



"IT JUST WORKS": RAY-TRACED REFLECTIONS IN 'BATTLEFIELD V'

Johannes Deligiannis
Jan Schmid
EA DICE

GAME DEVELOPERS CONFERENCE

MARCH 18-22, 2019 | #GDC19

TODAY WE PRESENT RAYTRACING

- Project background
- GPU Raytracing Pipeline
- Engine integration of DXR
- GPU Performance



BATTLEFIELD V

- FPS set in WWII
- Released Nov 2018
- Raytracing work began Dec 2017
- First DXR game released!

PROJECT BACKGROUND

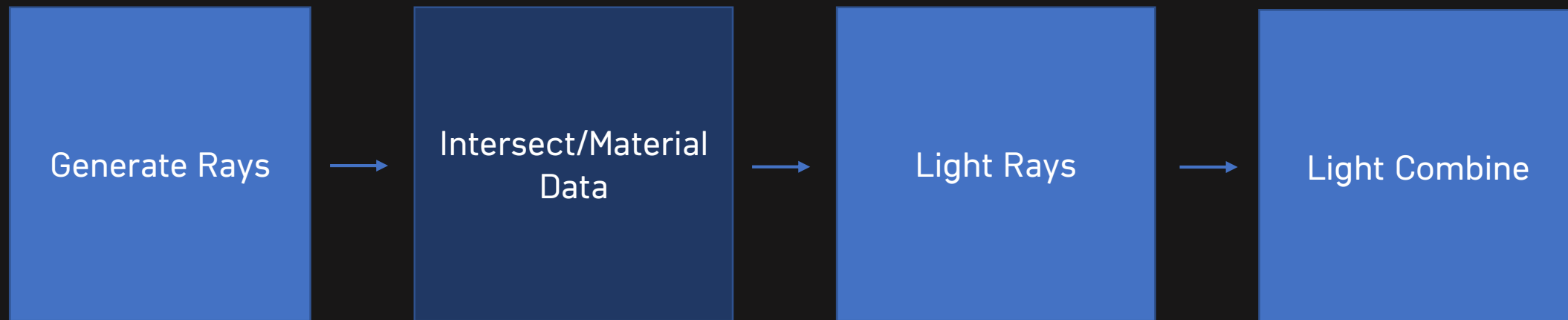
- ~10 months dev time
- Use DXR in Battlefield V
 - AO
 - GI
 - Shadows
 - Reflections
- Engineering
 - Yasin Uludag (EA DICE)
 - Johannes Deligiannis (EA DICE)
 - Jiho Choi (NVIDIA)
 - Pawel Kozlowski (NVIDIA)
- And a bunch of other people! 😊

MAIN CHALLENGES

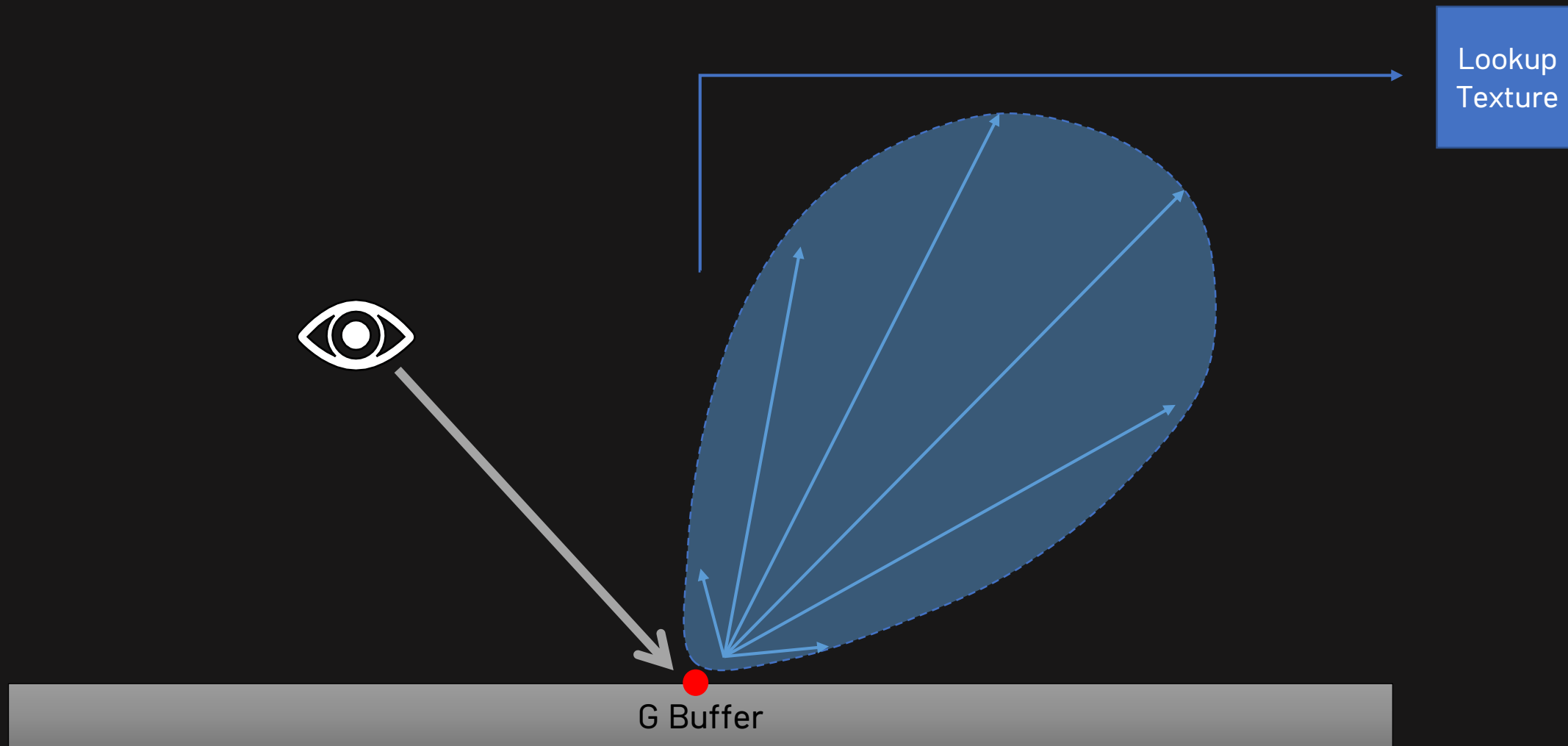
- Not a **Tech Demo**
 - Content is set
- Game in full production
- Scope of Engine changes
- Performance
 - Denoising vs Ray Count
 - No RTX cards
- Early adopter tax ☹️
 - API not final
 - Driver hang/bugs
 - BSoD
 - No capture tool (Nsight, Pix)
- But we shipped it 😊



(SIMPLE) RAYTRACING PIPELINE



GENERATE RAYS



*Tomasz Stachowiak and Yasin Uludag, Siggraph 2015. "Stochastic Screen-Space Reflections"

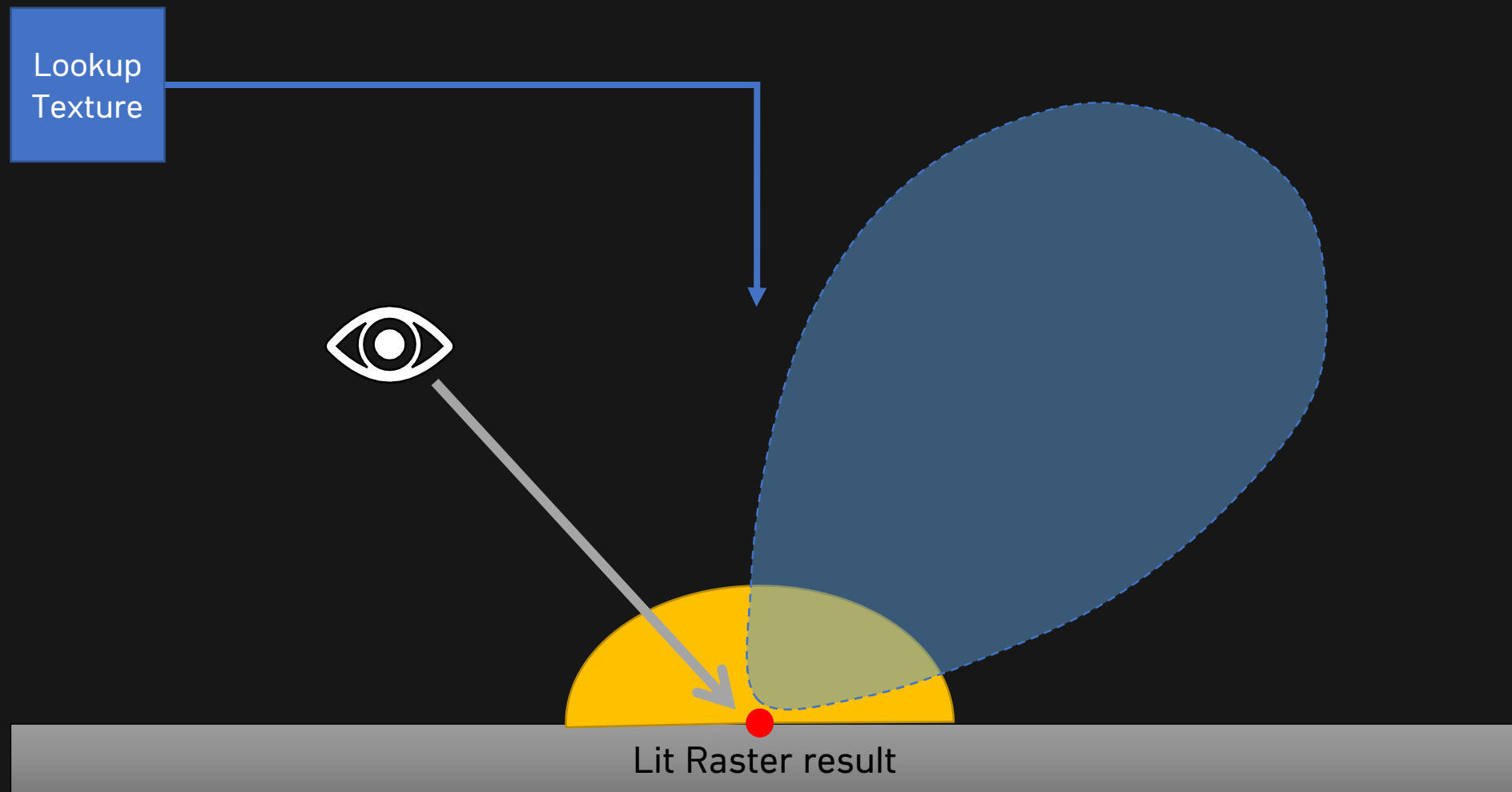
RAYTRACING

MAGIC

LIGHT RAYS

```
float4 light(MaterialData surfaceInfo, float3 rayDir) {  
    foreach (light : pointLights)  
        radiance += calcPoint(surfaceInfo, rayDir, light);  
  
    foreach (light : spotLights)  
        radiance += calcSpot(surfaceInfo, rayDir, light);  
  
    foreach (light : reflectionVolumes)  
        radiance += calcReflVol(surfaceInfo, rayDir, light);  
  
    ...  
}
```

