタフェース
 - ・ヘッダファイルの変更のみでFFTW使用のプログラムを高速化
 - ・徹底的に性能追求したNEC製FFTライブラリをカーネルで使用

モダンな数値解析アルゴリズムで高性能を実現

- メルセンヌ・ツイスタ擬似乱数
 - ・高品質な乱数を高速に生成
- ソボル準乱数
 - ・準モンテカルロ計算に必須なソボル準乱数にも対応
- Stencil Code Accelerator
 - ・任意形状ステンシル、1～4次元データの計算を
Vector Engine (VE) 用に最適化

VE-VH間オフロードにより計算時間を削減

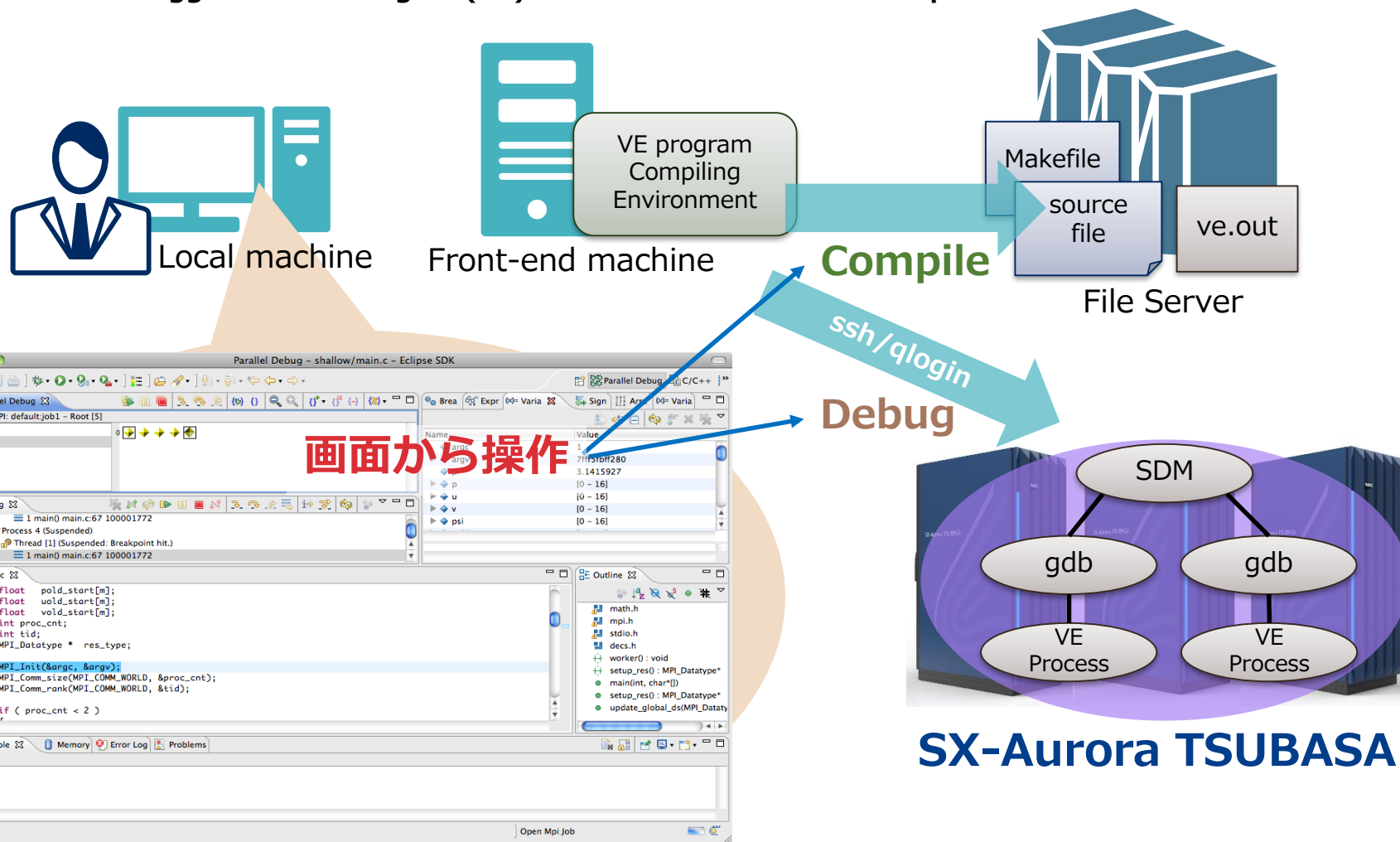
- ダイレクトスパースソルバ (HeteroSolver)
 - ・ベクトル向き／スカラ向きの処理をVE-VH間で役割分担
 - ・各アーキテクチャの長所を生かし、最速で求解



