



Enable AI & HPC to be Open, Safe and Accessible to All

# oneAPI と SYCL\* によりマルチベンダー・ アーキテクチャーをターゲットにする

Codeplay Software  
Rod Burns

HPSC 2023

## 会社概要

新しい AI 処理システム向けのハイパフォーマンス・ソフトウェア・ソリューションを実現するリーディング・カンパニー

オープン標準ベースのツールとミドルウェアで、最も強力なプロセッサを実現

2002年にスコットランドで設立、2022年にインテルに買収され、現在の従業員数は約 90 名



オープンで安全な  
誰でもアクセス  
できる AI と HPC  
を実現

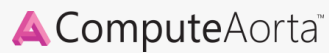
## サポートされるソリューション



アクセラレーター・アーキテクチャーを問わず、共通の開発者体験を提供する、業界を超えたオープンな SYCL\* ベースの統一された、マルチアーキテクチャー、マルチベンダーのプログラミング・モデル



TensorFlow\* などの画像処理とマシンラーニングに対応した SYCL\* オープン標準を利用した C++ プラットフォーム



OpenCL\*, SPIR-V\*, HSA\*, Vulkan\* に対応した Codeplay のコンピュート・テクノロジーの中核

## コラボレーション



SYNOPSYS®



ほか多数

## 市場

ハイパフォーマンス・コンピューティング (HPC)、  
車載用 ADAS、IoT、  
クラウド・コンピューティング、  
スマートフォンとタブレット、  
医療、産業、

テクノロジー: 人工知能 (AI)、  
画像処理、  
マシンラーニング、  
ビッグデータ・コンピューティング

# Codeplay について

- **Codeplay Software** は、長年にわたり**インテル**と共同でオープン標準に貢献してきましたが、2022年の買収により**インテルの子会社**となりました
- 引き続き、Codeplay Software として活動し、**SYCL\* エコシステム**、特に **oneAPI** を推進するため、**すべての関連業界**と幅広く協力していきます
- Codeplay は現在、インテルと共同で、**SYCL\* 標準**と **oneAPI** オープン・エコシステムをさらに進化させるべく取り組んでいます



# オープン標準のマルチベンダー・プログラミング

- オープン標準は自由と選択肢をもたらす
- oneAPI と SYCL\* はパフォーマンスの移植性を実現
- 使い慣れたハードウェアをオープン標準で使用
- 使用方法



Enable AI & HPC to be Open, Safe and Accessible to All

# オープン標準は 自由と選択肢をもたらす

HPSC 2023

# ベンダーによる縛りをなくす

## メリット

- ベンダーに関係なく、最高のハードウェアを使用
- ハードウェアの価格交渉を有利に進められる

解決策は、自由でオープン標準に準拠した製品を使用すること



出典: [https://en.wikipedia.org/wiki/Hazard\\_\(golf\)#/media/File:Road\\_hole\\_bunker.jpg](https://en.wikipedia.org/wiki/Hazard_(golf)#/media/File:Road_hole_bunker.jpg)

# C++ とオープン標準

- C++ が帰ってきた



🏆 Language of the Year: 2003, 2022

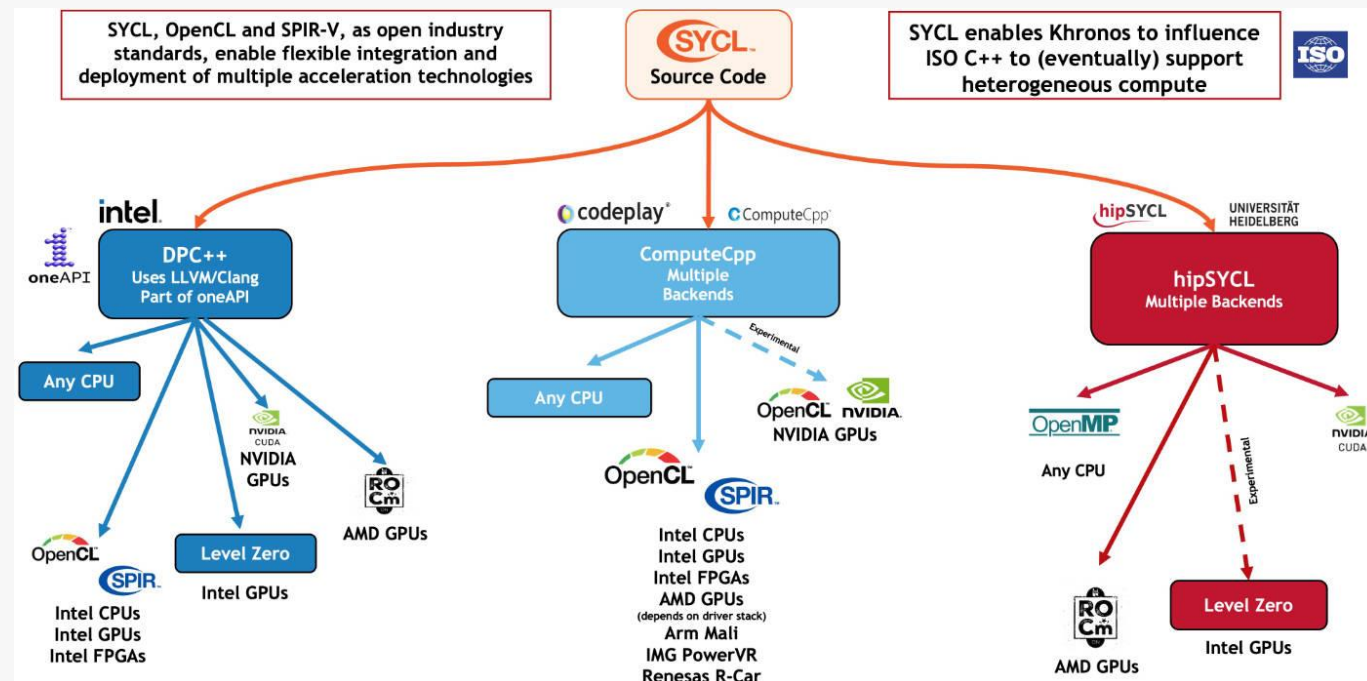
「上位 20 言語の中で最も急速に成長」

[www.tiobe.com](http://www.tiobe.com)

