

High quality mobile VR with Unreal Engine and Oculus

ARM

Daniele Di Donato

Senior Software Engineer, ARM

Rémi Palandri

Graphics Engineer, Oculus

Ryan Vance

Senior Rendering Programmer, Epic Games

GDC'17

3/1/20 17

Agenda

- VR Best practises on GearVR / UE4
- New VR features and rendering techniques in UE4
 - Monoscopic Far Field Rendering
 - Mobile Multiview
- New technologies in the horizon using Multiview
 - Foveated Rendering
- Debugging and Profiling on Mali
 - Mali Offline Shader Compiler
 - Mali Graphics Debugger
 - Streamline

Best practices

- Compared to PC and Console, Mobile has additional constraints
 - Most accessible development environment
 - Most challenging platform to ship on
 - Battery life and heat dissipation are primary concerns
 - Fast peak performance, but you can't run it pegged indefinitely
 - Optimization is more involved than PC and Console
 - Consistently making frame rate isn't enough
 - Android N sustained performance mode
 - Guaranteed to run indefinitely at a lower performance level

Best practices

- Asset budgets and suggestions
 - 50 - 60 k average number of triangles for the entire scene
 - Max 100 k
 - 50 draw calls per eye
 - Merge materials and meshes
 - Use instancing
 - Multiview improves this!
 - Aggressive LODs
 - Consider the memory impact

Best practices

- Use the stat system to profile scene complexity



The screenshot displays a game engine's statistics window, overlaid on a 3D scene. The window is divided into several sections: 'RHI [STATGROUP_rhi] Memory Counters', 'Counters', 'Init Views [STATGROUP_InitViews] Cycle counters (flat)', and 'Counters'. Each section contains a table of performance metrics. The background shows a dark, industrial-style environment with metallic structures and a bright light source.

RHI [STATGROUP_rhi]				
Memory Counters				
	UsedMax	Mem%	MemPool	Pool Capacity
Render target memory 2D	520.84 MB		GPU	
Texture memory 2D	275.85 MB		GPU	
Texture memory Cube	92.06 MB		GPU	
Vertex buffer memory	30.43 MB		GPU	
Render target memory 3D	12.25 MB		GPU	
Index buffer memory	7.94 MB		GPU	
Uniform buffer memory	0.81 MB		GPU	
Texture memory 3D	0.02 MB		GPU	

Counters				
	Average	Max		
Triangles drawn	291,327.00	291,327.00		
DrawPrimitive calls	558.00	558.00		
Lines drawn				

Init Views [STATGROUP_InitViews]					
Cycle counters (flat)					
	CallCount	InclusiveAvg	InclusiveMax	ExclusiveAvg	ExclusiveMax
View Visibility	1	0.14 ms	0.16 ms	0.01 ms	0.01 ms
View Relevance	1	0.10 ms	0.12 ms	0.00 ms	0.00 ms
Compute View Relevance	2	0.05 ms	0.06 ms	0.05 ms	0.06 ms
Static Mesh Relevance	2	0.04 ms	0.06 ms	0.04 ms	0.06 ms
Frustum Cull	1	0.01 ms	0.02 ms	0.00 ms	0.00 ms
Update Indirect Lighting Cache	1	0.01 ms	0.01 ms	0.00 ms	0.00 ms
Update Indirect Lighting Cache Prims	1	0.01 ms	0.01 ms	0.01 ms	0.01 ms
Occlusion Cull	1	0.01 ms	0.01 ms	0.00 ms	0.00 ms
Init dynamic shadows	1	0.01 ms	0.01 ms	0.00 ms	0.00 ms
GetDynamicMeshElements	1	0.01 ms	0.01 ms	0.01 ms	0.01 ms
Decompress Occlusion	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
Init Projected Shadow	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
Update Preshadow Cache	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
Update Indirect Lighting Cache Blocks	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
Update Fading	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
Update Indirect Lighting Cache Transitions	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms
GatherShadowPrimitives	1	0.00 ms	0.00 ms	0.00 ms	0.00 ms

Counters		
	Average	Max
Processed primitives	561.00	561.00
Visible static mesh elements	333.00	333.00
Frustum Culled primitives	258.00	258.00
Occluded primitives	96.00	96.00
Statically occluded primitives	96.00	96.00
Indirect Lighting Cache updates	4.00	4.00
Visible dynamic primitives	3.00	3.00

Output of stat rhi and initviews

Best practices

- Example output of our new automatic LOD generation (4.14)
 - Vertex data is maintained so materials and light maps can be shared



Best practices

- Asset budgets and suggestions
 - Materials should be no more than 125 instructions
 - No dynamic lights or shadows
 - Bake and fake
 - LDR
 - No post processing
- Create test levels with representative content
 - Profile on devices you intend to ship on to verify your budgets
 - Test for the duration of expected session time

Best practices

- Content suggestions
 - Remove triangles the user can't see
 - Remove back sides
 - Segment large models
 - Bake distant environment to a skybox
 - Use Oculus cube map layer for optimal sampling
 - Monoscopic can be used for middle ground
 - Fully rough materials
 - Fake environment reflection

Best practices

- Content suggestions
 - Don't render occluded objects
 - Wasted draw calls and primitive culling time
 - Design scenes to minimize draw distance
 - Use precomputed visibility volumes
 - Aggressive manual hiding of objects not in view
 - Take advantage of scene knowledge
 - Minimize transparent overdraw
 - Objects that are 100 % transparent are still drawn. Set visibility flag!

Best practices

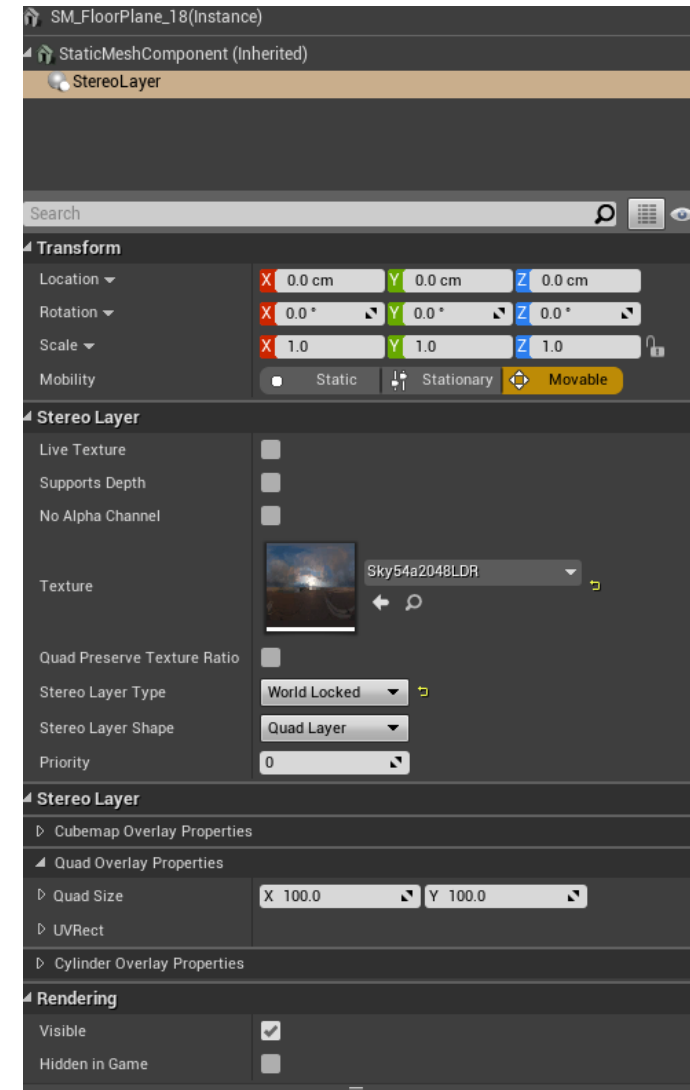
- Content suggestions
 - Use MSAA
 - At least 2x, and 4x when possible
 - Avoid post process anti-aliasing
 - Use ASTC for texture compression
 - Largest block size possible
 - Generate MIP maps
 - Avoid complex filtering options

Best practices

- Content suggestions
 - Track ticking object count
 - Don't tick if you don't need to
 - Spawning is extremely expensive
 - Spawn on load
 - Amortize over multiple frames
 - Consider building a manager to pool objects
 - Try blueprint nativization to reduce script VM overhead

Best practices

- Stereo Layers
 - Not rendered in-engine
 - Raytraced in the compositor
 - Single sampling!
 - Supports quads, cylinders, and cubemaps
 - Head-locked, tracker-locked, or world-locked
 - Stereo Layer Component
 - Works with UMG!

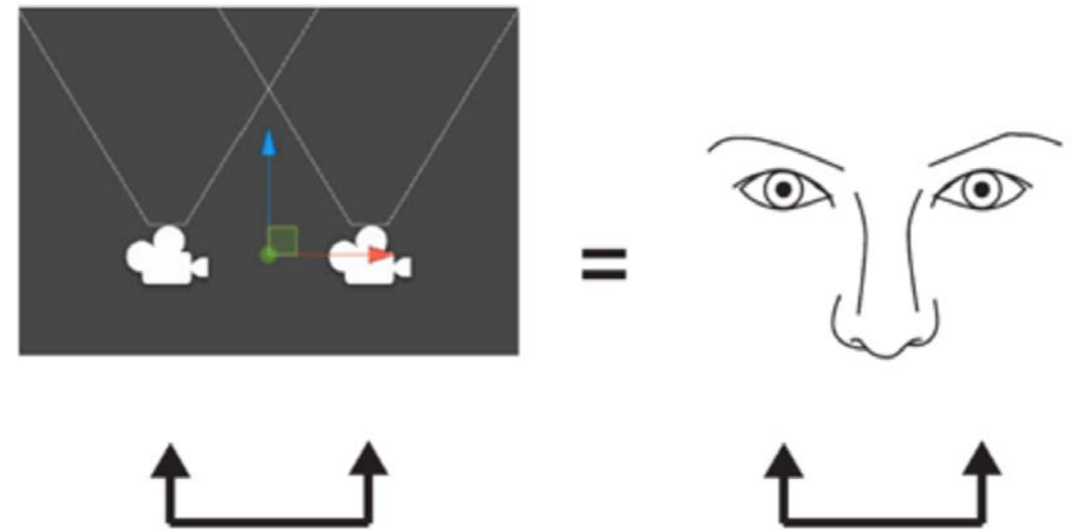


New VR features and rendering techniques in UE4

- Monoscopic Far Field Rendering
- Mobile Multiview

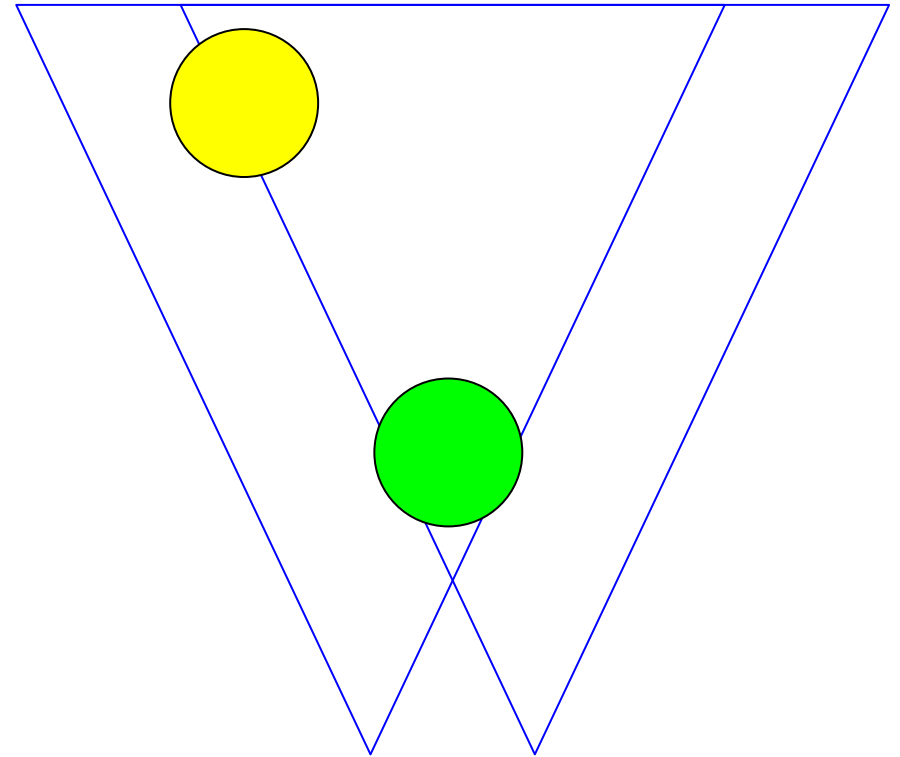
Monoscopic Rendering

- Rendering both eyes
 - Position difference creates binocular parallax
 - Projection difference creates binocular disparity
 - Depth!
- Performance issues
 - Double the CPU usage
 - Double the Vertex/Fragment usage
- Similarities?



