この資料の対象：ソフトウェア開発者

アプリケーションのパフォーマンス最適化に注目する開発者

- パフォーマンスのエキスパートである必要はなりません
- しかし、アプリケーション自体のことは熟知していることが前提です

インテル® グラフィックスやインテル® Iris™ グラフィックス向けに開発を行う

インテル® VTune™ Amplifier XE パフォーマンス・アナライザーを使用する

- ここで紹介するパフォーマンスに関する情報は、他のツール（PTU など）にも適用できますが、ここではインテル® VTune™ Amplifier XE に注目します

このスライドの使い方

一度スライドを読み通して、データ収集時に再度見てください

パフォーマンス解析は、何度かの繰り返しで達成されることを忘れないでください

ソフトウェアの最適化は、以下を行ってから始めてください：

- 任意のコンパイラー最適化オプションを適用（/O2、/QxAVX2 など）
- 適切なワークロードを選択
- 基準となるパフォーマンスの測定

インテル® HD グラフィックスと インテル® Iris™ グラフィックス向けの ソフトウェアをチューニングするため インテル® VTune™ Amplifier XE を使用する

Ver.1.0

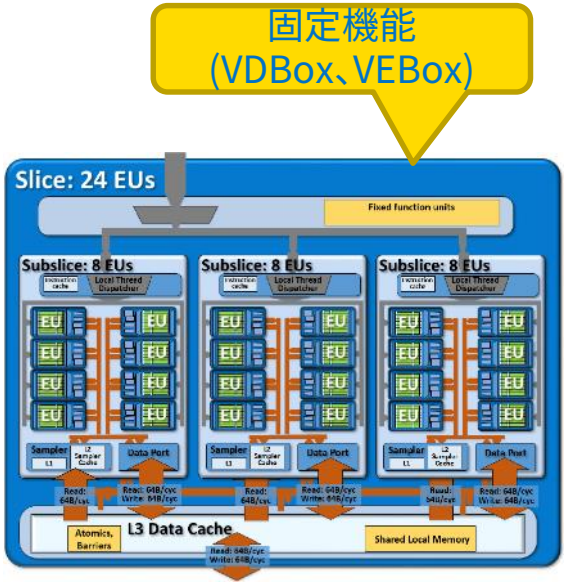
内容

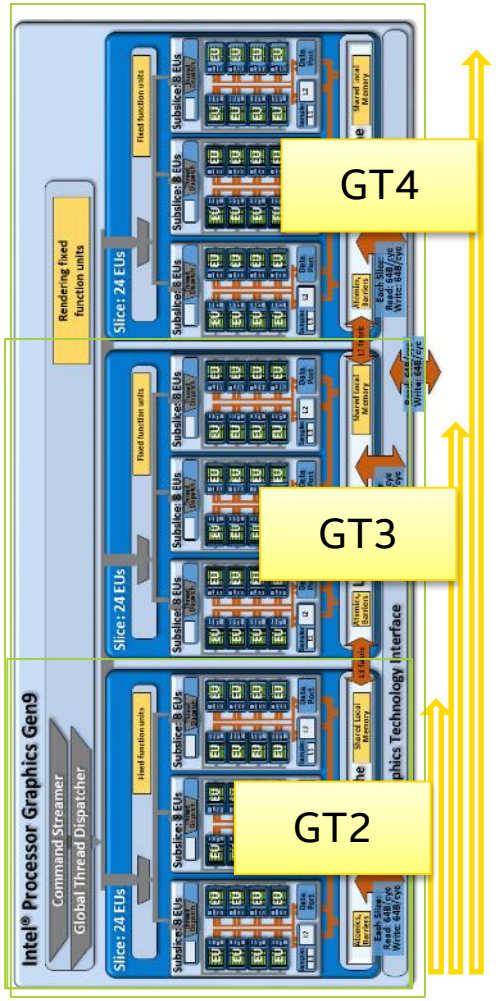
- インテル® HD グラフィックスとインテル® Iris™ グラフィックスの概要
- インテル® VTune™ Amplifier XE の GPU サポート
- ソフトウェアの最適化ステップ
  - CPU / GPU の並行性を調査
  - GPU の動作を確認
  - GPU の利用状況を確認
  - 時間軸で確認
  - コンテキスト固有のメトリックを調査
- OpenCL* カーネル実行の調査
- インテル® Media SDK タスクの実行を調査 (Linux* のみ)

インテル® プロセッサー・グラフィックス/ GPU の概要

主なグラフィックス・テクノロジー

- 実行ユニット (EU) = 汎用コア
- “スライス” には EU、サンプラー、キャッシュなどが含まれる
- 固定機能は “アンスライス” に含まれる
- eDRAM はキャッシュを追加し、帯域幅を増加





各プロセッサー・グラフィックス

	拡張	別名	要約
インテル® HD グラフィックス		GT2 “4+2”	良い
インテル® Iris™ グラフィックス	+ スライス + eDRAM	GT3 “2+3e”	さらに良い
インテル® Iris™ Pro グラフィックス	+ スライス + eDRAM	GT3e、GT4e “4+4e”	最良