- ハイパフォーマンス・コンピューティング・アプリケーションの開発に最適
- ISO 準拠のオープン標準仕様ベース

# SYCL\* について

- オープン標準仕様
- 標準 C++ のみを使用
- 並列実行が可能
- 多様なハードウェアに対応





# oneAPI と SYCL\*



- SYCL\* は oneAPI の中核
- オープンソース・コンパイラー・プロジェクト
- 共通ライブラリー

# SYCL\* と oneAPI は選択肢を提供

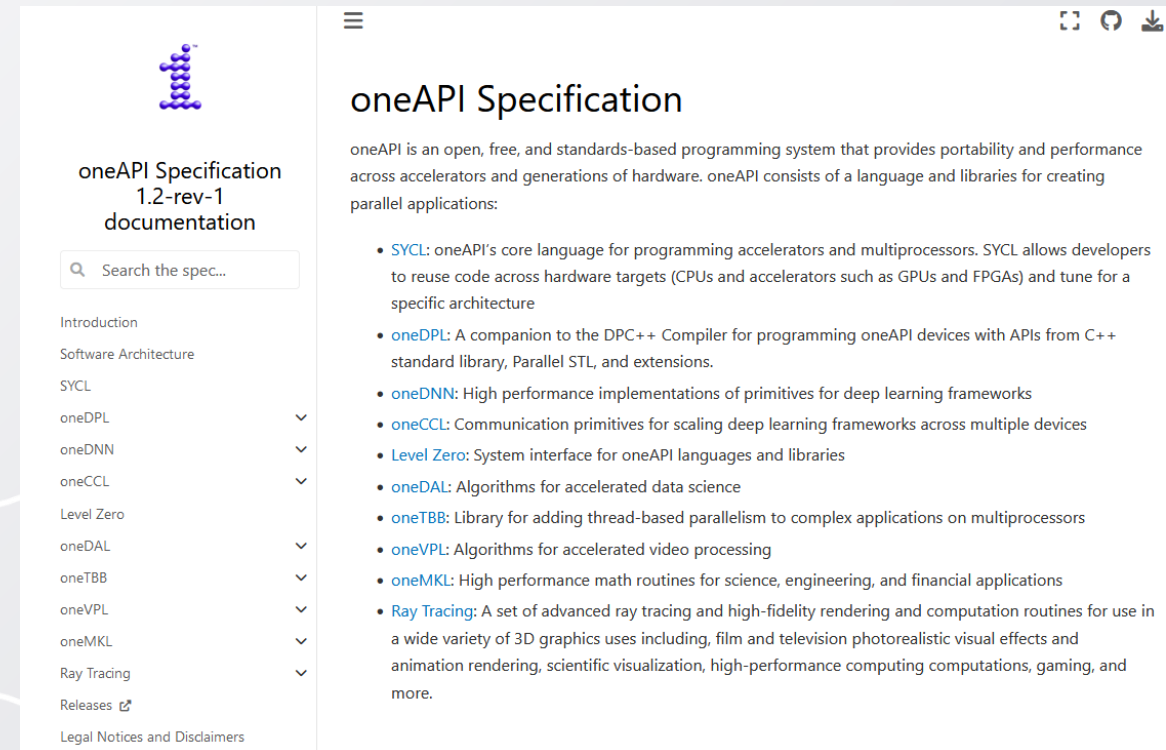
- 最速のスーパーコンピューターに導入されている
- 現在および次世代のスーパーコンピューターで動作可能なソフトウェア