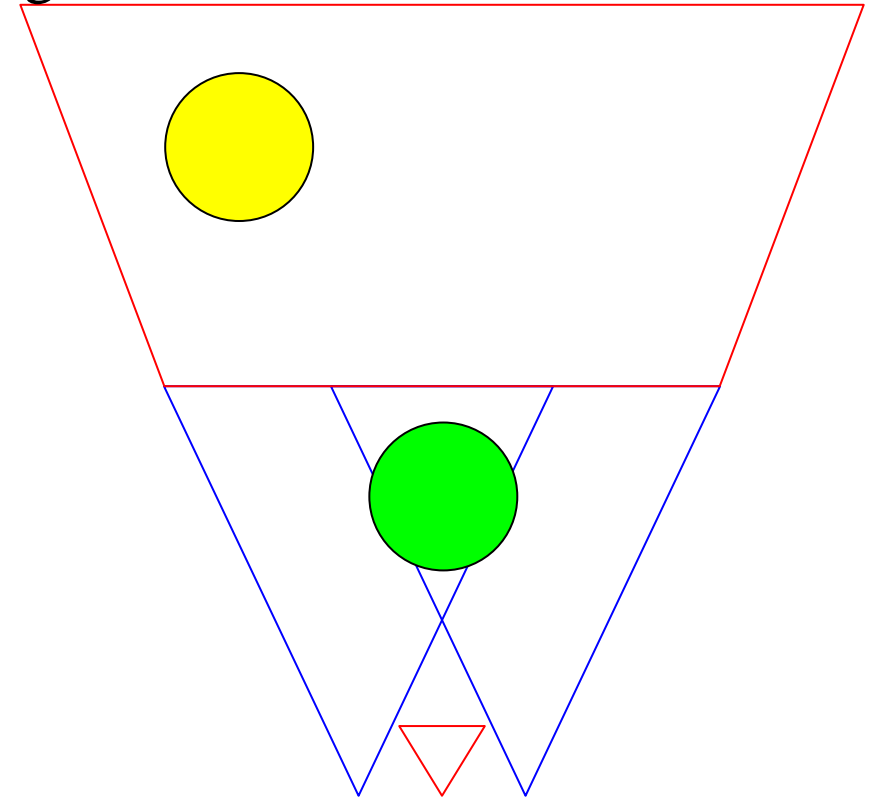
Monoscopic Rendering

- Position difference less significant as distances grow
- Adding a 3rd camera!
 - Two stereo cameras have a 30 ft far plane
 - Mono camera has a 30 ft near plane
 - Strict ordering of pixels
- New rendering pipeline

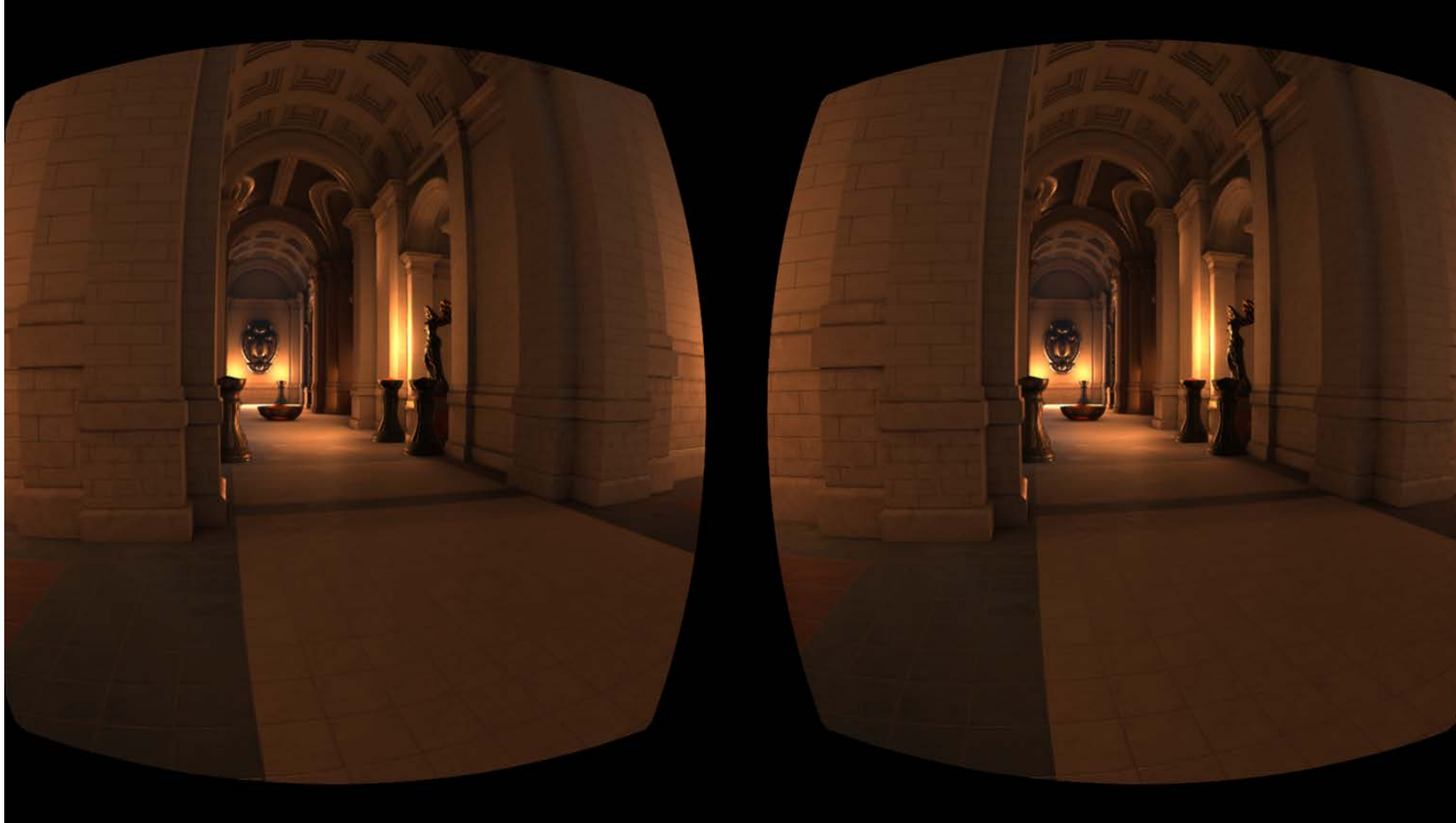


Monoscopic Rendering

- Position difference less significant as distances grow
- Adding a 3rd camera!
 - Two stereo cameras have a 30 ft far plane
 - Mono camera has a 30 ft near plane
 - Strict ordering of pixels
- New rendering pipeline