3) デバッグ支援 - NEC Parallel Debugger

VE対応GUIデバッガ(Eclipseプラグイン)

NEC Parallel DebuggerはVector Engine (VE) 対応のGUIデバッガであり、Eclipseプラグインとして使用可能です。

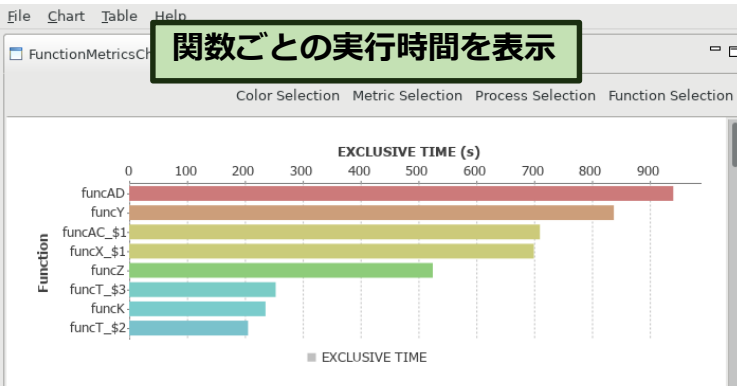


4) 性能解析支援ツール - NEC Ftrace Viewer

GUIベースの性能解析支援ツール

NEC Ftrace Viewerは、Vector Engine (VE) 向けのGUIベースの性能解析支援ツールです。

関数ごとの実行時間を表示



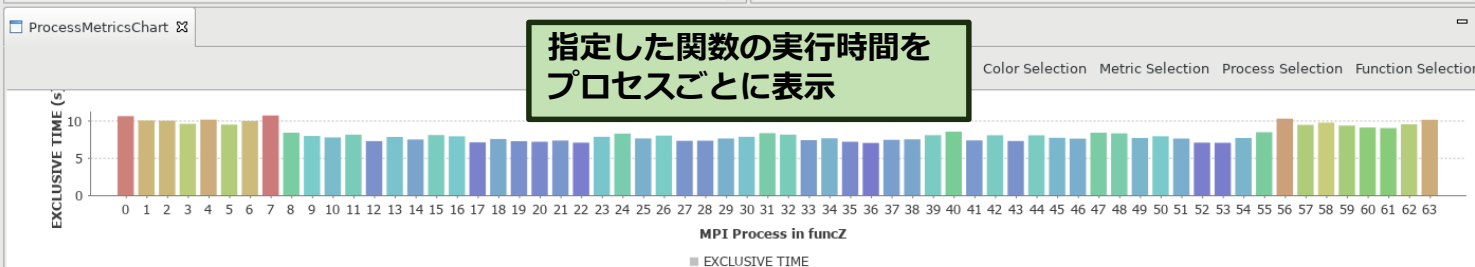
関数ごとに、プロセス間の実行時間のばらつき(最大・最少・標準偏差)を表示



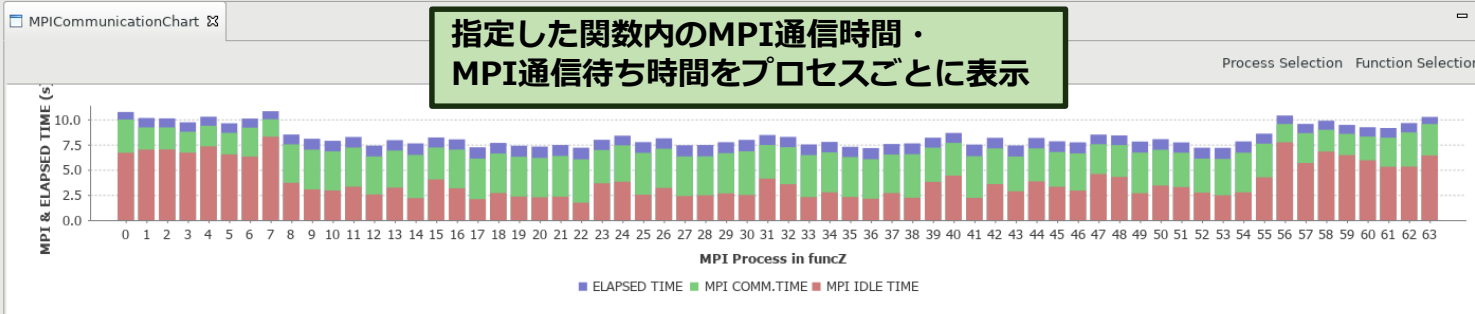
実行時間とベクトル化率を表形式で表示

PROC.NAME	EXCLUSIVE	V.OP.RATIO
Total	5242.59	95.39
funcAD	940.49	92.71
funcY	837.80	93.09
funcAC_\$1	710.00	98.19
funcX_\$1	699.58	98.14
funcZ	524.82	80.28
funcT_\$3	253.26	97.09
funcK	235.69	0.00
funcT_\$2	205.44	96.76
funcS	204.99	97.13
funcO	161.75	96.30
funcL	153.46	0.00
funcP_\$1	44.88	0.00
funcP_\$3	42.03	0.00
funcR_\$3	38.00	95.90
funcT_\$1	27.67	95.71
funcQ	27.30	0.18