# oneAPI オープン標準

## oneAPI 仕様

- SYCL\* 標準への依存
- 共通ライブラリーの定義
- オープンソース実装がある

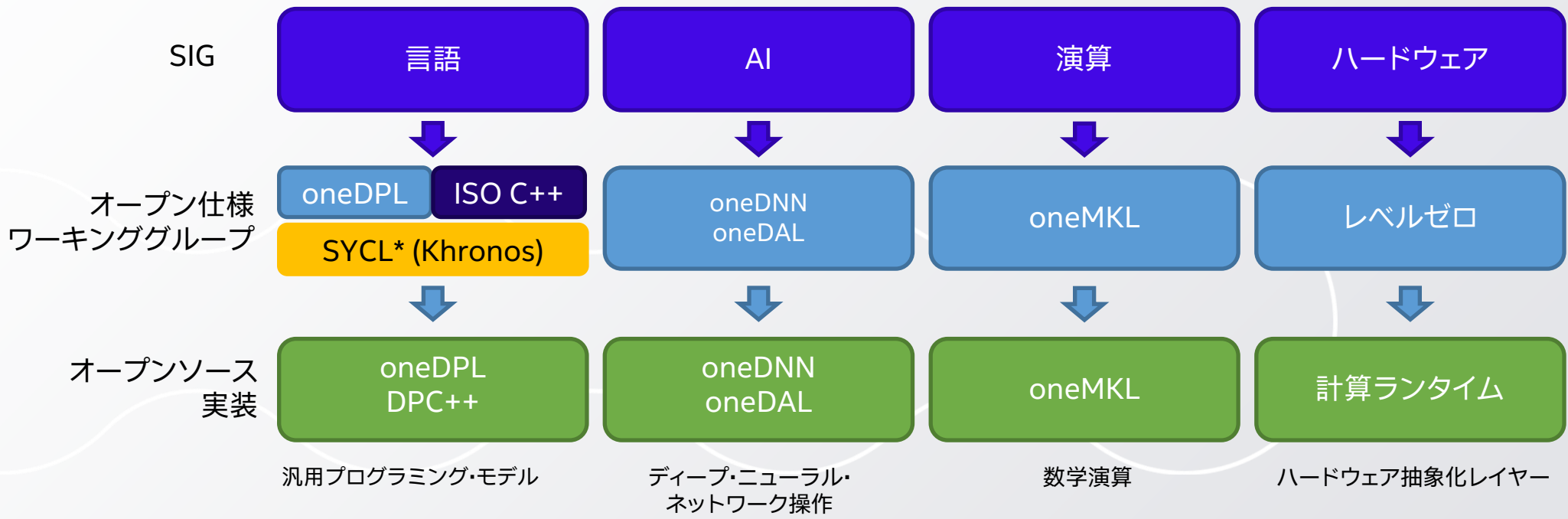


The screenshot shows the 'oneAPI Specification 1.2-rev-1 documentation' website. The page features a navigation menu on the left with items like Introduction, Software Architecture, SYCL, and various oneAPI libraries (oneDPL, oneDNN, oneCCL, Level Zero, oneDAL, oneTBB, oneVPL, oneMKL, Ray Tracing, Releases, and Legal Notices). The main content area is titled 'oneAPI Specification' and includes an introductory paragraph and a bulleted list of key components:

- **SYCL**: oneAPI's core language for programming accelerators and multiprocessors. SYCL allows developers to reuse code across hardware targets (CPUs and accelerators such as GPUs and FPGAs) and tune for a specific architecture
- **oneDPL**: A companion to the DPC++ Compiler for programming oneAPI devices with APIs from C++ standard library, Parallel STL, and extensions.
- **oneDNN**: High performance implementations of primitives for deep learning frameworks
- **oneCCL**: Communication primitives for scaling deep learning frameworks across multiple devices
- **Level Zero**: System interface for oneAPI languages and libraries
- **oneDAL**: Algorithms for accelerated data science
- **oneTBB**: Library for adding thread-based parallelism to complex applications on multiprocessors
- **oneVPL**: Algorithms for accelerated video processing
- **oneMKL**: High performance math routines for science, engineering, and financial applications
- **Ray Tracing**: A set of advanced ray tracing and high-fidelity rendering and computation routines for use in a wide variety of 3D graphics uses including, film and television photorealistic visual effects and animation rendering, scientific visualization, high-performance computing computations, gaming, and more.

# スペシャル・インタレスト・グループ (SIG)

SIG は仕様や実装に影響を与える



# oneAPI コミュニティ・フォーラムに貢献する

- SIG やワーキンググループに参加してリードする
- 技術的な議論をリードする
- 機能および変更に関する提案を提出する
- 提案に投票する

ヘテロジニアス・アーキテクチャー向け  
プログラミングの未来を切り拓く

<https://oneapi.io/community> (英語)

[oneapi@codeplay.com](mailto:oneapi@codeplay.com)