インテル® プロセッサー・グラフィックス

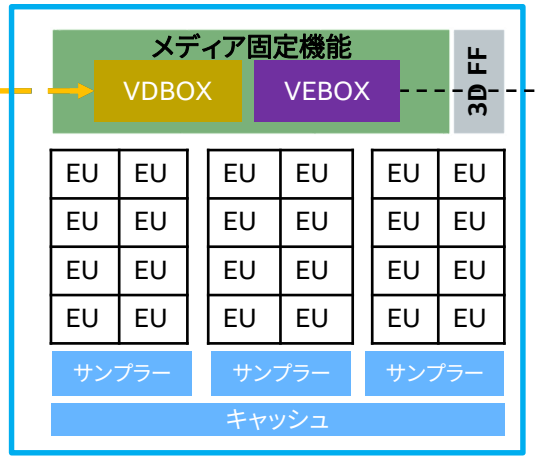
コーデックとフレーム処理は固定機能と実行ユニットを使用

ビデオ・エンコーディング

ENC= EU+VDBox VME (MB タイプ、動きベクトル、ビット配分/BRC)
PAK = VDBox (残差パッキングとエントロピー・コーディング)
VDENC = 省電力エンコード (第 6 世代インテル® Core™ プロセッサー・ファミリー以降)

ビデオ・デコーディング

BSD=VDBox デコード

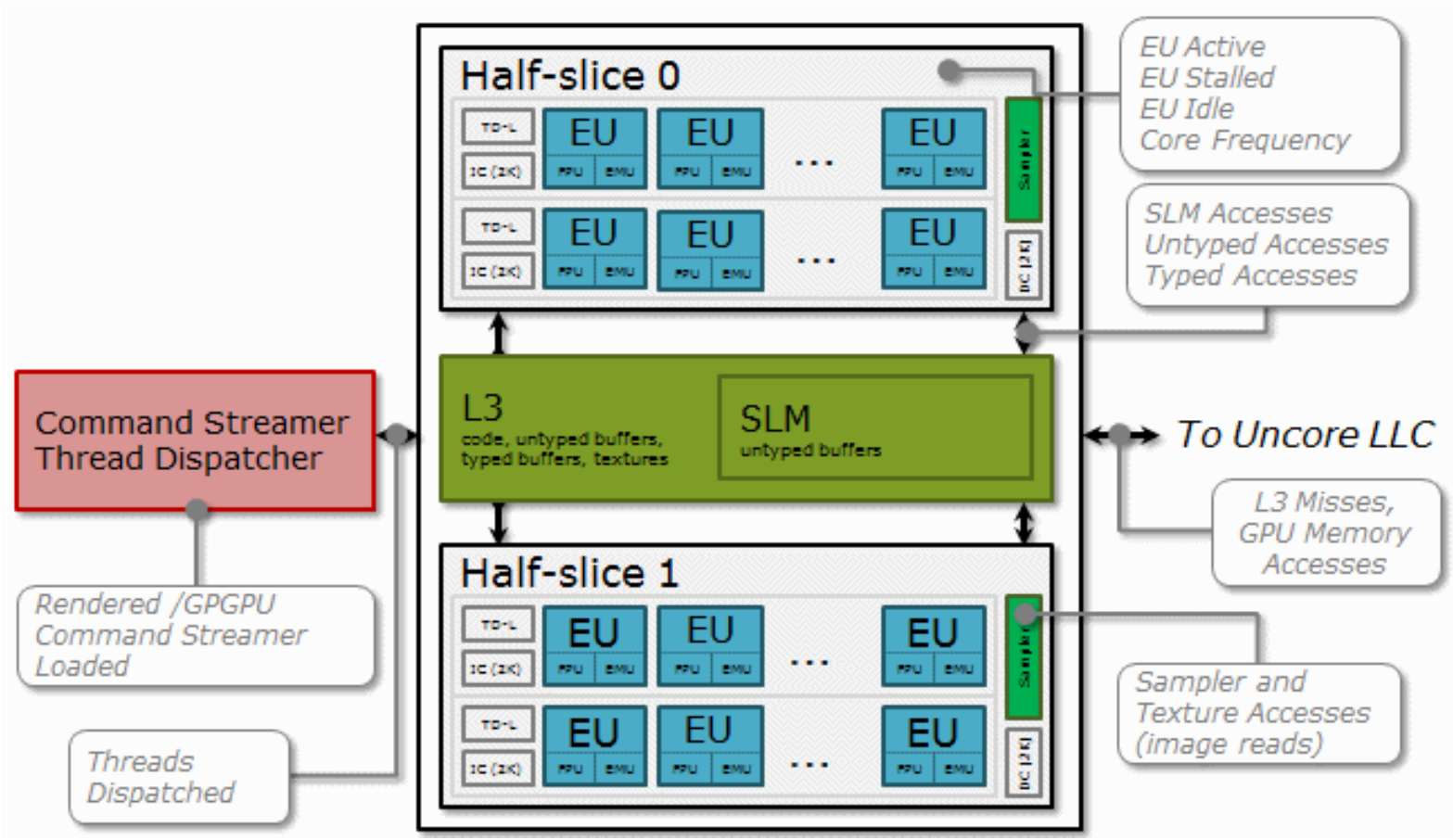


VPP

VPHal
ビデオ処理ハードウェア・
アクセラレーション・レイヤー

VEBox
- デインターレース
- ノイズ除去 (Luma/Chroma)
- フレームレート変換
- 色空間変換
- 合成/アルファ・ブレンディング
- スケーリング