LIGHT COMBINE

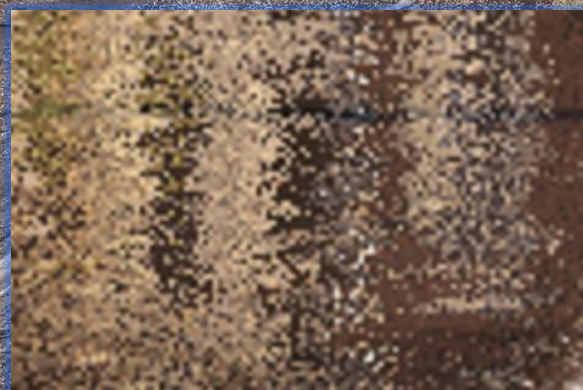




Rays Contribute
Less

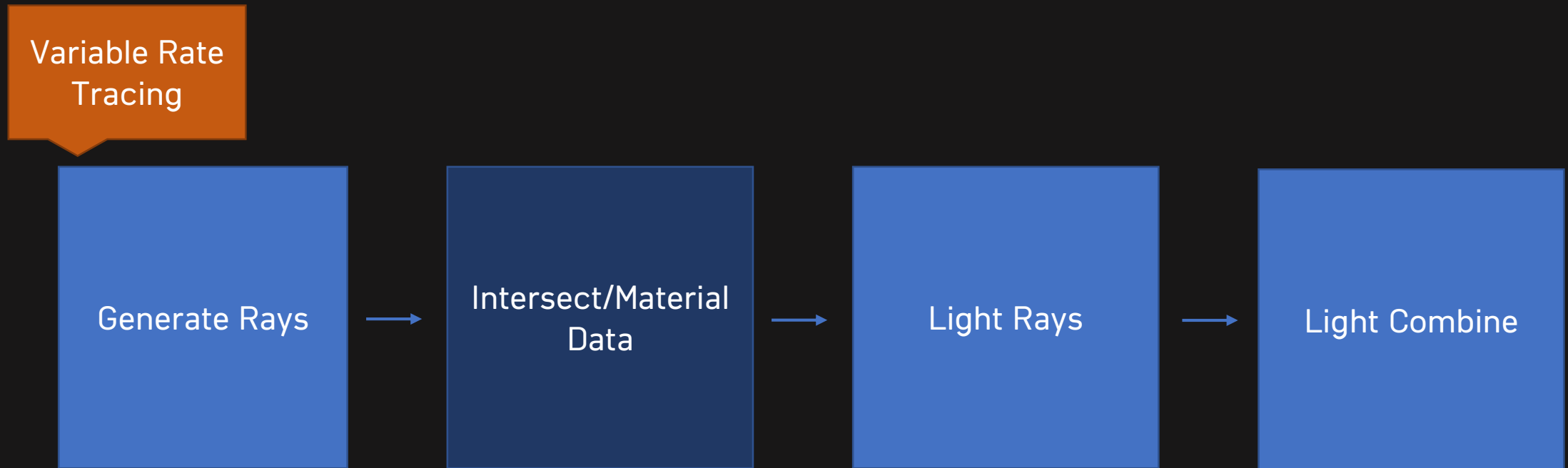
Slooooo

18,458ms

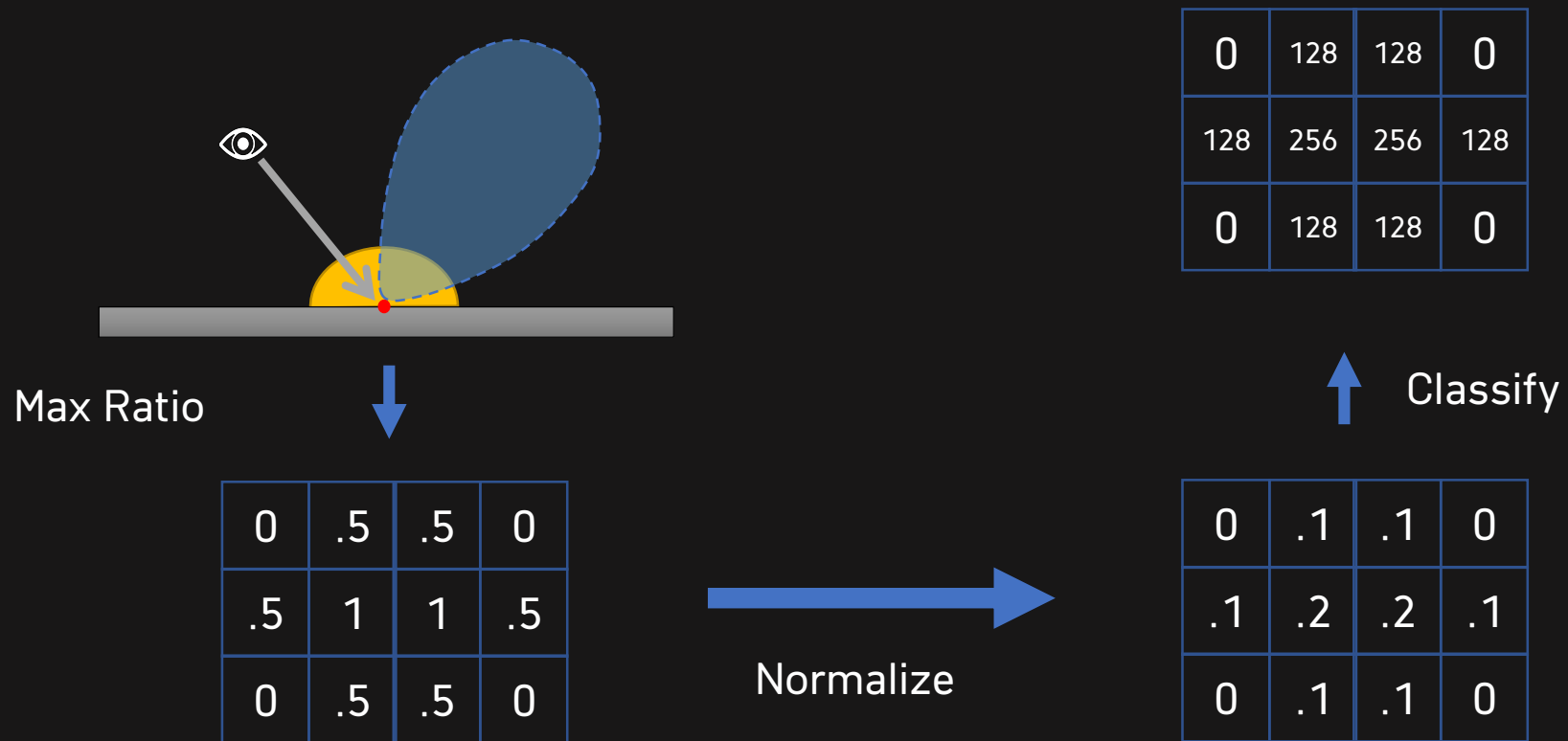


Very Noisy

IMPROVING RAYTRACING PIPELINE

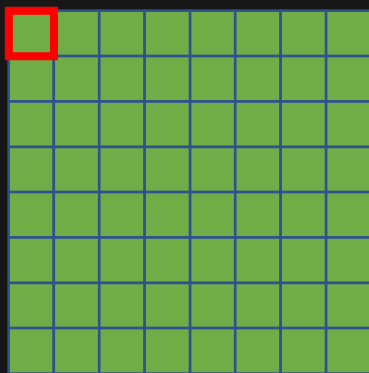


VARIABLE RATE TRACING

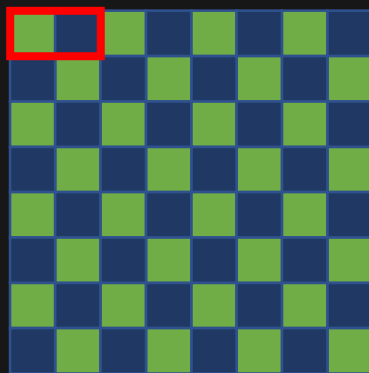


VARIABLE RATE TRACING

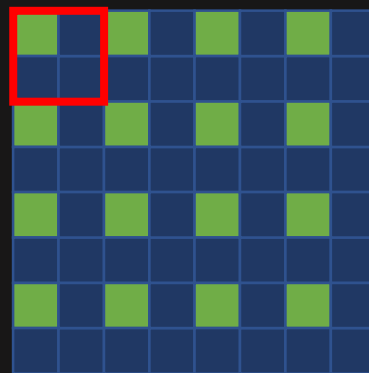
256 rays



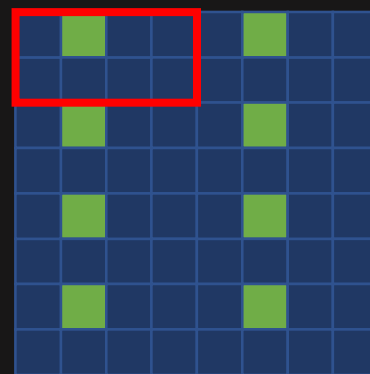
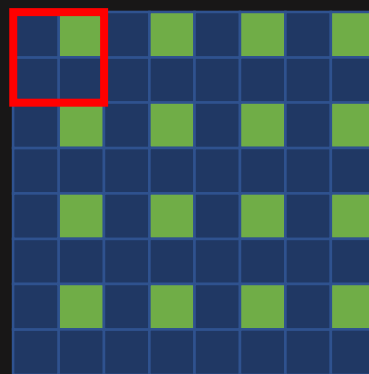
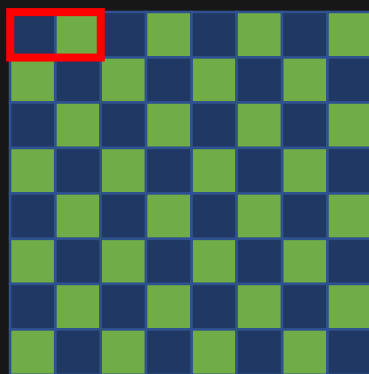
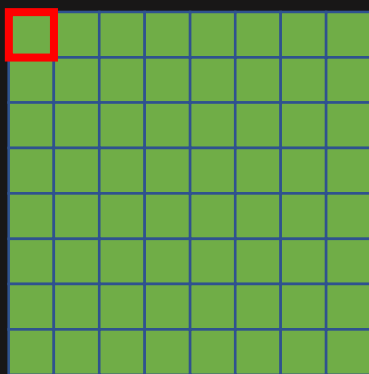
128 rays



64 rays



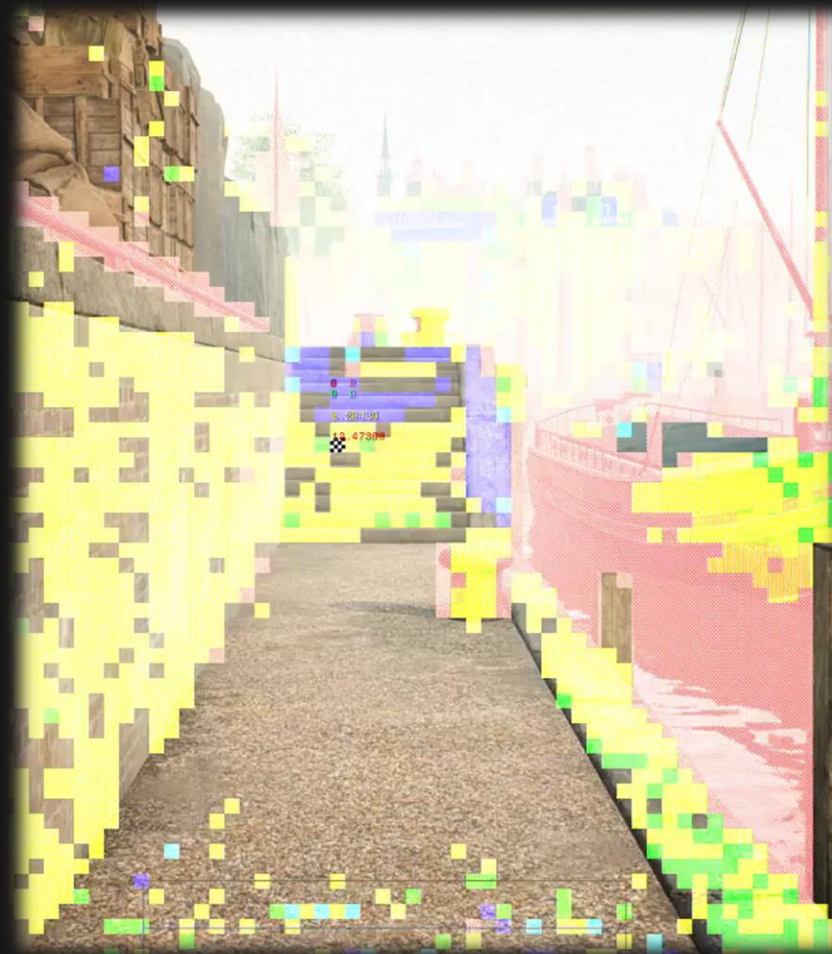
32 rays



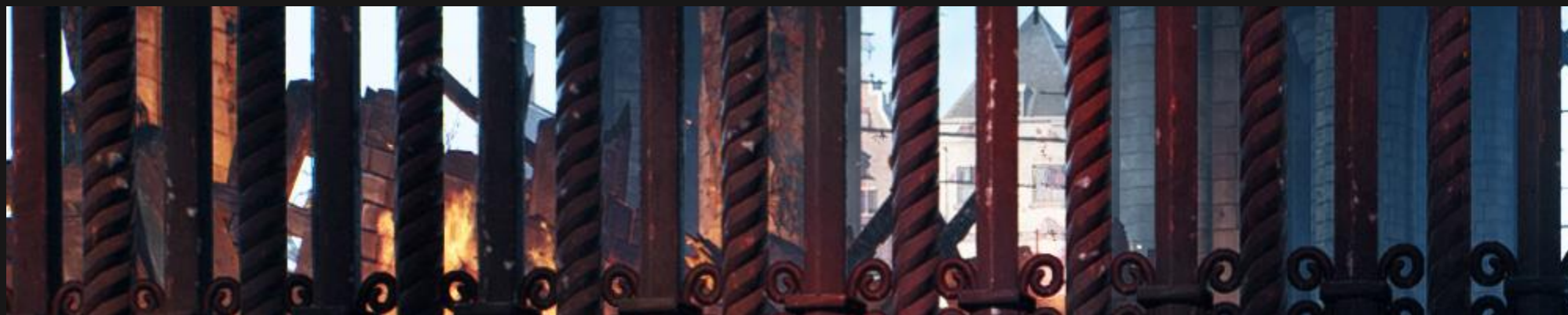
VARIABLE RATE TRACING

Success!

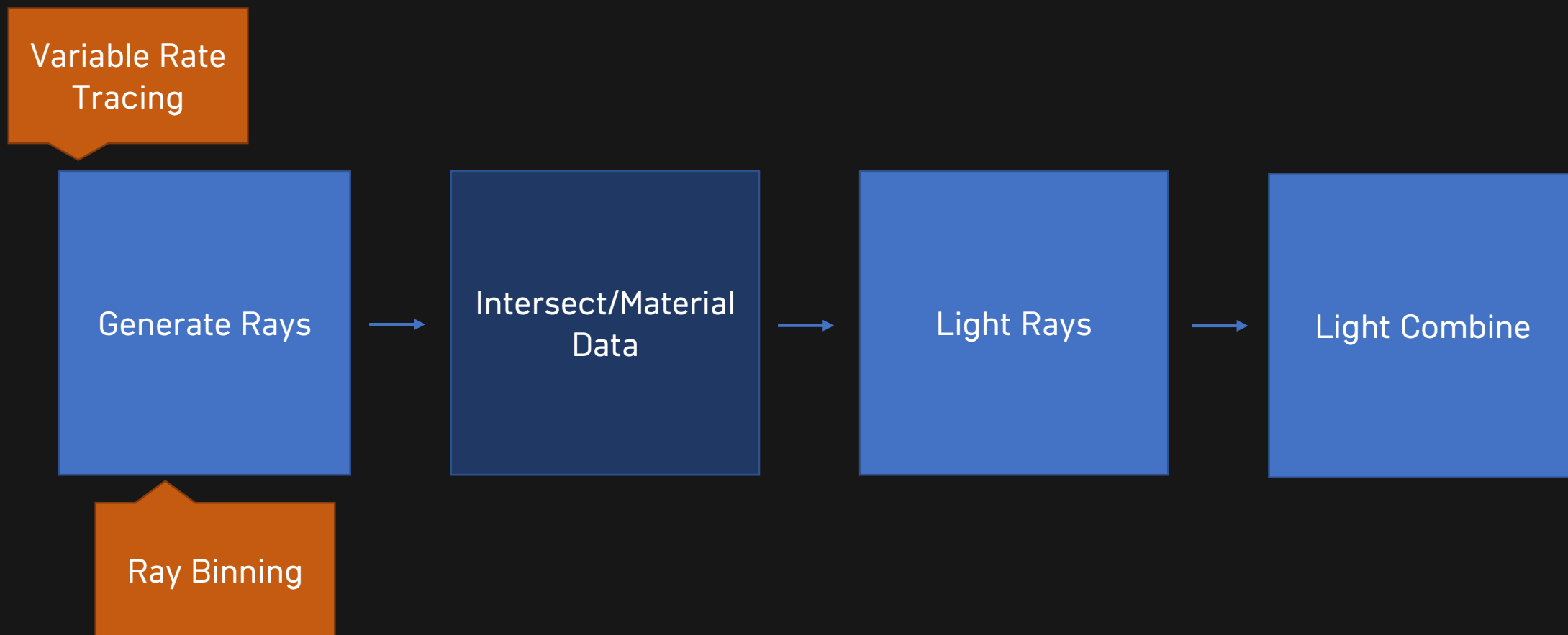
- More Rays on Water
- More Rays on grazing angles