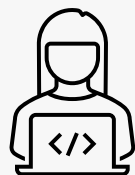


Enable AI & HPC to be Open, Safe and Accessible to All

# oneAPI と SYCL\* は パフォーマンスの移植性を実現

HPSC 2023

# oneAPI を NVIDIA\* GPU と AMD\* GPU 対応にする



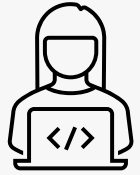
1つのコードを  
どこでも実行



パフォーマンスに  
妥協しない

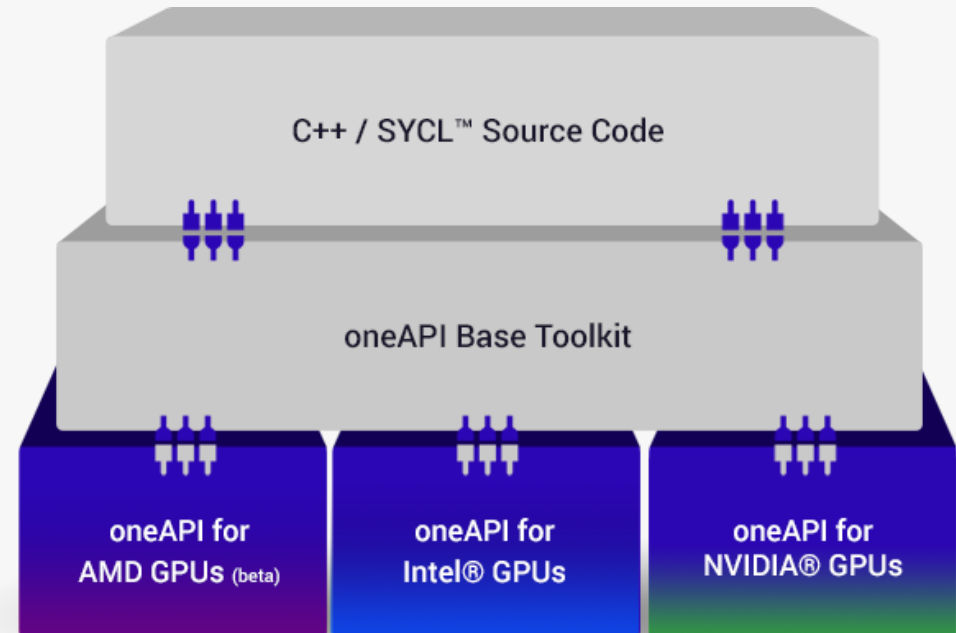


標準に関する  
オープンな業界を超えた  
コラボレーション



1つのコードを  
どこでも実行

**SYCL\*** でコードを記述して、  
インテル、NVIDIA、AMD の GPU で  
**自由に実行**







パフォーマンスに  
妥協しない

“

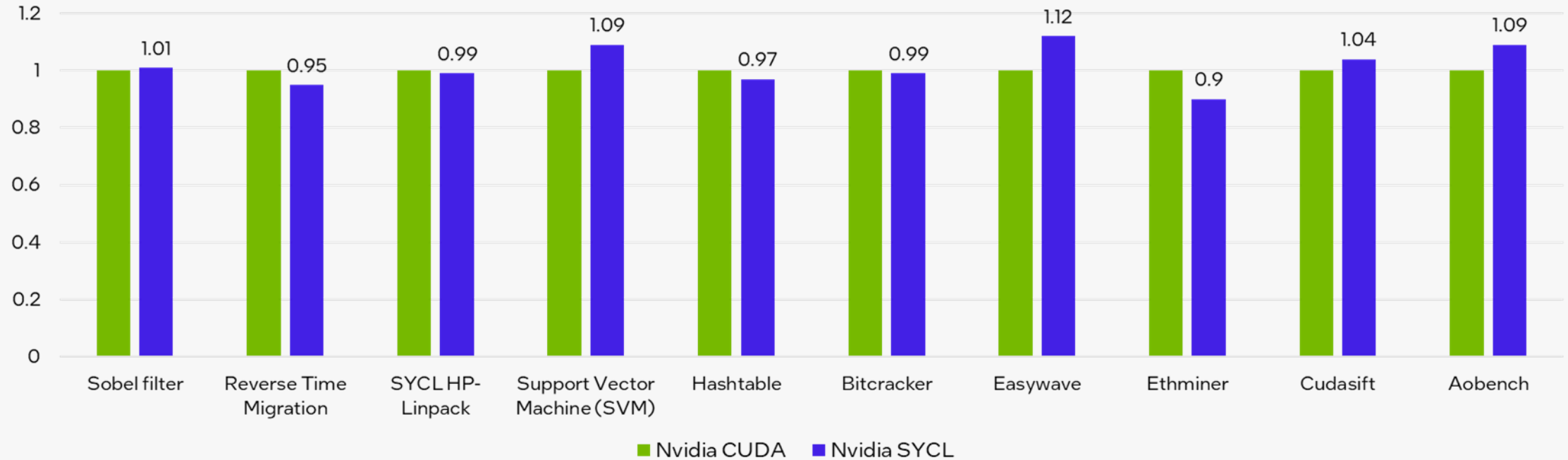
SYCL\* は NVIDIA と AMD のデバイス上で  
高速に動作し、多様なワークロードで  
CUDA\* や HIP\* のネイティブコードに  
匹敵するパフォーマンスを発揮

Codeplay  
CTO  
Ruyman Reyes

詳細: <https://codeplay.com/portal/blogs/2023/04/06/sycl-performance-for-nvidia-and-amd-gpus-matches-native-system-language> (英語)

性能の測定結果はシステム構成の日付時点のテストに基づいています。また、現在公開中のすべてのセキュリティ・アップデートが適用されているとは限りません。  
詳細は、システム構成を参照してください。絶対的なセキュリティを提供できる製品またはコンポーネントはありません。性能は、使用状況、構成、その他の要因によって異なります。  
詳細については、<http://www.intel.com/PerformanceIndex/> (英語) を参照してください。実際の費用と結果は異なる場合があります。

## Relative Performance: Nvidia SYCL vs. Nvidia CUDA on Nvidia-A100 (CUDA = 1.00) (Higher is Better)



**Testing Date:** Performance results are based on testing by Intel as of August 15, 2022 and may not reflect all publicly available updates.

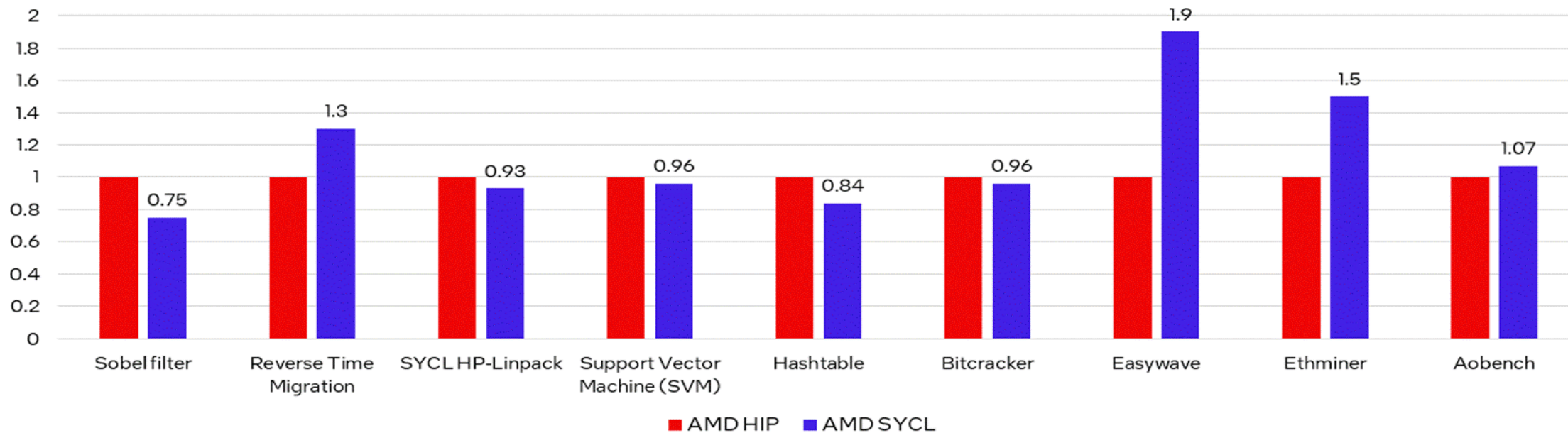
**Configuration Details and Workload Setup:** Intel® Xeon® Platinum 8360Y CPU @ 2.4GHz, 2 socket, Hyper Thread On, Turbo On, 256GB Hynix DDR4-3200, ucode 0x000363. GPU: Nvidia A100 PCIe 80GB GPU memory. Software: SYCL open source/CLANG 15.0.0, CUDA SDK 11.7 with NVIDIA-NVCC 11.7.64, cuMath 11.7, cuDNN 11.7, Ubuntu 22.04.1. SYCL open source/CLANG compiler switches: -fscycl-targets=nvptx64-nvidia-cuda, NVIDIA NVCC compiler switches: -O3 -gencode arch=compute\_80, code=sm\_80. Represented workloads with Intel optimizations.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more at [www.intel.com/PerformanceIndex](http://www.intel.com/PerformanceIndex). Your costs and results may vary.