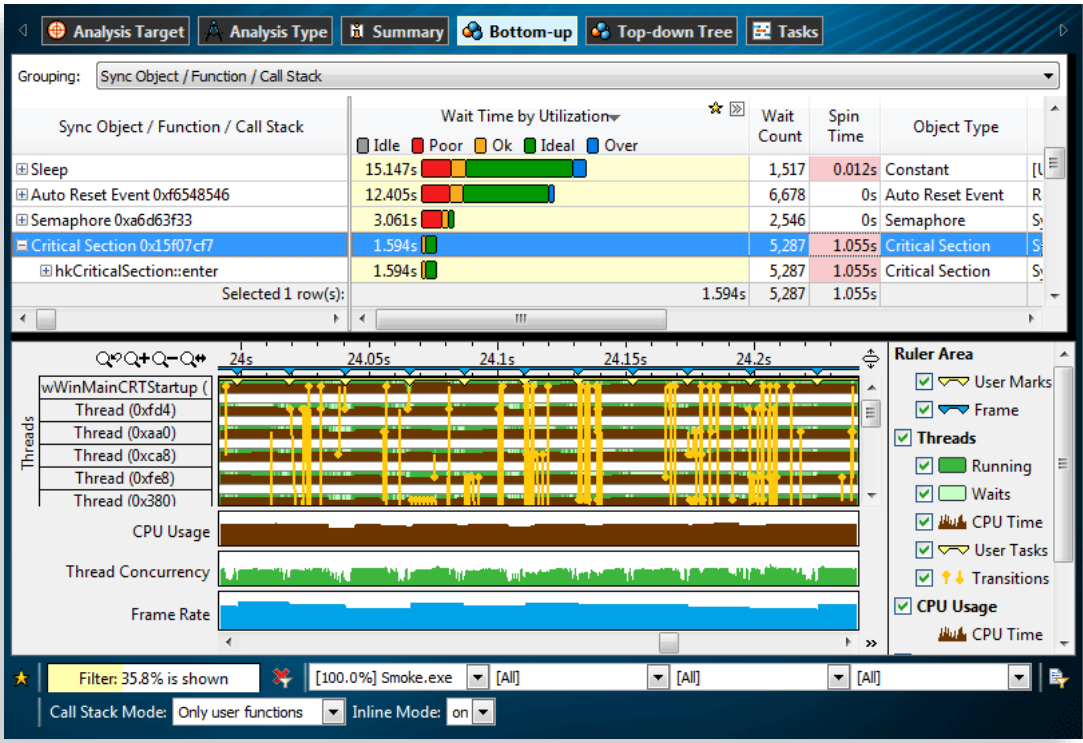
インテル® HD グラフィックスのイベント



インテル® VTune™ Amplifier XE

インテル® VTune™ Amplifier XE の機能：

- 複数の収集タイプ
  - ホットスポット
  - 帯域幅
  - イベント・ベース・サンプリング
- 全ての解析タイプを
  タイムライン表示へ統合
- ソースとアセンブリー表示
- C/C++、Fortran、Java、
  アセンブリー、.NET との互換性
- Visual Studio* への統合、
  コマンドラインでの利用、もしくは
  Windows® と Linux* 向けスタンドアロン・インターフェイスを利用



インテル® Vtune™ Amplifier の GPU 向けのオプション

オプション	オーバーヘッド	サポートされるターゲットシステム	サポートされるグラフィックス	サポートされる解析タイプ
GPU の利用解析	低	すべて	すべて	CPU/GPU 並行性 (デフォルト)、GPU ホットスポット (デフォルト)、カスタム解析
プロセッサー・グラフィックスのハードウェア・イベントを解析	中	Windows®、Linux* および Android*	インテル® HD グラフィックスおよびインテル® Iris™ グラフィックス (インテル® グラフィックス) のみ (管理者権限が必要)	CPU/GPU 並行性 (プリセットされた概要)、GPU ホットスポット、カスタム解析
OpenCL* とインテル® Media SDK プログラムをトレース	高	OpenCL カーネル解析: Windows® および Linux* インテル® Media SDK プログラム解析: Linux*	インテル® グラフィックスのみ	GPU ホットスポット (デフォルト)、カスタム解析

インテル® Vtune™ Amplifier で GPU の解析を行うには

レンダリング、ビデオ処理、および計算向けにグラフィックス処理ユニット (GPU) を使用するアプリケーションのプロファイルを行うためインテル® VTune™ Amplifier XE を使用します。インテル® VTune™ Amplifier XE は、CPU と GPU の両方のアクティビティーを監視し、解析して関連付けることができます

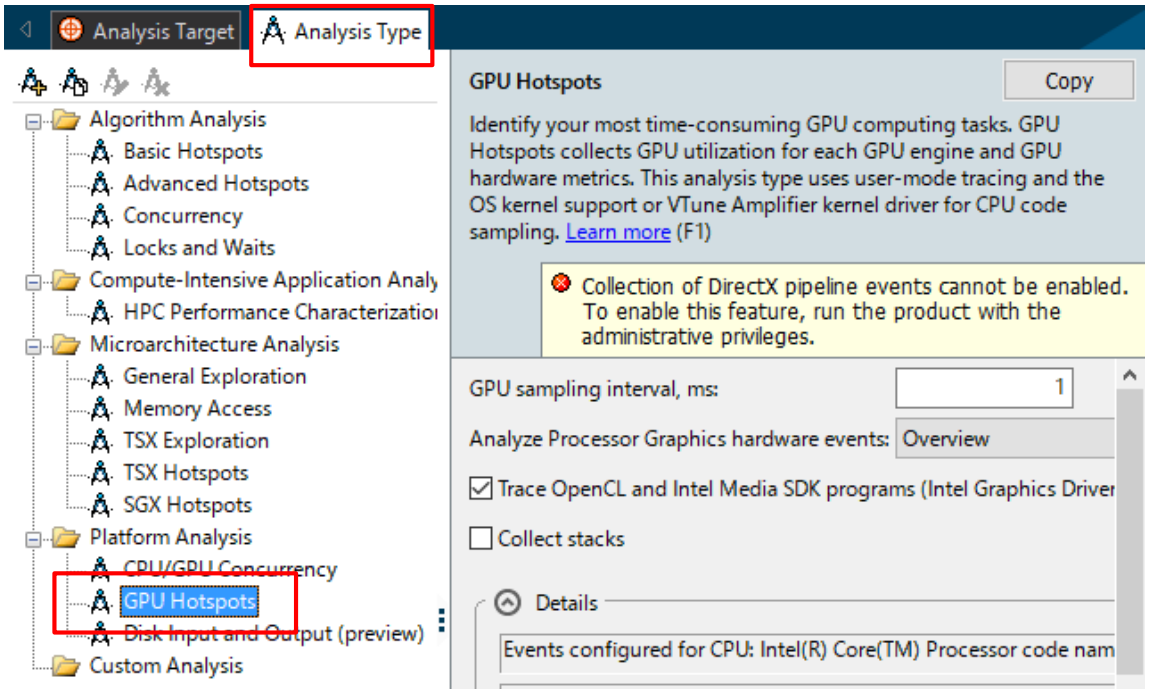
インテル® VTune™ Amplifire XE による GPU 解析向けには、次の手順が考えられます:

1. アプリケーションが GPU 依存であるかどうかを特定するため、CPU/GPU 並行性解析を実行します
2. インテル® Media SDK と OpenCL* ソフトウェア・テクノロジーを使用する GPU 依存のアプリケーションを詳しく解析するため、GPU Hotspots (GPU ホットスポット) 解析を実行します
   - GPU ハードウェア・メトリックの調査
   - OpenCL* カーネル実行の調査
   - インテル® Media SDK タスクの実行を調査 (Linux* のみ)

GPU 解析オプションを有効にする

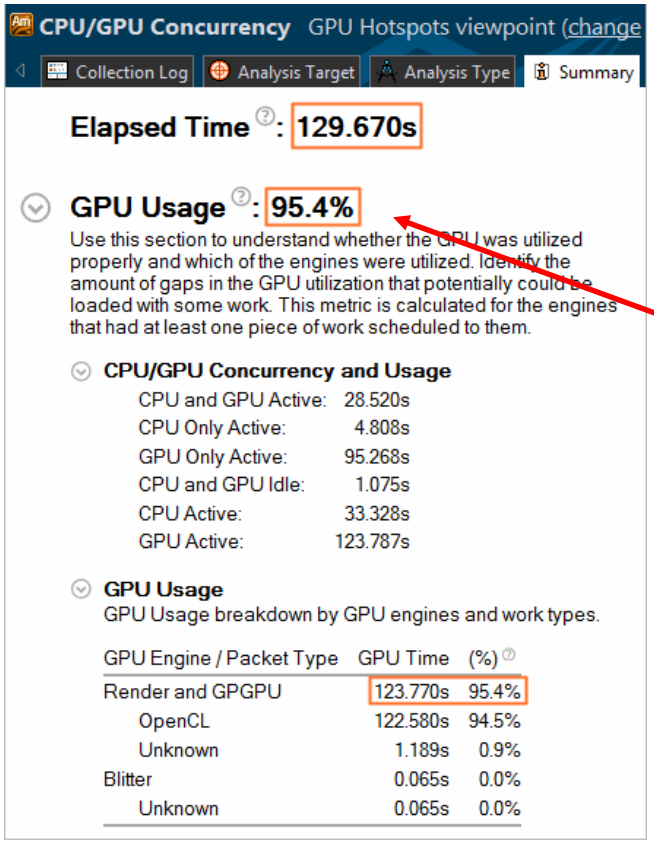
1. インテル® VTune™ Amplifier XE ツールバーの New Analysis ボタンをクリック
   Analysis Type 設定ウィンドウが開きます
2. 左フレームの Analysis ツリーから、該当する Analysis タイプを選択します
   例えば: GPU Hotspots、CPU/GPU Concurrency

右のペインは、選択された Analysis タイプの設定オプションで更新されます



CPU/GPU の並行性を調査

CPU/GPU 並行性解析 (英語) を実行して、時間軸における GPU 使用率を調査し、アプリケーションや処理フェーズの一部が CPU もしくは GPU 依存であるかを確認します。