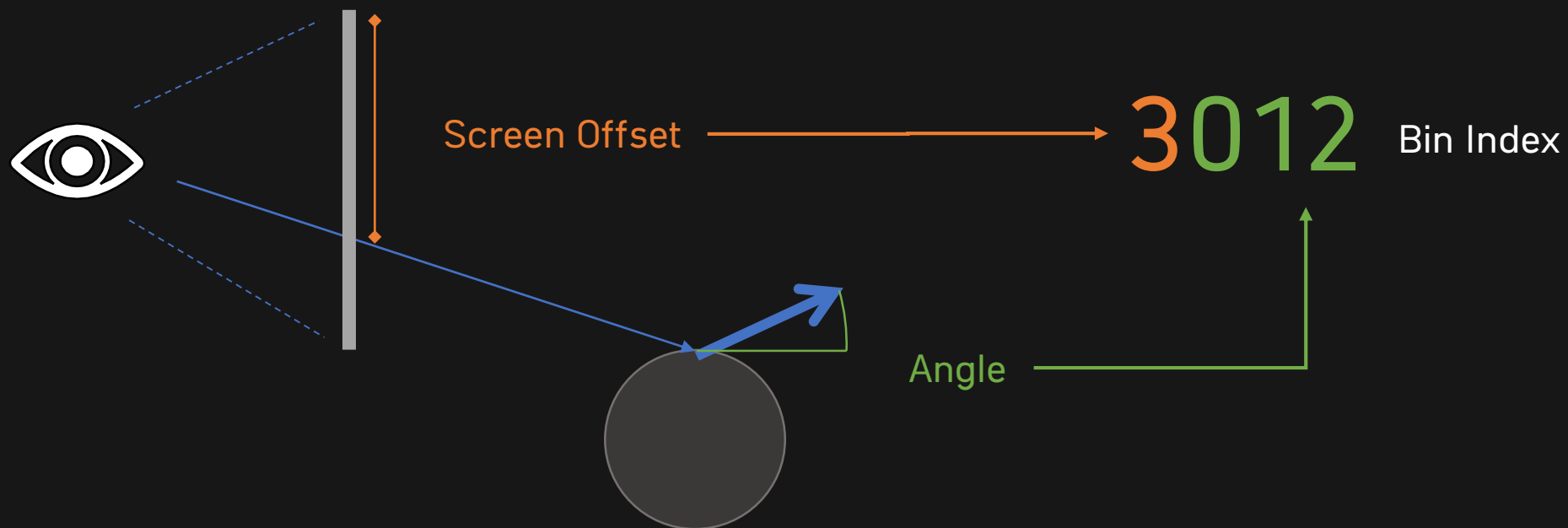
PROBLEM



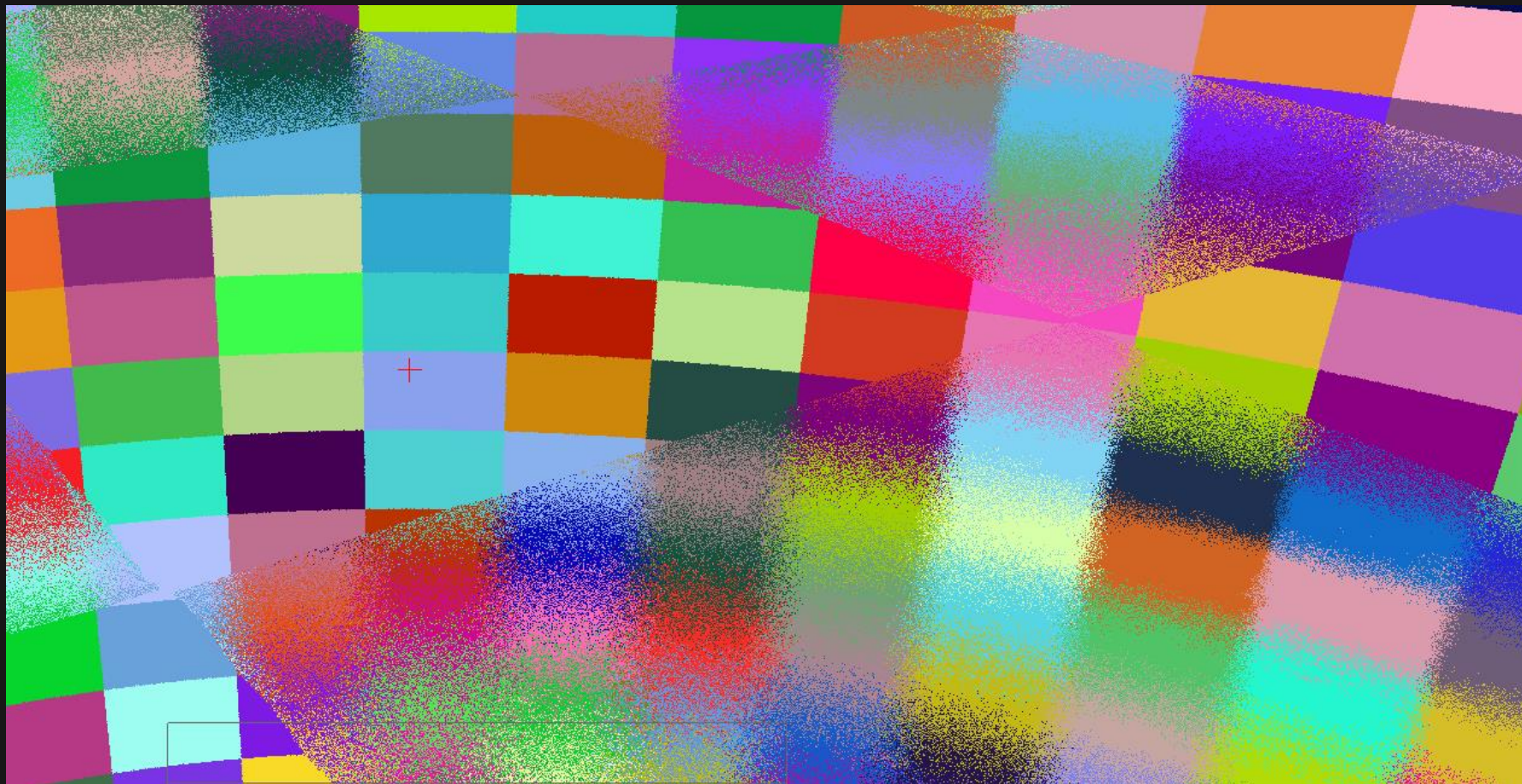
IMPROVING RAYTRACING PIPELINE



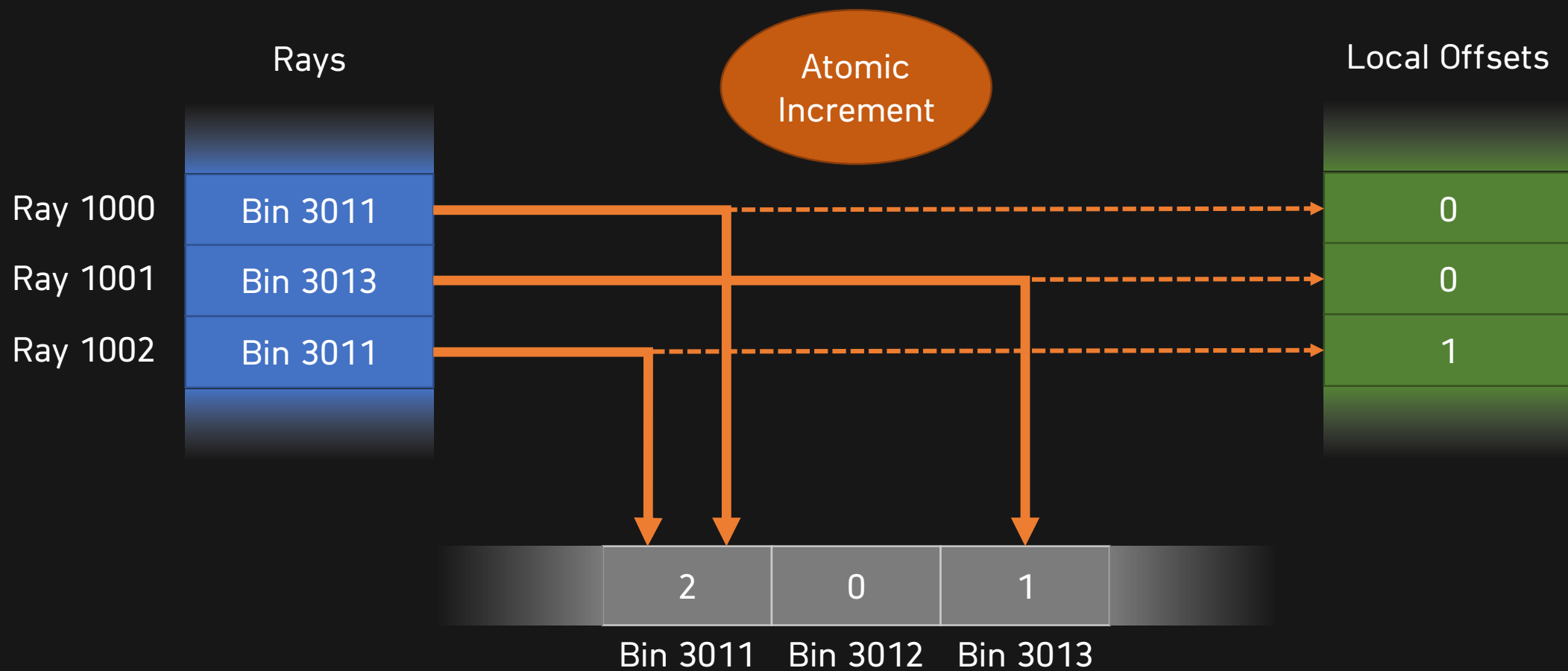
RAY BINNING



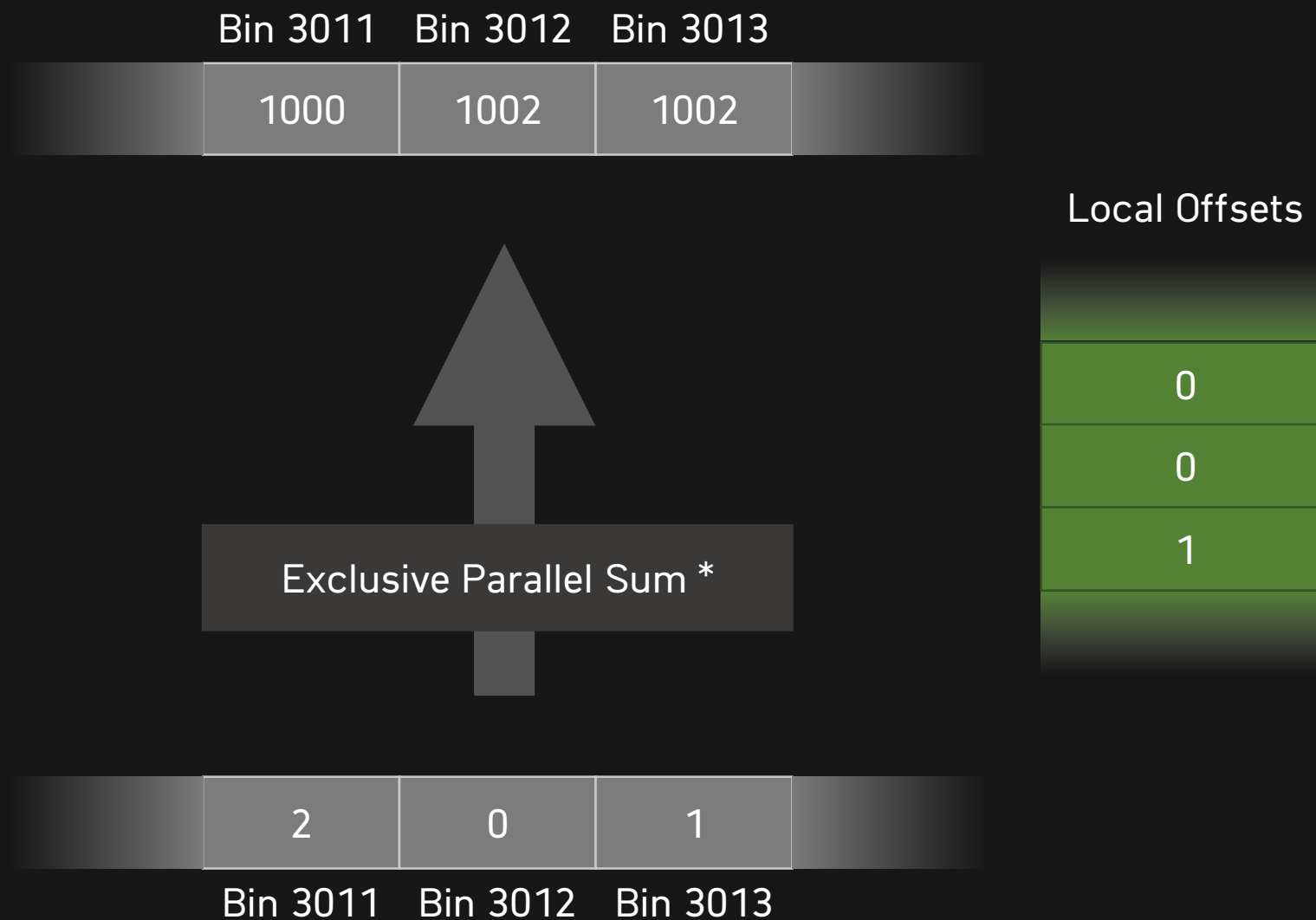