# Relative Performance: AMD SYCL vs. AMD HIP on AMD GPU

Relative Performance: AMD SYCL vs. AMD HIP on AMD Instinct MI100 Accelerator  
(HIP = 1.00)  
(Higher is Better)



**Testing Date:** Performance results are based on testing by Intel as of August 15, 2022 and may not reflect all publicly available updates.

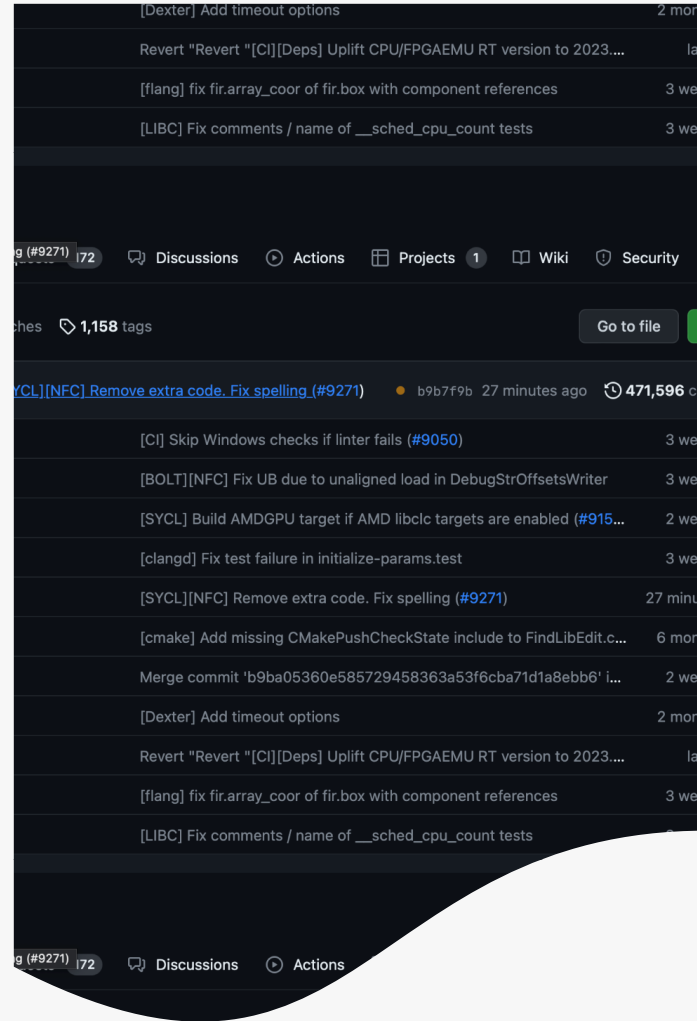
**Configuration Details and Workload Setup:** Intel® Xeon® Gold 6330 CPU @ 2.0GHz, 2 socket, Hyper Thread Off, Turbo On, 256GB Hynix DDR4-3200, ucode 0xd000363. GPU: AMD Instinct MI100, 32GB GPU memory. Software: SYCL open source/CLANG 15.0.0, AMD RoCm 5.2.1 with AMD-HIPCC 5.2.21152-4b155a06, hipSolver 5.2.1, rocBLAS 5.2.1, Ubuntu 20.04.4. SYCL open source/CLANG compiler switches: -fsycl-targets=amdgcn-amd-amdhsa -Xsycl-target-backend --offload-arch=gfx908, AMD-HIPCC compiler switches: -O3. Represented workloads with Intel optimizations.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See configuration disclosure for details. No product or component can be absolutely secure.

Performance varies by use, configuration, and other factors. Learn more at [www.Intel.com/PerformanceIndex](https://www.intel.com/PerformanceIndex). Your costs and results may vary.