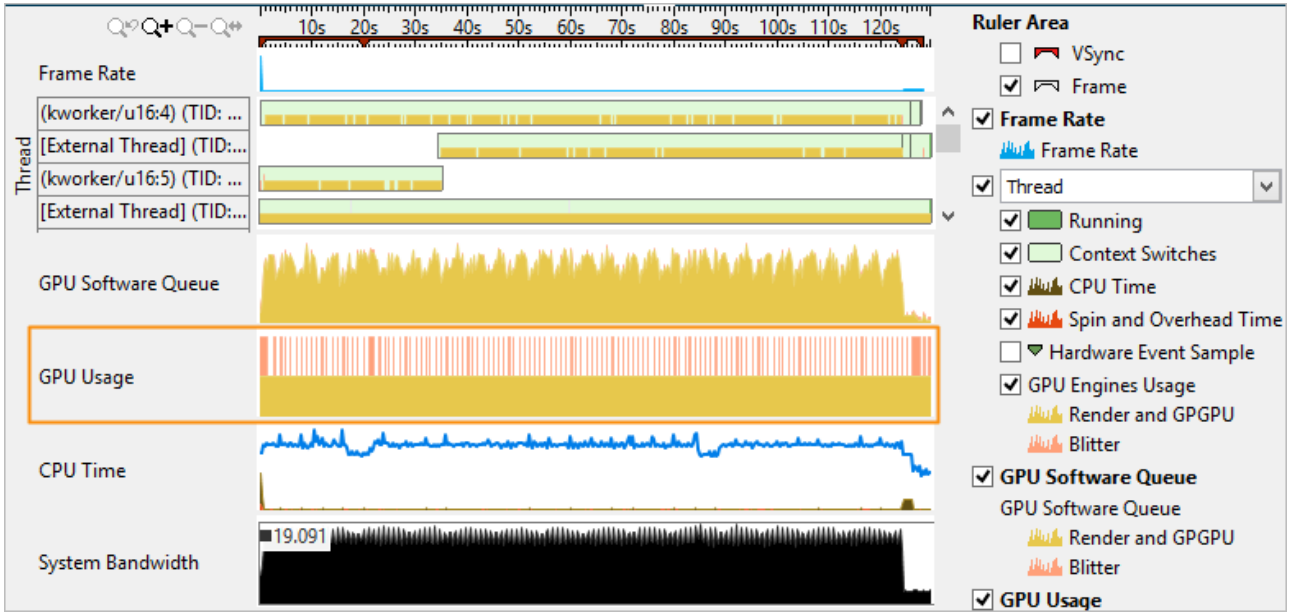
この例は、GPU 依存のアプリケーションの解析結果を示しています

[Summary] ウィンドウで、**Elapsed time (経過時間)** の大部分を **GPU Time (GPU 時間)** が占めていることが分かります

GPU の動作を確認

[Platform (プラットフォーム)] ウィンドウ (英語) に切り替えて、ソフトウェア・キュー上の GPU 使用率を解析する基本 CPU および GPU メトリックを開き、タイムライン上の CPU 使用率のデータと関連付けます

GPU が時間の経過とともにビジーになる場合、[Graphics (グラフィックス)] ウィンドウ (英語) に切り替えて、スレッドごとに実行されているワーク (レンダリングや計算) の種類を理解するため詳しく調査します



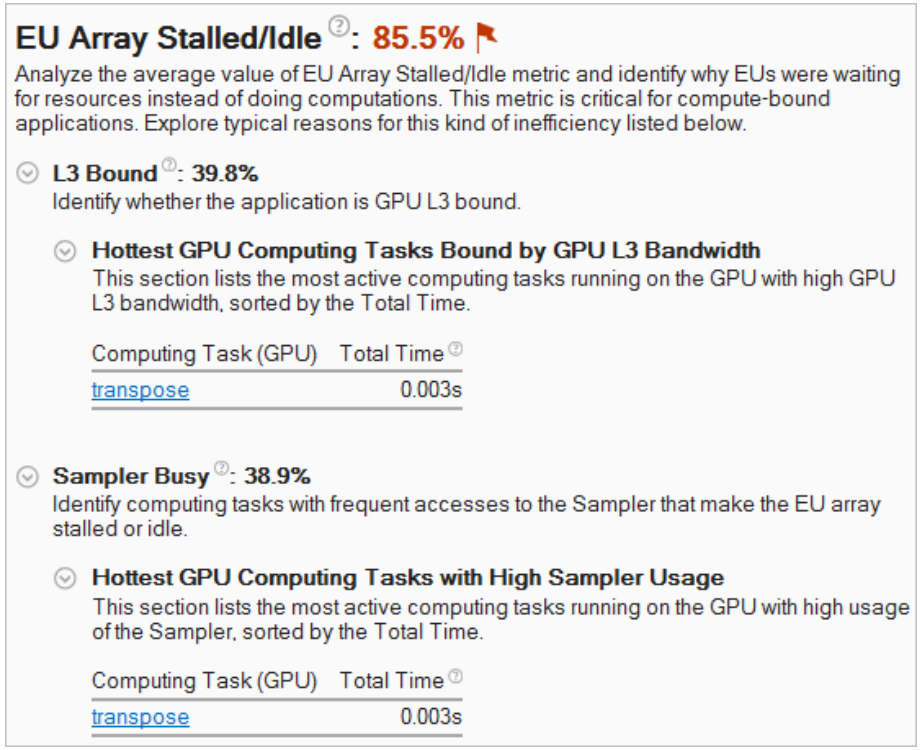
この例は、レンダーと GPGPU エンジン上のアクティビティー (黄色) と同様に、ブリッター (Blitter) エンジンのアクティビティー (ピンク) を示しています

GPU 依存のアプリケーションの GPU 使用率を解析

アプリケーションやその処理ステージが GPU 依存であることが判明している場合、GPU ホットスポット解析を実行して GPU エンジンが効率良く実行されているか、また改善の余地があるかを調査します。これは、インテル® VTune™ Amplifier XE のインテル® グラフィックスの**レンダーと GPGPU エンジン向け**のハードウェア・メトリックの収集により解析できます

データが収集されたら、**[Summary]** ウィンドウの **[EU Array Stalled/Idle (EU 配列ストール/アイドル)]** セクションを調査して、実行ユニットが待機している最も典型的な原因を特定します

[Sampler Busy (サンプラーがビジー)] には、頻繁にサンプラーにアクセスする GPU 計算タスクが、**[L3 Bound]** には、GPU L3 帯域幅に依存する最もホットな GPU 計算タスクのリストが表示されます



[Compute Basic] メトリック

インテル® VTune™ Amplifier XE は GPU 上の異なるタイプのデータアクセスを特定するメトリックを解析し、占有率の低い GPU タスクの識別に役立つ **[Occupancy (占有率)]** セクションを表示します

EU Array Stalled/Idle: **63.4%**

Analyze the average value of EU Array Stalled/Idle metric and identify why EUs were waiting for resources instead of doing computations. This metric is critical for compute-bound applications. Explore typical reasons for this kind of inefficiency listed below.

