



Arm[®] Frame Advisor

Version 1.4

User Guide

Non-Confidential

Copyright © 2023–2024 Arm Limited (or its affiliates). All rights reserved.

Issue 00

102693_1.4_00_en



Arm® Frame Advisor User Guide

This document is Non-Confidential.

Copyright © 2023–2024 Arm Limited (or its affiliates). All rights reserved.

This document is protected by copyright and other intellectual property rights.

Arm only permits use of this document if you have reviewed and accepted [Arm's Proprietary Notice](#) found at the end of this document.

This document (102693_1.4_00_en) was issued on 2024-09-05. There might be a later issue at <https://developer.arm.com/documentation/102693>

The product version is 1.4.

See also: [Proprietary notice](#) | [Product and document information](#) | [Useful resources](#)

Start reading

If you prefer, you can skip to [the start of the content](#).

Intended audience

This document is intended for software developers who want to use Arm® Frame Advisor for frame-based graphical analysis of Android applications.

Inclusive language commitment

Arm values inclusive communities. Arm recognizes that we and our industry have used language that can be offensive. Arm strives to lead the industry and create change.

We believe that this document contains no offensive language. To report offensive language in this document, email terms@arm.com.

Feedback

Arm welcomes feedback on this product and its documentation. To provide feedback on the product, create a ticket on <https://support.developer.arm.com>.

To provide feedback on the document, fill the following survey: <https://developer.arm.com/documentation-feedback-survey>.

Contents

1. Overview of Frame Advisor.....	5
1.1 Platform support.....	8
2. Get started with Frame Advisor.....	10
2.1 Setup tasks.....	10
2.2 Capture a frame burst.....	12
2.3 Analyze the results.....	17
3. Analyzing your frames.....	26
3.1 Navigating your frames.....	26
3.2 Analyze workloads using the Render Graph.....	28
3.3 Analyze object rendering.....	30
3.4 Analyze object complexity.....	35
3.5 Analyze model geometry.....	38
3.6 Content metrics in detail.....	40
3.6.1 Mesh complexity.....	40
3.6.2 Mesh locality.....	41
3.6.3 Mesh redundancy.....	43
3.6.4 Mesh memory layout.....	43
3.7 Analyze function calls.....	44
3.8 Analyze shader programs.....	45
4. How to get help.....	49
5. Troubleshooting Frame Advisor.....	50
5.1 My device is not listed in Frame Advisor.....	50
5.2 My application is not listed in Frame Advisor.....	50
5.3 The Framebuffer view is slow to load images.....	51
5.4 A timed out error message appears when I start a capture.....	52
5.5 An unsupported image format message is displayed in the Framebuffer.....	52
5.6 A no image data message is displayed in the Framebuffer.....	53
Proprietary notice.....	54

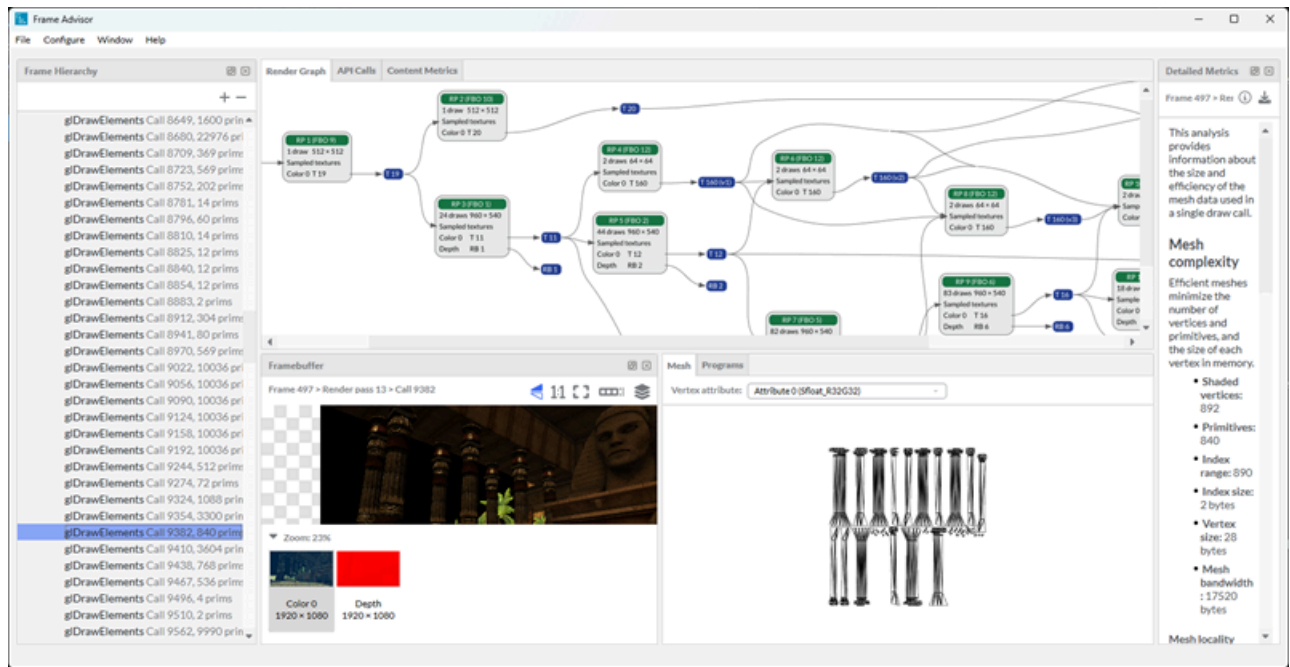
Product and document information.....56
Product status..... 56
Revision history.....56
Conventions..... 57

Useful resources.....59

1. Overview of Frame Advisor

Arm® Frame Advisor captures the API calls and GPU output for frames rendered by your application as it runs on a connected Android device. The analysis views provided by Frame Advisor show how your application is performing on Arm® GPUs. Evaluate the data and visual outputs to identify any frames that are causing performance problems with your device or application, save processing power, and improve use of memory bandwidth.

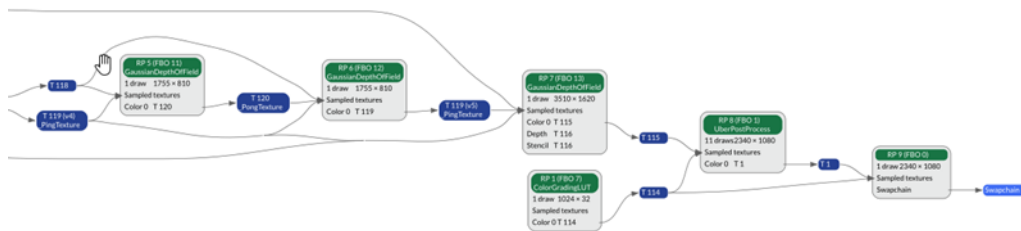
Figure 1-1: The Frame Advisor Analysis screen



Visualize the structure of your frame

Frame Advisor shows you a graph of the workloads and resources that make up a frame. You can see how render passes are processed by GPU workloads, the rendering commands in each workload, and the data flows between them. This information helps you to optimize render pass processing and memory bandwidth, both of which are essential for best performance using a tile-based Arm® GPU.

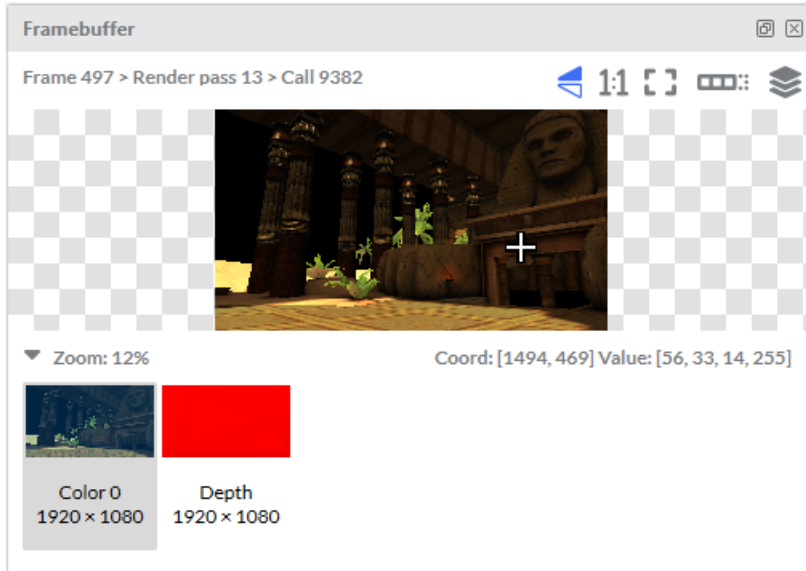
Figure 1-2: Render Graph



See the framebuffer output

See what the GPU rendered to the framebuffer for each draw call. Step through the draw calls and watch the visual output to see how your application constructed each frame. Identify draws are rendering in the correct order, or any draws that render no visual output. When used with capture modes, such as overdraw mode, a false-color visualization of useful key performance metrics is shown.

Figure 1-3: Framebuffer view



Inspect rendered object complexity

Step through the draw calls and look at the primitives to find problem objects that do not follow best practice standards, and reduce the processing cost of draw calls. Ensure the level of detail of the object is suitable based on its distance to the camera. Check if you can use a simpler mesh to improve performance.

Figure 1-4: Mesh view

Evaluate your model geometry

Use the content metrics provided in Frame Advisor together with other graphs and visual outputs to analyze your model geometry. Quickly find draw calls that are large or complex, and that use a lot of memory bandwidth to process. You can also use the content metrics to find draw calls that score poorly on one of the built-in efficiency metrics that Frame Advisor provides.

Perform a detailed analysis of the geometry and buffer memory layout used in a single draw call. Frame Advisor presents clear efficiency metrics, each associated with a specific optimization recommendation.

Evaluate graphics shaders

See where shader programs are run for each workload in different stages of the shader pipeline, and how efficiently they are processed by the GPU in your connected device. To help you identify which shader programs you can optimize to improve GPU performance, the [Programs view](#) shows how the shader programs use resources, as well as their cycle costs for the shortest and longest control flow paths. If you find a shader that requires additional investigation, you can open the OpenGL ES source or SPIR-V disassembly in the **Source** view to explore the shader source code.

Figure 1-5: Programs view

Program	Shader	Shortest path			Longest path			Total emitted			Occupancy
		A	LS	T	A	LS	T	A	LS	T	
376	377	0.08	1.00	0.00	0.08	1.00	0.00	0.08	1.00	0.00	
310	311	0.92	1.00	0.00	0.92	1.00	0.00	0.92	1.00	0.00	
370	371	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
346	347	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
337	338	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
331	332	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
295	296	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
280	281	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
265	266	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
250	251	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
235	236	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
220	221	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	

See [Analyze shader programs](#) to learn more about shader analysis.

Explore OpenGL ES or Vulkan API calls

See what functions and parameters the application requested and their return values. API calls are shown in the same order that they are seen by the GPU, at the point that workloads are submitted to a queue. This sequence makes it easier for you to see changes over time, and helps you to locate any problem calls in Vulkan applications.

See Frame Advisor in action

To see an example of Frame Advisor in action, watch our training video [Capture and analyze a problem frame with Frame Advisor](#).

To start capturing frames with Frame Advisor, see [Get started with Frame Advisor](#).

1.1 Platform support

Arm® Frame Advisor supports capturing profiles on a compatible [Android device](#).

Host support

Frame Advisor supports the following host OS versions:

- Windows 10 or later
- Ubuntu 20.04 or later
- macOS 10.15 (Catalina) or later

Device support

Frame Advisor supports the following device OS versions:

- OpenGL ES: Android 10 or later
- Vulkan: Android 9 or later

API support

Frame Advisor supports the following API versions:

- OpenGL ES 2.0 and 3.0-3.2
- Vulkan 1.0-1.2

GPU support

Frame Advisor supports Arm® Mali™ and Arm® Immortalis™ GPUs implementing the Midgard, Bifrost, Valhall, and 5th Generation architectures.

Image formats

We aim to increase the number of supported image formats in future releases as we continue to develop Frame Advisor. If you experience any problems with unsupported image formats, please contact the Arm® Performance Studio team. See [How to get help](#).

Minimum recommended hardware

For the best experience when using Frame Advisor, Arm recommends using a monitor with a minimum spec of 1080P 1920x1080 at 1:1 scaling.

2. Get started with Frame Advisor

Use Arm® Frame Advisor to quickly capture and analyze frames from a mobile game running on a connected Android device.

This tutorial helps you to learn how to interpret the data and find ways to optimize your application. It describes how to:

- Perform [Setup tasks](#) to prepare your computer and device.
- [Capture a frame burst](#) from a mobile game running on a connected device.
- [Analyze the results](#) to look for performance problems.



Note

Frame Advisor is a new tool, which is still in active development. If you encounter problems, or have a feature request, please send feedback to performancestudio@arm.com.

2.1 Setup tasks

Follow these steps to set up your computer and device so that you can analyze your application with Arm® Frame Advisor.

Before you begin

- Frame Advisor supports applications built with OpenGL ES versions 2.0 to 3.2, or Vulkan versions 1.0 to 1.2. For OpenGL ES applications, your device must be running Android 10 or later. For Vulkan applications, your device must be running Android 9 or later.
- Ensure you have installed Android Debug Bridge ([adb](#)). adb is available with the Android SDK platform tools, which are installed as part of [Android Studio](#). Alternatively, you can download them separately as part of the [Android SDK platform tools](#).
- [Download Arm Performance Studio for free](#) and follow the installation instructions in the [Arm Performance Studio Release Note](#).

Procedure

1. Connect your device to your computer through USB and ensure that the device is switched on.
2. Enable [developer mode](#) on your device.
3. On your device, go to **Settings > Developer Options** and enable **USB Debugging**. If your device asks you to authorize connection to your computer, confirm the connection. Test the connection by entering the `adb devices` command in a command terminal. If successful, the command returns the device ID.