funcAF	25.62	7.48
funcR_\$2	24.23	94.74
funcW	22.26	0.00
funcAB	21.72	0.00
funcA	16.14	56.04
funcP_\$5	8.75	0.00
funcP_\$7	8.56	0.00
funcP_\$4	1.26	91.91
funcP_\$3	0.99	97.06
funcB	0.97	0.51
funcU	0.89	52.88
funcP_\$2	0.80	92.08
funcM	0.64	97.13
funcP_\$8	0.63	95.08
funcN_\$2	0.57	97.00
funcP_\$6	0.47	95.08
funcJ	0.39	0.00

指定した関数の実行時間をプロセスごとに表示



指定した関数内のMPI通信時間・MPI通信待ち時間をプロセスごとに表示



5) NEC MPI

VE間のMPI通信を可能にし、大規模実行を可能にするMPIライブラリ

最新MPI仕様（MPI3.1）準拠

インターコネクトとしてInfiniBand（HDR）、Ethernetをサポート

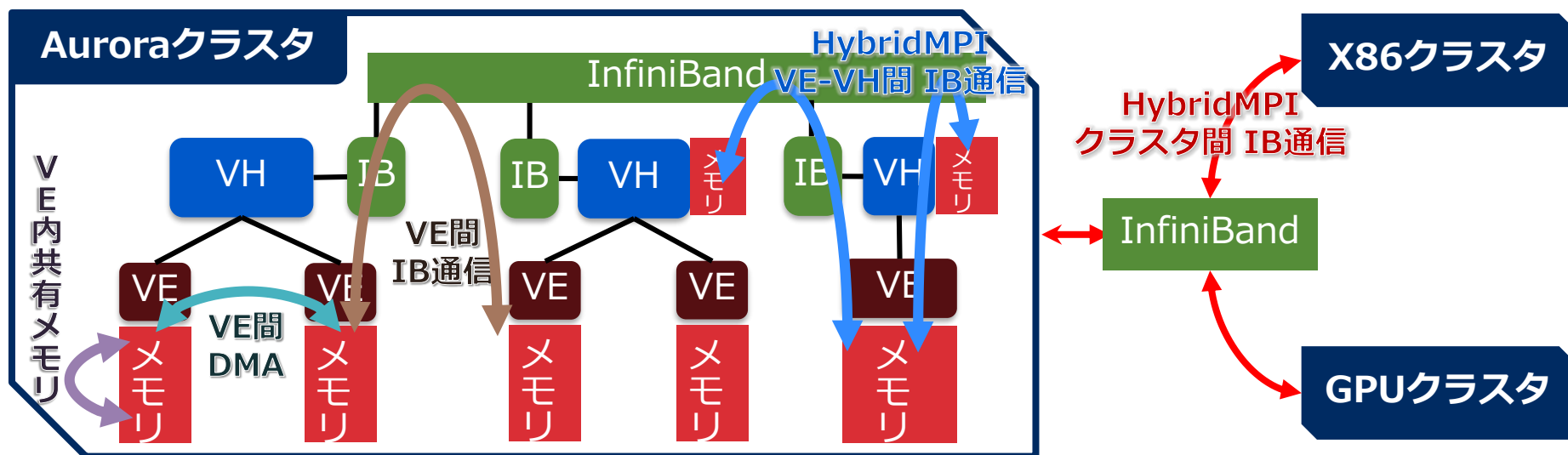
最適なプロセス間通信手段の自動選択&ゼロコピー転送による高速通信

- プロセス配置・HWトポロジ・転送サイズ等に応じて、最適な通信プロトコルと通信種別（共有メモリ/DMA/InfiniBand/Ethernet）を自動選択

HybridMPIにより計算処理の特性に合わせたリソース選択が可能

- VE-VHに跨ったアプリケーションの連携実行
- Auroraクラスタ/x86クラスタ/GPGPUクラスタに跨ったアプリケーションの連携実行

InfiniBand Adaptive Routingサポートによる高スループット通信



6) NLCPy (Python対応ライブラリ)