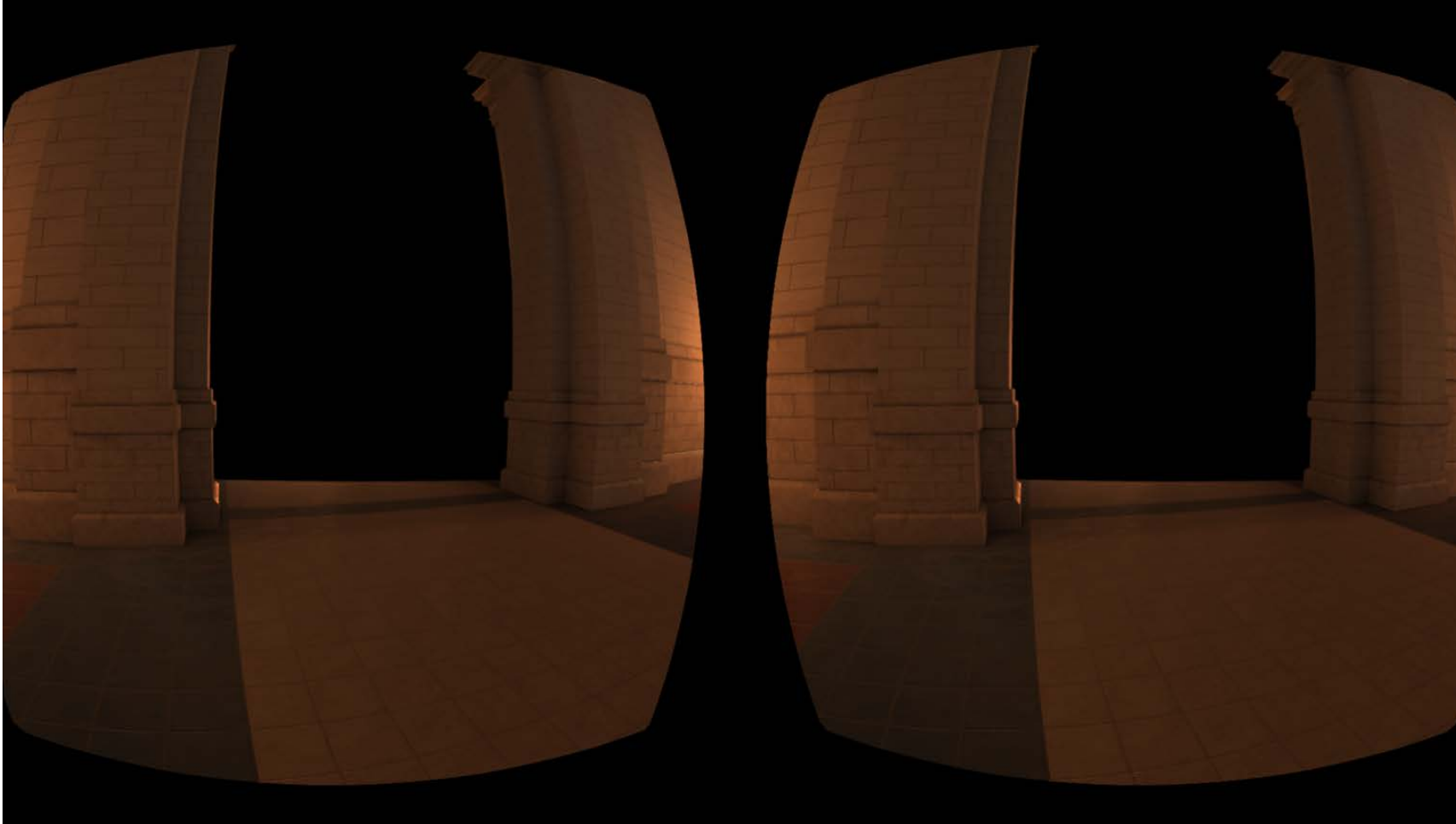


Monoscopic Rendering



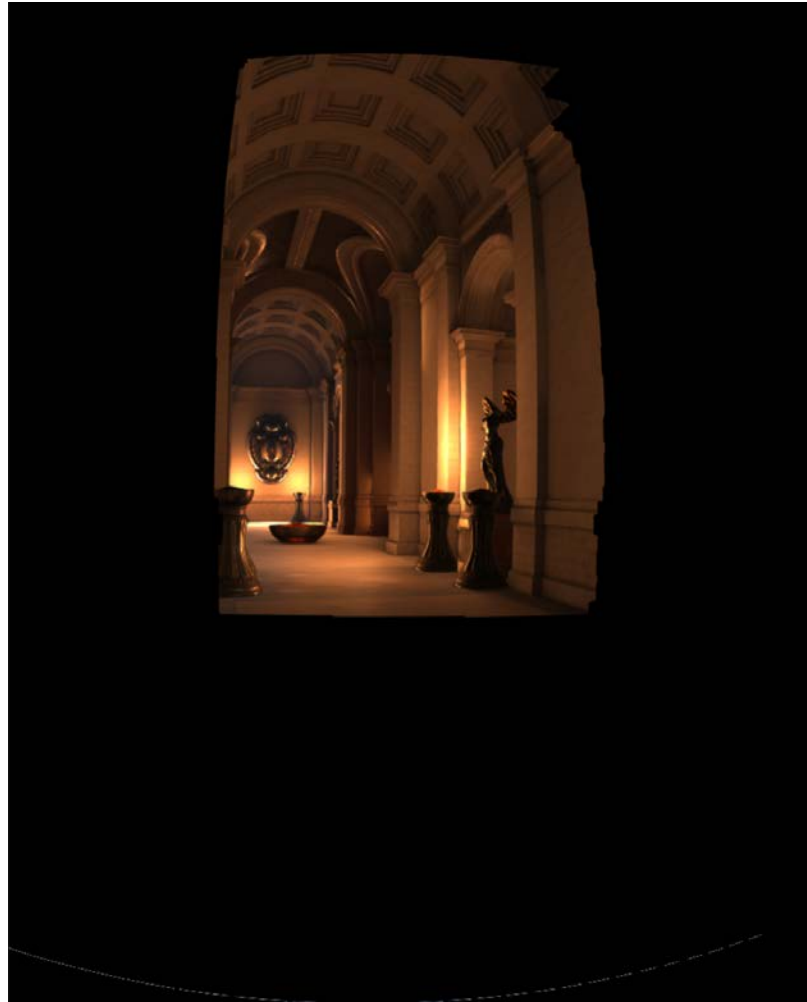
Original rendering

Monoscopic Rendering



Stereo cameras with 30ft far plane

Monoscopic Rendering

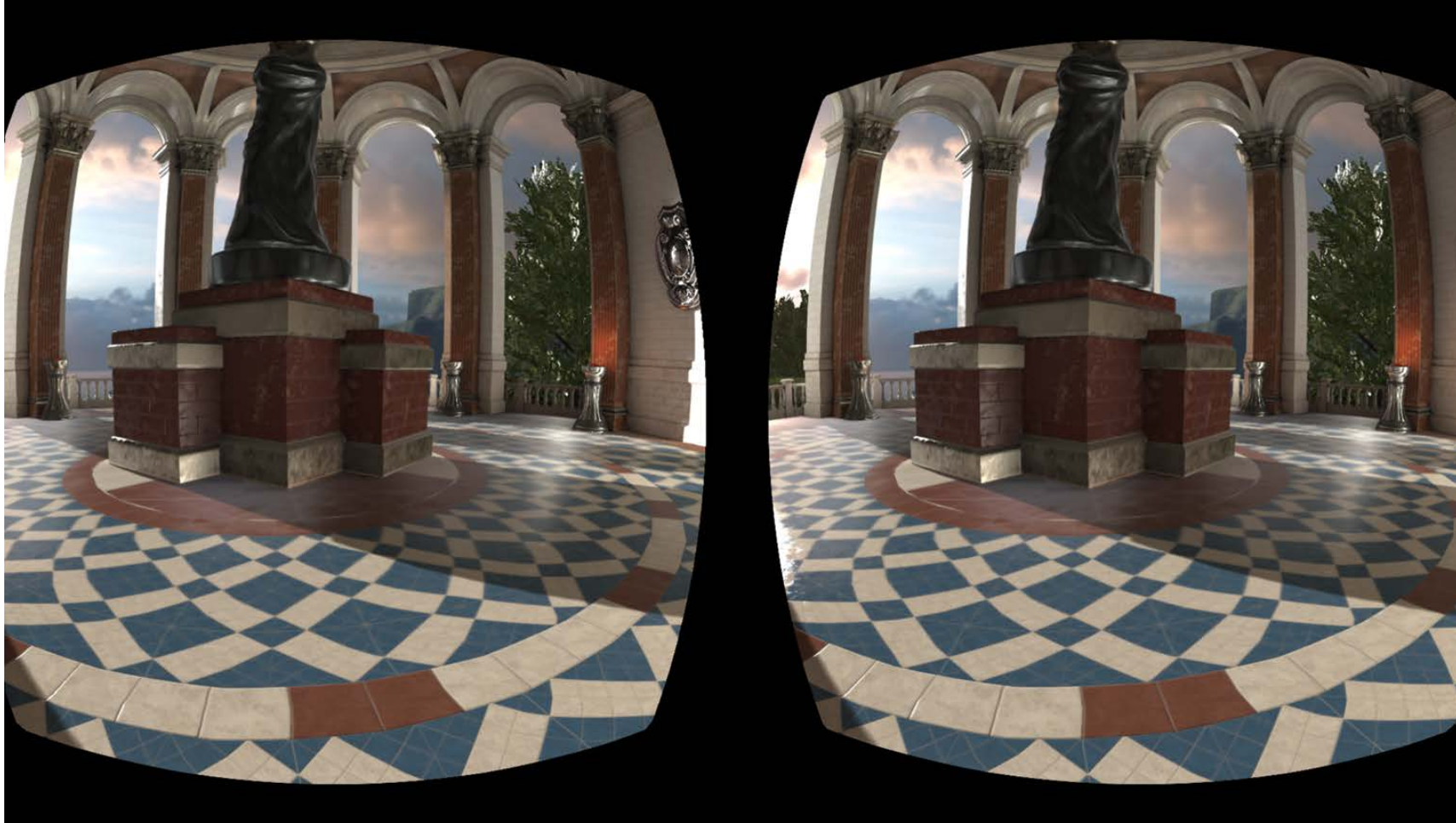


Mono camera with 30ft near plane

Monoscopic Rendering

- Issues with new pipeline
 - Monoscopic camera rendering unused pixels

Monoscopic Rendering



Standard SunTemple scene

Monoscopic Rendering



Close stereo cameras

Monoscopic Rendering



Mono camera with 30ft near plane

Monoscopic Rendering

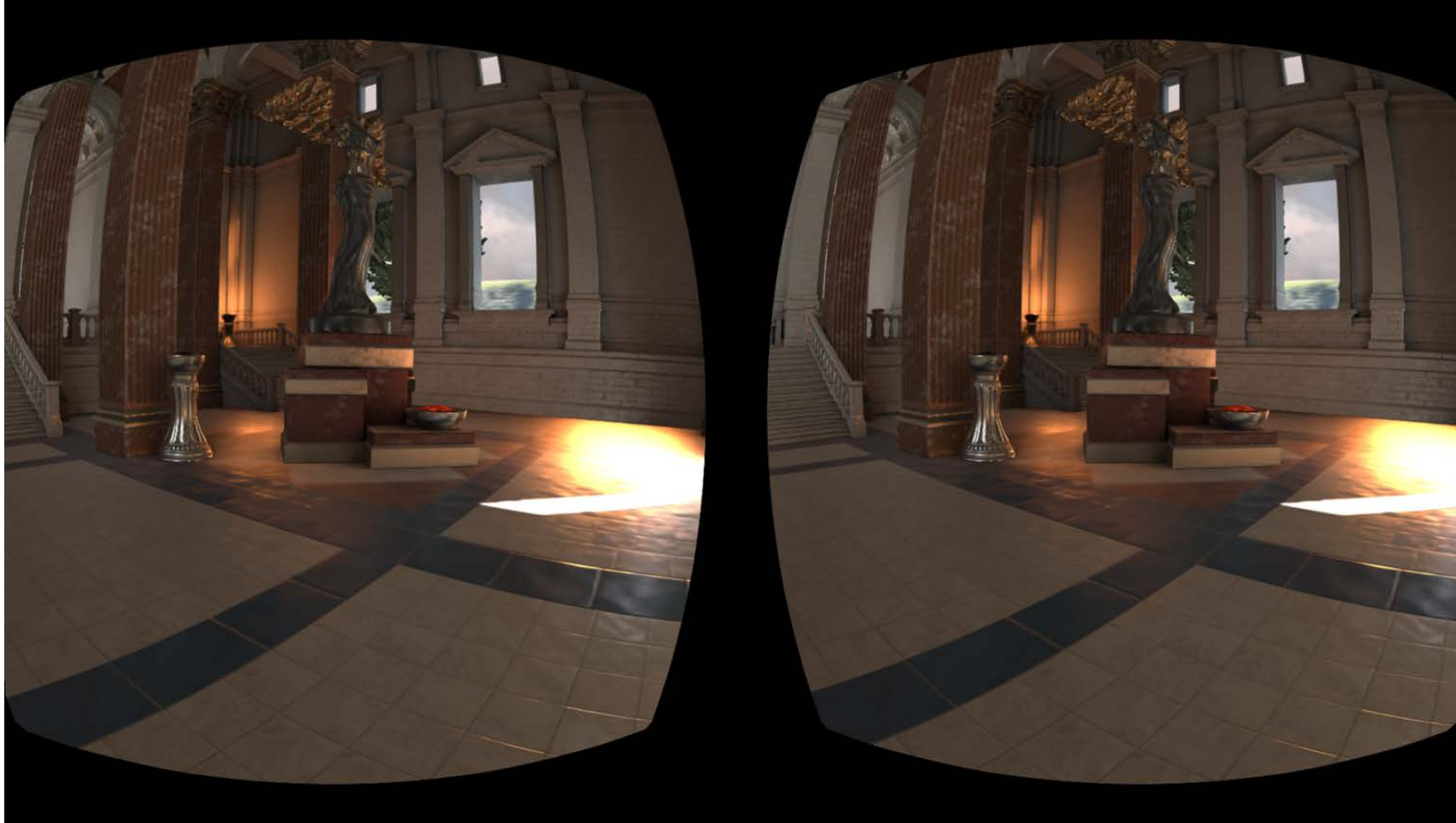


Mono camera with 30ft near plane and mask

Monoscopic Rendering

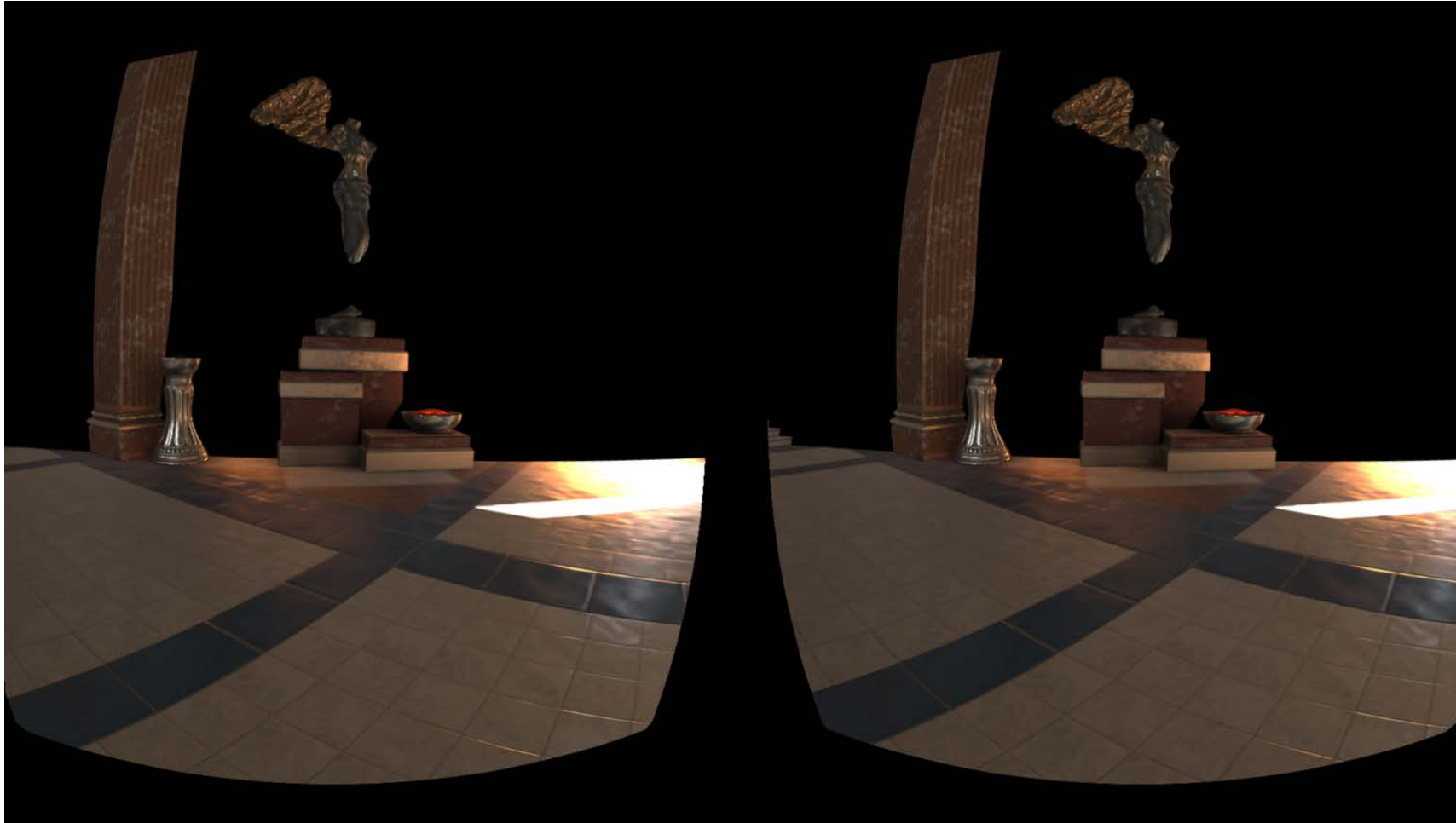
- Issues with new pipeline
 - Monoscopic camera rendering unused pixels
 - Stereo cameras drawing far object (frustum culling)