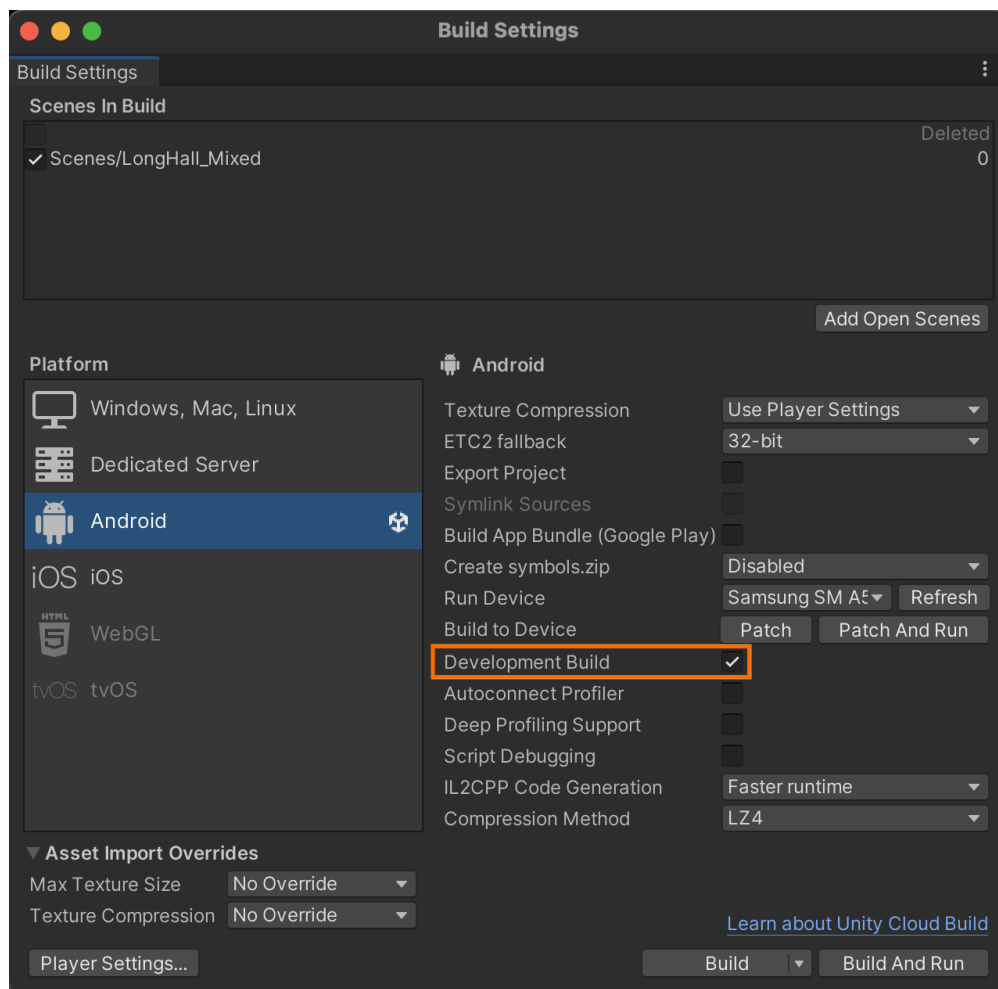
```
adb devices
List of devices attached
cel2345abdcfla1234    device
```

If you see that the device is listed as unauthorized, try disabling and re-enabling **USB Debugging** on the device, and accept the authorization prompt to enable connection to the computer.

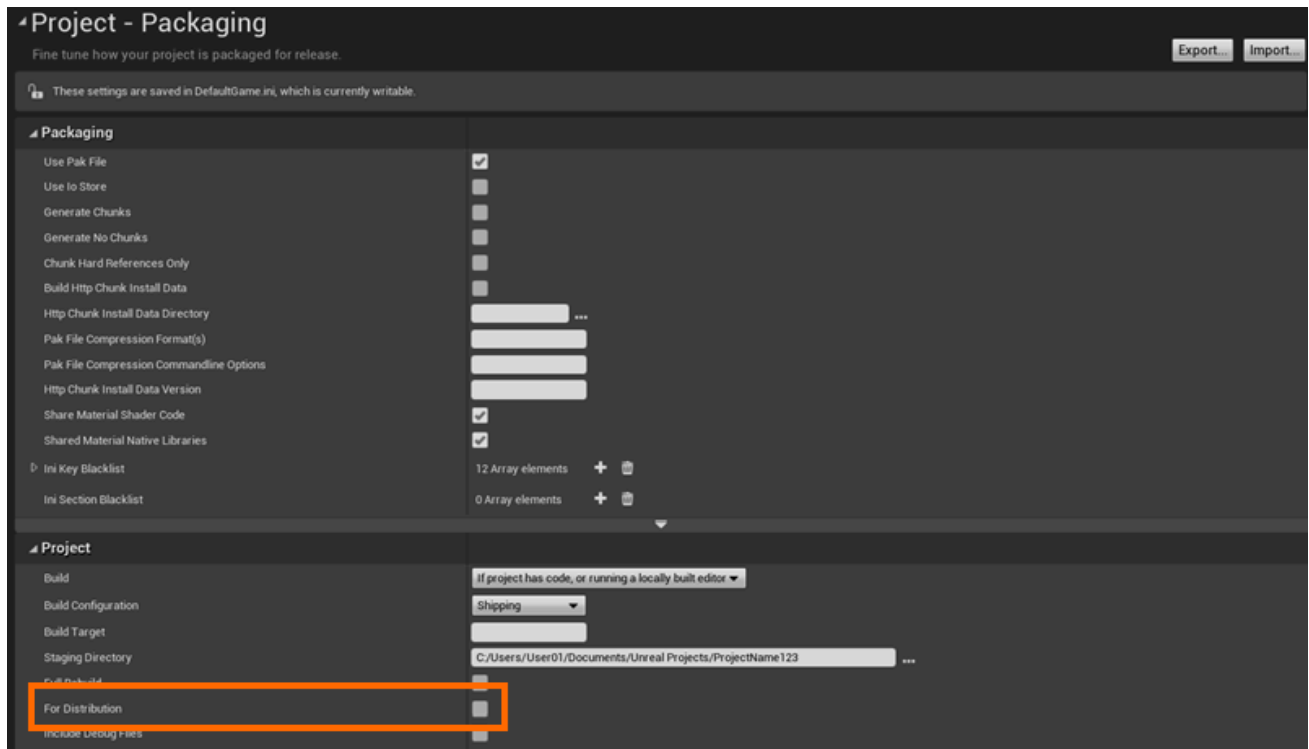
4. Disable permission monitoring. If your phone has **Settings > Developer options > Disable permission monitoring**, ensure that **Disable permission monitoring** is selected. If you do not have the option to disable permission monitoring, you can ignore this step.
5. Install a debuggable build of the application you want to profile on the device. In Android Studio, create a build variant that includes `debuggable true` in the build configuration. Or you can set `android:debuggable=true` in the application manifest file.

In Unity, select **Development Build** under **File > Build Settings** when building your application.

Figure 2-1: Unity Build Settings



In Unreal Engine, open **Project Settings > Project > Packaging > Project**, ensure that the **For Distribution** checkbox is not set.

Figure 2-2: Unreal Engine Build Settings

Next steps

Now that your computer and device are connected and set up, the next step is to [Capture a frame burst](#).

2.2 Capture a frame burst

Use Arm® Frame Advisor to capture frames from your application running on the connected device.

Procedure

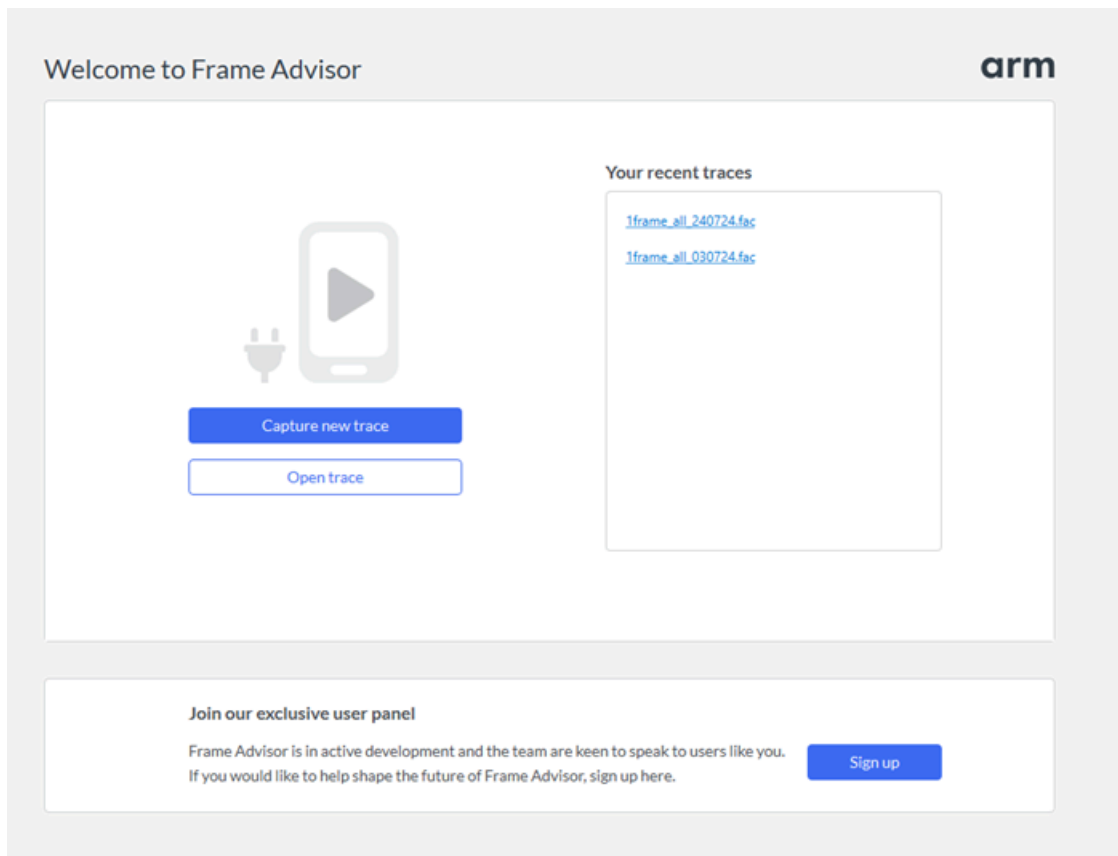
1. Open Frame Advisor:

- On Windows, from the **Start** menu, navigate to **Arm Performance Studio <version>** and select **Arm Frame Advisor <version>**.
- On macOS, navigate to the <install_directory>/frame_advisor folder, and double-click the FrameAdvisor-gui file.
- On Linux, navigate to the <install_directory>/frame_advisor directory in a terminal, and run the FrameAdvisor-gui file:

```
cd <install_directory>/frame_advisor
./FrameAdvisor-gui
```

2. When Frame Advisor launches, you can either capture a new trace or open an existing one.

Figure 2-3: Frame Advisor launch screen



Select **Capture new trace** to capture your problem frames.

3. Frame Advisor lists the connected devices and the applications installed.

Figure 2-4: Device connection screen

Select device

Filter search

Refresh

Status	Name	Serial
Connected	SM-A325F	RF8R41HN7DA

Select application

☒ Show debuggable only

Filter search

Refresh

Debug	Application	Activity
✓	com.android.gl2jni	.GL2JNIActivity
✓	com.Arm.ArmSampleAPK	com.unity3d.player.UnityPlayerActivity
✓	com.arm.pa.paretrace	.Activities.SelectActivity

Configure session

API settings

☒ OpenGL ES
 ☐ Vulkan

Application settings

☐ Pause on connect
 ☐ Terminate on disconnect

Next >

Cancel

Select your device, and the application that you want to evaluate. If your device or application is not listed, see [My device is not listed in Frame Advisor](#), or [My application is not listed in Frame Advisor](#) for help.

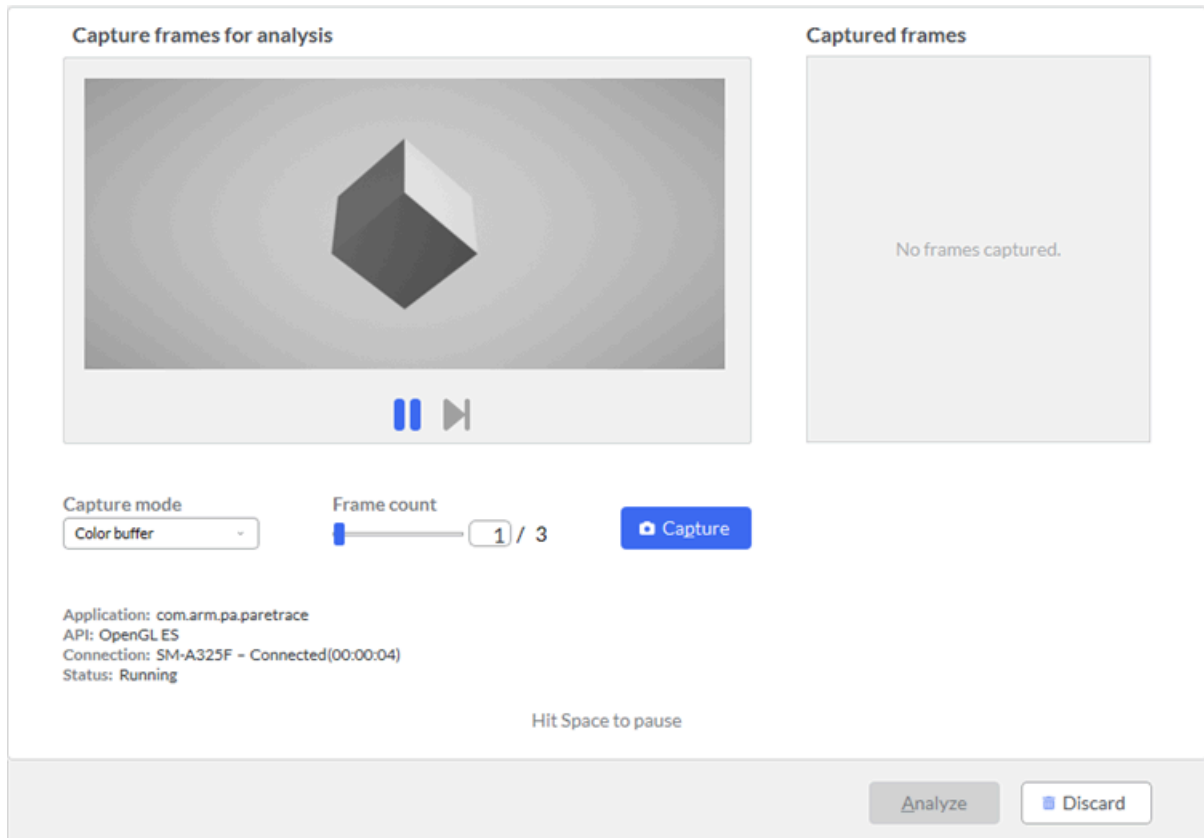
If your application uses the Vulkan API, change the selection in the API settings to **Vulkan**.

Optionally, you can use the **Application settings** to pause the application when Frame Advisor connects to it, and to stop the application when the capture is complete.