Numpy Python モジュールと互換性を持つ VE向けライブラリ

NLCPy

- NumPyスクリプトのモジュール名の変更だけでVector Engine (VE) の演算能力を利用
- 2020/8月プレビューリリース、GitHubおよびPyPIで公開済み
- 2021/4月正式リリース予定、FFTと線形代数機能をサポート

```
import numpy as np
```

```
a = np.random.rand(1000, 1000)  
b = np.random.rand(1000, 1000)
```

```
c = a + b  
d = np.dot(a, b)
```

Script

Linux/x86

```
import nlcpy as np
```

```
a = np.random.rand(1000, 1000)  
b = np.random.rand(1000, 1000)
```

```
c = a + b  
d = np.dot(a, b)
```

Script

Vector Host
(Linux/x86)

Library

Vector
Engine

Data transfer between VH and VE

Computation
on VE

5. アプリケーション事例

性能情報について

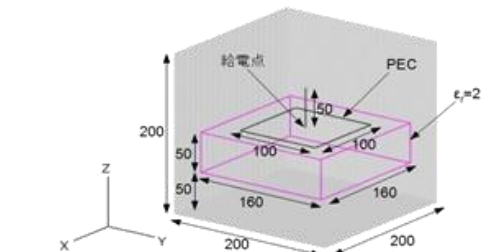
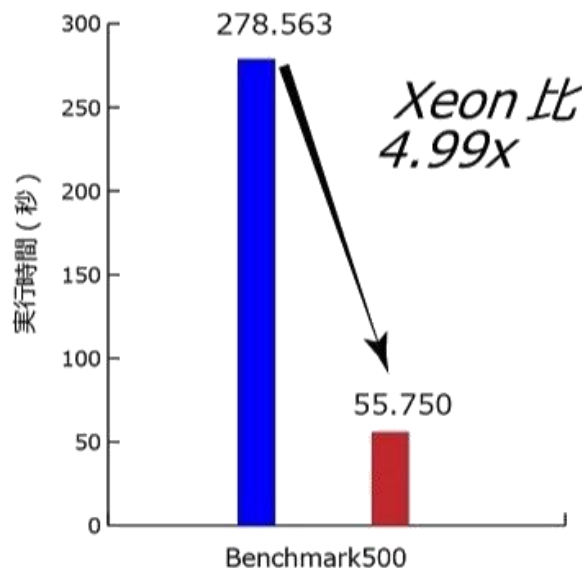
- 市販のアプリケーションをSX-Aurora TSUBASA 上で実行した際の、性能改善の事例を載せています。
- 下記サイトにも、事例を掲載していますので、合わせてご確認ください。
<https://jpn.nec.com/hpc/sxauroratsubasa/Application/index.html>

アプリケーション概要

- Maxwell方程式を、差分法により数値計算するFDTD法による電磁界シミュレータ
- 電子機器設計などで、最適な設計パターンの効率的探索、設計妥当性や現象確認に活用

高速化

- FDTD法の差分計算はベクトル化に適し、メモリアクセスがボトルネックになります。
- 高速メモリアクセス性能をもつSX-Aurora TSUBASAで、高速実行が可能です。
Xeonサーバでの計算に比べて、**解析計算時間が「約5分の1」に短縮**されます。



誘電体のブロックの上に完全導体（PEC）の板があり、
その中心にモノポールアンテナが立っている問題

【測定環境】

- Xeon Gold6148/2 ソケット /40 コア
 - SX-Aurora TSUBASA A300-8 Type10B/1VE/8 コア
- Benchmark500(500x500x500) で測定

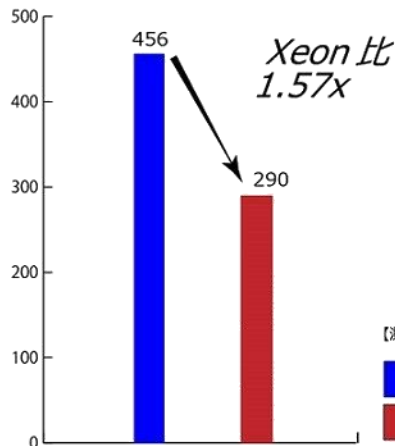
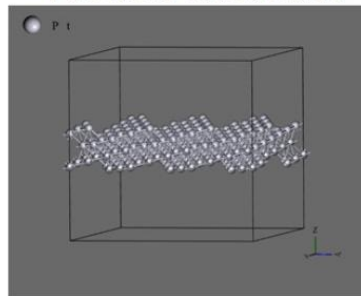
アプリケーション概要