Monoscopic Rendering



Standard SunTemple scene

Monoscopic Rendering



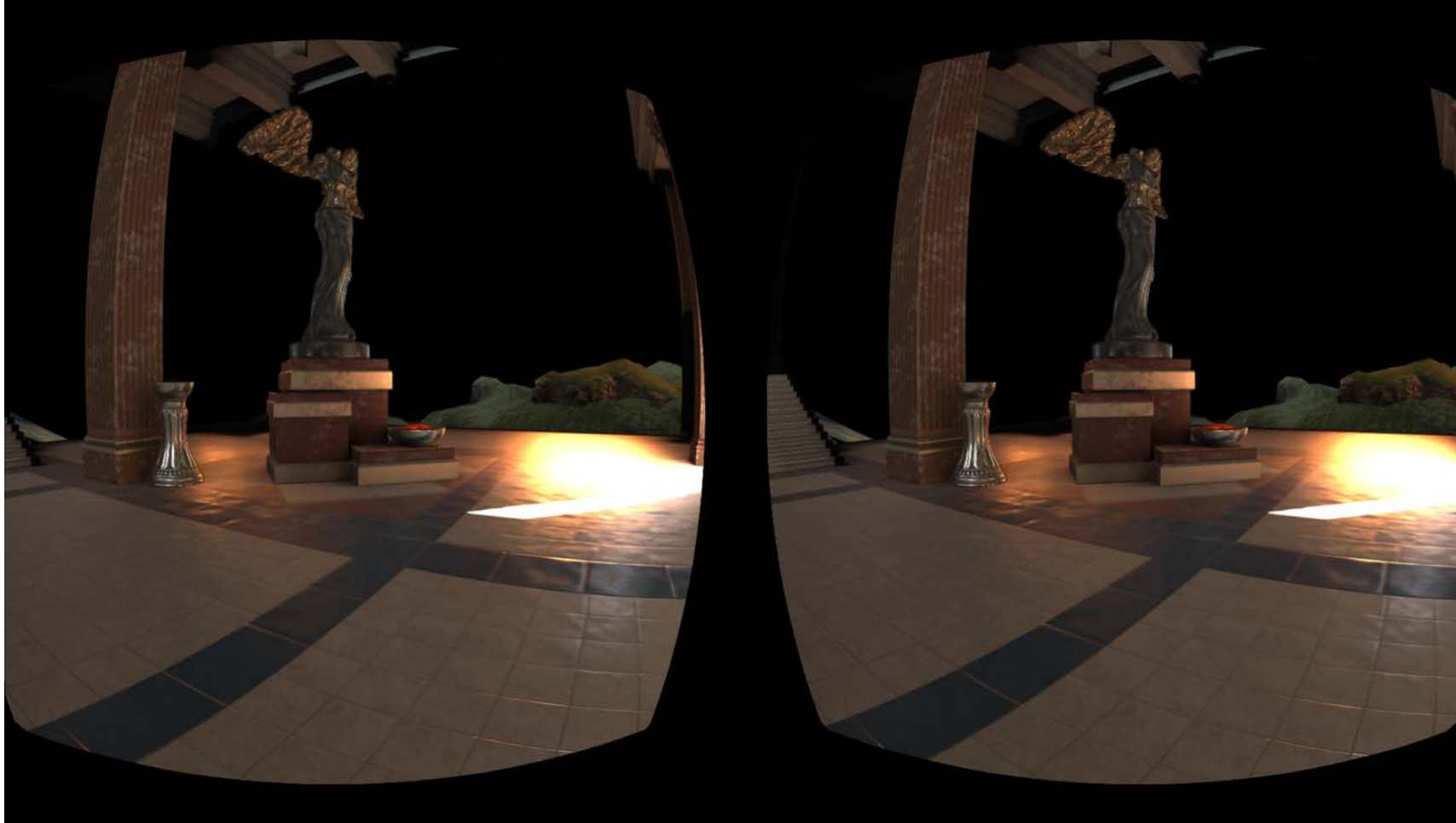
Close stereo cameras

Monoscopic Rendering



Stereo cameras without close depth clear
but with close frustum culling

Monoscopic Rendering



Stereo cameras without close depth clear
but with close frustum culling

Monoscopic Rendering

- Issues with new pipeline
 - Monoscopic camera rendering unused pixels
 - Stereo cameras drawing far object (frustum culling)
 - Compositing artifacts (transparency mainly)
 - Performance hits of running a third camera

Monoscopic Rendering

- Results
 - Performance is very environment-dependent
 - 20 +% increases in certain conditions
 - Both CPU and GPU implications
 - Can get performance decrease as well
 - Dynamic system, both on/off and view distance
 - *vr.FarFieldRenderingMode 0/1/2/3/4*

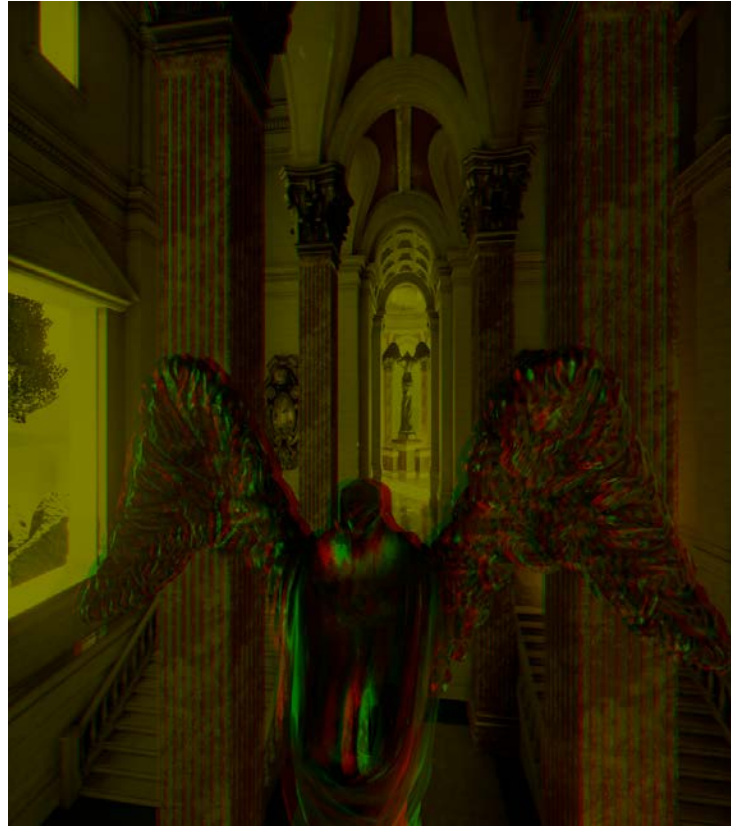
Mobile Multiview

- What problem are we solving?