RAY BINNING



RAY BINNING

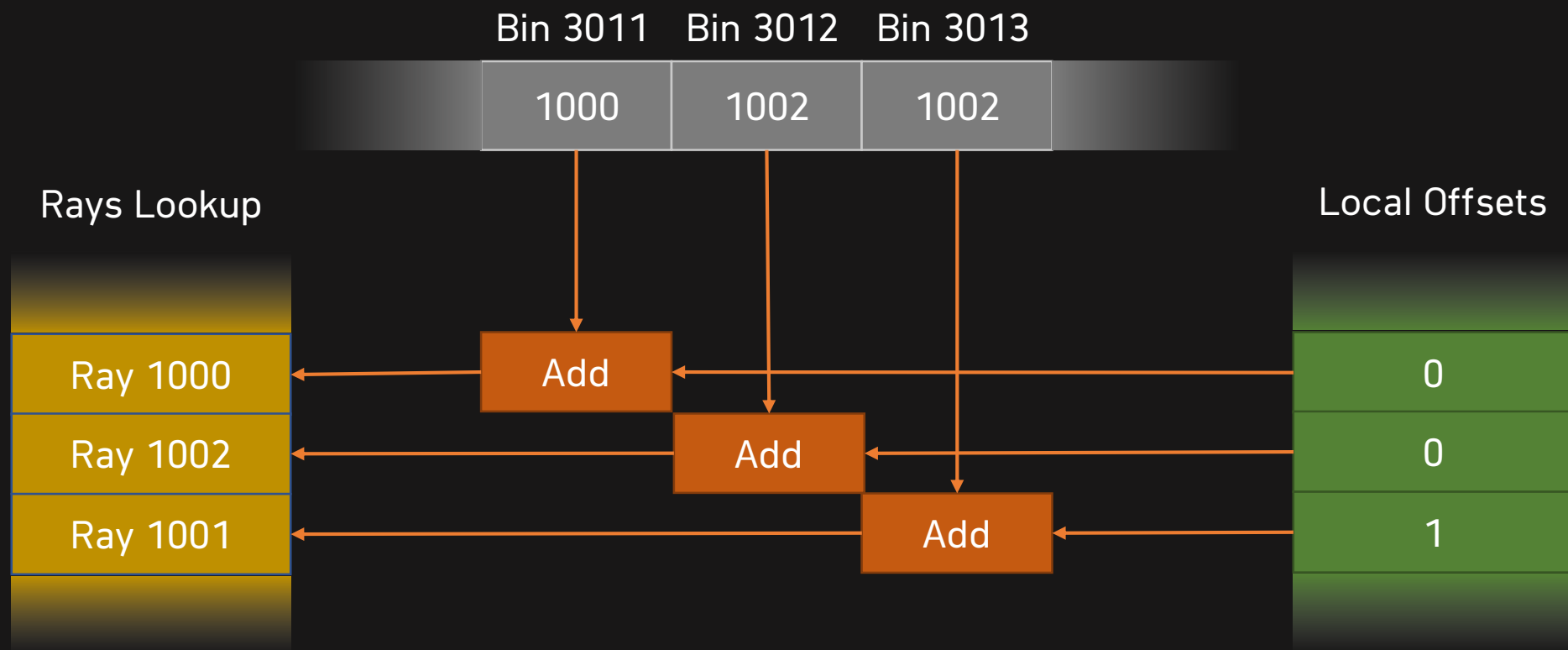


RAY BINNING

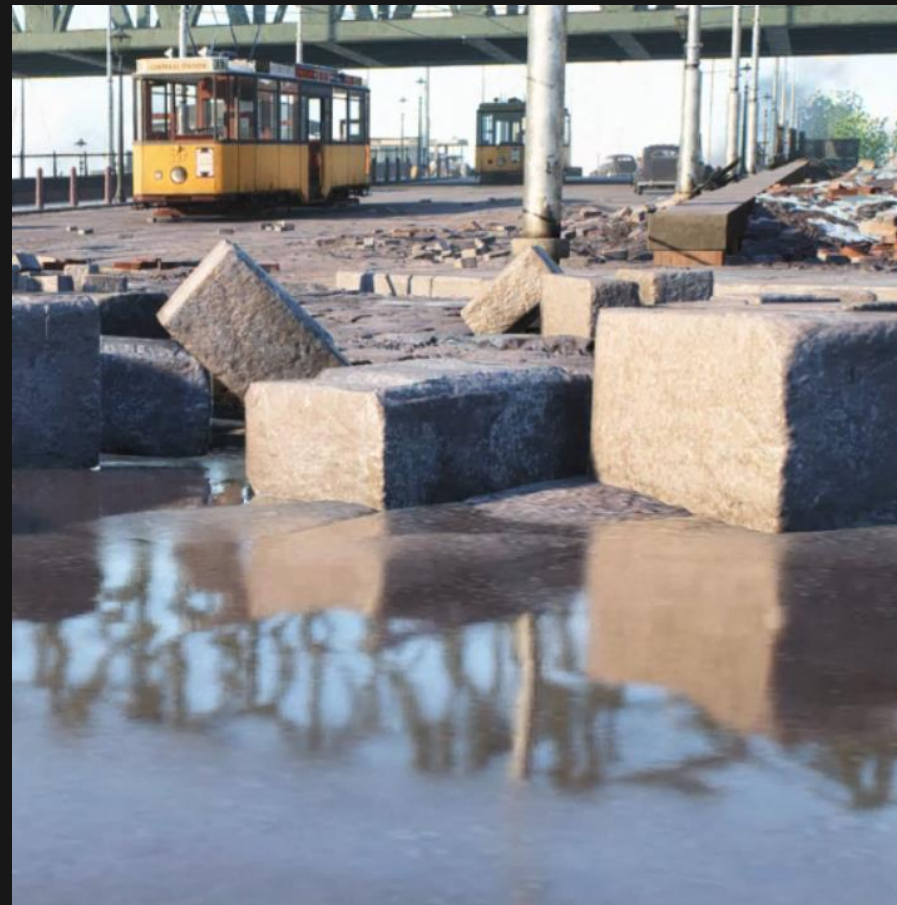


*Mark Harris, Shubhabrata Sengupta, and John Owens. "Parallel Prefix Sum (Scan) with CUDA"