- PWscf (Plane-Wave Self-Consistent Field) の電子状態計算をベースとした材料モデリングのためのOSS
- 基本プログラム以外に多数のパッケージ/プラグインを含み、多種多様な物理量が計算可能

高速化

- 第一原理計算の基本となるアルゴリズムは行列演算や高速フーリエ変換であり、ベクトルに適します。Xeonサーバの計算に比べ、**解析計算時間が「約2分の1」に短縮**されます。

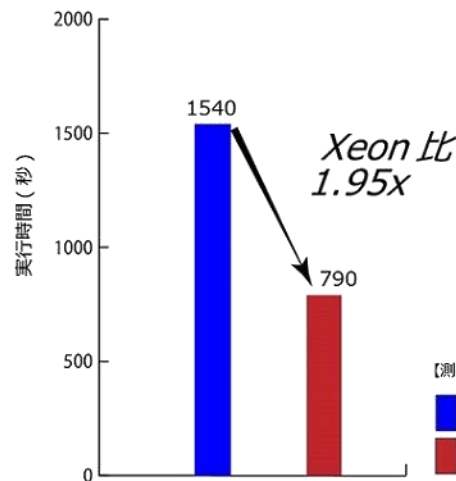
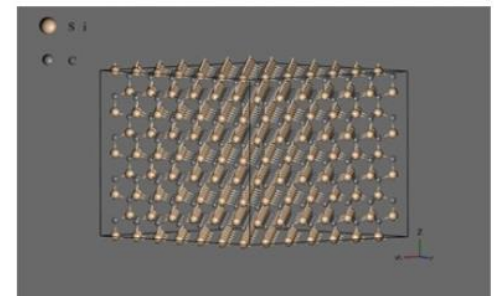
Pt(234)スラブモデル (216原子)



【測定環境】

- Intel Xeon Gold5218/2 ソケット / 32 コア
- SX-Aurora TSUBASA A300-8 Type10B/1VE/8 コア

SiCスーパーセルモデル (784原子)



【測定環境】

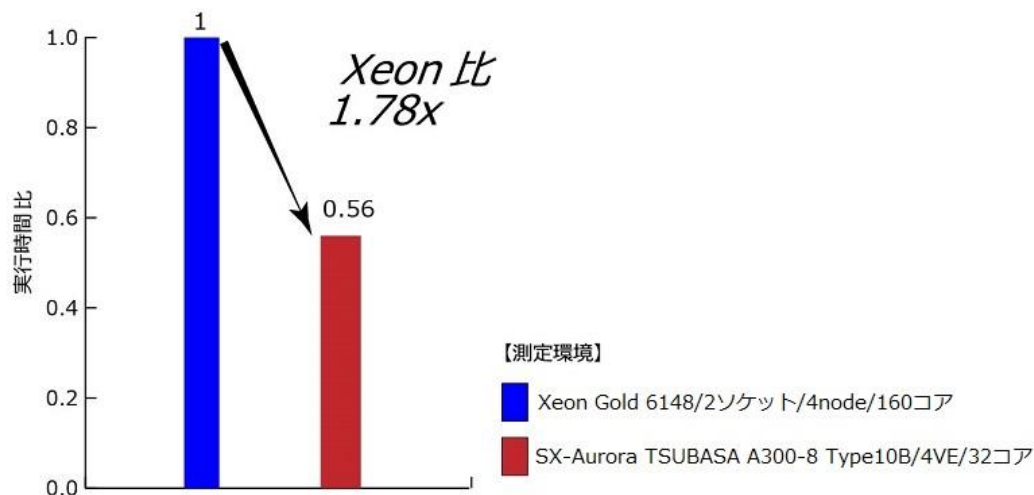
- Intel Xeon Gold5218/2 ソケット / 32 コア
- SX-Aurora TSUBASA A300-8 Type10B/1VE/8 コア

アプリケーション概要

- 第一原理量子力学計算を行うための計算化学ソフトウェア
- 電池、セラミックス、触媒、金属、半導体などの材料設計のために、世界中の1400以上の研究機関で広く使用されています

高速化

- 第一原理計算の基本となるアルゴリズムは行列演算や高速フーリエ変換であることからベクトル向きです。演算性能が最大1.78倍向上したことにより、Xeonサーバでの計算に比べて**解析計算時間が約45%短縮**されます。



※本SWはISV(有償)製品です。SX-Aurora版のVASPの提供は、お客様ご自身でVASPのライセンスをお持ちになっていることが条件になります。

 **Orchestrating** a brighter world

NEC