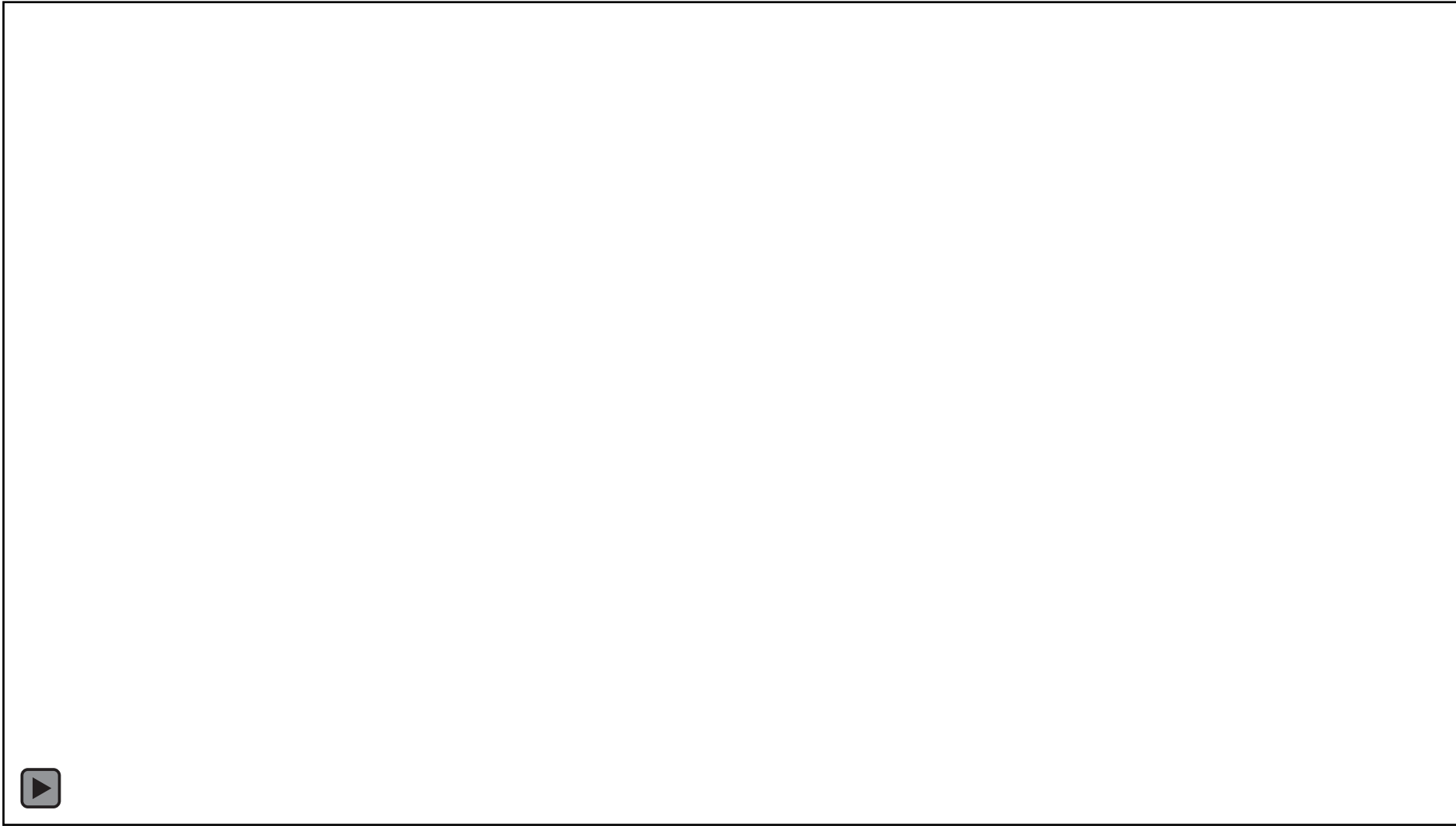


Mobile Multiview

- Minimal differences between views at primitive granularity

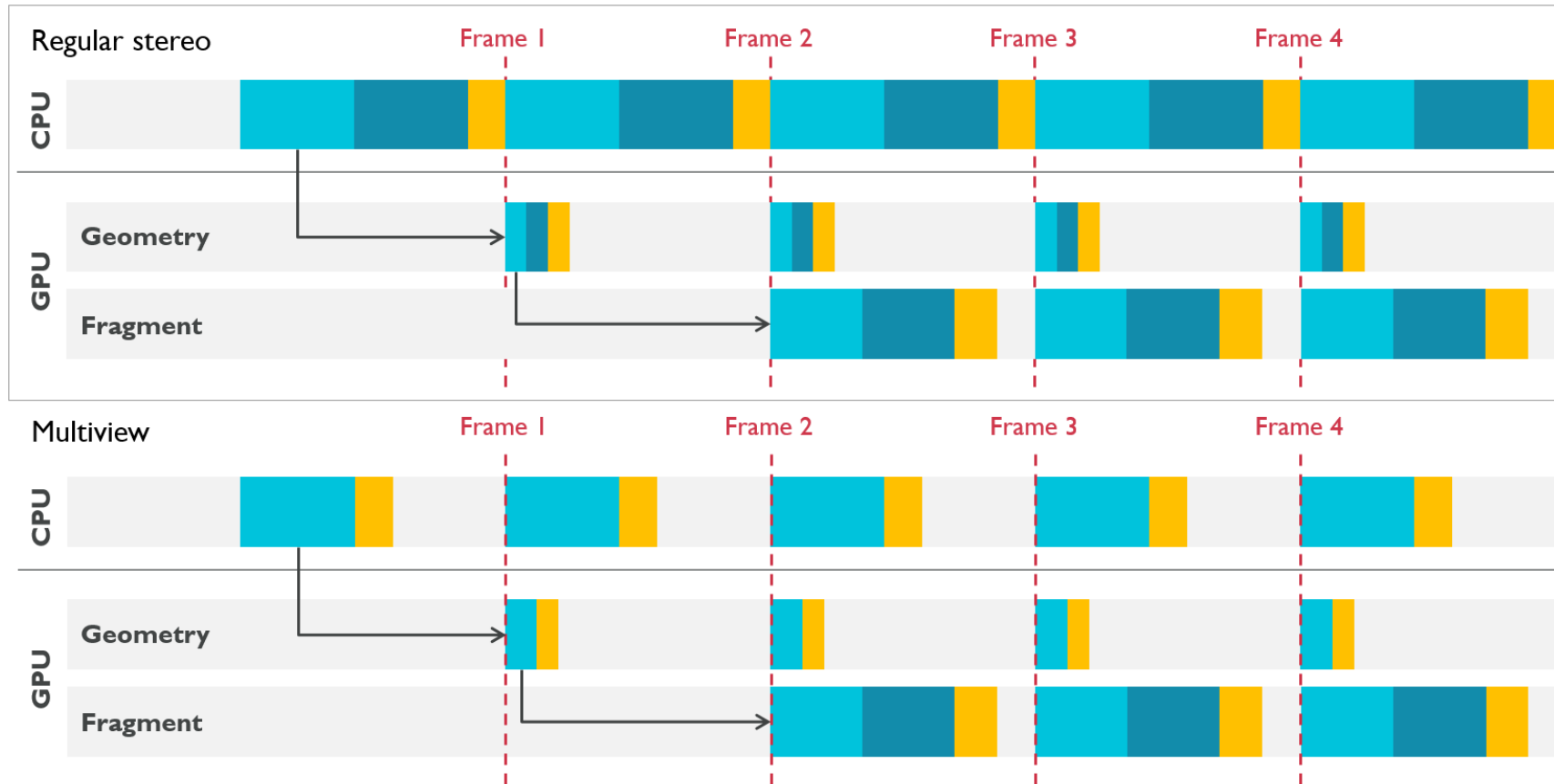


Mobile Multiview



Mobile Multiview

- Regular vs. multiview CPU-GPU timeline



Mobile Multiview

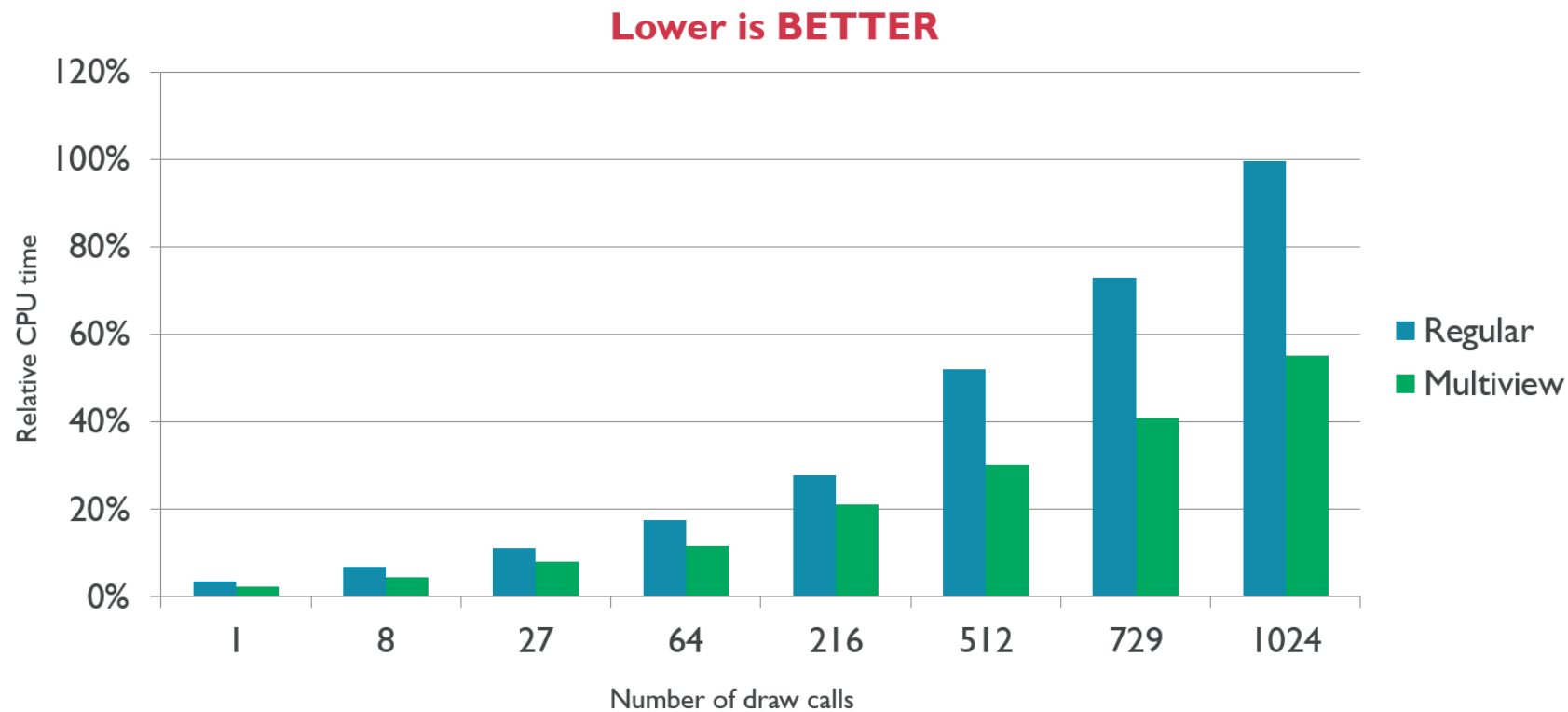
- Overloaded term
 - Instanced stereo for PC
 - Multiview extension of instanced stereo for PS4
 - Single pass stereo from Nvidia
 - DirectX 11 Multiview extension from AMD
 - Multiview OpenGL ES extension

Mobile Multiview

- UE4 implementation
 - PC/PS4 Instanced stereo and PS4 multiview
 - Standard graphics pipeline based
 - Instanced draw call, transform, culling, clipping vertex shader
 - Small extension for PS4 to reduce vertex shader work
 - Mobile multiview
 - Draw call instancing and vertex work done entirely by the driver
 - Leverages view uniform system from instanced stereo

Mobile Multiview

- Multiview: CPU performance



Mobile Multiview

- GL_OVR_multiview

GL_OVR_multiview

Restrict the use of gl_ViewID_OVR to the computation of gl_Position

GL_OVR_multiview2

No restricted usage of gl_ViewID_OVR, it can be used in fragment and vertex shader stage

OVR_multiview_multisampled_render_to_texture

Multiview version of
EXT_multisampled_render_to_texture

Mobile Multiview

- Vertex shader with multiview

```
#version 300 es
```

```
#extension GL_OVR_multiview2 : enable
```

```
precision highp float;
```

```
layout(num_views = 2) in;
```

```
in vec3 vertexPosition;
```

```
in vec2 UVCoordinates;
```

```
out vec2 texCoord;
```

```
uniform mat4 MVP[2];
```

```
void main(){
```

```
    gl_Position = MVP[gl_ViewID_OVR] * vec4(vertexPosition, 1.0f);
```

← This line is executed N times (e.g. 2x for 2 views)

```
    texCoord = UVCoordinates;
```

```
}
```

Mobile Multiview

- Using multiview in an application

```
// Create FBO
glGenFramebuffers(1, &FBO_ID);
glBindFramebuffer(GL_DRAW_FRAMEBUFFER, FBO_ID);

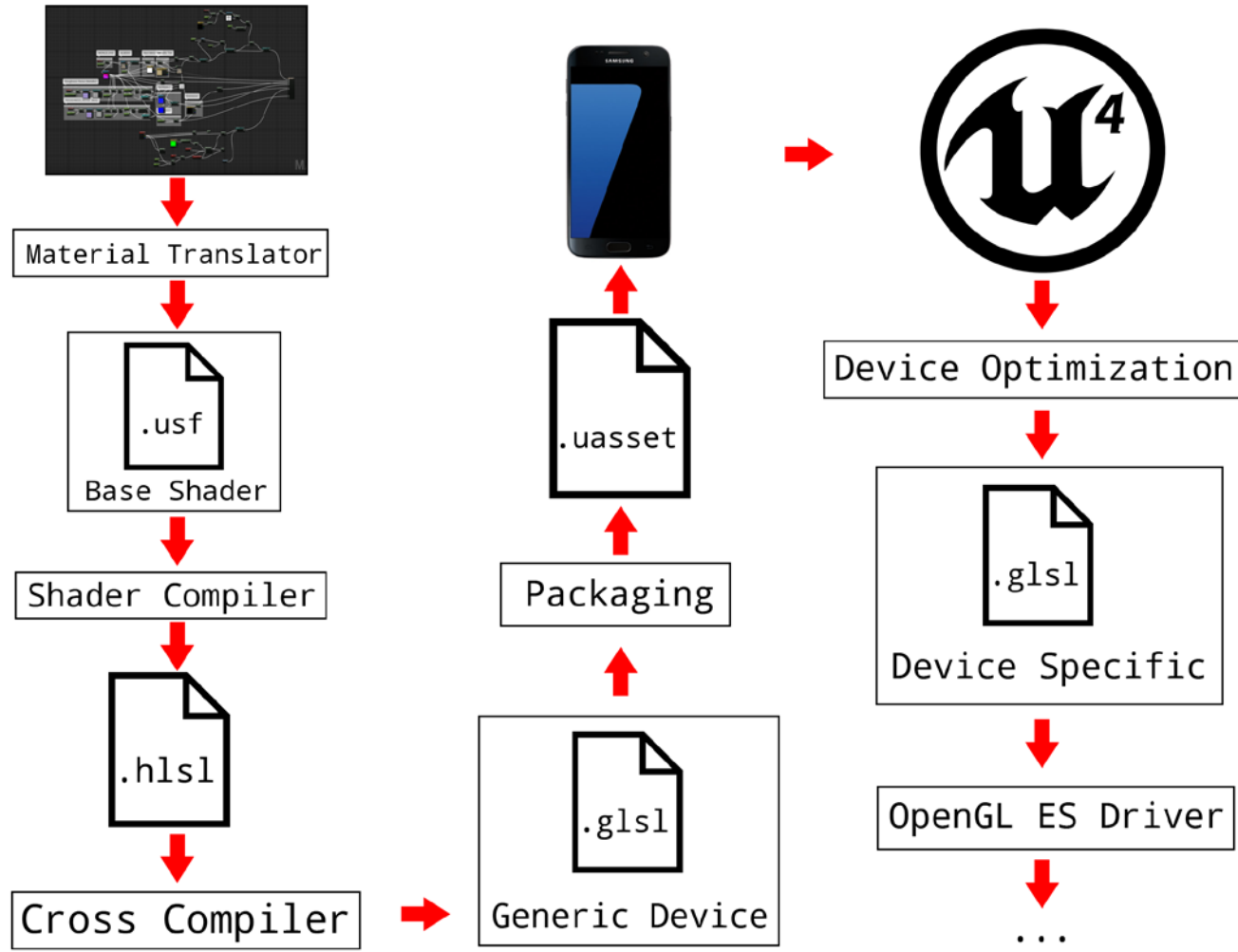
// Create Color textures
glGenTextures(1, &TextureColorID);
glBindTexture(GL_TEXTURE_2D_ARRAY, TextureColorID);
glTexStorage3D(GL_TEXTURE_2D_ARRAY, 1, GL_RGBA8, 1024, 1024, 4);

// Attach the color texture to the FBO, 4xMSAA
glFramebufferTextureMultisampledMultiviewOVR(GL_DRAW_FRAMEBUFFER, GL_COLOR_ATTACHMENT0, TextureColorID, 0, 4, 0, 2);

// Create Depth Textures
glGenTextures(1, &TextureDepthID);
glBindTexture(GL_TEXTURE_2D_ARRAY, TextureDepthID);
glTexStorage3D(GL_TEXTURE_2D_ARRAY, 1, GL_DEPTH_COMPONENT24, 1024, 1024, 4);

// Attach the depth texture to the FBO, 4xMSAA
glFramebufferTextureMultisampledMultiviewOVR(GL_DRAW_FRAMEBUFFER, GL_DEPTH_STENCIL_ATTACHMENT, TextureDepthID, 0, 4, 0, 2);
```