標準に関する  
オープンな業界を超えた  
コラボレーション



コードはすべてオープンソース



Codeplay ウェブサイトから  
無料のプラグインとして利用可能



Enable AI & HPC to be Open, Safe and Accessible to All

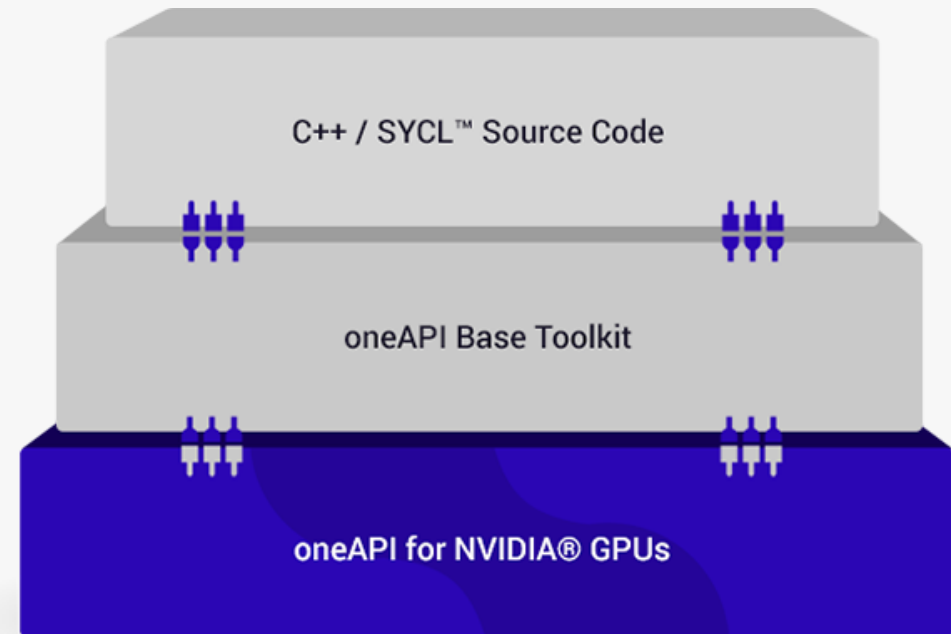
# 使い慣れたハードウェアを オープン標準で使用

HPSC 2023

# oneAPI for NVIDIA\* GPU

インテル® oneAPI ベース・  
ツールキットに NVIDIA\*  
GPU サポートを追加

SYCL\* でコードを開発して  
NVIDIA\* GPU で実行

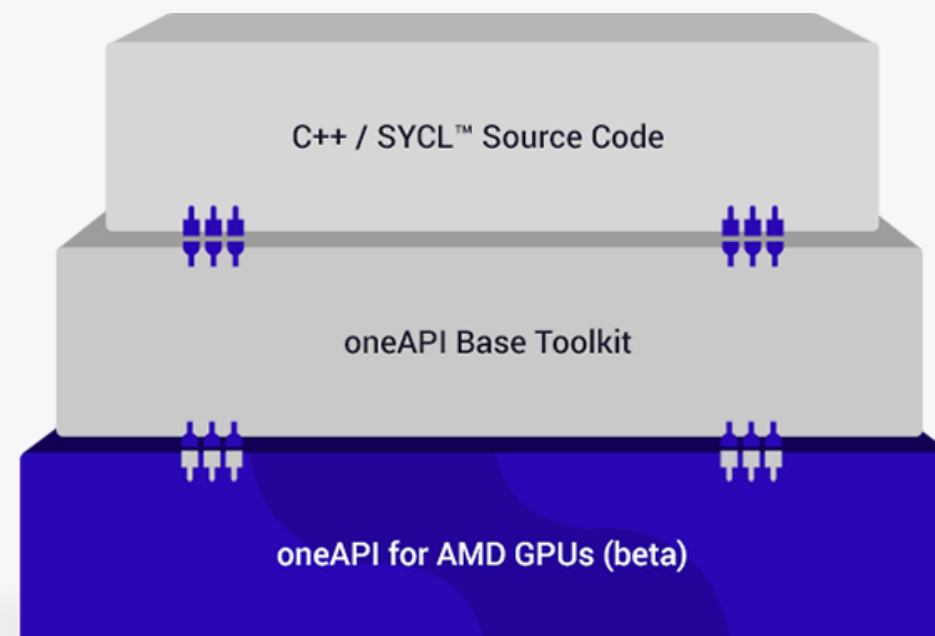


[developer.codeplay.com](https://developer.codeplay.com) (英語) からダウンロード可能

# ベータ版 oneAPI for AMD\* GPU

インテル® oneAPI ベース・  
ツールキットに AMD\* GPU  
サポートを追加

SYCL\* でコードを開発して  
AMD\* GPU で実行



[developer.codeplay.com](https://developer.codeplay.com) (英語) からダウンロード可能

## oneAPI for NVIDIA\* GPU

- Ubuntu\* 22.04
- CUDA\* SDK 12.0 以降
- sm\_50 以上の GPU
- 主に A100 GPU でテスト

## oneAPI for AMD\* GPU

- Ubuntu\* 22.04
- HIP 5.4.1
- GPU ドライバー 6.1.0
- 主に MI50 でテスト

無料のプラグイン  
完全にオープンソースベース





Enable AI & HPC to be Open, Safe and Accessible to All

# 使用方法

HPSC 2023

# 入手

必要条件を入手

CUDA\* または HIP 開発環境  
とドライバーをインストール

インテル® oneAPI  
ベース・ツールキット  
をダウンロード

インテルからツールキットを入手  
[intel.com/developer](https://intel.com/developer) (英語)

oneAPI for NVIDIA\*  
GPU と oneAPI for  
AMD\* GPU を  
ダウンロード

Codeplay からプラグインを入手  
[developer.codeplay.com](https://developer.codeplay.com) (英語)