Click **Next >** to continue.

Unless you chose the **Pause on connect** option in the **Device connection** screen, the application starts automatically on the device.

- The **Capture** screen provides options for your capture session.

Figure 2-5: Frame Advisor capture screen

The default **Capture mode**, **Color buffer**, captures only color framebuffer content. Optionally change the mode to **All attachments**, to capture color and depth framebuffer content, or capture only **Overdraw** content, if that is what you are interested in. You can also adjust the **Frame count** to the number of frames that you want to capture.

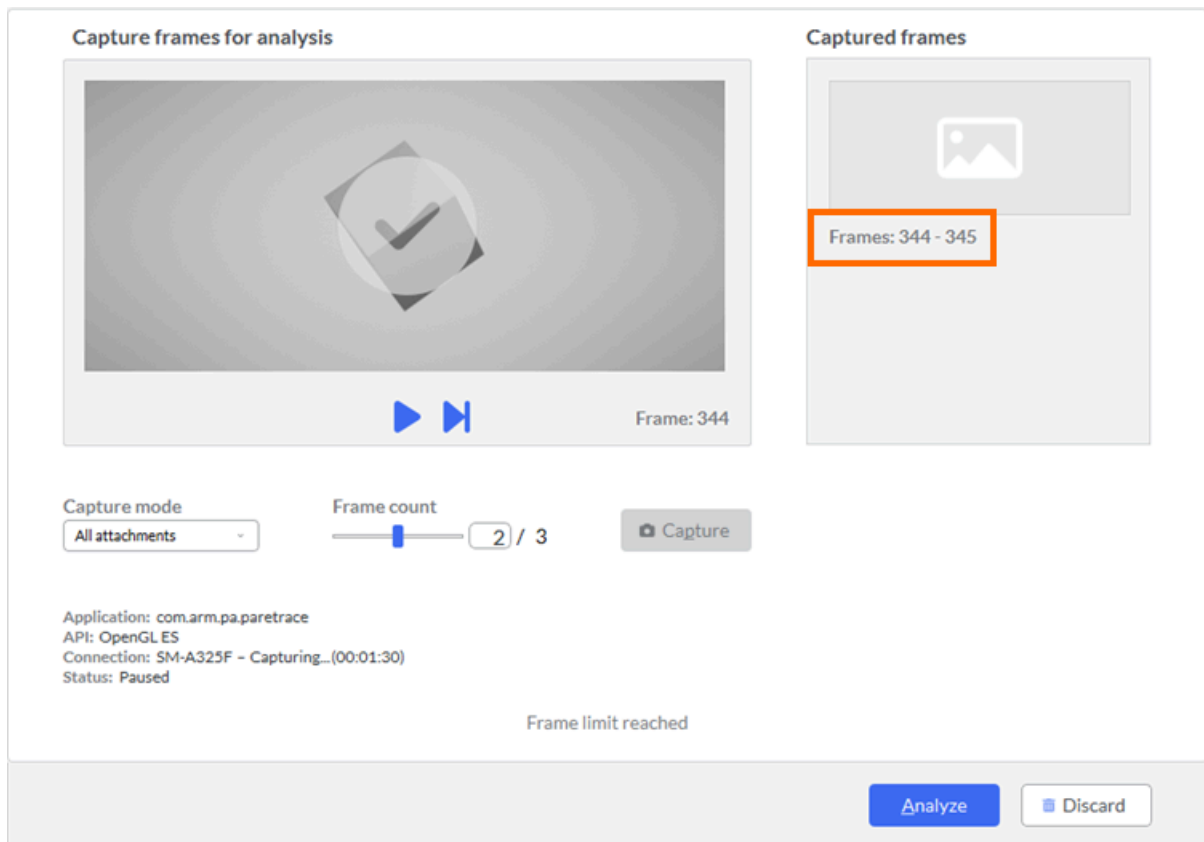


Note

You can capture up to 3 frames. The larger the frame count, the longer the time it takes for Frame Advisor to capture the frames.

5. When you get to the part of your game that you are interested in, optionally click the **Pause** button to pause the application just before you get to the problem frame. The current frame number is displayed.
When the application is paused, use the **Advance one frame** button to step through the frames more accurately.
6. Click the **Capture** button to start capturing the frames.
Frame Advisor starts capturing from the start of the next frame. This may take a few minutes to complete, depending on how many frames you are capturing.

When Frame Advisor finishes capturing the frames, the frame burst is displayed with the captured frame numbers.

Figure 2-6: Captured frame numbers

7. Click the **Analyze** button to see the results. It may take a few minutes to analyze the data.

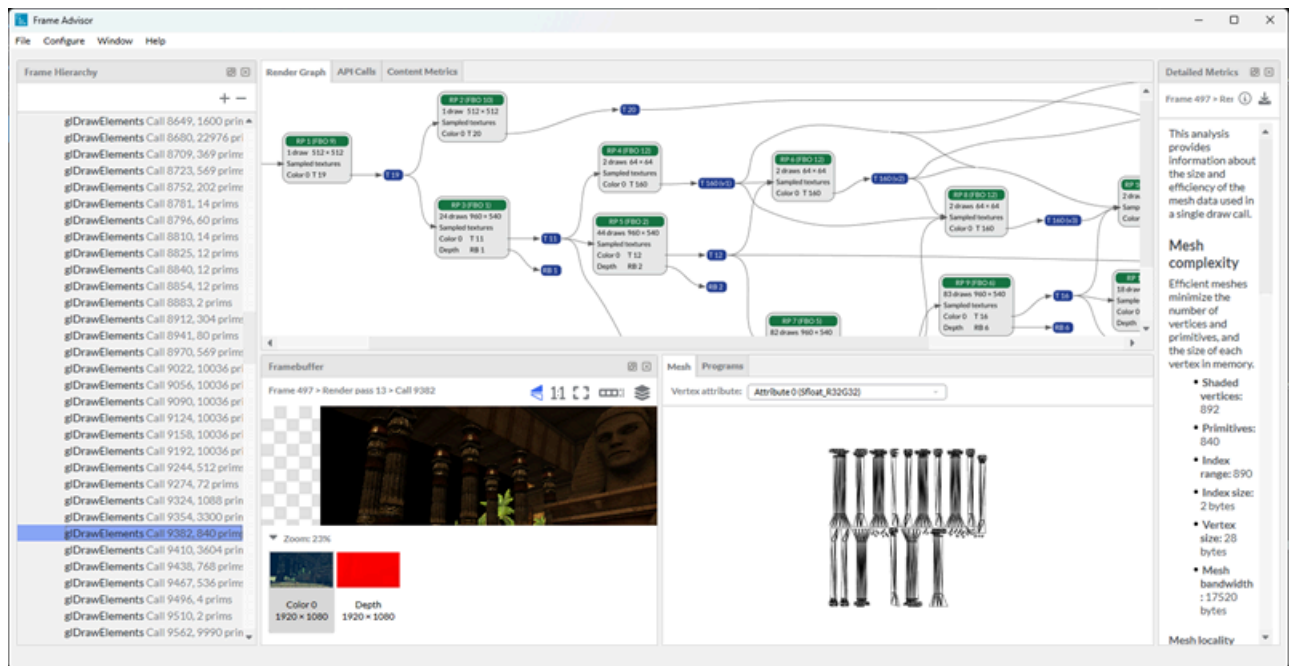
Next steps

Now that you have captured a frame burst, it's time to look at the results. See [Analyze the results](#).

2.3 Analyze the results

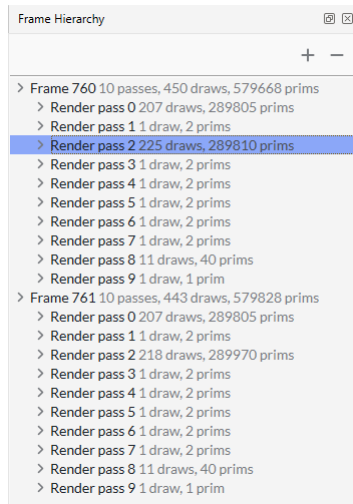
When you have captured some frames, click **Analyze**. Arm® Frame Advisor processes the data and presents it in the **Analysis** screen. Explore each frame to evaluate how efficiently they were rendered on the device.

Figure 2-7: Example Analysis screen



Frame Hierarchy view

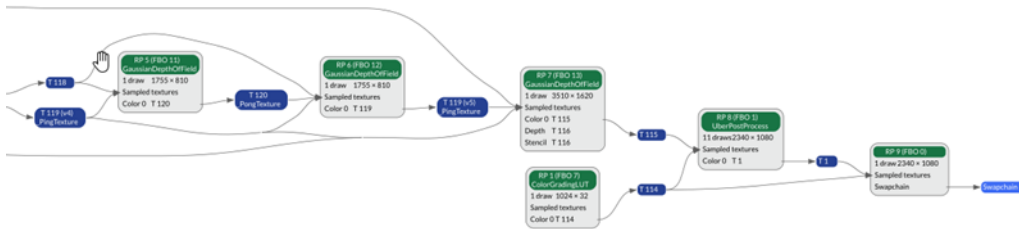
The **Frame Hierarchy** view lists the captured frames in a tree structure. Expand a frame to view the render passes and draw calls within it. Select an item in the hierarchy to navigate to it, and to show its output data for the item in other views of the **Analysis** screen.

Figure 2-8: Frame Hierarchy view

Contextual information helps you to find frames that need further investigation, such as those with the highest number of draws or primitives. See [Navigating your frames](#).

Render Graph view

The **Render Graph** shows an overview of the processing operations that are performed to create the final rendered frame. See the data flow between render passes in the frame, and how resources, such as textures, are produced and consumed.

Figure 2-9: Render Graph

Data is processed from the left side of the render graph to the right side of the render graph. The **Swapchain** node represents the end of the rendering process, and shows the final image displayed on the screen of your device. Render passes are the main processing blocks in the rendering pipeline, which are identified when either:

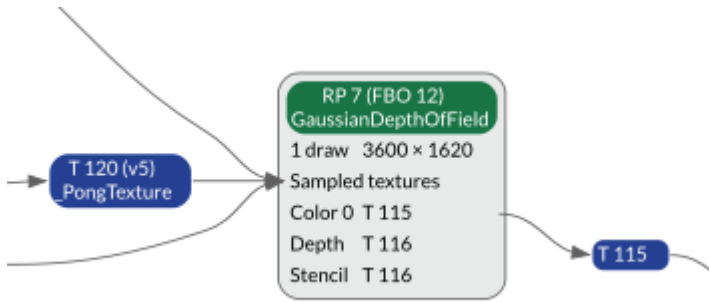
- An API call changes the framebuffer, or when queued rendering is flushed in the OpenGL ES API.
- A `VkRenderPass` is specified in the Vulkan API.

A render pass node in the render graph provides useful information about the render pass:

- Labels that show the render pass name, API name and user label
- Draw count and resolution

- Output attachments

Figure 2-10: Render graph label

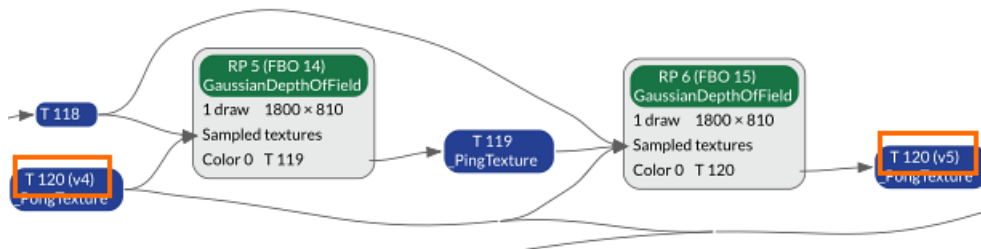


Short form names are used on the label:

- RP: Render pass
- FBO: Framebuffer object
- T: texture
- RB: Renderbuffer

Render passes consume resources and input attachments to render an output image to a set of attachments. Resources, represented by the dark blue nodes, are textures and renderbuffers that are consumed and modified by render passes. When a resource is modified by a renderpass, a version number, which increments each time the resource is modified is displayed on the label. Resources that are not modified do not show a version number.

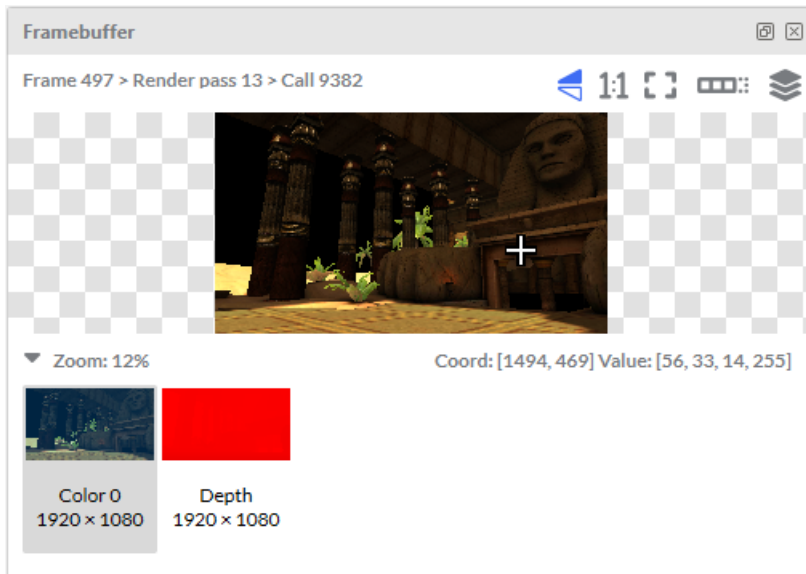
Figure 2-11: Render graph resource version



Use the **Render Graph** to identify expensive parts of a frame that process workloads inefficiently, and see where you can improve workload performance by minimizing the number of external memory accesses. See [Analyze workloads using the render graph](#).

Framebuffer view

The **Framebuffer** view visualizes the rendered output for your captured frames. See how your frames are composed from the GPU rendered output by stepping through each draw call.

Figure 2-12: Framebuffer view

To help you analyze rendering efficiency, use the **Framebuffer** view to check object ordering, and levels of overdraw and shading. See [Analyze object rendering](#).

Mesh view

The **Mesh** view shows the mesh of objects in your rendered output. Use the **Mesh** view with the **Framebuffer** view to check complexity of objects in relation to their size and position on screen. Use the **Vertex attribute** menu to visualize different attributes of your mesh.

Figure 2-13: Mesh view

This view helps you to identify very small triangles, long and thin triangles, and draw calls that submit the same geometry, all of which use a lot of processing power. See [Analyze object complexity](#).