Mobile Multiview

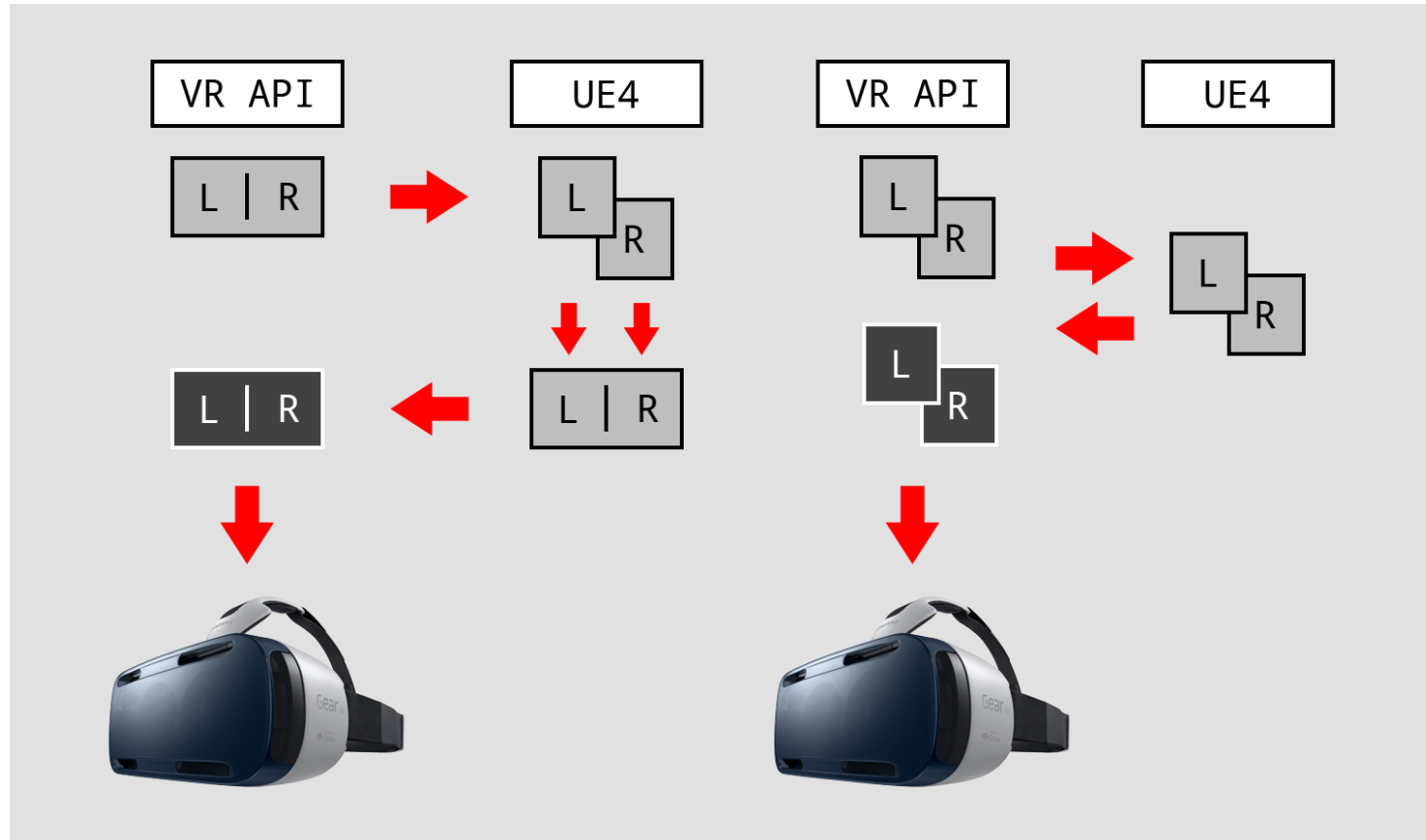


Mobile Multiview

- Driver support landscape
 - Multiple GPU vendors
 - Many driver bugs in initial implementations across all vendors
 - Long delay between driver updates and availability on end user devices
 - We strip out multiview code from the shader during application initialization if the device is known to have issues to ensure driver bugs don't break your application
 - Samsung Galaxy S6, Samsung Galaxy S7 Mali (Android M and N), S7 Adreno (Android N)

Mobile Multiview

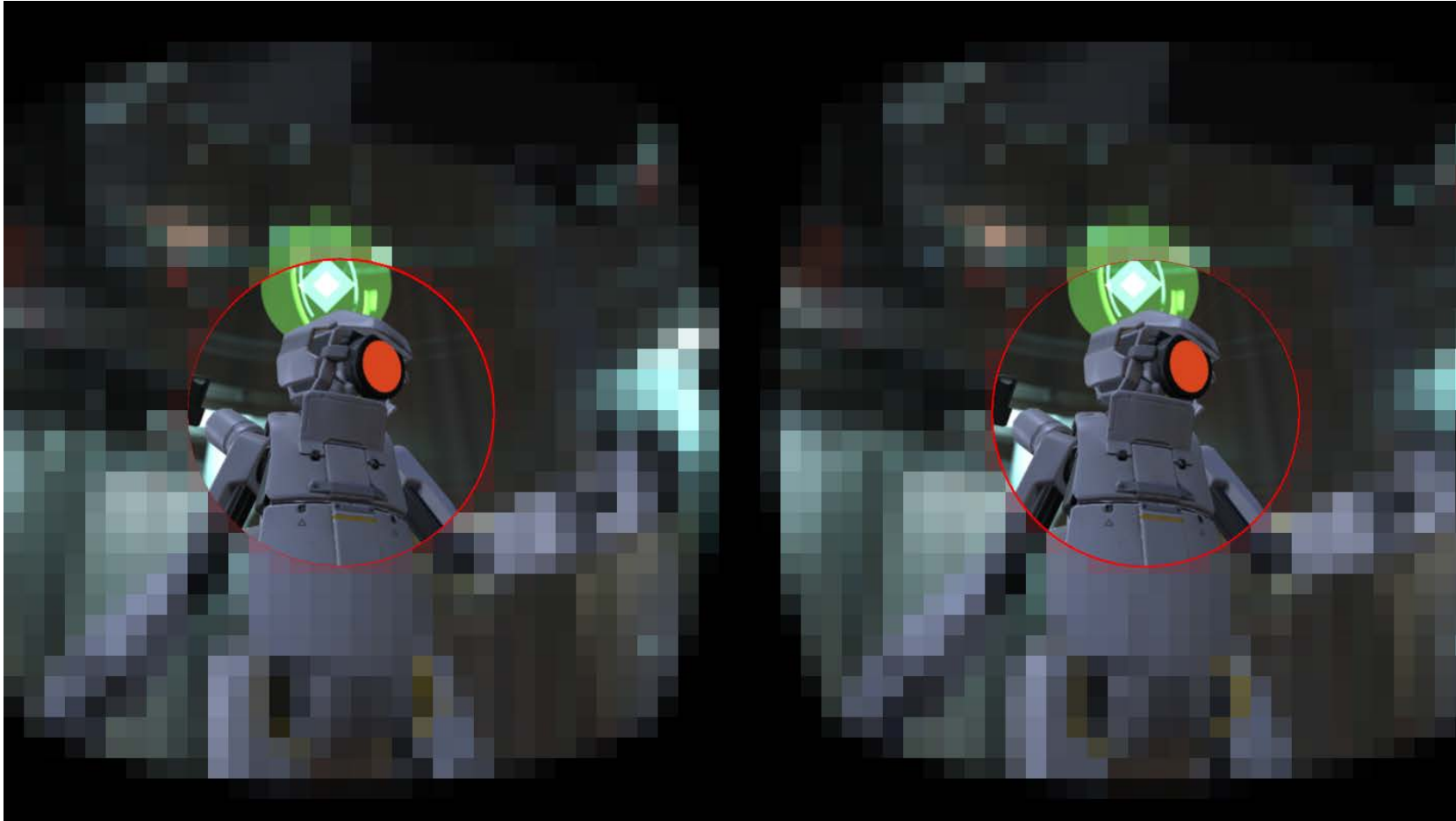
- Current work in development



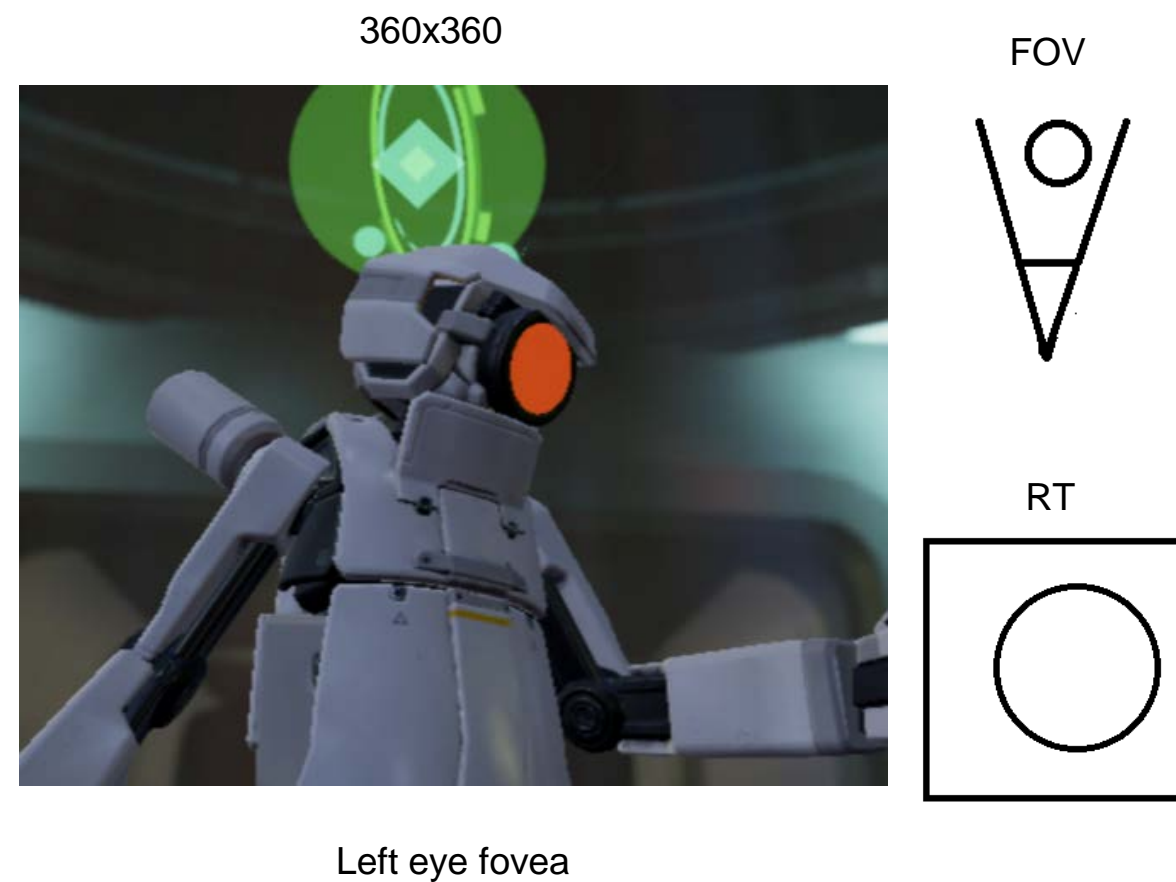
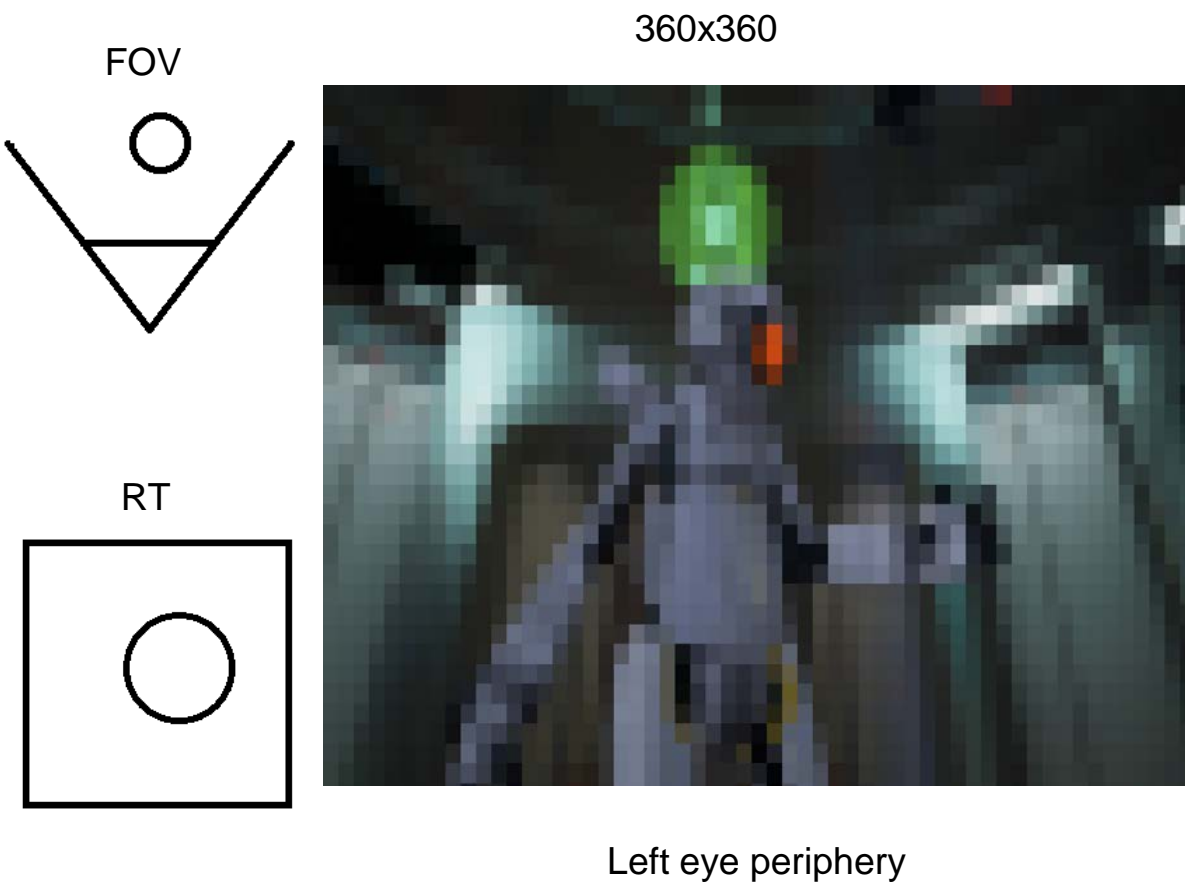
New technologies in the horizon using Multiview