L3 Bound: **8.2%**

Occupancy: **14.2%**

Identify too large or too small computing tasks with low occupancy that make the EU array idle while waiting for the scheduler. Note that frequent SLM accesses and barriers may affect the maximum possible occupancy.

Hottest GPU Computing Tasks with Low Occupancy

This section lists the most active computing tasks running on the GPU with a low Occupancy, sorted by the Total Time.

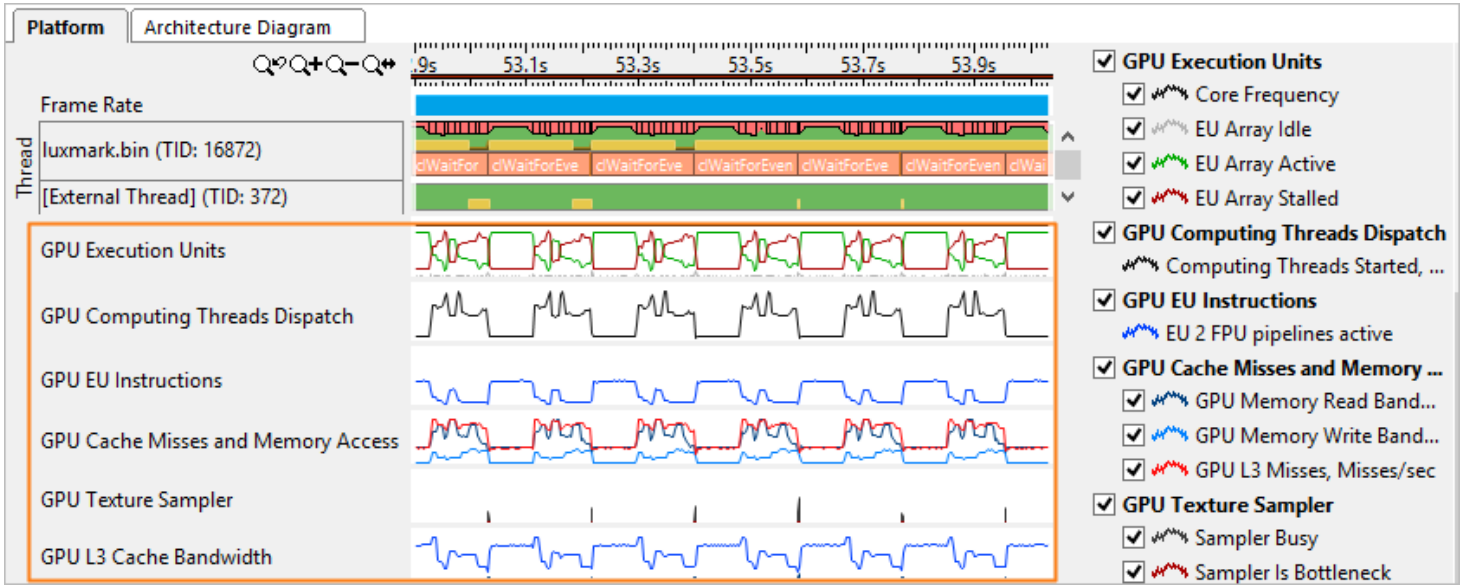
Computing Task (GPU)	Total Time	Global Size	Local Size	SIMD Width
workload	7.612s	53760		32

占有率 (occupancy) がアプリケーションの問題であると通知されている場合、タスクが大きすぎるか、または小さすぎることで、EU 配列がアイドルになることが考えられるため、計算タスクのサイズを変更することを検討してください

時間経過軸で確認する

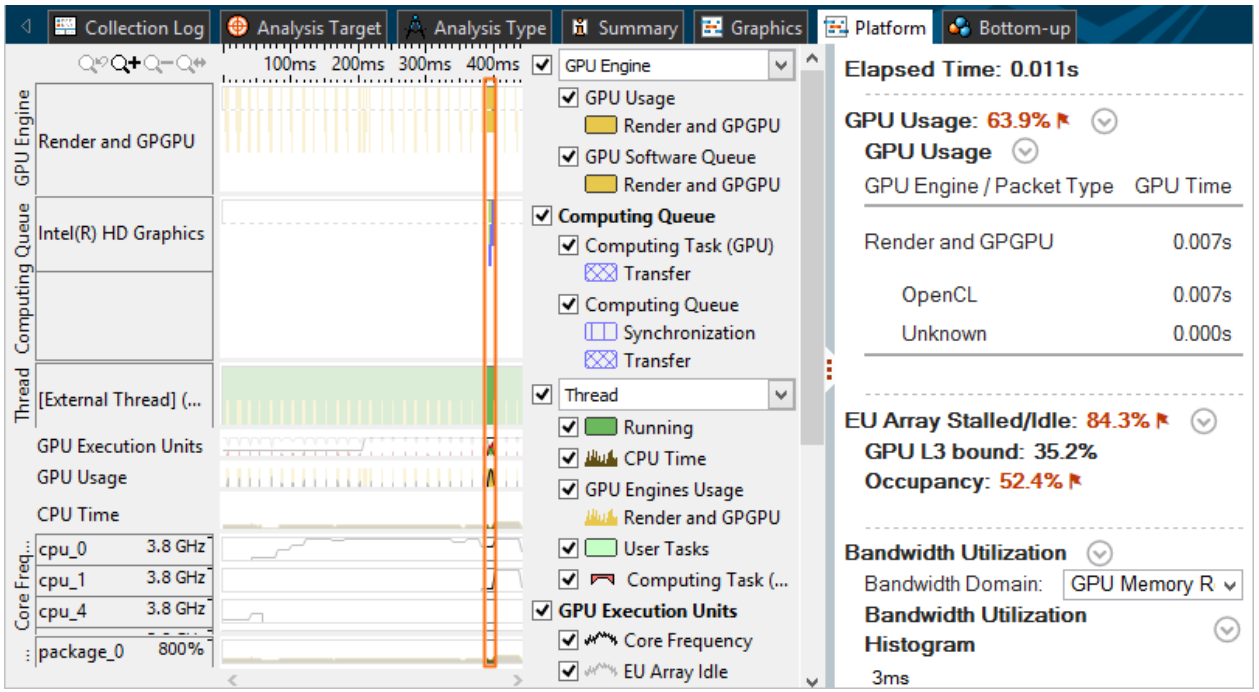
時間経過における HW メトリックごとの GPU パフォーマンス・データを解析するには、**[Graphics]** ウィンドウを開き、**[Timeline]** ペインに注目します。**[Graphics]** ウィンドウに表示される GPU メトリックのリストは、解析設定で選択されたハードウェア・イベントの事前定義に依存します