Programs view and Source view

The **Programs** view shows a table of performance metrics for shader programs that are used for a selected workload and stage in the shader pipeline. Use the **Programs** view to find which shader programs you can optimize to ensure the shader cores are kept busy with work to process, and help the GPU run more efficiently.

Figure 2-14: Programs view

Program	Shader	Shortest path			Longest path			Total emitted			Occupancy
		A	LS	T	A	LS	T	A	LS	T	
376	377	0.08	1.00	0.00	0.08	1.00	0.00	0.08	1.00	0.00	
310	311	0.92	1.00	0.00	0.92	1.00	0.00	0.92	1.00	0.00	
370	371	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
346	347	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
337	338	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
331	332	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
295	296	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
280	281	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
265	266	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
250	251	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
235	236	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	
220	221	1.08	1.00	0.00	1.08	1.00	0.00	1.08	1.00	0.00	

To find computationally expensive shader programs, and see how efficiently your shader programs are processed by the GPU in the connected device, use the filter to focus your attention on a specific workload or shader stage. The table of metrics provides a breakdown of cycle cost for the functional units in the shader core, as well as resource usage.

Sort the metrics to identify parts of the shader code that you can simplify to improve processing. Double-click anywhere along a row that is of interest to open the **Source** view, where you can review the associated shader code.

Figure 2-15: Source view

```

#version 300 es
#ifdef GL_ES
precision highp float;
#endif
#ifdef GL_ES
#define highp
#define mediump
#define lowp
#endif
//TODO: szetbontani 2 pass-ra, 1 init feedback shader, 1 advect feedback shader

layout (location = 0) in vec3 in_Pos; // particle instance's position
layout (location = 1) in vec3 in_Age01_Speed_Accel;

layout (location = 2) in vec3 in_Amplitude;
layout (location = 3) in vec3 in_Phase;
layout (location = 4) in vec3 in_Frequency;
layout (location = 5) in vec3 in_T;
layout (location = 6) in vec3 in_B;
layout (location = 7) in vec3 in_N;
layout (location = 8) in vec4 in_Velocity;

//ifdef USE_UBOs
//
//include emitterAdvectConsts;
//
#else
#define MAX_SUBSTEPS 5 //MUST match with kcl_particlesystem2.h/KCL::_emitter::max_num_substeps
!!!!
uniform ivec4 particleBufferParamsXYZ_pad[ MAX_SUBSTEPS ]; //startBirthIdx,
endBirthIdx, noOverflow

uniform mediump_vec4 emitter_apertureXYZ focusdist:

```

For example, to help keep the GPU busy and working efficiently, see if the shaders are spilling variables to stack memory and where you need to reduce precision to 16-bit. To determine if you

can improve object meshes or rendering to reduce processing cost, compare the metrics in the **Programs** view with the visualizations in the **Framebuffer** view and **Mesh** view.

See [Analyze shader programs](#) to learn more about shader analysis.

Content Metrics view

The **Content Metrics** view shows a table of calculated rendering metrics broken down by frame, render pass, and draw call.

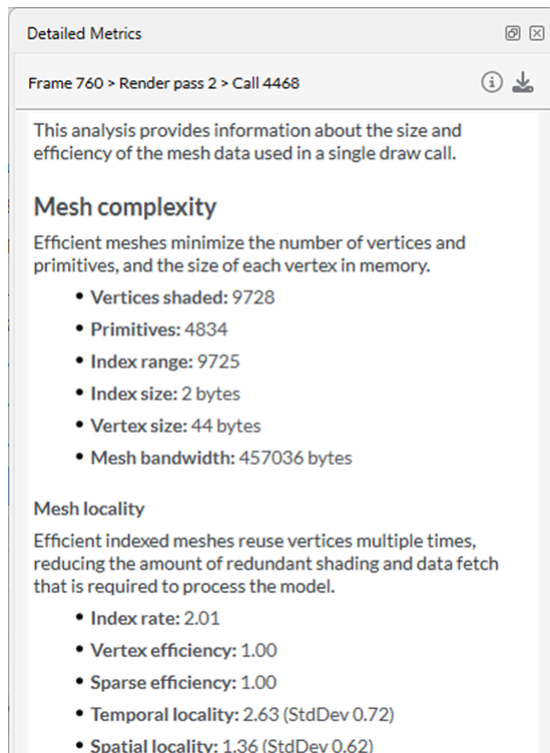
Figure 2-16: Content Metrics view

Render Graph		API Calls	Content Metrics					
Frames	Render passes	Draws						
Frame 497 > Render pass 13 ×								
Frame	Render pass	Call	Prims	Unique indices	Vert size	VSE	VME	
497	13	9274	72	152	28	0.45	0.97	
497	13	9324	1088	2176	28	0.91	0.93	
497	13	9354	3300	2371	28	0.98	1.00	
497	13	9382	840	890	28	0.96	1.00	
497	13	9410	3604	4955	28	0.71	1.00	
497	13	9438	768	940	28	0.99	1.00	
497	13	9467	536	336	28	0.97	1.00	
497	13	9496	4	8	36	1.00	0.70	

To find computationally expensive draw calls, filter the data to narrow it down to a range of interest. Sort the metrics to identify large draw calls, or calls that show signs of inefficiency. For example, to find the most complex draws in the render pass, filter and sort the table by the number of primitives. See [Analyze model geometry](#).

Detailed Metrics view

The **Detailed Metrics** view helps you to understand the efficiency of a mesh for a selected draw call, and its impact on memory bandwidth.

Figure 2-17: Detailed Metrics view

Complicated meshes containing [many triangles](#) are expensive for the GPU to process and, depending on the object's position and size on screen, may get no benefit from the added complexity. Reducing the number of triangles in a mesh dramatically reduces shader cost. If the object's silhouette is accurately drawn, you can use textures or normal maps to efficiently create the detail within the silhouette.

Just as important as complexity, is whether the object's mesh is drawn with good [index reuse](#), to minimize the number of unique vertices required per triangle. See [Content metrics in detail](#).

API Calls view

The **API Calls** view shows the function calls that were made by the application, and what values were returned from the graphics system. When you select a draw call in the **Frame Hierarchy** view, the API call that requested the draw is highlighted in the **API Calls** view.

Figure 2-18: API Calls view

Ctx	Call	Return val	Error statu	Function call
0x76...	4460		GL_NO_ER...	glUniform4fv(location=63, count=16, value=[64 items...])
0x76...	4461		GL_NO_ER...	glBindBuffer(target=GL_COPY_WRITE_BUFFER, buffer=1
0x76...	4462		GL_NO_ER...	glBufferSubData(target=GL_COPY_WRITE_BUFFER, offse
0x76...	4463		GL_NO_ER...	glBindBufferBase(target=GL_UNIFORM_BUFFER, index=C
0x76...	4464		GL_NO_ER...	glBindBuffer(target=GL_ARRAY_BUFFER, buffer=2959)
0x76...	4465		GL_NO_ER...	glVertexAttribPointer(index=0, size=3, type=GL_FLOAT, nc
0x76...	4466		GL_NO_ER...	glVertexAttribPointer(index=1, size=3, type=GL_FLOAT, nc
0x76...	4467		GL_NO_ER...	glVertexAttribPointer(index=2, size=4, type=GL_FLOAT, nc
0x76...	4468		GL_NO_ER...	glDrawElements(mode=GL_TRIANGLES, count=14502, ty

Search the function calls table for mis-used rendering commands, and check for error messages for each draw call that might indicate issues that affect performance. See [Analyze function calls](#).

Related information

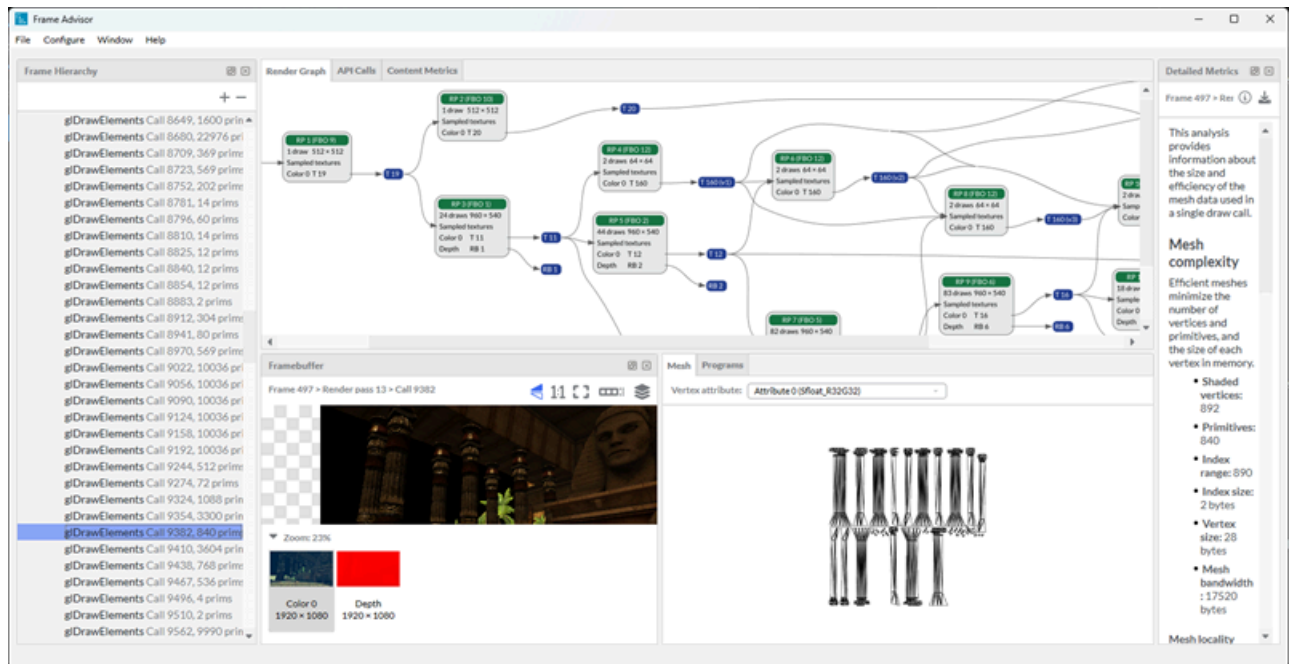
- [Analyzing your frames](#)
- [How to get help](#)

3. Analyzing your frames

The **Analysis** screen presents analysis data and visualizations in a set of distinct views, which enable you to investigate different aspects of your captured frames. Navigate between the views to identify problem areas in your application, or areas where you can improve performance.

The **Analysis** screen opens after you have captured your frames and clicked the **Analyze** button. To analyze an existing trace file, open the **Landing** screen and click **Open trace**.

Figure 3-1: Example Analysis screen



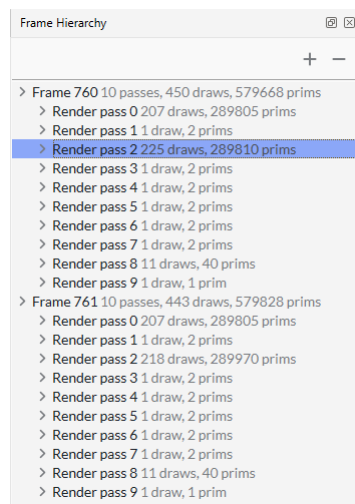
3.1 Navigating your frames

The **Frame Hierarchy** view lists the frames that you captured from your application in a tree structure. Expand the frames to see the render passes and draw calls contained within them. Select an item to update the information shown in other views of the **Analysis** screen.

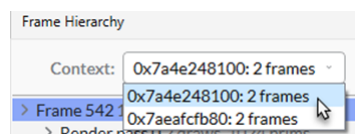


You must select or undock the **API Calls** and **Content Metrics** views to see them.

To help you to find items that require further analysis, each item in the **Frame Hierarchy** view is presented with contextual information. For example, you may want to select a frame that contains the most draws or primitives to investigate it further.

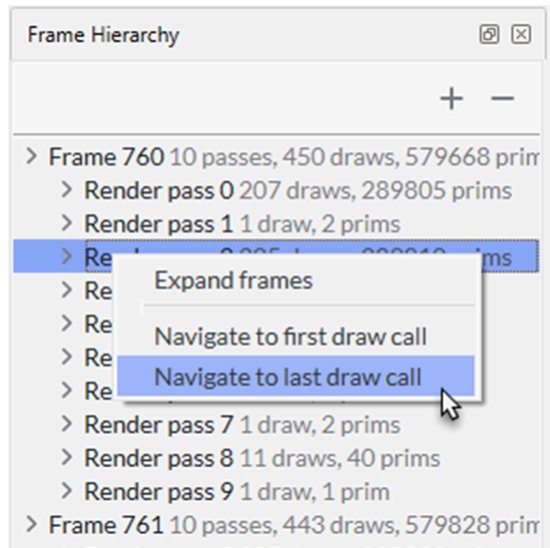
Figure 3-2: Frame Hierarchy view

If your trace uses multiple contexts, you can choose the context from the drop-down menu. If your trace uses one context only, the drop-down menu is not visible.

Figure 3-3: Context selection

As you select a frame, the **Render Graph**, **Framebuffer**, and **API Calls** views update to show the output from the last draw call in the frame. Use these views with the **Frame Hierarchy** to narrow down the problem area.

You can use the right-click menu to navigate directly to the first or last draw call in the render pass, which is useful if a render pass contains a high number of draw calls. Alternatively, use the mouse or the arrow keys to step through the draw calls in a render pass.

Figure 3-4: Navigate to the last draw call

Here are the kinds of problems to look for:

- Look for render passes with high numbers of draw calls. Draw calls are expensive for the CPU to process, so it is important to reduce them where possible. Check if these draw calls actually render visible changes to the framebuffer. If not, look at [software culling techniques](#) to see if they can be eliminated. Draws could be outside of the frustum, or behind other objects.
- Look for instances where many identical objects are each being drawn individually. There could be an opportunity to reduce the number of draw calls by [batching multiple objects](#) into a single draw.