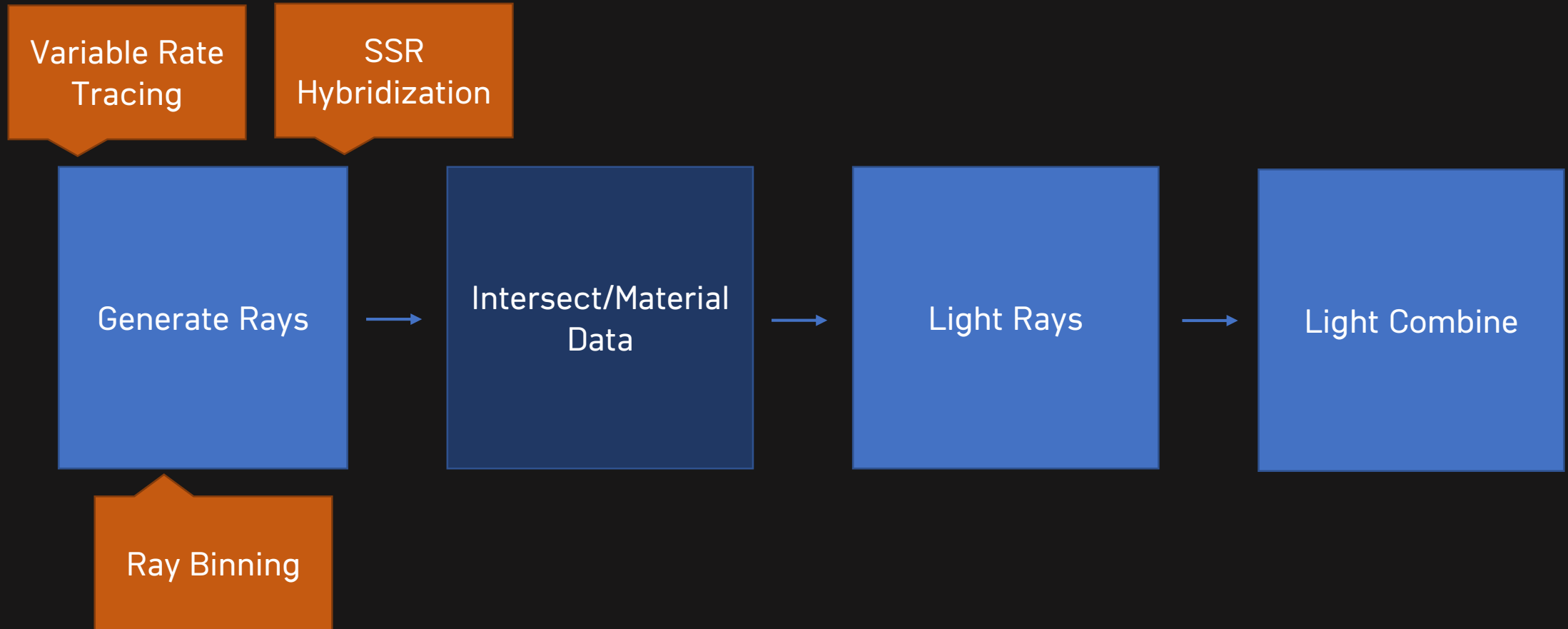
RAY BINNING



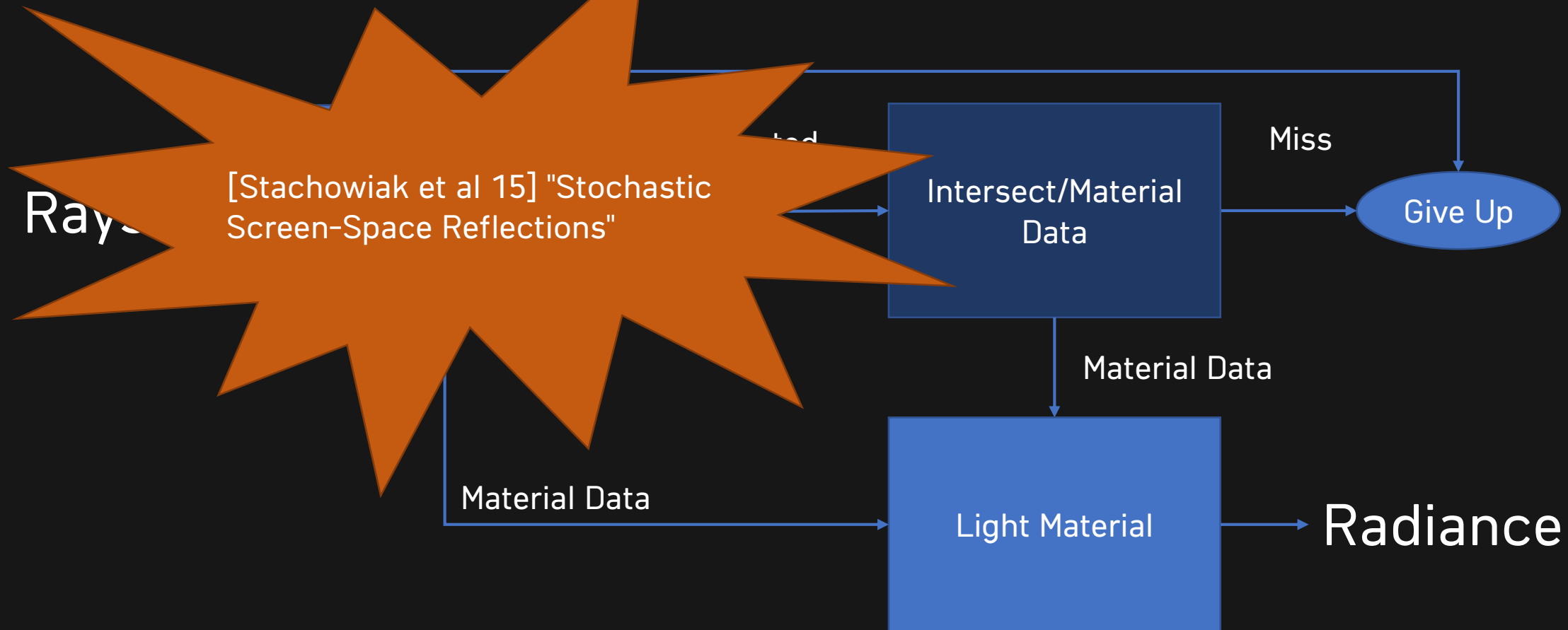
PROBLEM



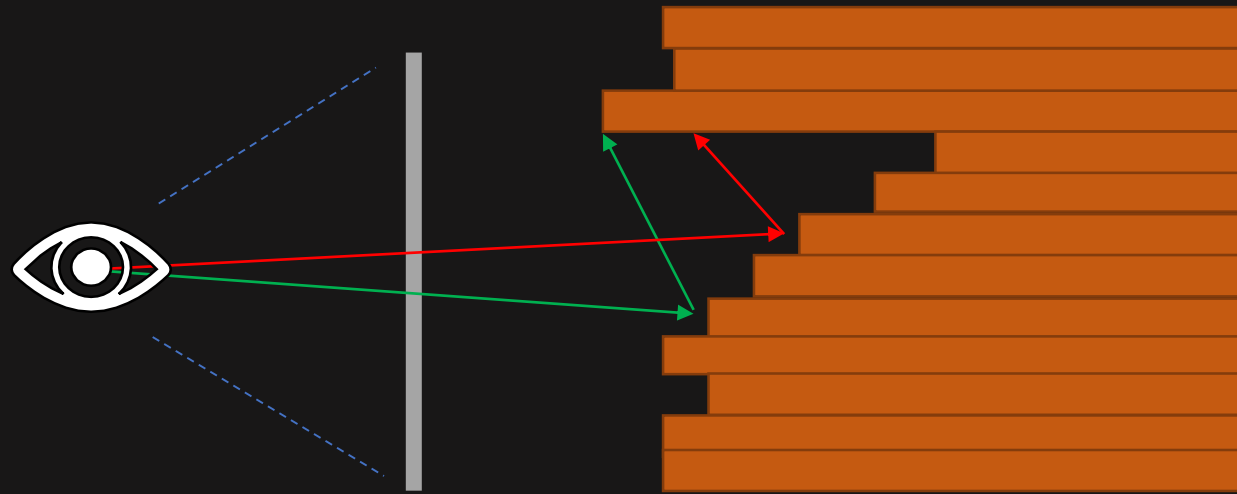
IMPROVING RAYTRACING PIPELINE



SS-HYBRIDIZATION



SS-HYBRIDIZATION



SS-HYBRIDIZATION



PROBLEM

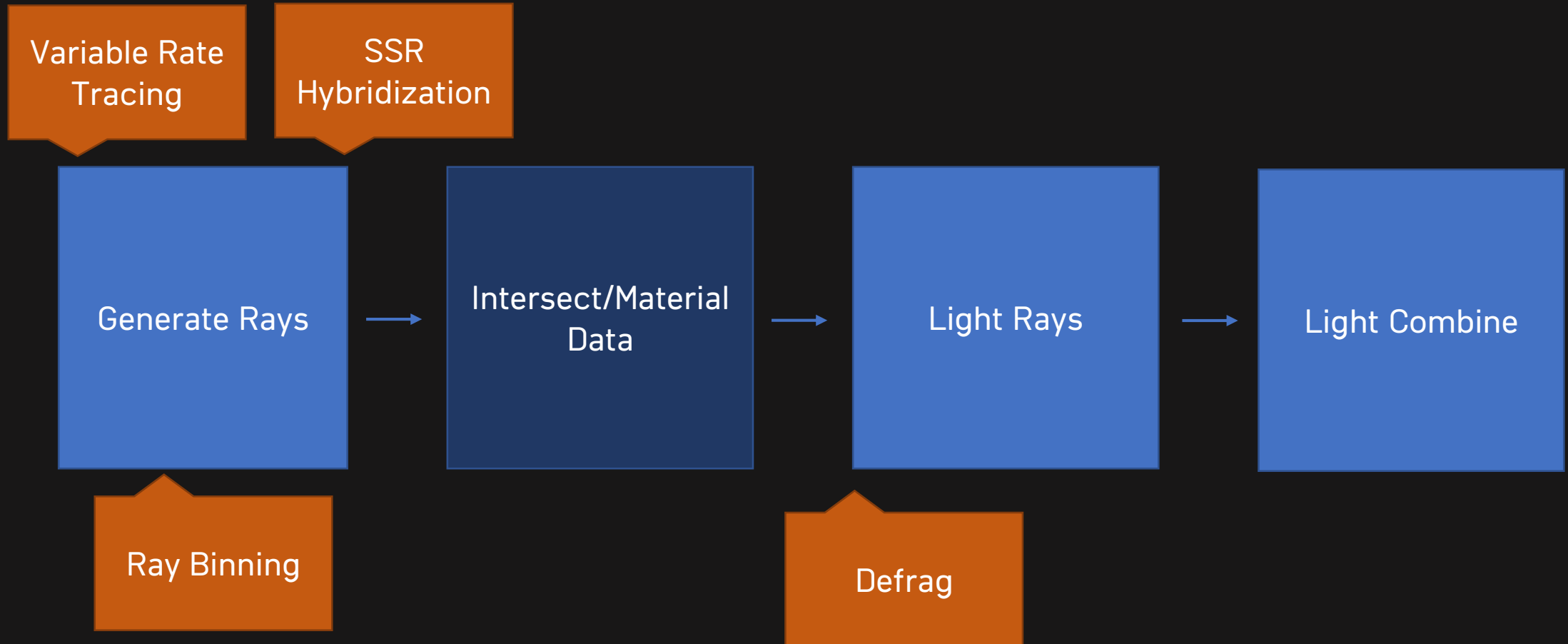
| | | | |
|------|------|-----|------|
| Hit | Miss | Hit | Miss |
| Miss | Hit | Hit | Miss |
| Miss | Hit | Hit | Miss |
| Miss | Miss | Hit | Hit |

Raytrace

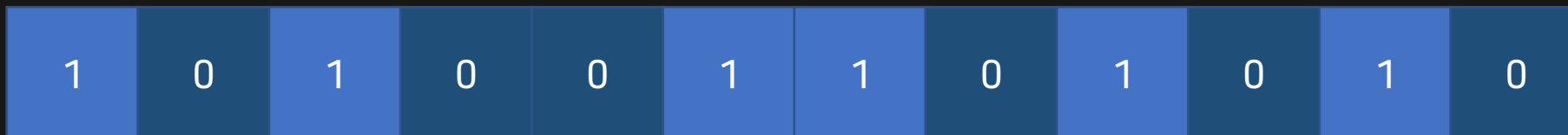
| | | | |
|------|------|------|------|
| Busy | Idle | Busy | Idle |
| Idle | Busy | Busy | Idle |
| Idle | Busy | Busy | Idle |
| Idle | Idle | Busy | Busy |

Light Shader Wavefront

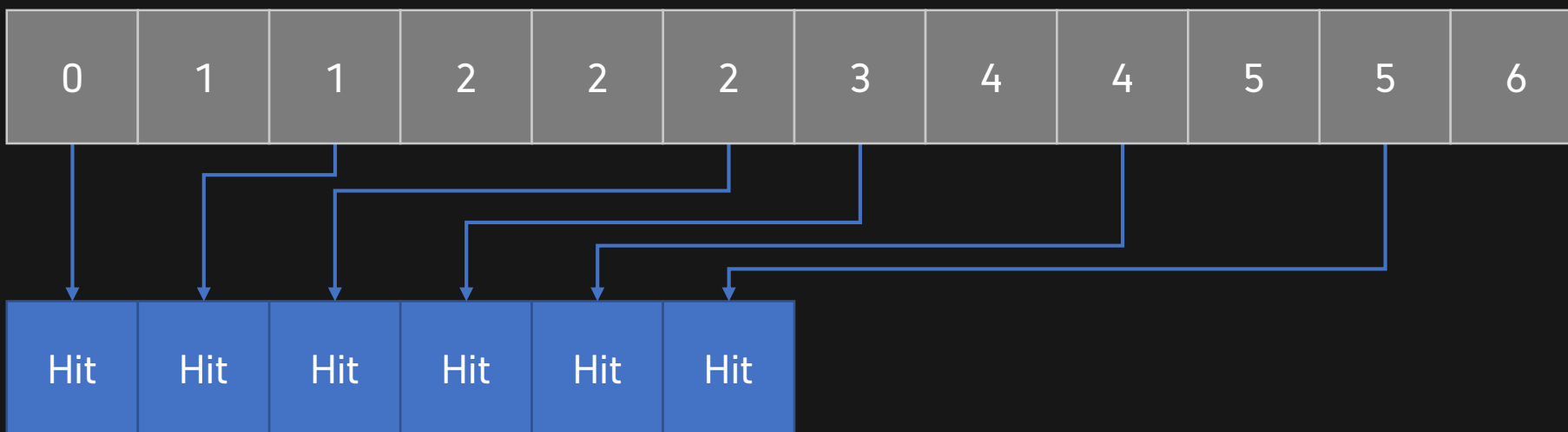
IMPROVING RAYTRACING PIPELINE



DEFRAG



Exclusive Parallel Sum *



*Mark Harris, Shubhabrata Sengupta, and John Owens. "Parallel Prefix Sum (Scan) with CUDA"

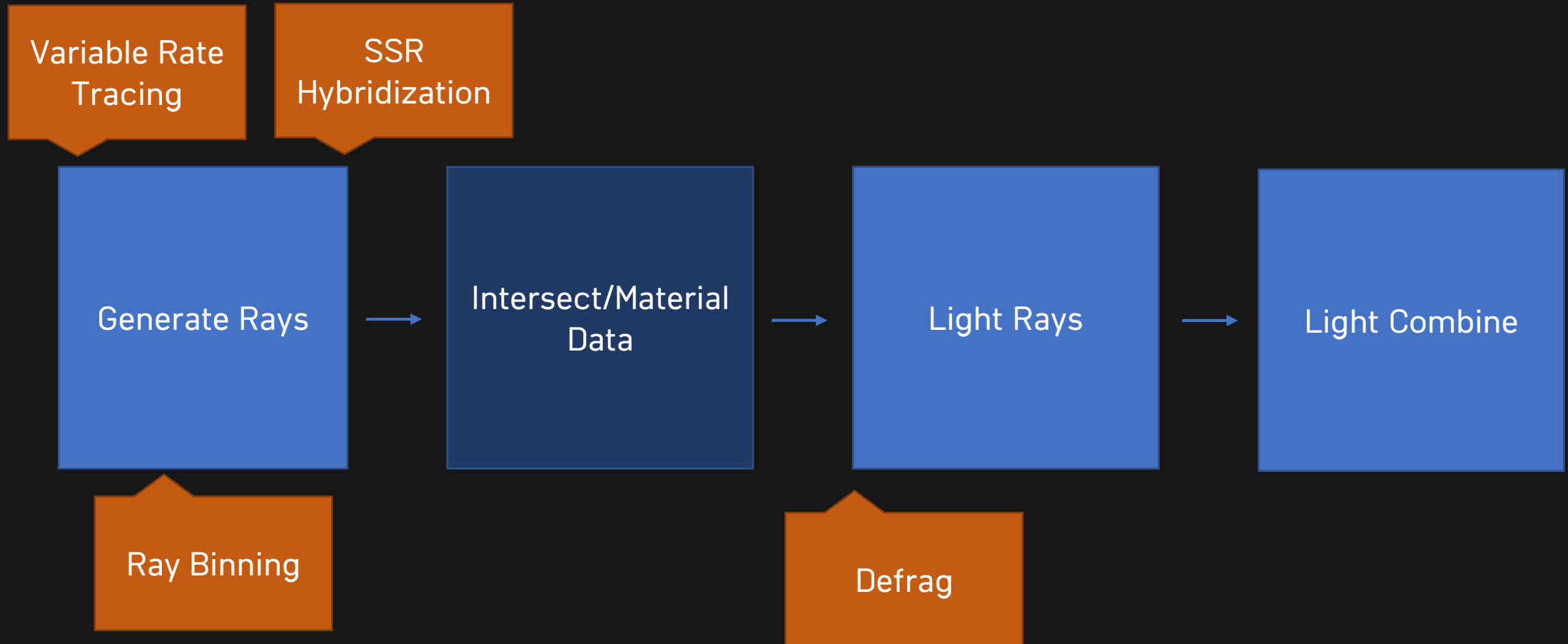
PROBLEM

| | | | |
|------|------|------|------|
| Busy | Busy | Busy | Busy |
| Busy | Busy | Busy | Busy |
| Busy | Busy | Busy | Busy |
| Busy | Busy | Busy | Busy |