# コンパイル

```
icpx -fsycl -fsycl-targets=nvptx64-nvidia-cuda sycl-app.cpp -o sycl-app
```

SYCL\*  
コンパイラー  
を使用

NVIDIA 向け  
にコンパイル

ソースファイル

バイナリー

バイナリーの実行

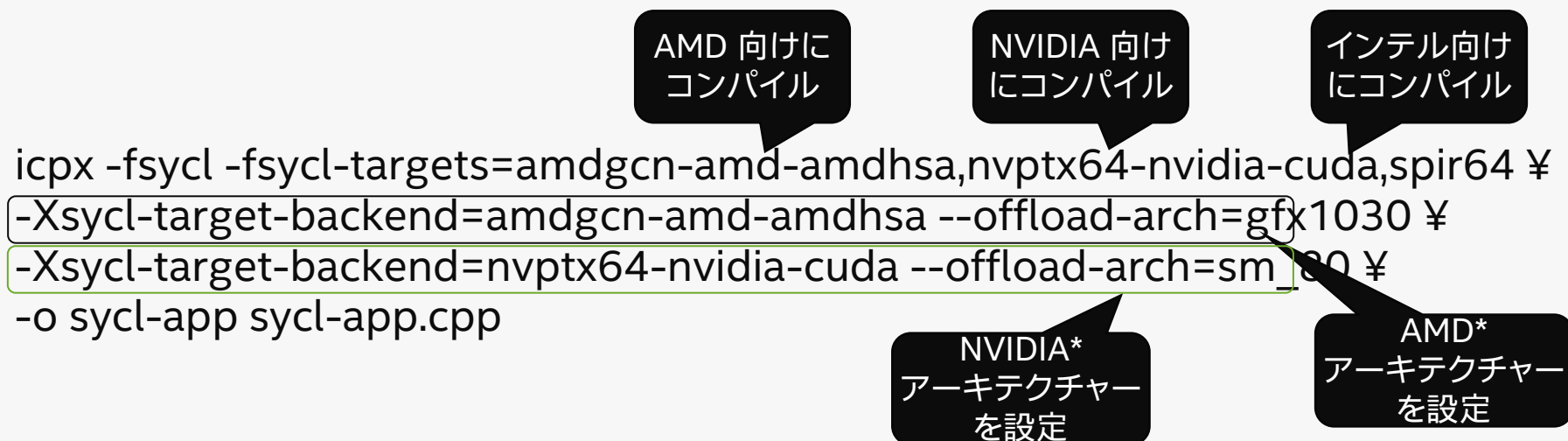
```
ONEAPI_DEVICE_SELECTOR=cuda:* SYCL_PI_TRACE=1 ./simple-sycl-app
```

ランタイムに  
NVIDIA\* GPU  
の使用を指示

警告などを表示  
するように出力  
トレースを設定

# マルチターゲット・コンパイル

単一バイナリーを複数のターゲット向けにコンパイル



# マルチターゲット・コンパイル

ターゲット・ハードウェアでバイナリーを実行

ランタイムに  
NVIDIA\* GPU  
の使用を指示

```
ONEAPI_DEVICE_SELECTOR=cuda:* SYCL_PI_TRACE=1 ./simple-sycl-app
```

コードでデバイス  
セクターを使用  
して行うことも  
できる

```
ONEAPI_DEVICE_SELECTOR=hip:* SYCL_PI_TRACE=1 ./simple-sycl-app
```

ランタイムに  
AMD\* GPU の  
使用を指示

# デバッグ

- 標準のデバッグツールを使用
- gdb と NVIDIA\* 固有の CUDA-GDB
- VS Code 統合

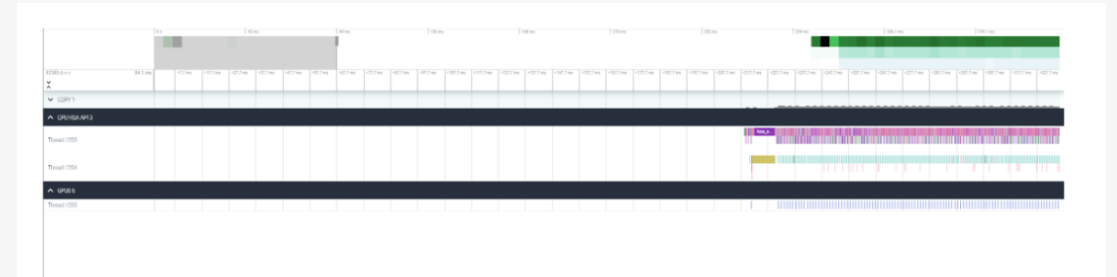
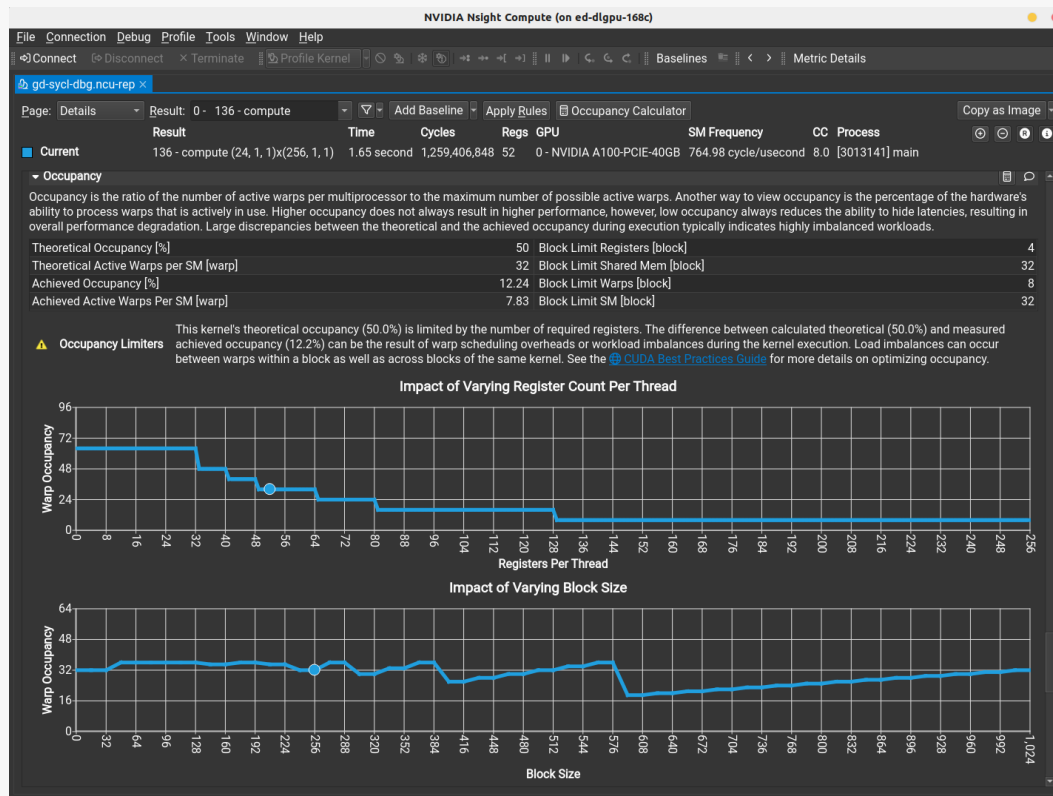
```
[Switching focus to CUDA kernel 1, grid 4, block (5,0,0), thread (32,0,0), device 0, sm 10, warp 0, lane 0]

Thread 1 "main" hit Breakpoint 1, main::{lambda(sycl::_V1::handler&)#5)::operator()(sycl::_V1::handler& const
::{lambda(sycl::_V1::nd_item<1>)#1)::operator()(sycl::_V1::nd_item<1>) const (this=0x7fffa3fffb68, item=...) a
t main.cpp:115
115         float v = sycl::log(1+sycl::exp(-1*A_y_label[i]*xp));
(cuda-gdb) info cuda kernels
Kernel Parent Dev Grid Status          SMs Mask GridDim BlockDim Invocation
*      1      -  0   4 Active 0x00000000000000005555555555555555 (24,1,1) (256,1,1) typeid name for main::{lam
bda(sycl::_V1::handler&)#5)::operator()(sycl::_V1::handler& const)::compute()
(cuda-gdb) list
110         for( int j = A_row_ptr[i]; j < A_row_ptr[i+1]; ++j){
111             xp += A_value[j] * x[A_col_index[j]];
112         }
113
114         // compute objective
115         float v = sycl::log(1+sycl::exp(-1*A_y_label[i]*xp));
116         auto atomic_obj_ref = atomic_ref<float,
117             memory_order::relaxed, memory_scope::device,
118             access::address_space::global_space> (total_obj_val[0]);
119         atomic_obj_ref.fetch_add(v);
(cuda-gdb) print i
$1 = 1312
(cuda-gdb) print xp
$2 = 0.0494509786
(cuda-gdb) next
118         access::address_space::global_space> (total_obj_val[0]);
(cuda-gdb) print v
$3 = 0.668727338
(cuda-gdb) continue
Continuing.
[Switching focus to CUDA kernel 1, grid 4, block (0,0,0), thread (0,0,0), device 0, sm 0, warp 3, lane 0]

Thread 1 "main" hit Breakpoint 1, main::{lambda(sycl::_V1::handler&)#5)::operator()(sycl::_V1::handler& const
::{lambda(sycl::_V1::nd_item<1>)#1)::operator()(sycl::_V1::nd_item<1>) const (this=0x7fffa3fffb68, item=...) a
t main.cpp:115
115         float v = sycl::log(1+sycl::exp(-1*A_y_label[i]*xp));
(cuda-gdb) print xp
$4 = 0.0241964087
(cuda-gdb) □
```