Tip

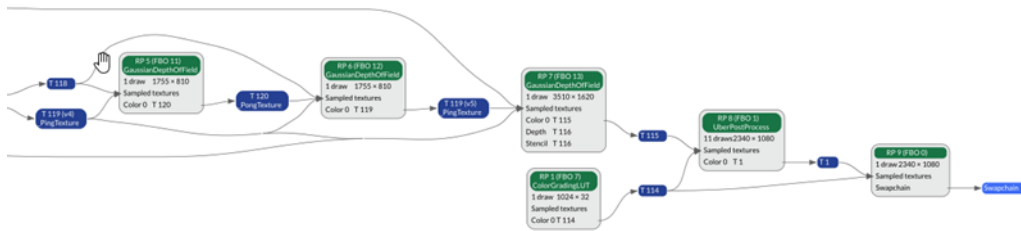
To learn more about how to avoid common API problems when building graphics for mobile, watch *Episode 2.4* of the *Mali GPU training series* [Engine and API best practices](#)

3.2 Analyze workloads using the Render Graph

The **Render Graph** shows the rendering operations performed for the frame selected in the **Frame Hierarchy** view. See how your GPU workloads are processed in the API sequence by analyzing the data flow between render passes, as well as the resources that are produced and consumed by the render passes. You can then use this information to identify expensive parts of a frame that process workloads inefficiently, and improve workload performance.

Procedure

1. In the **Frame Hierarchy** view, select the frame you are interested in exploring further. The **Render Graph** visualization updates to show the render passes and resources for the selected frame.

Figure 3-5: Example render graph

Click on a render pass in the graph to navigate to it, or right-click and navigate to the first or last draw call of that render pass.

2. In the **Render Graph** view, use your mouse to zoom and pan around the graph. See how workloads and resources are processed to create the rendered output for the selected frame. Identify any render passes that are doing the most work, which you can optimize.



Right-click in the graph area and select **Fit to window** for an overall view of rendering operations performed for the selected frame. Click **Reset view** to go back to the initial zoom level.

To help you understand what is happening for each render pass, contextual information is provided about the workload performed, such as:

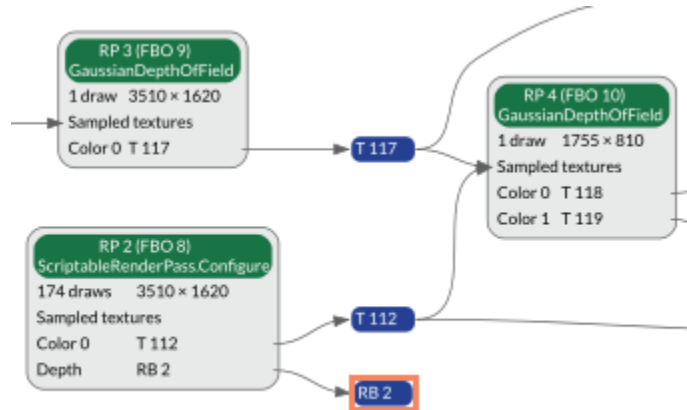
- Index of the render pass, and the framebuffer object that it relates to.
 - Number of draw calls in the render pass.
 - Rendered resolution
 - Input and output attachments
3. Click a render pass to navigate to it and update information shown in the other views of the **Analysis** screen.
 4. Look for render passes that are consuming bandwidth and accessing memory unnecessarily, such as:
 - Render passes where the attachment of one render pass is immediately input to another without being used. Merging these render passes can prevent excess memory reads and writes.
 - Render passes that do not clear or invalidate input attachments at the start of the render pass. Reading these input attachments from external DRAM at the start of the render pass consumes read memory bandwidth. If you do not want to use the rendered output from a previous frame, ensure that you clear or invalidate all input attachments at the start of the render pass, before any draw calls are made.

You can see if clears and invalidates exist within a render pass in the **Frame Hierarchy** view, and in the **API Calls** view.

- Render passes that do not invalidate output attachments at the end of the render pass. Writing these output attachments to external DRAM at the end of a render pass consumes

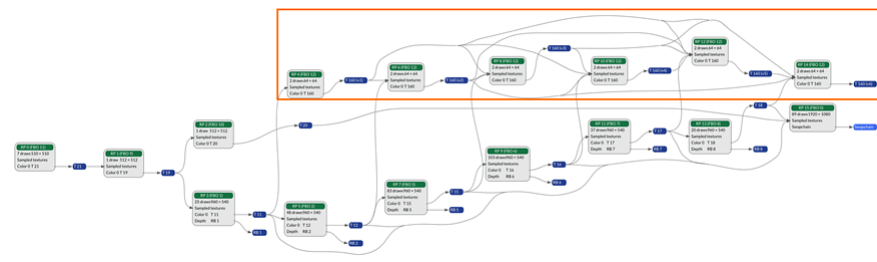
write memory bandwidth. If the output attachment is no longer needed after the render pass, ensure that you invalidate it so that it is discarded at the end of the render pass.

Figure 3-6: Output attachment that requires discarding



- Render passes where output attachments are not used in the final rendered output. Ensure that all render passes that are submitted to the GPU are actually useful rendering operations, with no redundant rendering. Discard any unused attachments at the end of the render pass so that they do not consume memory bandwidth.

Figure 3-7: Unused nodes

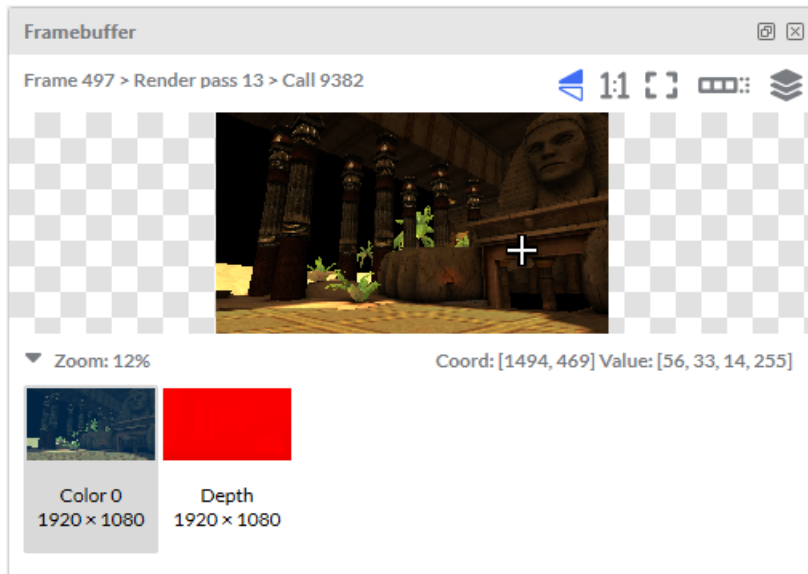


3.3 Analyze object rendering

The **Framebuffer** view shows a visualization of the rendered output of your captured frames. The visualization enables you to review what was rendered to the screen for each draw call. Review the framebuffer output to identify problem objects that do not follow best practice standard, and see where you can improve the performance cost of draw calls.

Procedure

- In the **Frame Hierarchy** view, expand a frame that you are interested in exploring further to see the render passes and draw calls within it.
The **Framebuffer** visualization updates to show the rendered output after the final draw call of the selected frame.

Figure 3-8: Framebuffer view

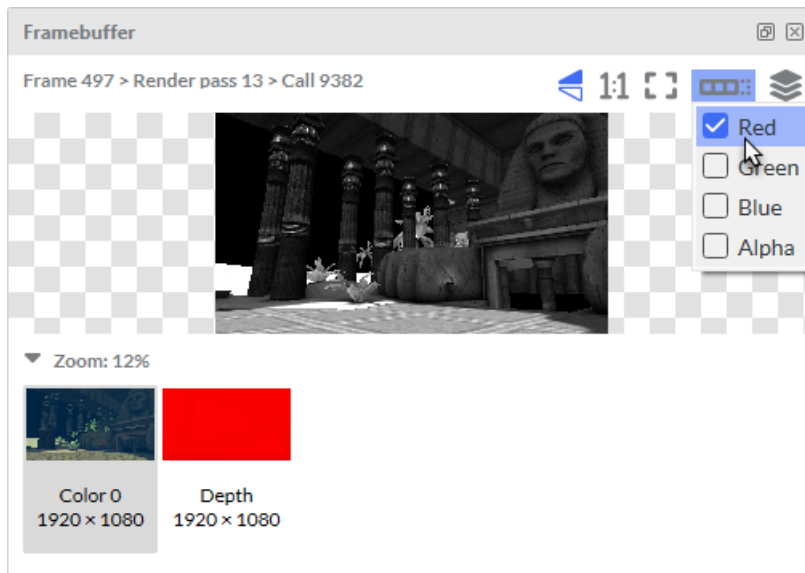
If required, use the **Flip vertically** button to adjust the vertical orientation so that it appears the correct way around.


2. Expand a render pass in the **Frame Hierarchy** view, then use the arrow keys or mouse to step through each draw call in sequence.
See how your frame is composed by watching how the framebuffer changes as you step through the draw calls. Any color, depth and overdraw attachments used in the framebuffer are shown as thumbnails underneath the framebuffer visualization. Click the thumbnail to show the rendered attachment.



You can temporarily hide the thumbnails of the framebuffer attachments to provide more space for the framebuffer visualization.

3. To see more detail, use the mouse controls to zoom and pan around the visualization. To quickly rescale the image, use the appropriate button:
 - **Zoom to 1:1** button to resize the image to the original size.
 - **Fit to window** button to resize the image inside the visualization window.
4. If you store texture information in RGBA channels, you can select any combination of the channels so that you can look at just the data that you care about.
Click the **Select RGBA channels** button, and select the channels that you want to show.

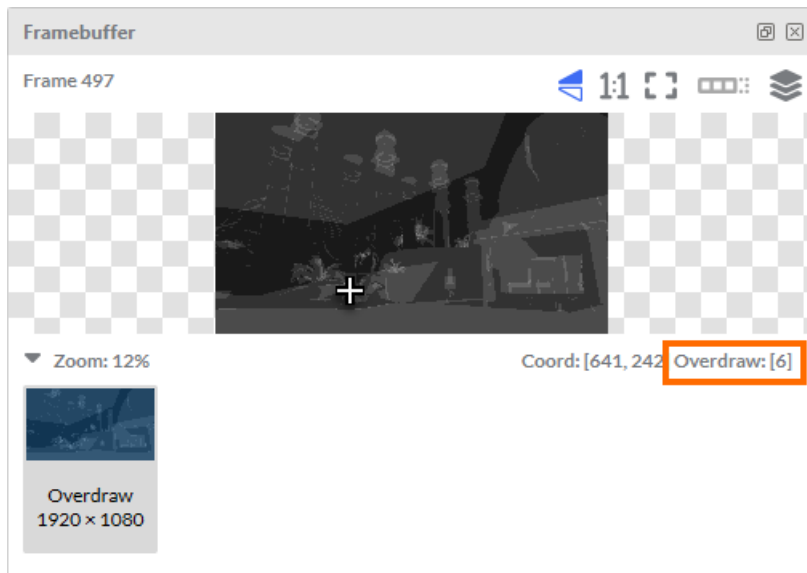
Figure 3-9: Select RGBA channel

The RGBA color values, and the image coordinates, for the pixel under the cursor are displayed in the framebuffer by default. Click the **Select pixel information** button  and clear the check box if you do not want to see this information.

5. Check the ordering of objects to avoid high levels of overdraw.
First ensure that opaque objects are rendered in order from front-to-back, starting with objects closest to camera. This order enables the GPU to use [early ZS testing](#) to prevent shading of occluded objects, which reduces overdraw and unnecessary processing before fragment shading.

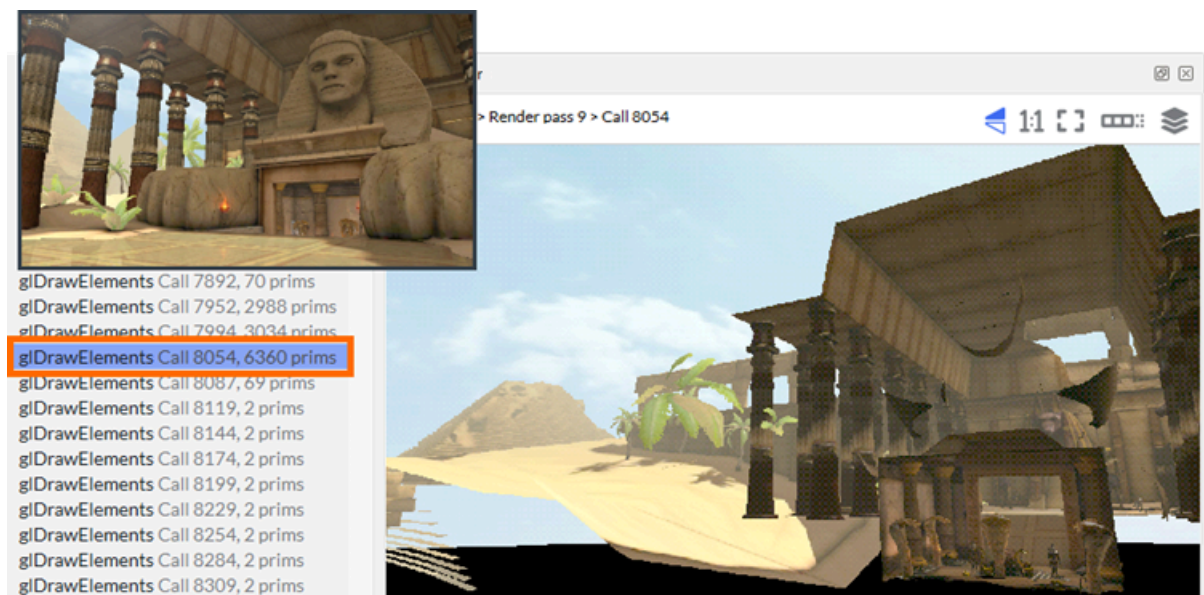
Then render transparent objects in a back-to-front order, starting with the objects that are furthest away from the camera. This order helps blending because it ensures that objects that are further away from the camera are already in the framebuffer when the transparent layer is rendered.

6. If you captured frames using **Overdraw mode**, click the overdraw thumbnail to show it in the visualization area. Move the cursor around to see the levels of overdraw in the rendered output. Every time a pixel is rendered to the framebuffer, the overdraw value increases, and the lighter it appears in the final image.

Figure 3-10: Framebuffer overflow view

Overdraw occurs when the same pixel is shaded on many layers. High levels of overdraw affects performance negatively because shading pixels, which are then overwritten before they are shown, wastes GPU resources by unnecessary processing.