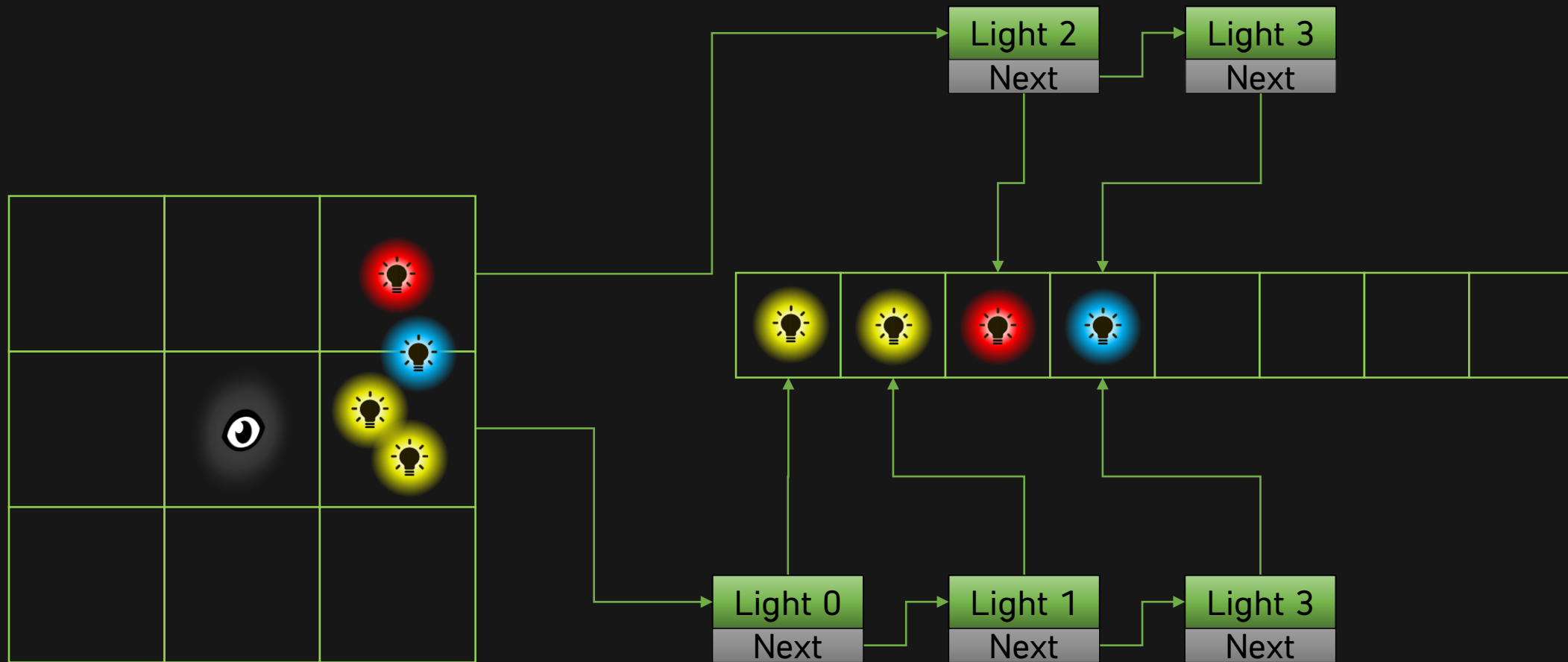
Light Shader

2.0ms

IMPROVING RAYTRACING PIPELINE



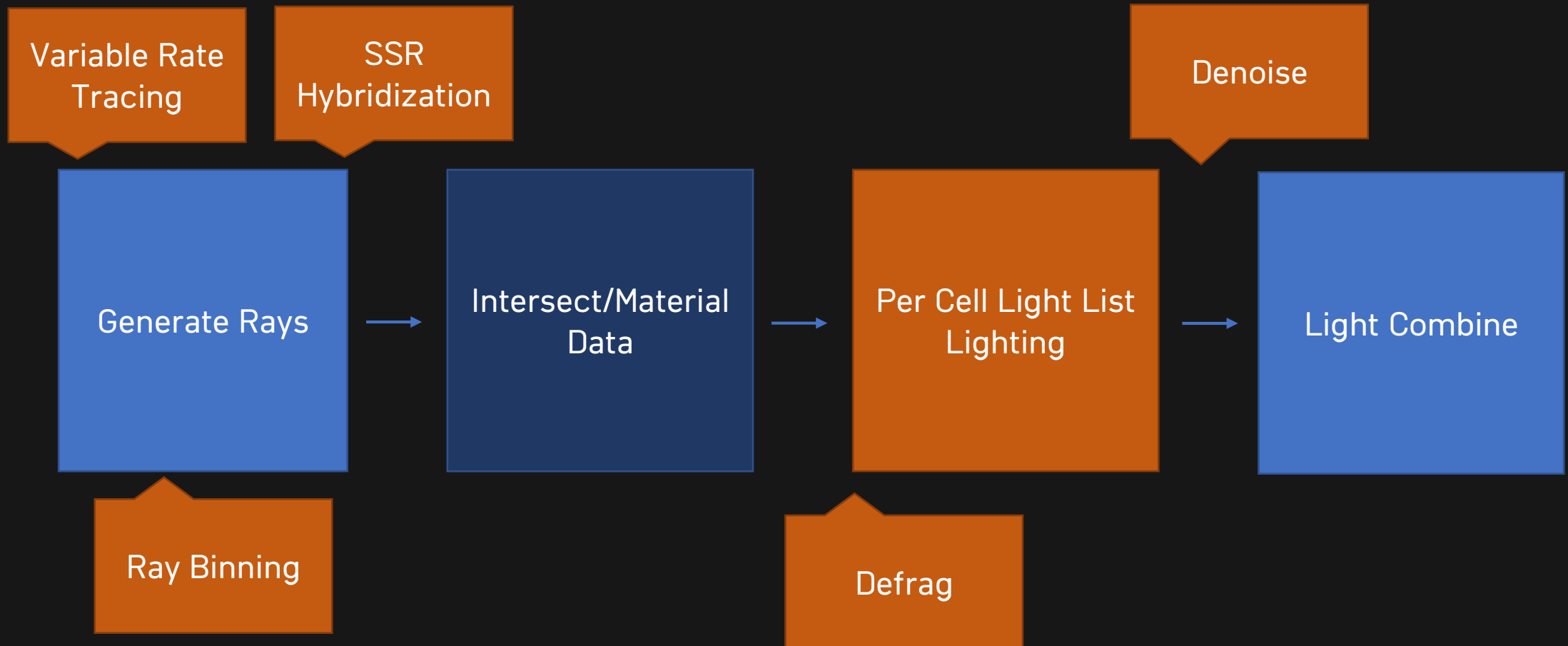
PER CELL LIGHT LISTS



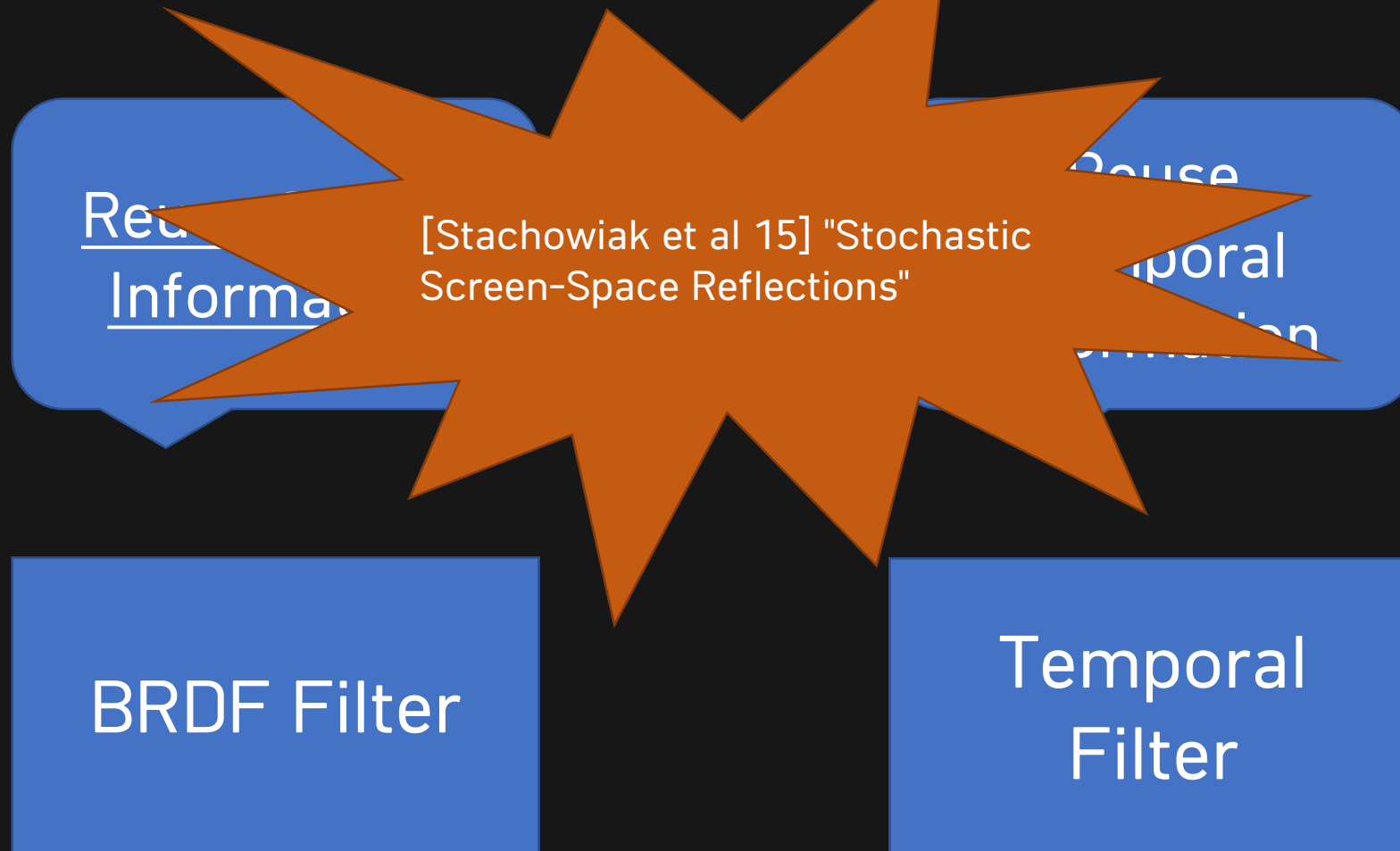
PROBLEM



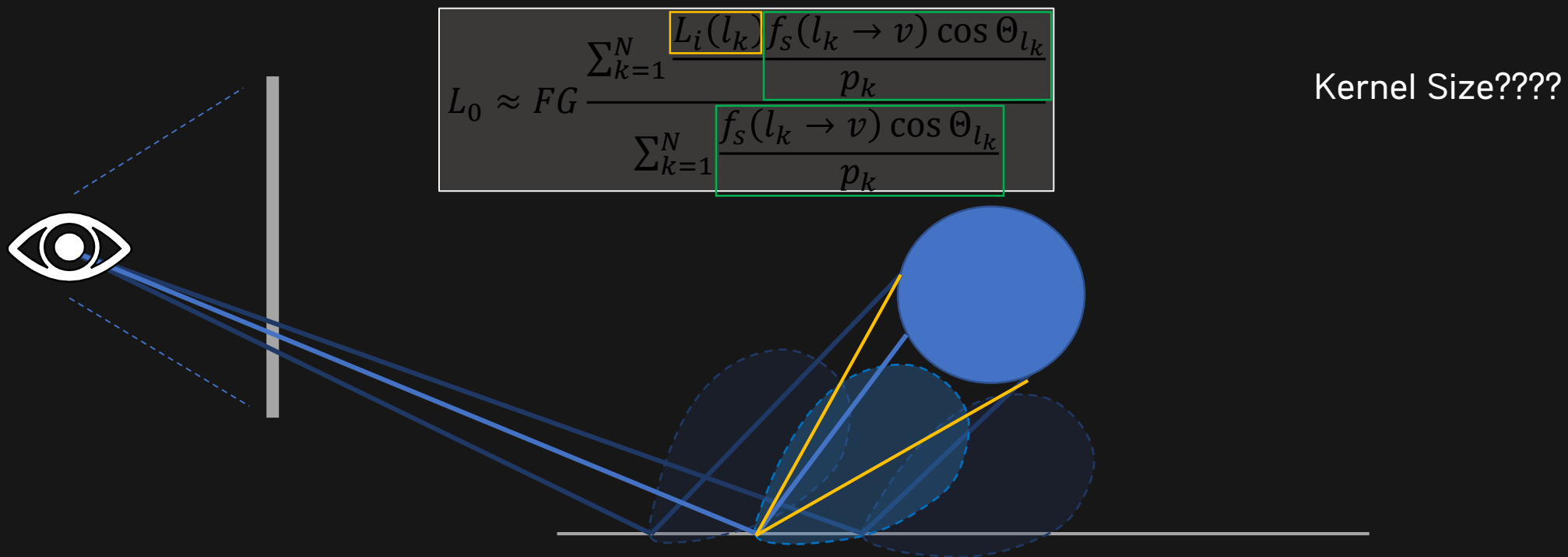
IMPROVING RAYTRACING PIPELINE



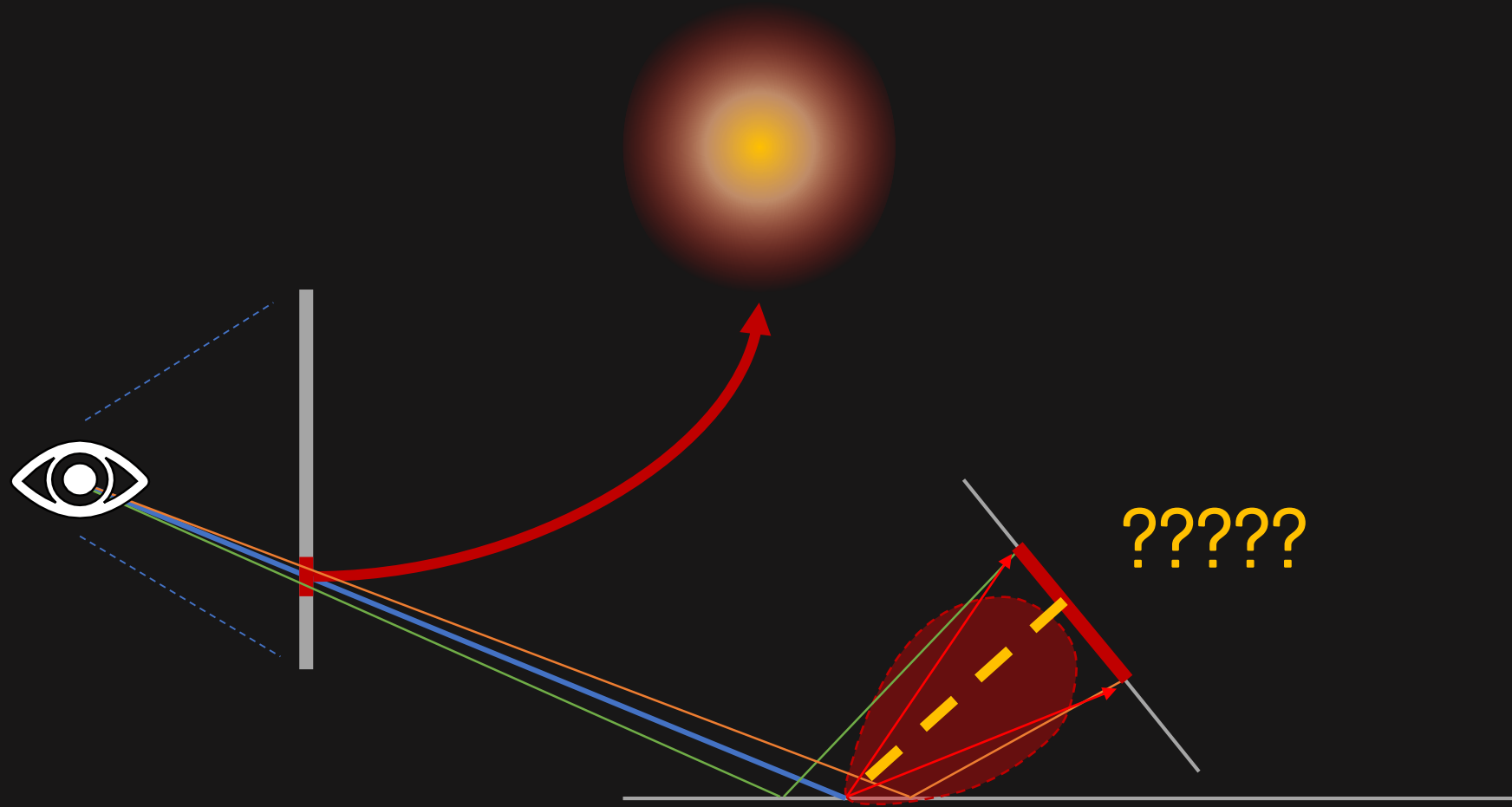
DENOISING



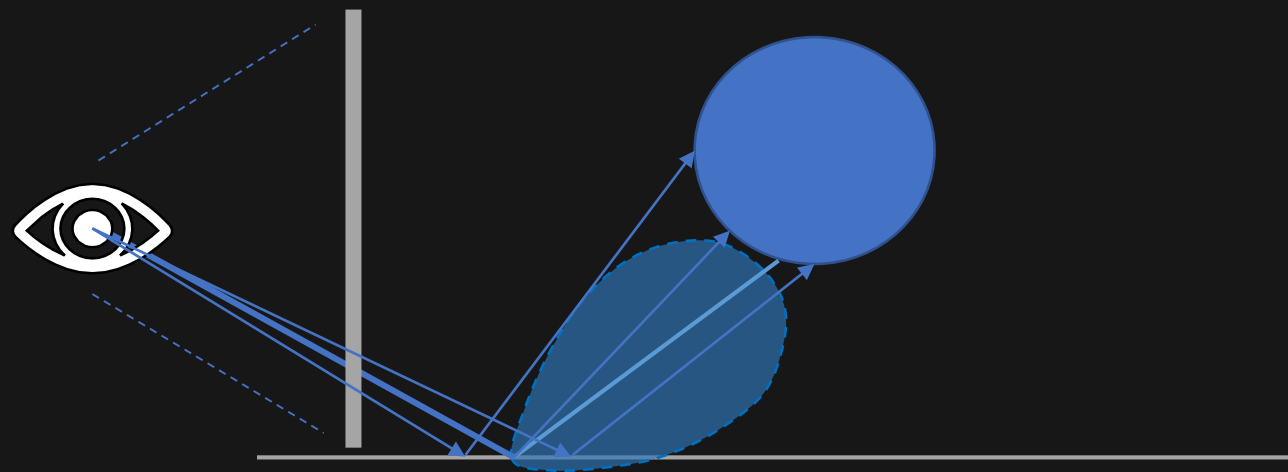
BRDF DENOISE FILTER



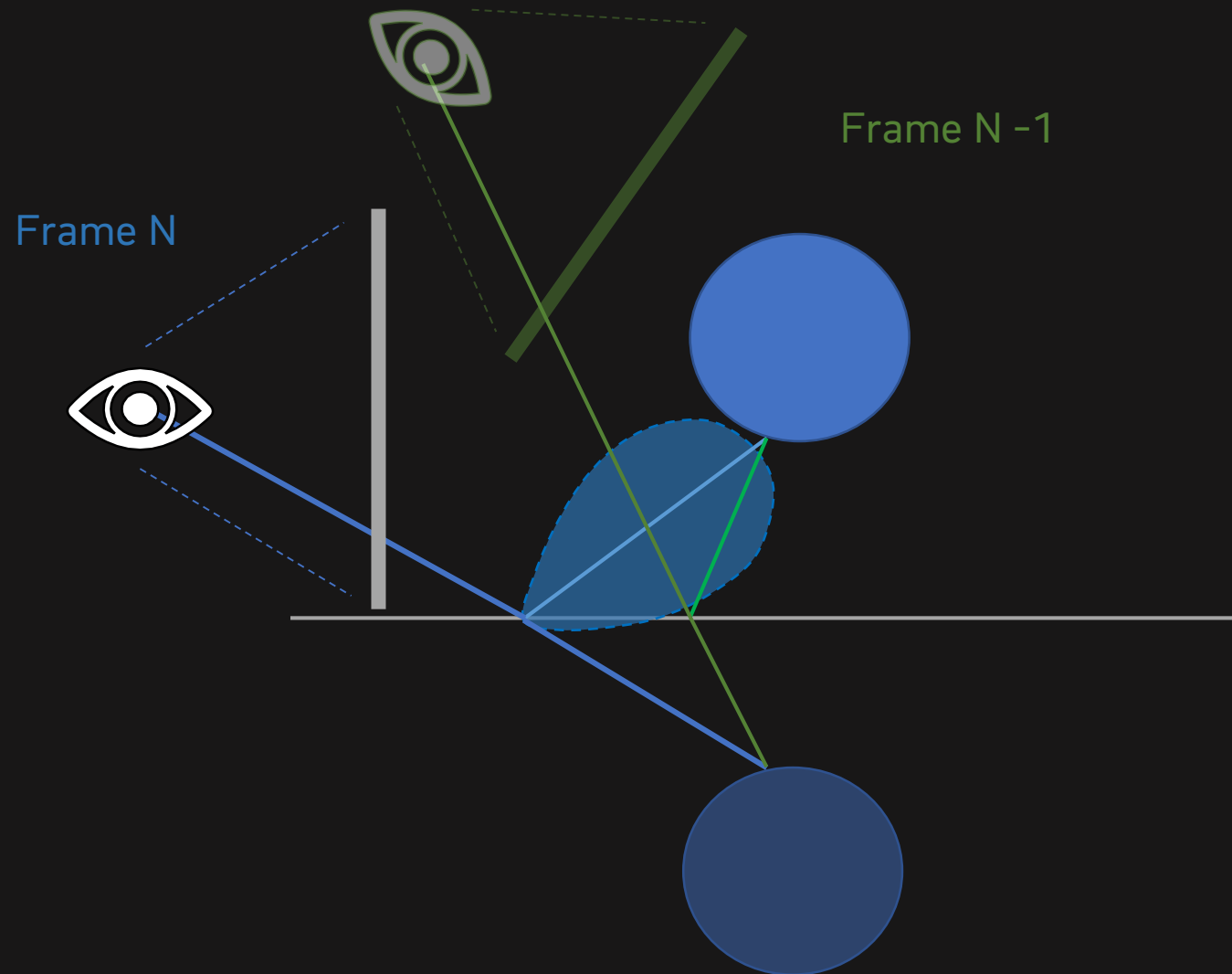
BRDF DENOISE FILTER

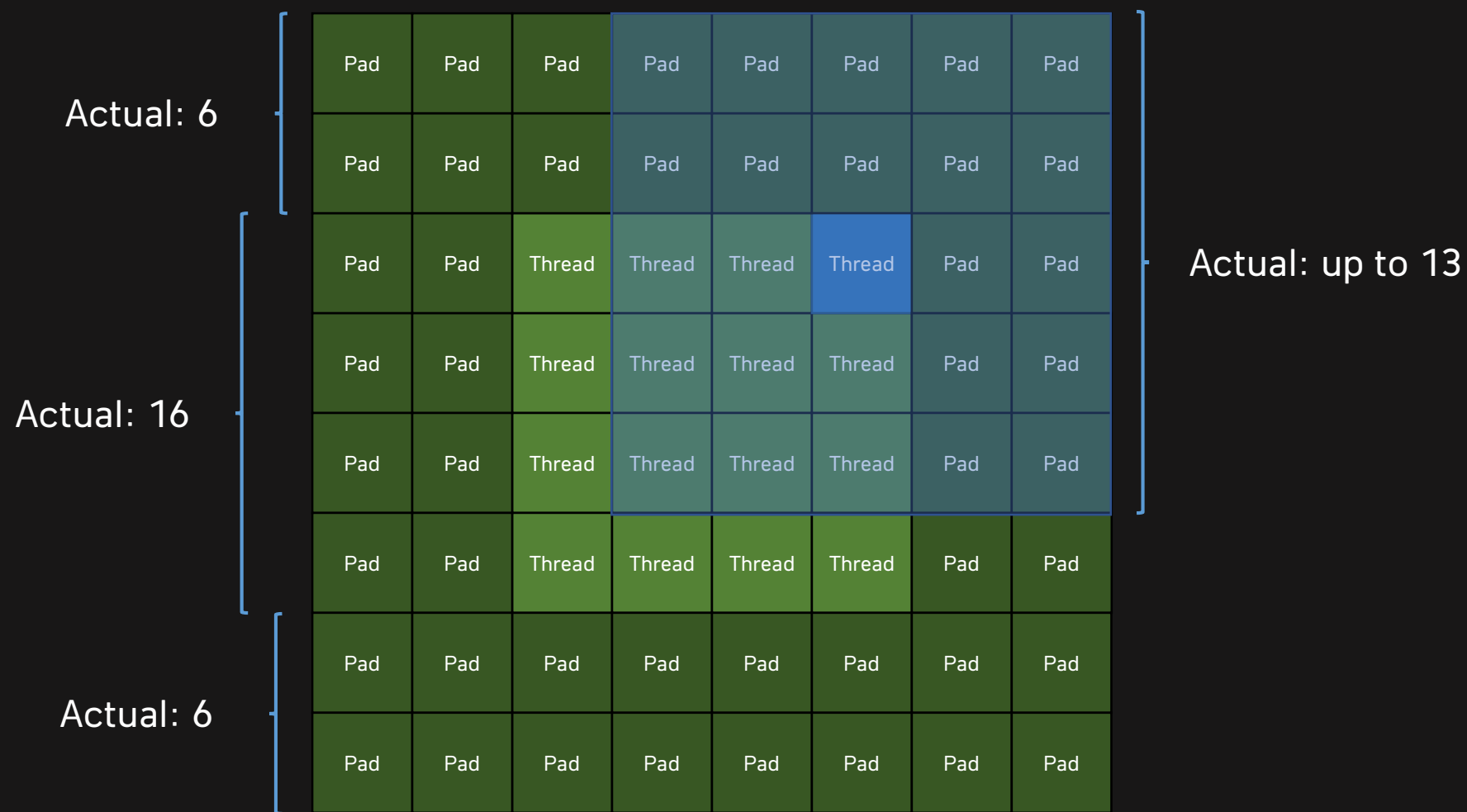


BRDF DENOISE FILTER



BRDF DENOISE FILTER





TEMPORAL DENOISE FILTER