次の例は、GPU 依存のアプリケーションで収集された **[Overview]** メトリックのグループです



コンテキスト固有の GPU メトリックを調査

[GPU Hotspots vewpoint] で、**[Platform]** タブに切り替えて、CPU データ、メモリー帯域幅、割り込み (収集されていれば) などがどのように GPU メトリックデータに関連するか解析します。タイムライン上で対象とする領域を選択して右クリックし、コンテキスト・メニューから **[Filter In by Selection (選択でフィルター)]** を選択して、右にあるペインでコンテキスト固有の GPU メトリックを調査します



この例では、選択された範囲で使用されている GPU 実行ユニットごとの統計が表示されています。赤く強調された値は、占有率が低いため GPU 時間の大部分がアイドルであることを示しています。この原因は、非効率なワークのスケジューリングであると考えられます

OpenCL* カーネル実行の調査

アプリケーションが、OpenCL* ソフトウェア・テクノロジーを使用している場合、**[Graphics]** ウィンドウの **[Timeline]** ペインで **[GPU Computing Threads Dispatch (GPU 計算スレッドのディスパッチ)]** メトリックを使用して、解析を続行しインテル® グラフィックス上で実行されている OpenCL* カーネルの情報を取得します

[Summary] ビューでは、**[Hottest GPU Computing Tasks (最もホットな GPU 計算タスク)]** セクションで GPU 上で実行される OpenCL* カーネルを表示し、パフォーマンス上クリティカルなカーネルを通知します。カーネル名をクリックして、**[Computing Task (GPU) / Instance (計算タスク (GPU) / インスタンス)]** でグループ化された **[Graphics]** ウィンドウを開きます

インテル® VTune™ Amplifier XE は、次の計算タスクを識別します:
Compute (カーネル)、Transfer (OpenCL* ルーチンがホストから GPU へのデータ転送を行います)、および Synchronization (例えば、clEnqueueBarrierWithWaitList)

最初に、実行時間が最も長いホットなカーネルを解析し最適化します。ホットなカーネルには、平均実行時間が長い、あるいは平均実行時間は短くても頻繁に呼び出されるといった特徴があります。どちらの場合も注目すべきです

GPU Hotspots GPU Hotspots viewpoint (change)

Collection Log | Analysis Target | Analysis Type | Summary | Graphics | Platform | Bottom-up

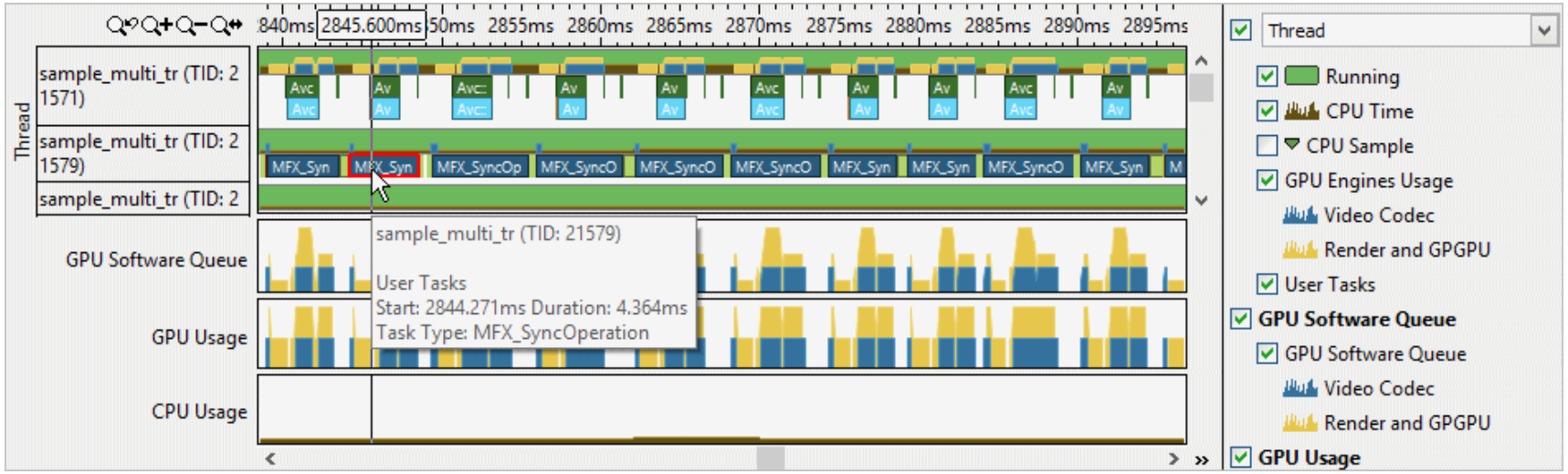
Grouping: Computing Task (GPU) / Instance