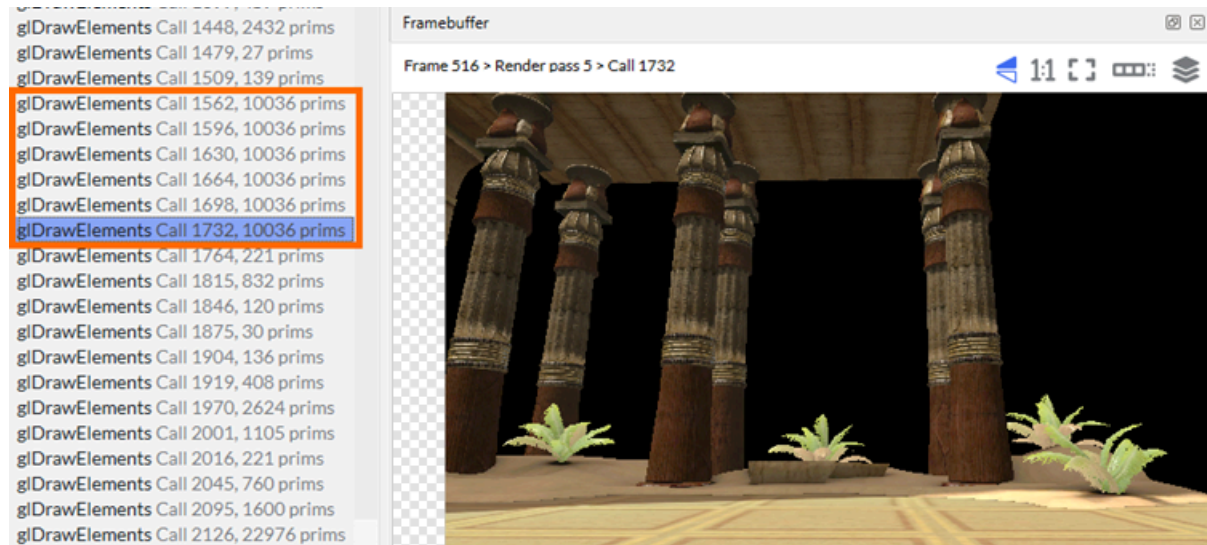
7. As you step through the draw calls, identify any draw calls that do not make visible changes to the framebuffer output.
Check if your application is drawing objects that are either not visible on screen, or are occluded by other objects. Ensure your GPU is **culling** any primitives that are not visible on screen.

Figure 3-11: Unused objects in the Framebuffer

In this example, you can see the output of the selected draw call in the framebuffer visualization. Only a small part of this output is shown in the final rendered frame, displayed inset.

8. As you step through the draw calls, check for multiple identical objects that are drawn individually, instead of being processed as a batch in a draw call.

Figure 3-12: Multiple identical objects in the Framebuffer



In this example, the draw calls in the **Frame Hierarchy** show that each pillar is drawn as an individual object in separate draw calls. Processing individual objects is expensive. To help improve performance, merge the identical objects into a single draw call to reduce the draw call count.

9. If your captured frames contain HDR images, the **Framebuffer** displays a range slider that you can use to adjust the exposure level in the visualization. The minimum and maximum values of the range are determined by data from the image. To reveal details of interest in the darkest or brightest areas of the HDR image, move the minimum and maximum slider controls to a suitable range.



Note

The adjusted minimum and maximum values are maintained as you continue to step through draw calls.

The following example shows the same HDR image, but the range has been adjusted in the second image to reveal more details and lighting effects in the rendered output.

Figure 3-13: HDR image comparison

10. To check if the complexity of objects is appropriate for their size on screen, compare the framebuffer output with the **Mesh** view as you step through the draw calls. Very fine-detailed objects that are further away from the camera can negatively impact performance because of the [triangle density](#) in the mesh of an object. Large numbers of very small triangles are expensive for the GPU to process, often with no visual benefit. See [Analyze object complexity](#).

Use mesh LODs to select simpler geometry as objects move further away from the camera. Check the efficiency of your mesh in the **Detailed Metrics** view. See [Content metrics in detail](#).

Next steps

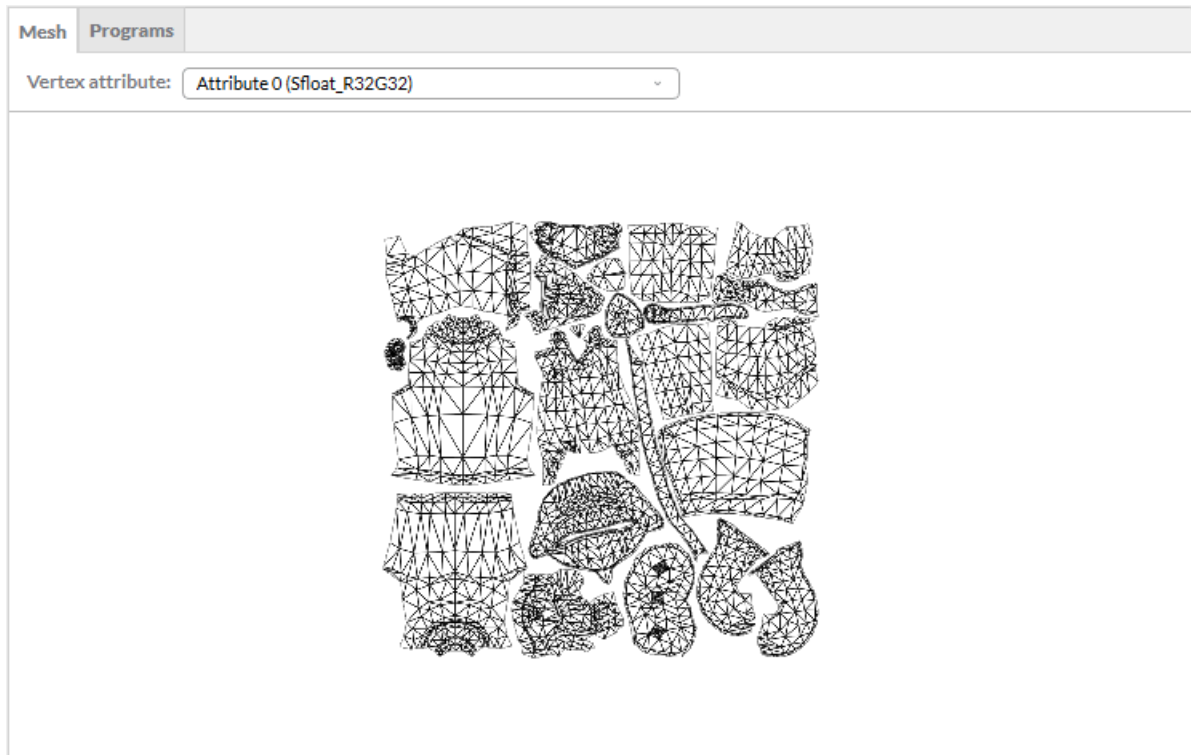
Investigate the efficiency of your mesh further in the **Detailed Metrics** view. See [Analyze object complexity](#) and [Content metrics in detail](#) for more information.

3.4 Analyze object complexity

The **Mesh** view shows a visualization of the primitives drawn by the selected draw call. The visualization enables you to see how an object was constructed, and whether objects are the right level of detail for their size and position on screen. Compare the **Mesh** view with data in the **Framebuffer** and the **Detailed Metrics** views to identify problem objects that do not follow best practice standards, and reduce the processing cost of draw calls.

Procedure

1. In the **Frame Hierarchy** view, expand a frame that you are interested in exploring further to see the render passes within it.
2. Expand a render pass in the **Frame Hierarchy** view to see the draw calls within it, then step through the draw calls to look for objects that have degenerate primitives, or draw calls that have lots of primitives compared to other draw calls.
When you select a draw call, the **Mesh** view shows a visualization.

Figure 3-14: Mesh view

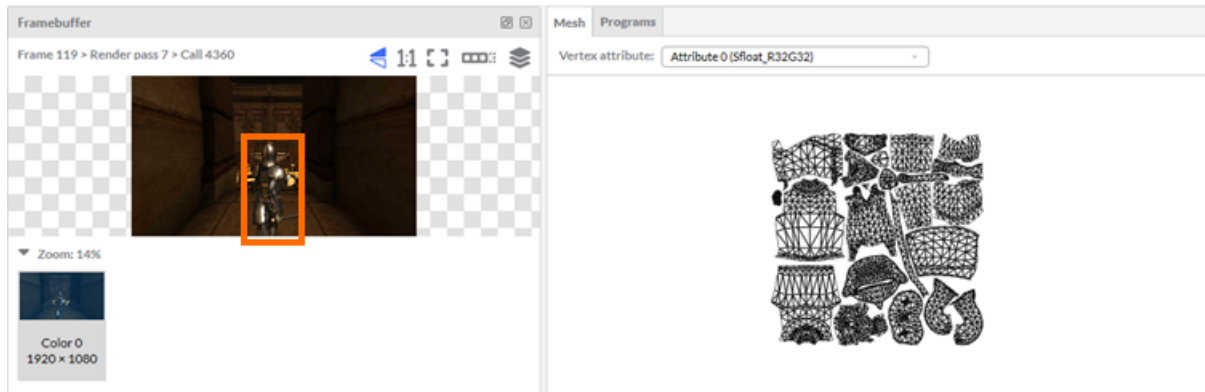
You can select different attributes in **Vertex Attribute** menu.

Use the mouse controls to move around the visualization:

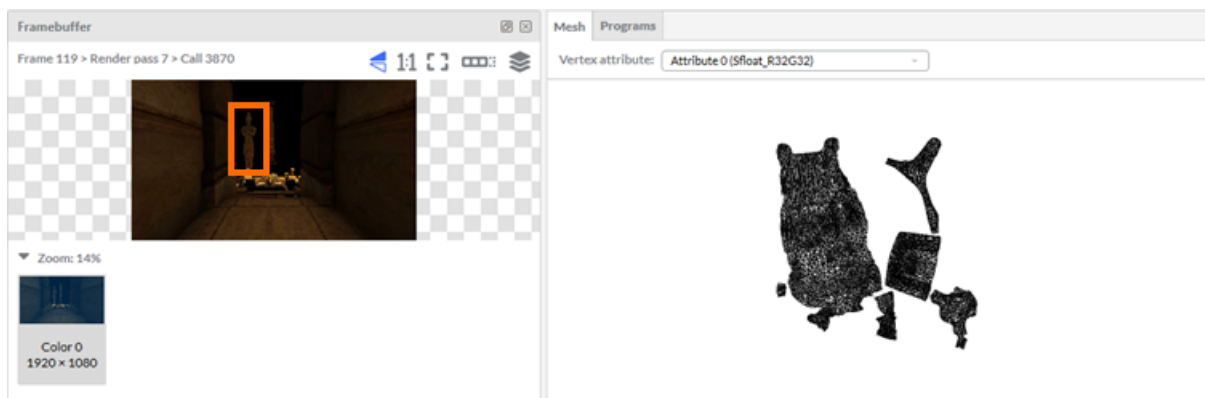
- Drag to rotate
- Shift+click, or hold middle mouse button, to pan
- Scroll the wheel button to zoom
- Right-click and select **Reset view** from the context menu to return the visualization to the default position

3. Look at the visualization in the **Mesh** view to see if all of the primitives are required. Check that the level of detail of the object is suitable based on its distance to the camera. Objects that have a high level of detail, and are further away from the camera, can negatively impact performance because of the [triangle density](#). Large numbers of very small triangles are expensive for the GPU to process, often with no visual benefit. Reduce the level of detail of objects as they move further away from the camera.

The following image shows an appropriate level of detail for the object, considering its distance to the camera.

Figure 3-15: An example of a good draw call

The following image shows an excessive level of detail for the object, considering its distance to the camera.

Figure 3-16: An example of a complex draw call**Tip**

Remove unseen primitives. If your mesh has primitives that are never seen, they do not need to be part of the mesh. Processing occluded primitives wastes bandwidth, even though they are culled in a later step.

- To see where the draw call appears in the **Framebuffer** view, use the arrow keys or mouse to step through the draw calls in the **Frame Hierarchy** view. If you cannot see the drawn mesh appear in the **Framebuffer** view, it is possible that you do not need it. Consider culling the mesh before it is submitted to the GPU.

**Tip**

Merge draw calls that submit the same geometry. Arm® recommends drawing the same geometry in a single instanced draw call to improve efficiency.

5. Look at the **Detailed metrics** view to see a range of metrics about the mesh, including the number of primitives, how many vertices were actually shaded, and how much bandwidth the GPU used to read the mesh. See **Content metrics in detail** for more information.

Figure 3-17: Mesh complexity in Detailed metrics

Mesh complexity

Efficient meshes minimize the number of vertices and primitives, and the size of each vertex in memory.

- Vertices shaded: 45952
- Primitives: 22976
- Index range: 45952
- Index size: 2 bytes
- Vertex size: 28 bytes
- Mesh bandwidth: 781184 bytes

Related information

[Arm GPU best practices](#)

3.5 Analyze model geometry

The **Content Metrics** view shows you data about all the frames, render passes, and draw calls in your trace. Filter and sort the metrics to help you identify opportunities to optimize your model geometry. For example, find out if you have any duplicated vertices or degenerative primitives. From any draw call, you can navigate to the corresponding API call so that you can see exactly which function call to modify.

About this task

In this example, we are looking at the data of a capture to see if there are any vertex shading efficiency (VSE) issues.

Procedure

1. In the **Content metrics** view, select a captured frame. In this example, we captured one frame only.
2. Click the **Render passes** tab.
3. You can sort the columns in the tables to help you find issues. To sort the tables, click a column heading. In this example, we click the **Draw calls** column heading to sort the table so that the render pass with the highest number of draw calls is at the top. Looking at a large number of draw calls gives us a higher chance to find issues.

Figure 3-18: Sorted draw calls column

Frame 332 x				
Frame	Render pass	Draw calls	Prims	Unique indice
332	2	192	247991	247991
332	0	163	247984	247984
332	8	11	80	80
332	1	1	4	4

- Select the top row of the table, which should be the render pass with the highest number of draw calls.
- To see all the draw calls in this render pass, click the **Draws** tab.
- Sort the table so that the draw call with the lowest **VSE** is shown in the top row. A low VSE is caused by high average temporal locality, duplicate vertices, or low index sparseness efficiency metrics.

In this example, you can see that **Call 3084** has a **VSE** of 0.00.

Figure 3-19: Sorted VSE column

Frame 553 > Render pass 2 x							
Frame	Render pass	Call	Prims	Unique indice	Vert size	VSE	VME
332	2	3084	4834	9725	4	0.00	1.00
332	2	3333	5	7	24	0.88	1.00
332	2	2193	45	110	32	0.90	1.00
332	2	2233	45	110	32	0.90	1.00

To find out what is causing the low VSE, open the **Detailed Metrics** view.

- To open the **Detailed Metrics** view for the chosen call, right-click on that call in the table, then select **Navigate to call**. The call is now also selected in the **Frame Hierarchy**, **Mesh**, and the **API Calls** views, and the **Detailed Metrics** view is populated with data about that call.
- In the **Detailed Metrics** view, you can see that nearly all the vertices are duplicated and they are all being shaded. This is why the **Vertex Shading Efficiency** is 0.00.