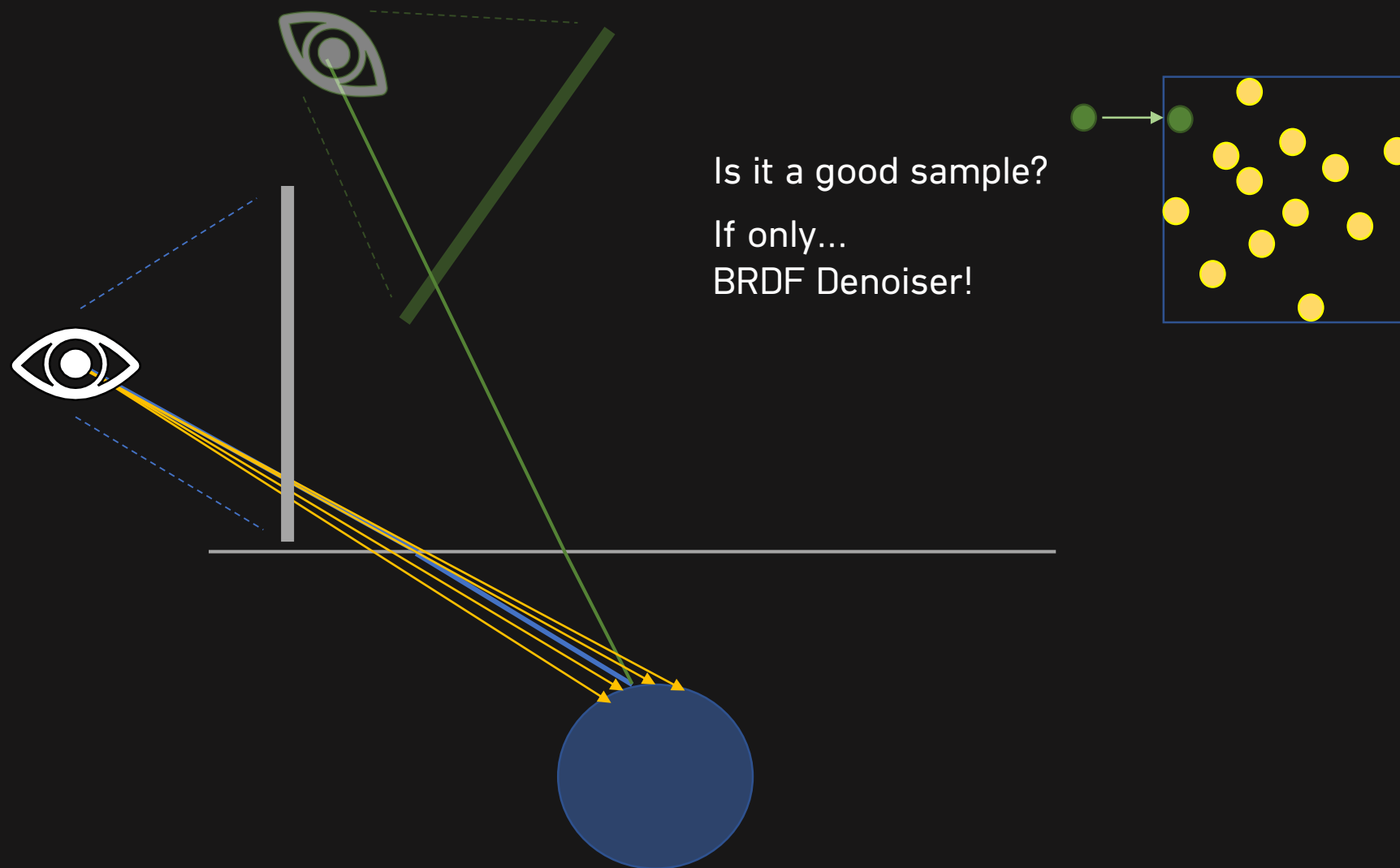




IMAGE DENOISE FILTER

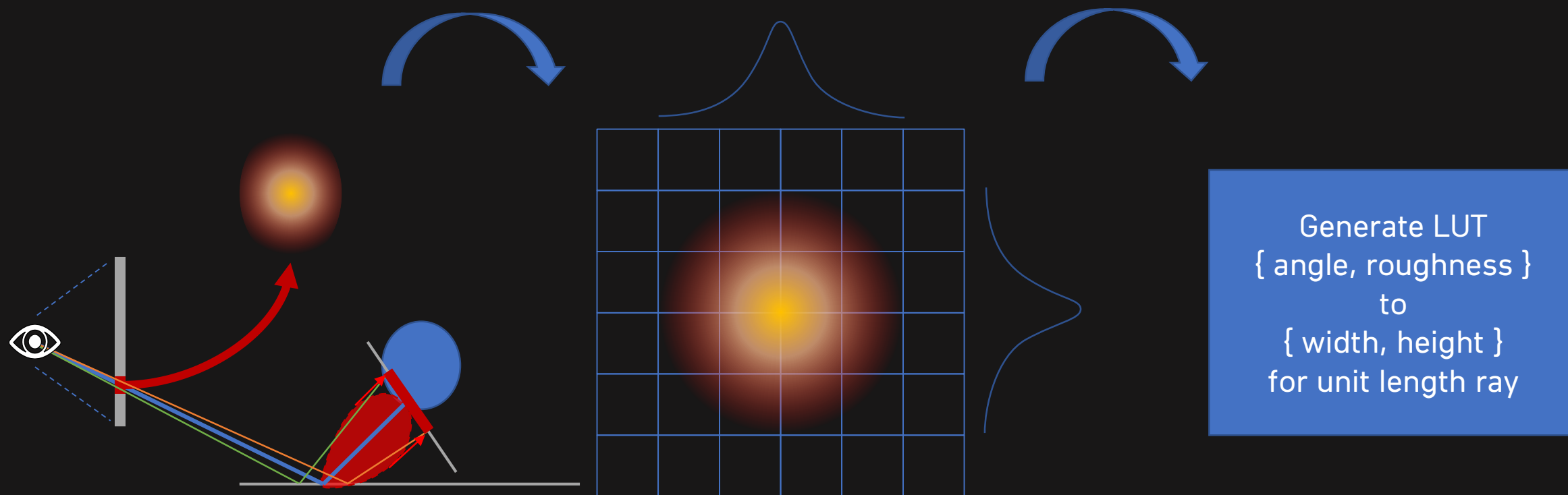


IMAGE DENOISE FILTER



IMAGE DENOISE FILTER

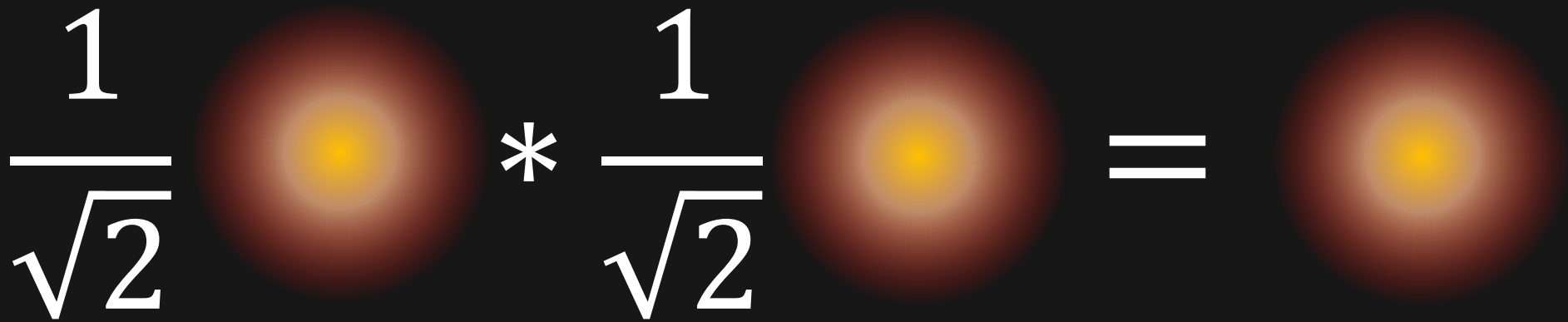
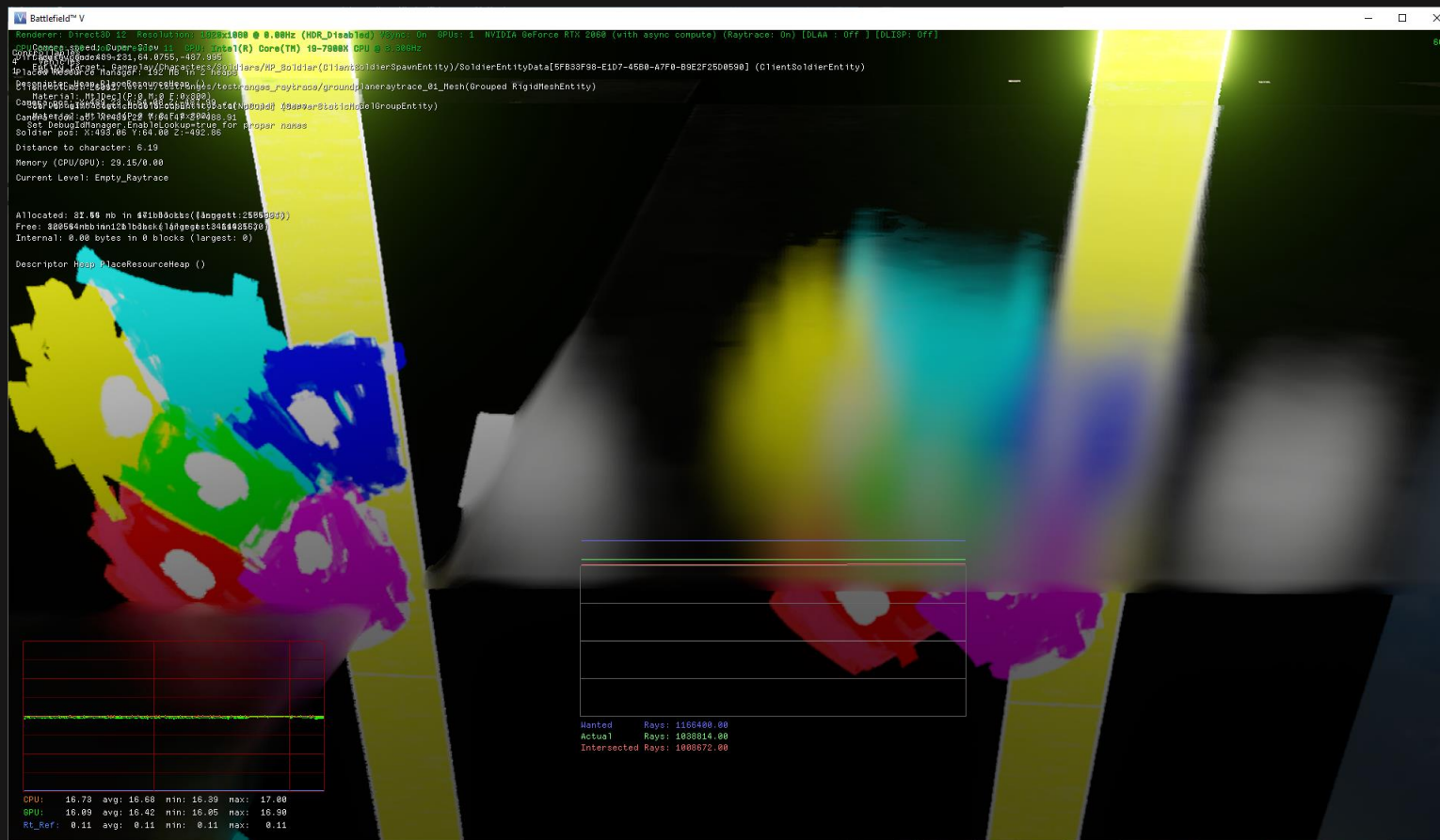
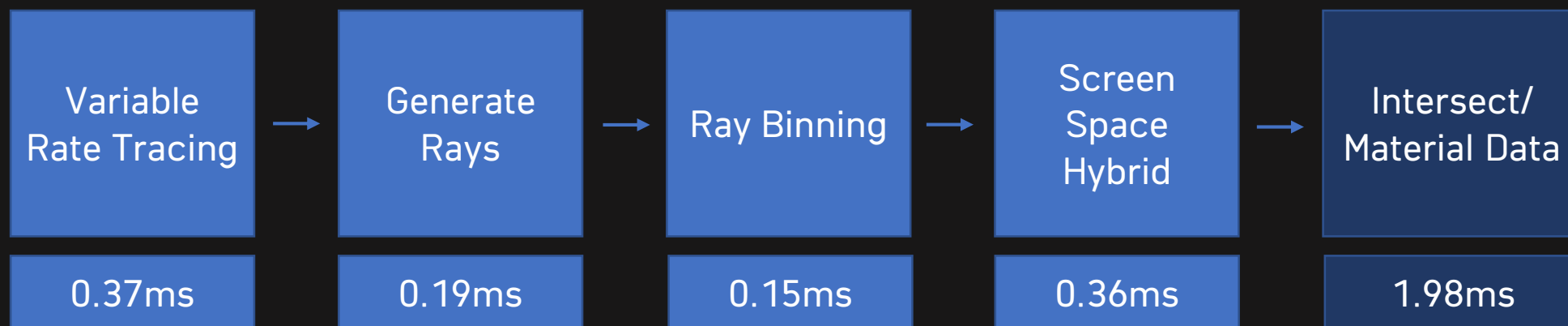
$$\frac{1}{\sqrt{2}} \text{ (Gaussian Kernel)} * \frac{1}{\sqrt{2}} \text{ (Gaussian Kernel)} = \text{Gaussian Kernel}$$
The diagram illustrates the convolution of two Gaussian kernels. On the left, the first kernel is represented by the coefficient $\frac{1}{\sqrt{2}}$ followed by a Gaussian kernel (a