# プロファイル

- 標準の NVIDIA\* プロファイルツールを使用
- 標準の AMD\* プロファイルツールを使用




# 一般的な最適化

- 共有メモリーを注意して使用 (malloc\_shared)
- 最適なワークグループ・サイズを見つける

<https://codeplay.com/portal/blogs/2020/01/09/sycl-performance-post-choosing-a-good-work-group-size-for-sycl.html> (英語)

- インデックス・スワップ
- インライン展開
- ループアンロール



パフォーマンス・ガイド:  
developer.codeplay.com  
(英語)



# NVIDIA\* プラグインのサポート



## エンタープライズ・サポート

大規模なチーム向けの  
最高レベルのサポート

Codeplay のエンジニアと  
エキスパートに直接アクセス可能

要件に合わせたカスタム・  
サポート・プランを提供

<https://codeplay.com/company/contact/> (英語)



## 優先サポート

小規模なチームや個人向け  
チケット制のサポートデスクに  
アクセス可能

質問や要求に対して  
迅速な対応が得られる



## フォーラムサポート

Codeplay のエンジニアが  
モデレーターを務める  
パブリックフォーラム

無料で利用可能

oneAPI コミュニティーや Codeplay  
のエンジニアと交流可能

<https://support.codeplay.com> (英語)

# まとめ

- オープン標準の oneAPI と SYCL\* を使用することで選択肢が拡大
- oneAPI と SYCL\* の方向性に影響を与えることができる
- oneAPI と SYCL\* によりパフォーマンスの移植性を実現
- Codeplay のプラグインにより NVIDIA と AMD の GPU をターゲットにできる

# プラグインをお試してください

---



無料で入手可能:

<https://developer.codeplay.com> (英語)

### 法務上の注意書き

性能は、使用状況、構成、その他の要因によって異なります。

性能の測定結果はシステム構成の日付時点のテストに基づいています。また、現在公開中のすべてのセキュリティー・アップデートが適用されているとは限りません。構成の詳細は、補足資料を参照してください。絶対的なセキュリティーを提供できる製品またはコンポーネントはありません。

実際の費用と結果は異なる場合があります。

インテルのテクノロジーを使用するには、対応したハードウェア、ソフトウェア、またはサービスの有効化が必要となる場合があります。

© Codeplay Software Ltd. Codeplay、Intel、インテル、Intel ロゴ、その他のインテルの名称やロゴは、Intel Corporation またはその子会社の商標です。\* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

We're  
Hiring!

[codeplay.com/careers/](https://codeplay.com/careers/)



Enable AI & HPC to be Open, Safe and Accessible to All

# お問い合わせ先



[@codeplaysoft](https://twitter.com/codeplaysoft)



[info@codeplay.com](mailto:info@codeplay.com)



[codeplay.com](https://codeplay.com)