Figure 3-20: Vertex shading efficiency**Mesh redundancy**

Efficient meshes minimize the amount of redundant data in their models.

- Degenerate primitives: 0
- Duplicate vertices: 9681
- Duplicate vertex percentage: 99.55%
- Vertex padding size: 0 bytes

To improve the VSE and the performance of the application, remove all duplicate vertices.

There are many other metrics you can look at to improve the performance of your application. We are now going to add a column to show degenerate primitives.



9. To add and remove columns of data in the **Content Metrics** view, ensure you are on the **Draws** tab, then click the **Edit columns** icon .
10. Select **Degenerative primitives** from the list of available columns.
11. Sort the **Deg prims** column to see the draw calls ordered by the highest number of degenerate primitives first.
12. Select a draw call that has degenerative primitives.

Figure 3-21: Degenerate primitives

Frames Render passes Draws								
Frame 553 > Render pass 2 × 								
Frame	Render pass	Call	Prims	Unique indice	Vert efficienc	Vert size	Deg prims	
553	2	3793	4834	9725	1.00	44	624	



Note

To clear the filters and see data for all the frames in your trace, click the filters.

13. To locate which API call is creating the degenerate primitives, right-click a draw call that has degenerative primitives, and select **Navigate to call**. The corresponding API call is highlighted in the **API Calls** view.

3.6 Content metrics in detail

The **Detailed Metrics** view helps you to understand the cost of meshes and find opportunities to improve the efficiency of meshes. The recommended mesh memory layout splits the input attributes into two streams. One stream for the attributes needed to compute position, and one stream for the remaining attributes. The recommended mesh memory layout also repacks the attributes to remove any padding, unless it is required for type alignment.

To see detailed metrics about a draw call, right-click a row in the table on the **Draw** tab in the **Content metrics** section, then select **Navigate to call**.

3.6.1 Mesh complexity

Efficient meshes minimize the number of vertices and primitives, and minimize the size of each vertex in memory.

Vertices shaded

The number of vertices shaded by the GPU, factoring in any reshading or overshading effects.

Primitives

The total number of non-degenerate primitives in all instances.

Index range

The difference between the minimum index and maximum index.

Index size

The size of a single index in bytes.

Vertex size

The actual size of a vertex in bytes.

Mesh bandwidth

An estimate of the GPU bandwidth used to read the mesh using the current memory layout.

3.6.2 Mesh locality

Efficient indexed meshes reuse vertices multiple times, which reduces the amount of redundant shading and data fetch required to process the model.

Index rate

The number of unique indices per primitive.

The **Index rate** is a measure of how many vertices are shared across multiple primitives.

Sharing vertices reduces memory usage, which improves efficiency. An **Index rate** of 3 indicates that no vertices are shared, so aim to get your **Index rate** as low as possible.

To reduce your **Index rate**, share as many vertices as possible.

Vertex shading efficiency

The ratio of vertex shader invocations to the number of useful input vertices.

An efficiency of 1 indicates optimal shading efficiency, with 1 shader invocation per useful input vertex. An efficiency of 0.5 indicates 2 shader invocations per useful input vertex.

Low shading efficiency is caused by several issues:

- Referenced indices that are shaded multiple times because of non-optimized temporal locality in the index buffer.
- Referenced indices that duplicate the data payload of other indices that are shaded.
- Non-referenced indices that are shaded because they are in the same group of 4 indices as a referenced index. Arm® GPUs always shade indices in groups of 4 consecutive indices, even if those indices are not used in the draw call.

To maximize vertex shading efficiency, improve the **average temporal locality**, **duplicate vertices**, and **index sparseness efficiency** metrics.

Vertex memory efficiency

The ratio of data that is fetched to how much is needed for processing this mesh.

An efficiency of 1 indicates optimal memory layout. An efficiency of 0.5 indicates that twice as much data is fetched because of an unoptimized memory layout.

Memory inefficiency is caused by:

- Padding between fields in a packed vertex.
- Padding between packed vertices.
- Attributes using an interleaved layout, that causes non-position data to load during position shading.

To maximize vertex memory efficiency, improve the **Mesh memory layout**, and optimize the [precision](#) and [layout](#) attributes as detailed in the [Arm GPU Best Practices Developer Guide](#).

Index sparseness efficiency

The ratio of vertex indices shaded to the number of unique indices in the index buffer.

Arm® GPUs always shade indices in groups of 4 aligned indices, even if those indices are not used in the model. For example, if you want to shade indices 0, 1, and 2, then index 3 is also shaded. If you want to shade indices 1, 6, and 9, then all indices from 0 to 11 are shaded. That is not very efficient because of the unnecessary shading. In this case, try to reconfigure the structure of your mesh so only 0, 1, 2, and 3 are shaded.

An efficiency of less than 1 indicates that the model is overshading because of unused indices in these groups of 4.

To maximize efficiency, minimize overshading by tightly packing meshes so as many indices between the minimum and maximum index value are used.

Average temporal locality

The number of indices, mean and standard deviation, between reuse of an index value.

For example, in the sequence of vertices 0,1,2,0,5,0, the mean is 1.5 because the average number of vertices between the reused vertex of 0 is 1.5.

Arm® Frame Advisor models a 1024-byte entry post-transform cache for computing vertex reshading.

To improve cache efficiency of the post-transform cache during vertex shading, reduce the average temporal locality so that it does not exceed 1024 bytes. To reduce your average temporal locality, reduce the number of indices between reuse of an index value.

Average spatial locality

The index difference, mean and standard deviation, between neighboring indices.

To improve memory access and cache efficiency during vertex shading, reduce the average spatial locality between neighboring indices. To reduce your average spatial locality, reduce the index difference between neighboring indices.

3.6.3 Mesh redundancy

Efficient indexed meshes reuse vertices multiple times, which reduces the amount of redundant shading and data fetch required to process the model.

Instance degenerate primitives

The number of primitives that have zero area.

Degenerate primitives occur when the vertices of a primitive do not form a triangle. Processing degenerate primitives wastes bandwidth because those primitives are not displayed. A significant percentage of degenerate primitives might indicate a mesh encoding issue.

Remove as many degenerate primitives as possible from your model. If you use degenerate primitives for encoding spatial jumps in the mesh, use primitive restart instead.

Duplicate vertices

The number of vertices that have identical data to another vertex in the model.

Processing duplicate vertices wastes bandwidth because the system processes the same vertex more than once.

To reduce the number of duplicate vertices, use an index buffer to ensure each vertex is uploaded to the GPU only once. If that is not possible, remove as many duplicate vertices as possible.

Vertex padding size

The number of bytes of unused padding in each vertex, accounting for space between attributes or between vertices.

When fetching vertex attribute data from memory, the GPU also fetches any unused bytes between attributes or vertices. These padding bytes waste bandwidth.

To improve memory performance, ensure that vertex data is as tightly packed as possible.

3.6.4 Mesh memory layout

Efficient meshes optimize their attribute stream memory layout to reduce the amount of redundant data fetched from main memory. Arm® GPUs, and many other mobile GPUs, compute position and perform culling before running the remaining part of the vertex shader.

The recommended mesh memory layout splits the input attributes into two streams. One stream for the attributes needed to compute position, and one stream for the remaining attributes. This layout is more efficient for the GPU because it looks at data in each stream that is relevant to the process it is running. The ideal layout also repacks the attributes to remove any padding, unless it is required for type alignment.

Mesh bandwidth

An estimate of the GPU bandwidth used to read the mesh using the current memory layout.

Ideal mesh bandwidth

An estimate of the GPU bandwidth that would be used with the recommended mesh layout.

3.7 Analyze function calls

The **API Calls** view shows you every call that was made for all the frames that are in your trace. You can search the data to help you identify opportunities for optimization or find errors in your API calls. See return values for applicable calls, and find out if the CPU is the limiting factor for a call.

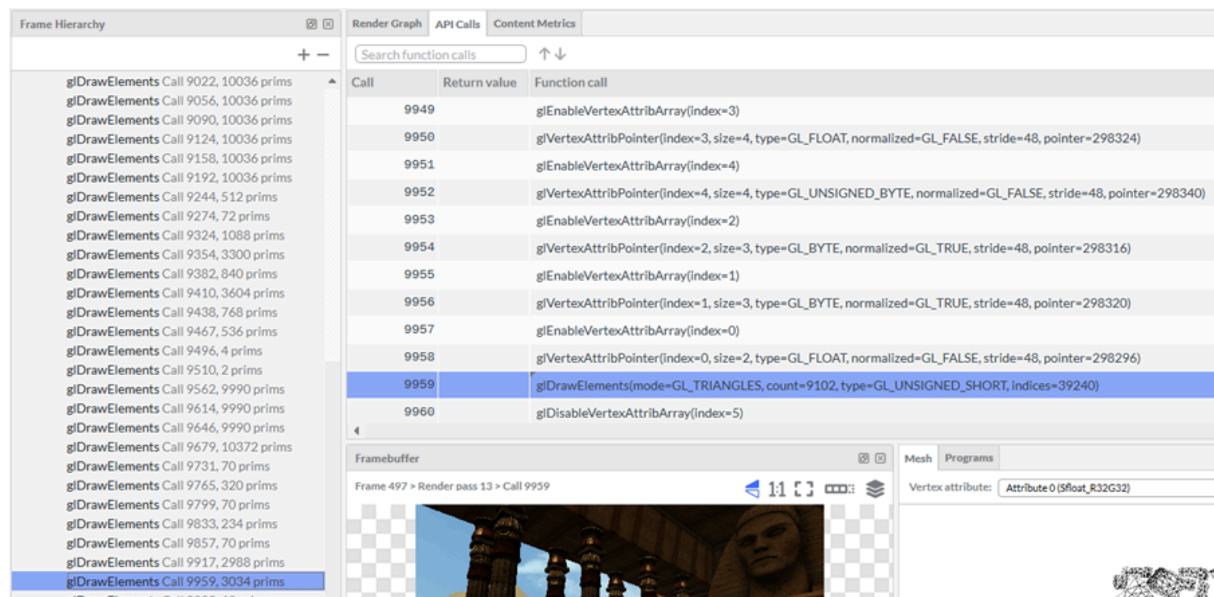
About this task

Here is an example of how you could use the **API Calls** view to identify where an issue is.

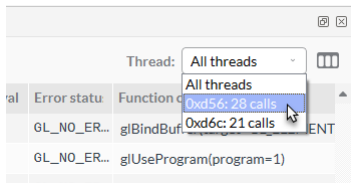
Procedure


1. To see how the draw calls render in the **Framebuffer** view, step through the draw calls in the **Frame Hierarchy** view.
You can see that the draw call selected in the **Frame Hierarchy** view is also selected in the **API Calls** view.

Figure 3-22: Draw calls selected



If your trace uses multiple threads, you can choose the thread from the drop-down menu. If your trace uses one thread only, the drop-down menu is not visible.

Figure 3-23: See function calls for a specific thread

2. To add and remove columns showing different data for your function calls, click the **Edit columns** icon .
3. In the **API Calls** view, look for errors or other values that are not correct for that draw call. Alternatively, use the search box in the **API Calls** view to find inefficient uses of the API. For example, opaque draws that have blending enabled, or if back-face culling is disabled.

Next steps

If you see an error, or an unexpected return value or function call, review the code in your application.

3.8 Analyze shader programs

Frame Advisor reports information about the shader programs used in your application. Follow these steps to discover how your shaders were used in the captured frame. Explore which shaders are in use for each frame, render pass or draw call, find the most expensive shaders and look for best practice violations.

Before you begin

- Before capturing a trace, you need to set the path to the Mali Offline Compiler directory so that Frame Advisor can generate and display metrics for shaders in the **Programs** view. To do this, select **Configure** → **Open preferences** in the Frame Advisor menu and click **Browse** to locate the directory. This directory is normally in the Arm Performance Studio install directory.

Figure 3-24: Frame Advisor Preferences

Frame Advisor Preferences

Preferences

Path to adb Browse

Logging directory Browse
 Current logging directory: /Users/username/Library/Application Support/Frame Advisor/logs

Path to Mali Offline Compiler directory Browse
 Enter path to Mali Offline Compiler directory, or leave this field empty to reset to default.

Device data timeout (seconds)

Max framebuffer cache size (GiB)

Apply and close Close

Procedure

1. [Capture a frame burst](#) for a frame that you want to analyze.
2. Select a frame or render pass, or step through the draws in the **Frame Hierarchy** view. Look at the **Programs** view to see which shaders are used in the currently selected object, and review their performance characteristics in the table of metrics. Use the filter buttons to show different shader stages.

Figure 3-25: Frame Advisor Programs View

Mesh

Programs

Filter by render pass

Stage:

Vert position

Vert main

Fragment

Program	Shader	Shortest path			Longest path			Total emitted			Occupancy	Work regs	Uniform regs	Stack spills	FP16 usage
		A	LS	T	A	LS	T	A	LS	T					
99	100	2.42	17.00	0.00	2.42	17.00	0.00	2.42	17.00	0.00	50	48	62	0	0
45	46	0.00	3.00	0.00	0.00	3.00	0.00	0.00	3.00	0.00	100	7	24	0	100
39	40	2.50	3.00	0.00	2.50	3.00	0.00	2.50	3.00	0.00	100	17	64	0	0
21	22	0.00	3.00	0.00	0.00	3.00	0.00	0.00	3.00	0.00	100	7	42	0	100
15	16	0.00	3.00	0.00	0.00	3.00	0.00	0.00	3.00	0.00	100	7	38	0	100



For Vulkan applications the metrics are generated without using the application pipeline state, so the numbers might not exactly match the runtime driver.

3. The **Programs** view table includes an approximate cycle cost breakdown for the major functional units in the shader core, the arithmetic unit, the load/store unit, the varying unit, and the texture unit. Sort the table by highest cost to find the shaders that are the best candidates for optimization. Look for the functional unit with the highest cycle cost in either or both of the shortest and longest path cycles. Consider how you could optimize the shader to reduce cost for that functional unit first.

Shortest path

An estimate of the number of cycles for the shortest control flow path through the shader program.

Longest path

An estimate of the number of cycles for the longest control flow path through the shader program.

Total emitted

The cumulative number of cycles for all instructions that are generated for the program, irrespective of program control flow.



The table shows the highest cost per thread. However, Frame Advisor does not know how many threads were executed in total, therefore this number does not represent the total shader cost per frame.