Computing Task (GPU) / Instance	Work Size		Computing Task					Data Transfer...		EU Array		
	Global	Local	Total Time ▼	Average ...	Instance ...	SIMD ...	SVM ...	Size	Total, ...	Active	Stalled	Idle
▶ Intersect	65536	64	73.997s	0.007s	10,273	8			0.000	72.2%	27.1%	0.7%
▶ AdvancePaths	65536	64	16.383s	0.002s	10,273	8			0.000	42.6%	50.8%	6.7%
▶ Sampler	65536	64	12.252s	0.001s	10,273	16			0.000	83.6%	12.8%	3.5%
▶ clEnqueueReadBuffer			0.831s	0.000s	2,276			6 G...	8.308	7.7%	85.9%	6.4%
▶ Init	65536	64	0.003s	0.003s	1	16			0.000	45.2%	52.6%	2.2%
▶ InitFrameBuffer	362432	64	0.000s	0.000s	1	32			0.000	0.0%	0.0%	100...
▶ [Outside any task]				0s					0.000	4.4%	6.2%	89.4...

[Graphics] または **[Platform]** ウィンドウの **[Timeline]** ペインにある、**[Computing Queue (計算キュー)]** データを調査します

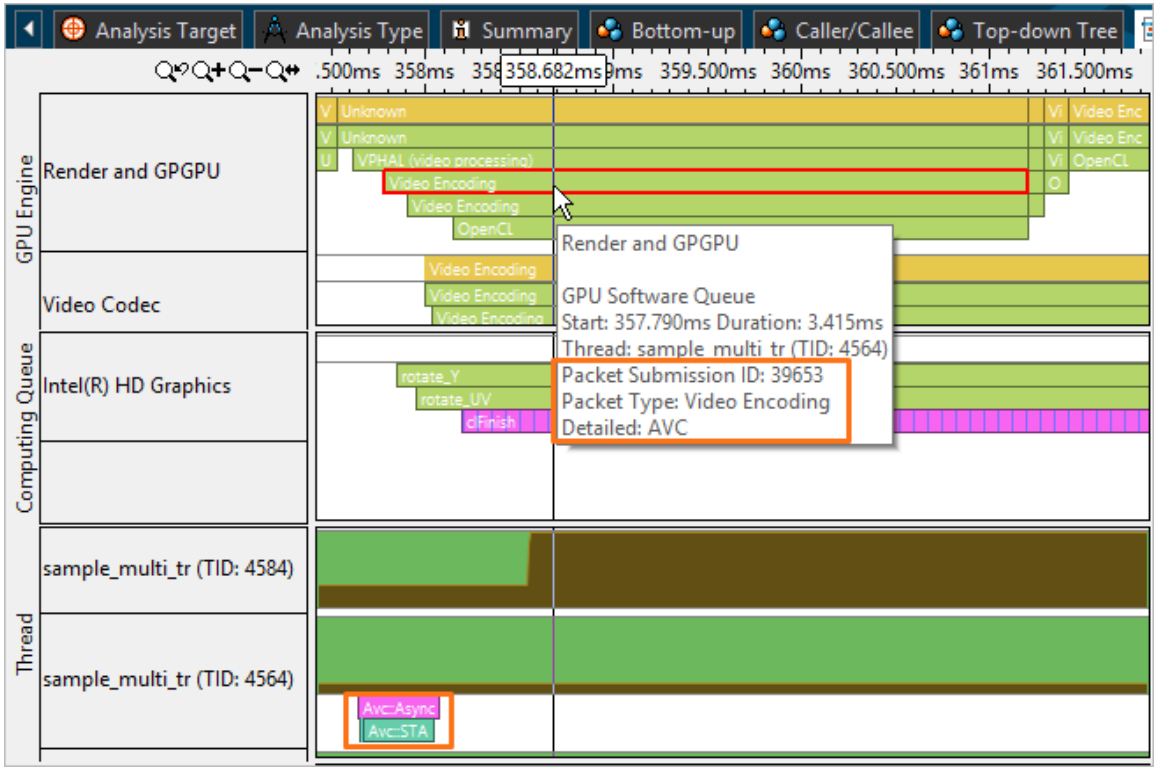
インテル® Media SDK タスクの実行を調査 (Linux* のみ)

[Analyze GPU usage] と **[Trace OpenCL and Intel Media SDK programs]** オプションの両方をオンにして、インテル® Media SDK プログラム解析を行う場合、**[Graphics]** ウィンドウを使用してインテル® Media SDK タスクの実行データと GPU ソフトウェア・キューのデータを関連付けます



GPU ソフトウェア・キューと GPU パケット転送の詳細を表示 (続き)

[Platform] ウィンドウ (英語) ウィンドウに切り替えて、**[GPU Engine (GPU エンジン)]** 領域を調査することで、インテル® Media SDK アプリケーション向けの GPU ソフトウェア・キューと GPU パケット転送の詳細を表示できます



ありがとうございます！
より詳しい情報について:

インテル® VTune™ Amplifier XE 向けビデオ、フォーラムおよびリソース:
http://www.isus.jp/intel-vtune-amplifier-xe/

インテル® 64 アーキテクチャーおよび IA-32 アーキテクチャー・ソフトウェア開発者マニュアル:
http://www.intel.com/products/processor/manuals/index.htm

他のマイクロ・アーキテクチャー向けのインテル® VTune™ Amplifier XE のチューニング・ガイド:
http://www.isus.jp/products/vtune/processor-specific-performance-analysis-papers/

iSUS