- Foveated Rendering

Foveated Rendering



Foveated Rendering: 4-view multiview



Foveated Rendering: 4-view multiview

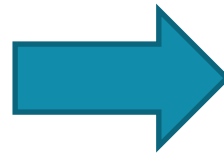
Pipeline

- Current 2-view Multiview in UE4.14

Scene Rendering into
Texture array of $2 \times 1024 \times 1024 = 2.09 \text{ MPx}$



Blit



Side-by-Side texture 2048×1024



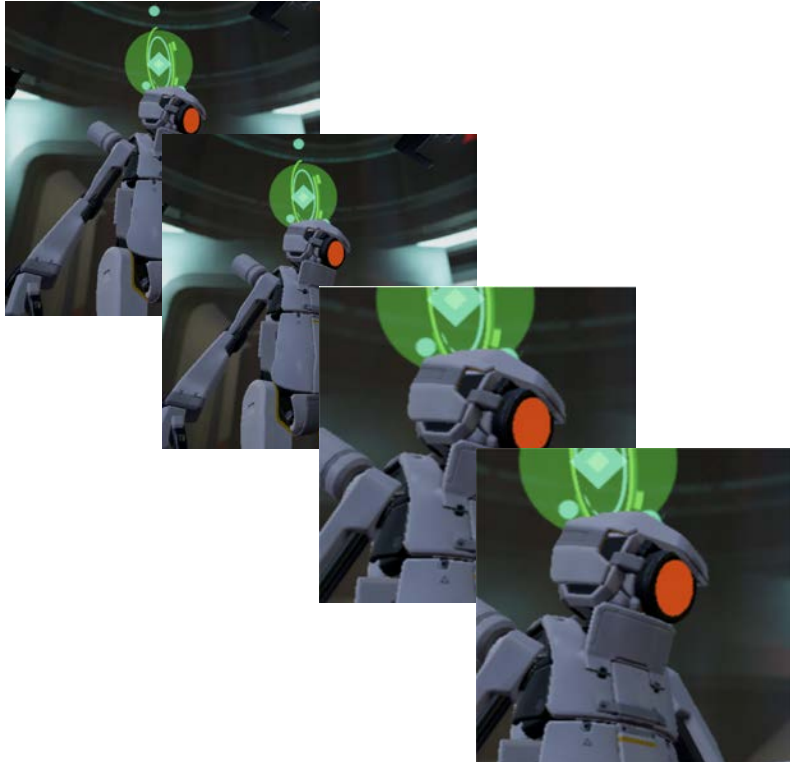
Foveated Rendering: 4-view multiview

Pipeline

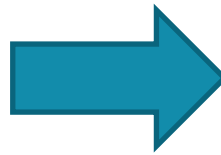
- Foveated Rendering with 4-view Multiview (65% reduction)

Scene Rendering into
Texture array of 4x360x360 = **0.52 MPx**

Side-by-Side texture 2048x1024

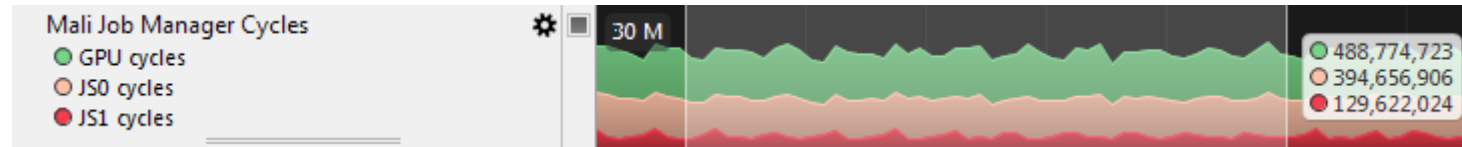


Composing



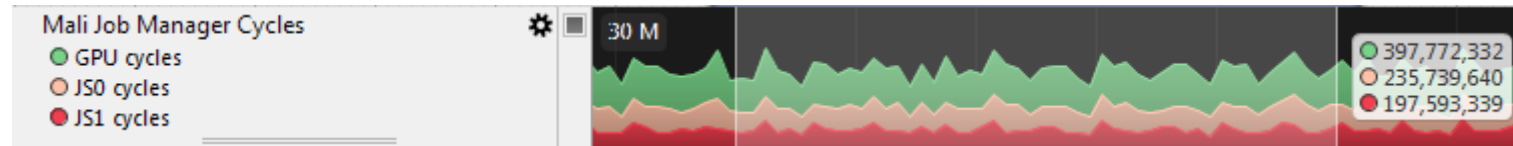
Foveated Rendering: Results

Multiview



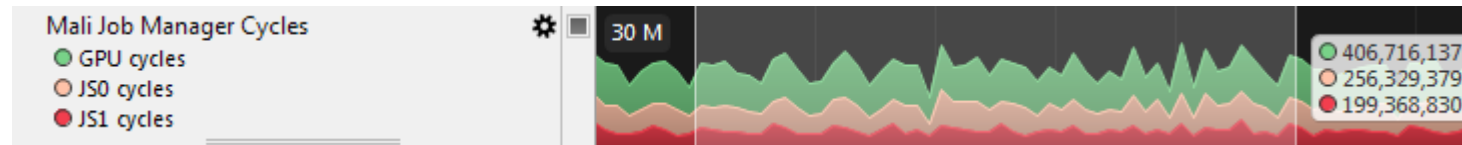
Total: 488 Mcycles
Vertex: 129 Mcycles
Fragment: 394 Mcycles

Foveated (35% original framebuffer size)



Total: 397 Mcycles (-20%)
Vertex: 197 Mcycles (+52%)
Fragment: 235 Mcycles (-40%)

Foveated 8xMSAA (35% original framebuffer size)



Total: 406 Mcycles (-17%)
Vertex: 199 Mcycles (+53%)
Fragment: 256 Mcycles (-35%)

Debugging and Profiling on Mali

- Mali Offline Shader Compiler
- Mali Graphics Debugger (MGD)
- Streamline
 - Result for Foveated Rendering and Circuit VR

Mali Offline Shader Compiler

- Mali Offline Compiler
- Analyze shader performance
- Command line tool. Easy to integrate.
- Number of cycles
- Registers utilization



Mali Offline Shader Compiler use-case

- Shows how many cycles the shortest and longest path takes:
 - Arithmetic pipeline
 - Load/Store pipeline:
 - Texture Pipeline

```
D:\CircuitUR\circuit-vr\TEMP>malisc -c Mali-T880 --vertex Vertex.glsl
ARM Mali Offline Compiler v5.6.0
(C) Copyright 2007-2017 ARM Limited.
All rights reserved.

No driver specified, using "Mali-T600_r13p0-00rel0" as default.
No core revision specified, using "r2p0" as default.

16 work registers used, 8 uniform registers used, spilling used.

Instructions Emitted:  A      L/S      T      Bound
Shortest Path Cycles: 11     15       0      L/S
Longest Path Cycles: 15     22       0      L/S

A = Arithmetic, L/S = Load/Store, T = Texture
```

Mali Offline Shader Compiler: Getting the UE4 shaders

- Enable shader dumps and shader development in ConsoleVariables.ini
- Invalidate the shader cache with r.InvalidateCachedShaders 1
- Restart the editor
 - This will dump the shaders in
 <ProjectFolder>/Saved/ShaderDebugInfo/PCD3D_SM[4|5]
- To generate the shaders for mobile:
 - Package the game for a mobile platform or
 - Activate the mobile preview
- Shaders will be in
 - <ProjectFolder>/Saved/ShaderDebugInfo/GLSL_ES[2|3]



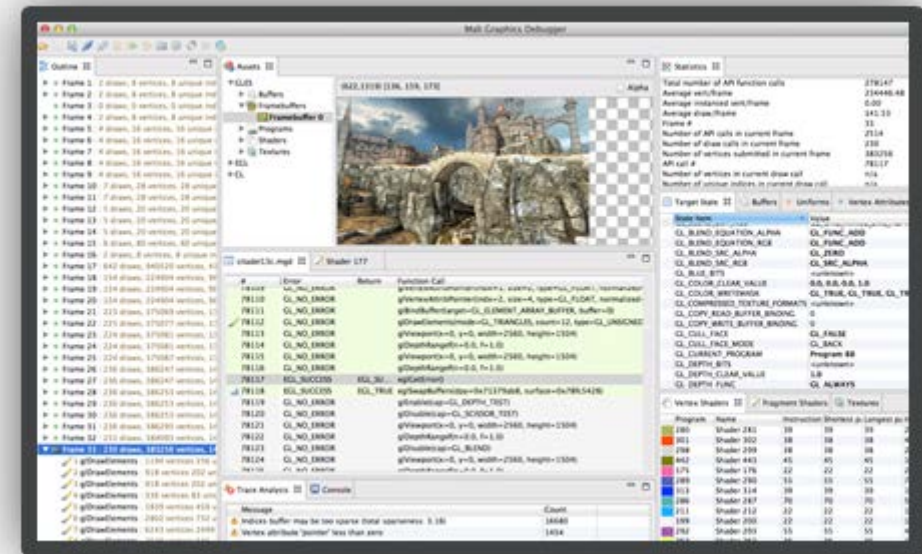
Generic Device



Mali Graphics Debugger (MGD)

- Runtime API Trace and resources analysis
- OpenGL ES, OpenCL
- Debug and improve performance at frame level

- Available in UE4.15!! 😊
- But currently not fully working with VR 😞
- Still useful to debug a No-VR version

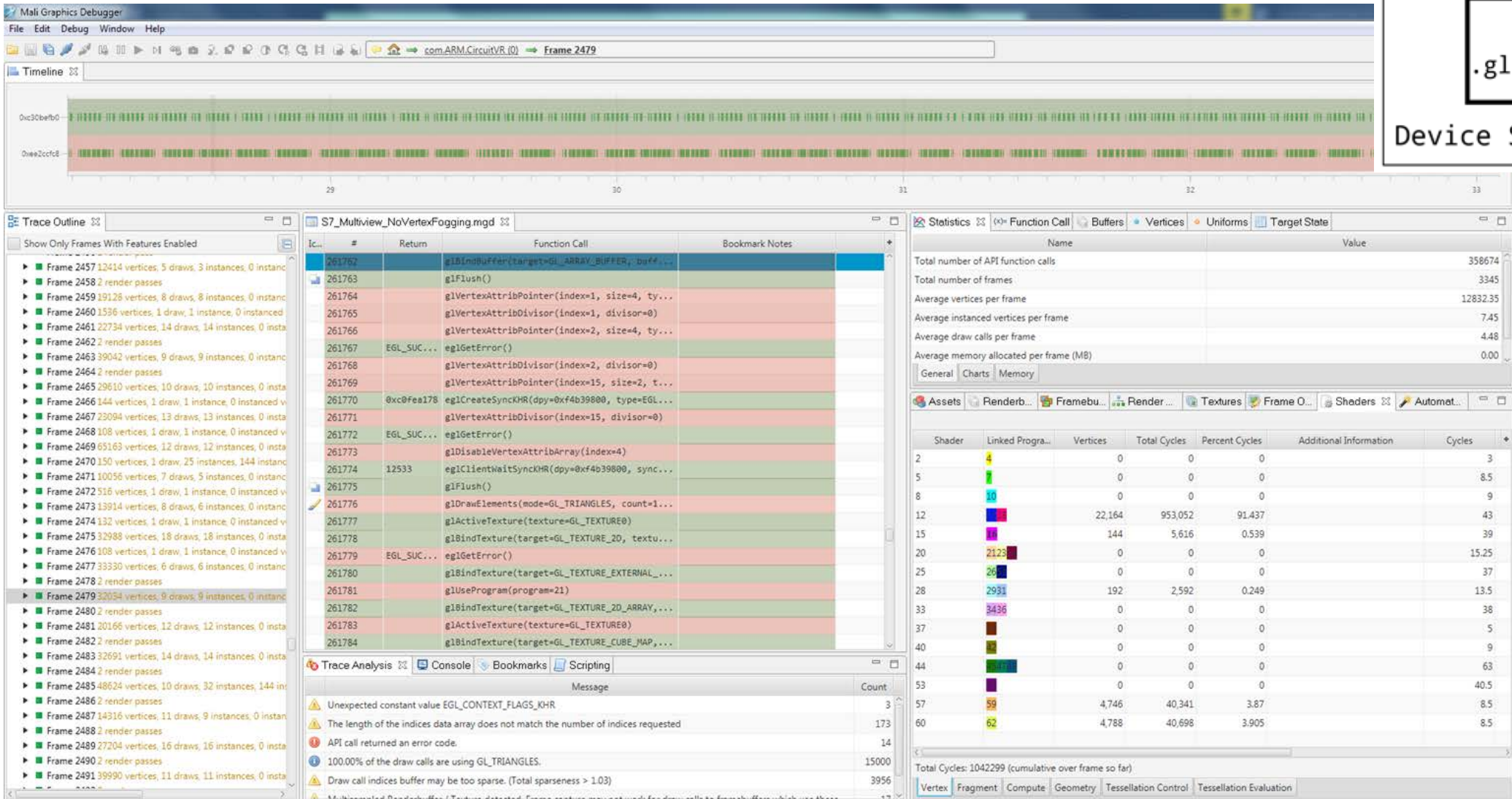


Mali Graphics Debugger (MGD)