4. Sort the table by shaders with the most **Stack spills**. Shaders that spill to stack are expensive for a GPU to process, so try to reduce register pressure by:
 - a. Reducing variable precision.
 - b. Reducing the live ranges of variables.
 - c. Simplifying the shader program.
5. The **FP16 usage** column reports the percentage of arithmetic operations that are performed at 16-bit precision or lower for each shader program. A higher number here is better, because 16-bit precision is twice as fast as 32-bit precision. Sort the table by the lowest use of FP16 to find shaders to optimize by reducing precision.
Reducing precision also reduces both energy consumption and register pressure, and can double the performance. There are situations where highp is always required, such as for position and depth calculations. However, in many cases there is little noticeable difference on-screen when reducing precision to 16-bit.

For OpenGL ES applications, set precision to `mediump`. For Vulkan, use `RelaxedPrecision`. Alternatively, use explicit 16-bit type extensions to set precision.

6. Work registers are general purpose read-write registers that are allocated to each running thread. Sort the table by the highest number of work registers in use, and look for shaders using 32 or more.

The available physical register pool is divided among the shader threads that are executing. Therefore, reducing work register usage can increase the number of threads that can be executed simultaneously, helping to keep the GPU busy. Aim to reduce work register usage to below 32 to improve occupancy. Try reducing precision from highp (32-bit) to mediump (16-bit), to enable the GPU to store twice as many variables per register.

7. Uniform registers are read-only registers that are allocated to each running program. Uniform registers are used to store uniform and literal constant values. Shaders that run out of uniform storage need to fall back to per-thread memory loads for additional values. Look for shaders that use the highest numbers of uniform registers and aim to reduce them by using 16-bit data types, or by reducing the number of uniforms and constants in the shader program.
8. To view the source code for a shader program, double-click anywhere along a row to open the **Source** view.

Figure 3-26: Source view

```
#version 300 es
#ifdef GL_ES
precision highp float;
#endif
#ifdef GL_ES
#define highp
#define mediump
#define lowp
#endif
//TODO: szetbontani 2 pass-ra, 1 init feedback shader, 1 advect feedback shader

layout (location = 0) in vec3 in_Pos; // particle instance's position
layout (location = 1) in vec3 in_Age01_Speed_Accel;

layout (location = 2) in vec3 in_Amplitude;
layout (location = 3) in vec3 in_Phase;
layout (location = 4) in vec3 in_Frequency;
layout (location = 5) in vec3 in_T;
layout (location = 6) in vec3 in_B;
layout (location = 7) in vec3 in_N;
layout (location = 8) in vec4 in_Velocity;

#ifdef USE_UBOs
//
#include emitterAdvectConsts;
//
#else
#define MAX_SUBSTEPS 5 //MUST match with kcl_particlesystem2.h/KCL::_emitter::max_num_substeps
!!!!
uniform ivec4 particleBufferParamsXYZ_pad[ MAX_SUBSTEPS ]; //startBirthIdx,
endBirthIdx, noOverflow

uniform mediump_vec4 emitter_apertureXYZ focusdist;
```

Next steps

- See the Arm GPU best practices guide for information on [vertex shading](#) and [fragment shading](#)
- To analyze a single shader program and how it would perform on any Arm GPU, use [Mali Offline Compiler](#).

4. How to get help

Arm® provides lots of resources for you to find more information about Arm® Frame Advisor or other Arm® Performance Studio tools. You can also engage with other users, or ask us a question directly.

Here are the various ways for you to find more information or to get in touch:

- See the latest discussions, learn from experienced users, or ask a question in the [Graphics, Gaming, and VR community forum](#).
- Find information and resources for Arm® Performance Studio on the [Developer website](#).
- To ask a question directly, you can email the Arm® Performance Studio team at performancestudio@arm.com.
- To give feedback about Frame Advisor, fill in this [feedback form](#). You can also access this form in Frame Advisor. Select **Help > Send feedback**.

Your activity data in Frame Advisor is saved to a log file in your user directory. Arm® recommends that you include the log file when you contact the Support team for help with unexpected issues. To change the path to your Frame Advisor log file, click **Configure -> Open preferences** in the Frame Advisor menu.

5. Troubleshooting Frame Advisor

Find answers to common problems that might occur when capturing or analyzing data in Arm® Frame Advisor.

5.1 My device is not listed in Frame Advisor

When starting a new trace, my device is not listed on the connection screen.

You may not have set up your computer and device correctly

The [Setup tasks](#) explain how to set up your computer and device to use Arm® Frame Advisor.

Solution

1. Check that the device is connected to your computer via USB.
2. Check that the device is set to [Developer mode](#).
3. Check that the device has USB debugging enabled in Settings > Developer options. When enabling USB debugging, your device may ask you to authorize connection to your computer. Confirm this.
4. Ensure you have installed Android Debug Bridge (adb).
5. In a shell terminal, run the `adb devices` command. This command returns the ID of all connected devices. For example, with one device connected:

```
adb devices
List of devices attached
RZ8MC03VVEW device
```

If the device is listed as unauthorized, this means that your device has USB debugging enabled, but the computer it is connected to has not been given authority to access it.

6. Go to Settings > Developer Options, then disable and re-enable USB debugging. You can also try revoking USB debugging authorizations here. When re-enabling USB debugging, your device asks you to authorize access from your computer.

5.2 My application is not listed in Frame Advisor

When starting a new trace, my application is not listed on the connection screen.

You may not have set up your application correctly

Ensure that adb is authorized, and that your application is built properly.

Solution

1. In a shell terminal, run the `adb devices` command. This command returns the ID of all connected devices. For example, with one device connected:

```
adb devices
List of devices attached
RZ8MC03VVEW device
```

If the device is listed as unauthorized, this means that your device has USB debugging enabled, but the computer it is connected to has not been given authority to access it.

2. Check the Android manifest file to ensure your application is set to debuggable.
3. Check that the activity is marked as main when you build your application in the Android manifest.

5.3 The Framebuffer view is slow to load images

When you click through the draw calls in your captured frame, the **Framebuffer** view is slow to load images.

Increase your Framebuffer cache

Arm® Frame Advisor stores recently accessed framebuffer images. To enable Frame Advisor to store more images, increase the size of the cache so that the images load faster in the **Framebuffer** view when they are accessed again.

Solution

1. To open the **Preferences** dialog, click **Configure -> Open preferences**.
2. Increase the **Max framebuffer cache size (GiB)**, then click **Apply and close**.



Note

Do not allocate all of your available memory in the **Max framebuffer cache size (GiB)**. Using too much of your available memory in the framebuffer cache can adversely affect the performance of your system.

3. Restart Frame Advisor.
4. Click some draw calls in the **Frame Hierarchy** view. The images in the **Framebuffer** view load faster.

5.4 A timed out error message appears when I start a capture

After you click the **Capture** button, the *Timed out waiting for a response from the daemon.* message is displayed.

Your application is not sending API calls

Arm® Frame Advisor shows this message when your application is not sending API calls. If you think that your application is sending API calls, but the timeout is happening too soon, you can increase the time that Frame Advisor checks for API calls.

Solution

1. In Frame Advisor, open the **Preferences** dialog. Click **Configure -> Open preferences**.
2. For **Device data timeout (seconds)**, increase the number of seconds for Frame Advisor to perform the API calls check. Then click **Apply and close**.
3. Restart Frame Advisor.
4. Capture a new trace.

If the timed out message is displayed again:

- Check your code for any obvious problems.
- If you have checked your code, see [How to get help](#).

5.5 An unsupported image format message is displayed in the Framebuffer

As you click through draw calls in the **Framebuffer** view, the 'Unsupported image format' message is displayed instead of the visualized output.

The image format used cannot be interpreted

Arm® Frame Advisor shows this message when the version of Frame Advisor that you are using does not support the image formats used.

Solution

For more information about supported image formats, email the Arm® Performance Studio team at performancestudio@arm.com. Include your log file in the email so that the team can check which image formats you are using. To find where your log file is saved, click **Configure -> Open preferences** in the Frame Advisor menu.

5.6 A no image data message is displayed in the Framebuffer

As you click through draw calls in the **Framebuffer** view, the 'No image data' message is displayed instead of the visualized output.

The image data cannot be interpreted

Arm® Frame Advisor shows this message when it cannot interpret the image data.

Solution

Email your Frame Advisor log file to the Arm® Performance Studio team at performancestudio@arm.com, to help the team to identify the problem with your image data. To find where your log file is saved, click **Configure -> Open preferences** in the Frame Advisor menu.

Proprietary Notice

This document is protected by copyright and other related rights and the use or implementation of the information contained in this document may be protected by one or more patents or pending patent applications. No part of this document may be reproduced in any form by any means without the express prior written permission of Arm Limited ("Arm"). No license, express or implied, by estoppel or otherwise to any intellectual property rights is granted by this document unless specifically stated.

Your access to the information in this document is conditional upon your acceptance that you will not use or permit others to use the information for the purposes of determining whether the subject matter of this document infringes any third party patents.

The content of this document is informational only. Any solutions presented herein are subject to changing conditions, information, scope, and data. This document was produced using reasonable efforts based on information available as of the date of issue of this document. The scope of information in this document may exceed that which Arm is required to provide, and such additional information is merely intended to further assist the recipient and does not represent Arm's view of the scope of its obligations. You acknowledge and agree that you possess the necessary expertise in system security and functional safety and that you shall be solely responsible for compliance with all legal, regulatory, safety and security related requirements concerning your products, notwithstanding any information or support that may be provided by Arm herein. In addition, you are responsible for any applications which are used in conjunction with any Arm technology described in this document, and to minimize risks, adequate design and operating safeguards should be provided for by you.

This document may include technical inaccuracies or typographical errors. THIS DOCUMENT IS PROVIDED "AS IS". ARM PROVIDES NO REPRESENTATIONS AND NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY, SATISFACTORY QUALITY, NON-INFRINGEMENT OR FITNESS FOR A PARTICULAR PURPOSE WITH RESPECT TO THE DOCUMENT. For the avoidance of doubt, Arm makes no representation with respect to, and has undertaken no analysis to identify or understand the scope and content of, any patents, copyrights, trade secrets, trademarks, or other rights.

TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL ARM BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF ARM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Reference by Arm to any third party's products or services within this document is not an express or implied approval or endorsement of the use thereof.

This document consists solely of commercial items. You shall be responsible for ensuring that any permitted use, duplication, or disclosure of this document complies fully with any relevant

export laws and regulations to assure that this document or any portion thereof is not exported, directly or indirectly, in violation of such export laws. Use of the word “partner” in reference to Arm’s customers is not intended to create or refer to any partnership relationship with any other company. Arm may make changes to this document at any time and without notice.

This document may be translated into other languages for convenience, and you agree that if there is any conflict between the English version of this document and any translation, the terms of the English version of this document shall prevail.

The validity, construction and performance of this notice shall be governed by English Law.

The Arm corporate logo and words marked with ® or ™ are registered trademarks or trademarks of Arm Limited (or its affiliates) in the US and/or elsewhere. Please follow Arm’s trademark usage guidelines at <https://www.arm.com/company/policies/trademarks>. All rights reserved. Other brands and names mentioned in this document may be the trademarks of their respective owners.

Arm Limited. Company 02557590 registered in England.

110 Fulbourn Road, Cambridge, England CB1 9NJ.

PRE-1121-V1.0

Product and document information

Read the information in these sections to understand the release status of the product and documentation, and the conventions used in Arm documents.

Product status

All products and services provided by Arm require deliverables to be prepared and made available at different levels of completeness. The information in this document indicates the appropriate level of completeness for the associated deliverables.

Product completeness status

The information in this document is Final, that is for a developed product.

Revision history

These sections can help you understand how the document has changed over time.

Document release information

The Document history table gives the issue number and the released date for each released issue of this document.

Document history

Issue	Date	Confidentiality	Change
0104-00	5 September 2024	Non-Confidential	New document for v1.4
0103-00	7 June 2024	Non-Confidential	New document for v1.3
0102-00	18 April 2024	Non-Confidential	New document for v1.2
0101-00	15 February 2024	Non-Confidential	New document for v1.1
0100-00	21 November 2023	Non-Confidential	New document for v1.0

Change history

For information about the technical changes to the Frame Advisor User Guide, see the [Arm® Performance Studio Release Notes](#).

Conventions

The following subsections describe conventions used in Arm documents.

Glossary

The Arm Glossary is a list of terms used in Arm documentation, together with definitions for those terms. The Arm Glossary does not contain terms that are industry standard unless the Arm meaning differs from the generally accepted meaning.

See the Arm Glossary for more information: developer.arm.com/glossary.

Typographic conventions

Arm documentation uses typographical conventions to convey specific meaning.

Convention	Use
<i>italic</i>	Citations.
bold	Interface elements, such as menu names. Terms in descriptive lists, where appropriate.
monospace	Text that you can enter at the keyboard, such as commands, file and program names, and source code.
monospace <u>underline</u>	A permitted abbreviation for a command or option. You can enter the underlined text instead of the full command or option name.
<and>	Encloses replaceable terms for assembler syntax where they appear in code or code fragments. For example: <div>MRC p15, 0, <Rd>, <CRn>, <CRm>, <Opcode_2></div>
SMALL CAPITALS	Terms that have specific technical meanings as defined in the <i>Arm® Glossary</i> . For example, IMPLEMENTATION DEFINED , IMPLEMENTATION SPECIFIC , UNKNOWN , and UNPREDICTABLE .



We recommend the following. If you do not follow these recommendations your system might not work.



Your system requires the following. If you do not follow these requirements your system will not work.



You are at risk of causing permanent damage to your system or your equipment, or harming yourself.



This information is important and needs your attention.



A useful tip that might make it easier, better or faster to perform a task.



A reminder of something important that relates to the information you are reading.

Useful resources

This document contains information that is specific to this product. See the following resources for other useful information.

Access to Arm documents depends on their confidentiality:

- Non-Confidential documents are available at developer.arm.com/documentation. Each document link in the following tables goes to the online version of the document.
- Confidential documents are available to licensees only through the product package.

Arm product resources	Document ID	Confidentiality
Android performance triage with Streamline Tutorial	102540	Non-Confidential
Arm® GPU Best Practices Developer Guide	101897	Non-Confidential
Arm® Mali GPU Training	Arm® Mali GPU Training	Non-Confidential
Arm® Performance Studio Release Notes	107649	Non-Confidential
Understanding Render Passes	102479	Non-Confidential

Non-Arm resources	Document ID	Organization
Android Debug Bridge (adb)	Android Debug Bridge adb	Android Developers
Android Studio	Android Studio	Android Developers
Configure on-device developer options	Configure on-device developer options	Android Developers