Timeline

Drawcalls

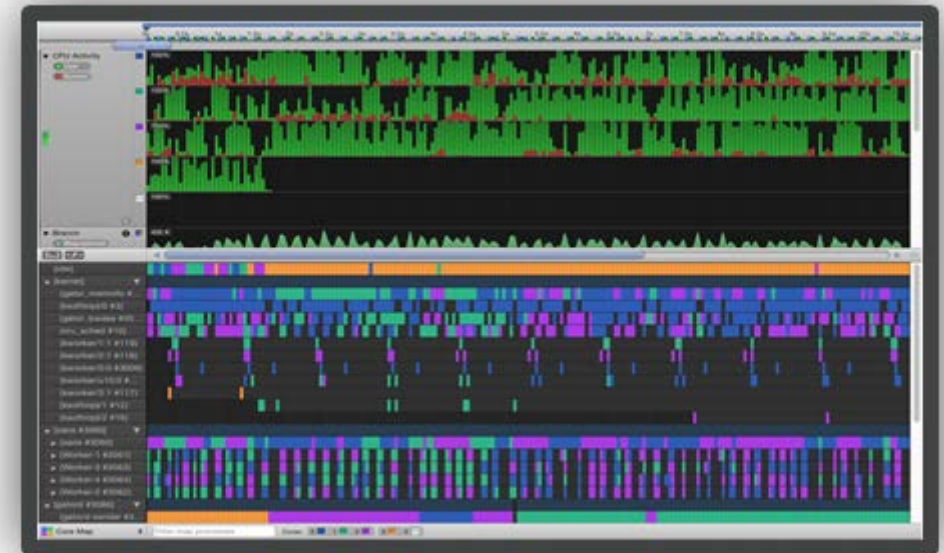


- Statistics
- Buffers
- Vertices
- Uniforms
- TargetState

- Shaders
- FBs
- Textures
- RBs
- Assets
- Etc...

Streamline

- Profile CPUs and Mali GPUs
 - Timeline
 - HW Counters
 - OpenCL visualizer
-
- New version in April which shows Mali counters without rooting the device.
 - Rooting needed for precise CPU load analysis

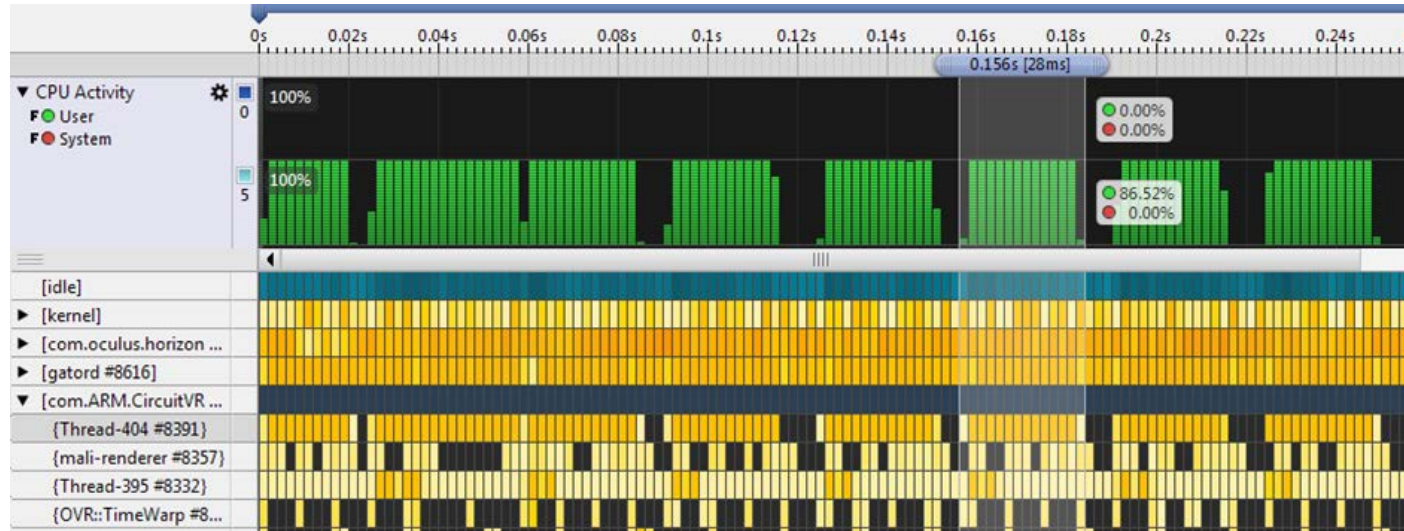


Streamline use-case: Multiview On/Off CPU

- No-Multiview

UE4 Render thread:

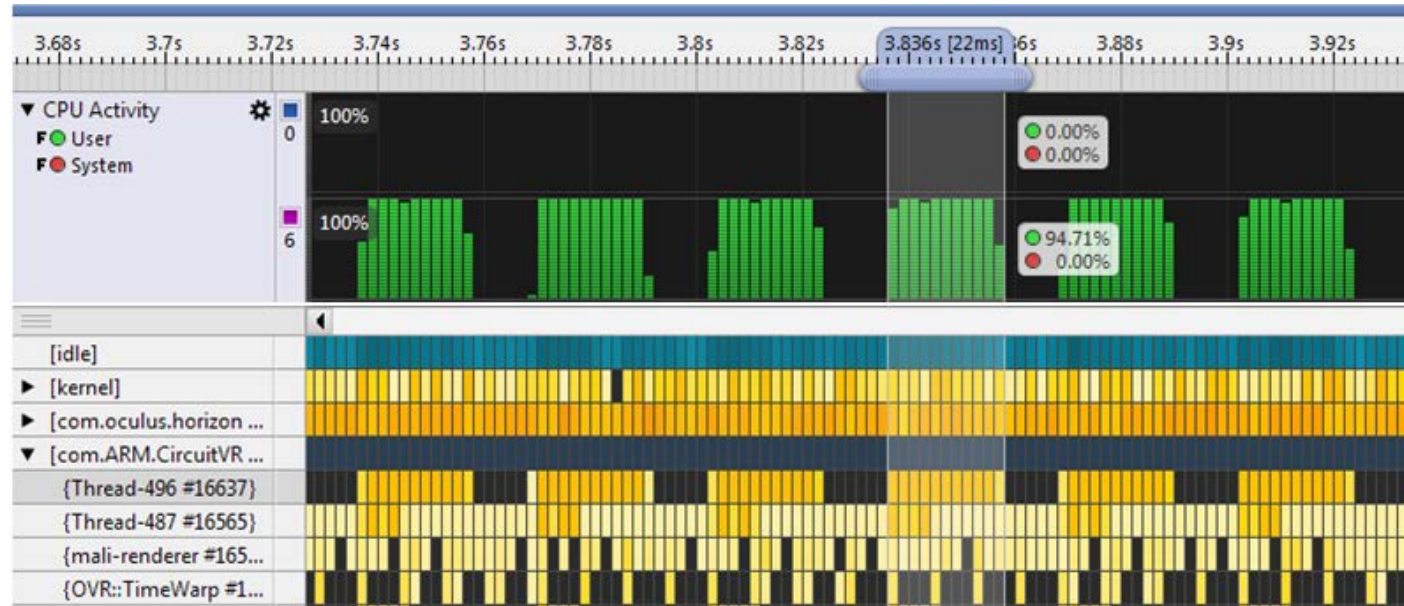
28ms



- Multiview

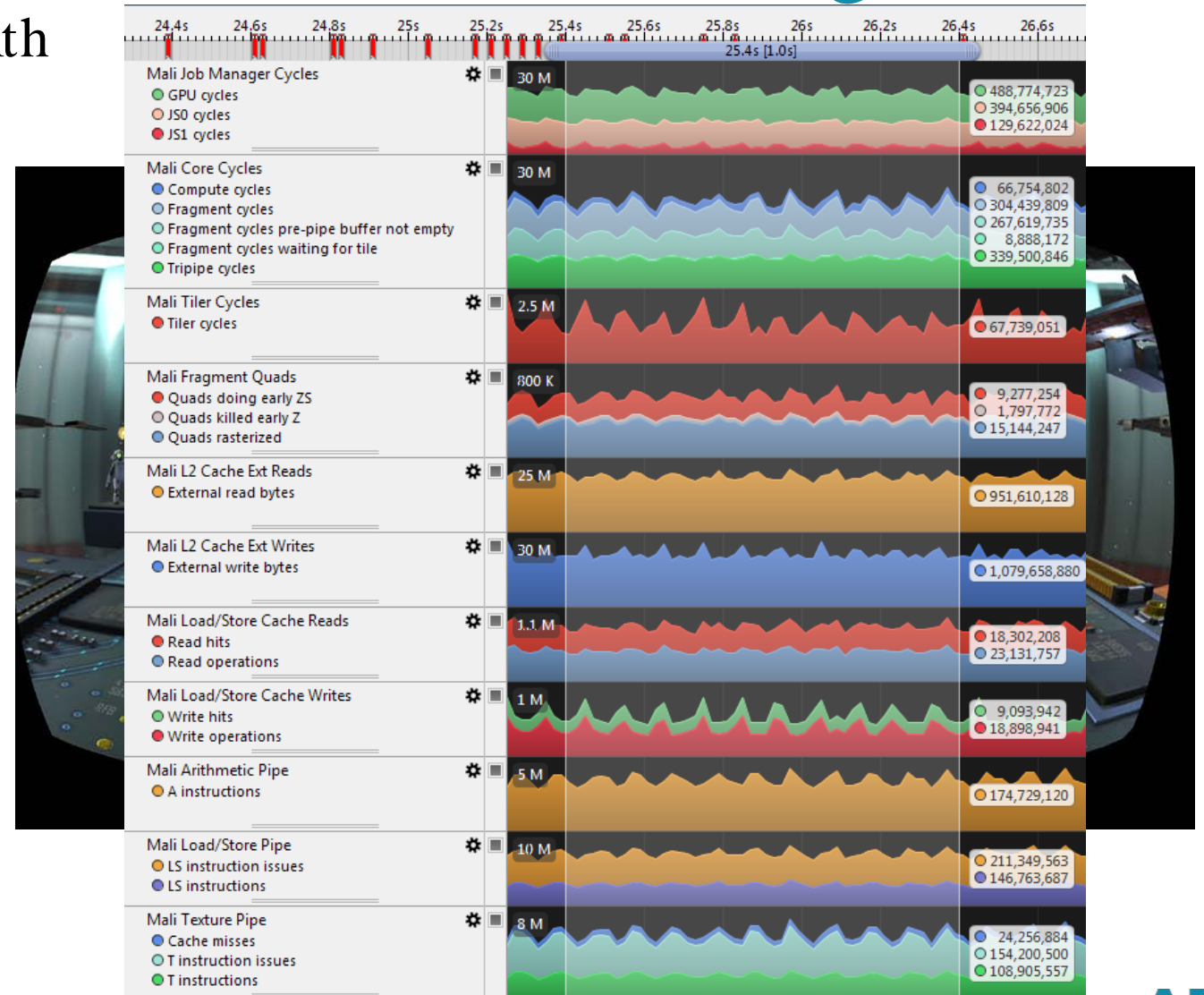
UE4 Render thread:

22ms (~23% reduction in CPU load)



Streamline use-case: Foveated rendering

- Useful metrics measurable with streamline:
 - GPU active cycles
 - Separated for vertex/fragment
 - Overdraw/Early-Z
 - Bandwidth
 - Cache hit/miss for textures and load/store
 - GPU Utilization
 - Separated for Arith, L/S and Texture
 - Average CPI, Cycles per vertex/fragment
 - Much more!



Streamline: GPU references

Description and optimization tips for tiled based gpus:

<https://community.arm.com/graphics/b/blog/posts/the-mali-gpu-an-abstract-machine-part-1---frame-pipelining>

Description of GPU counter available:

Midgard

<https://community.arm.com/graphics/b/documents/posts/mali-midgard-family-performance-counters>

Bifrost

<https://community.arm.com/graphics/b/documents/posts/mali-bifrost-family-performance-counters>

ARM

Q & A

The trademarks featured in this presentation are registered and/or unregistered trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. All rights reserved. All other marks featured may be trademarks of their respective owners.

Copyright © 20 17 ARM Limited