yellow center fading to dark red). This is followed by a multiplication symbol $*$. The second kernel is also represented by the coefficient $\frac{1}{\sqrt{2}}$ followed by an identical Gaussian kernel. This is followed by an equals sign $=$. On the right side of the equals sign is a single Gaussian kernel, which is visually identical to the two on the left, representing the result of the convolution.

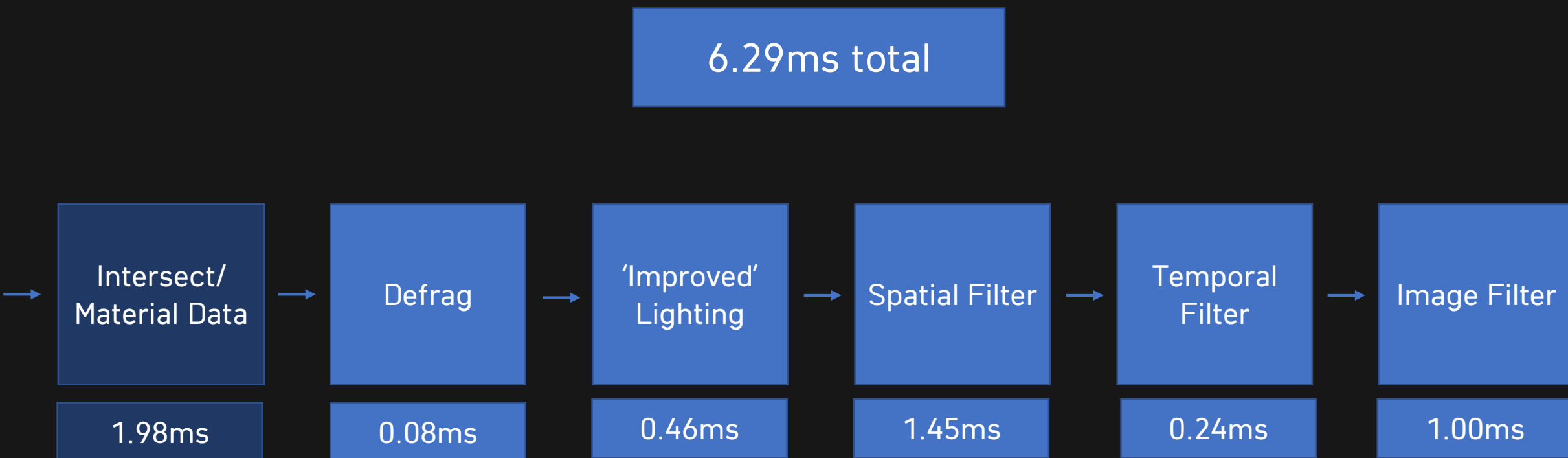
IMAGE DENOISE FILTER



NEW PIPELINE



NEW PIPELINE

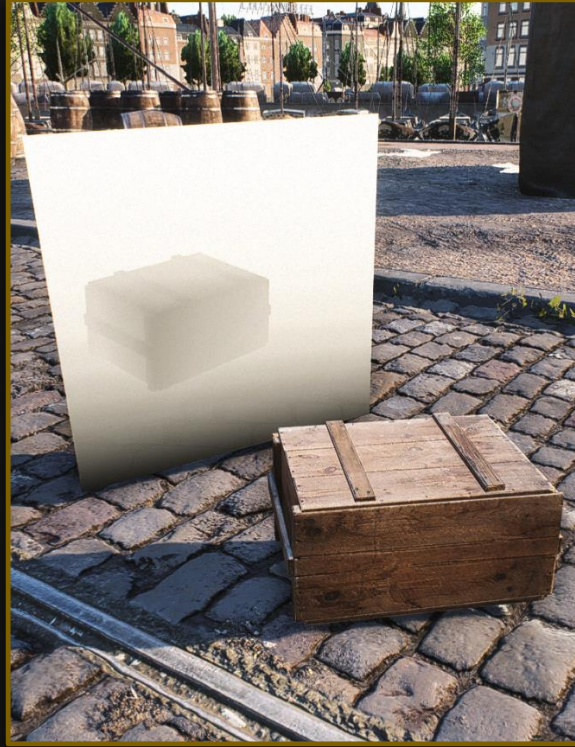




DXR – a.k.a "BLACK BOX"



No DXR



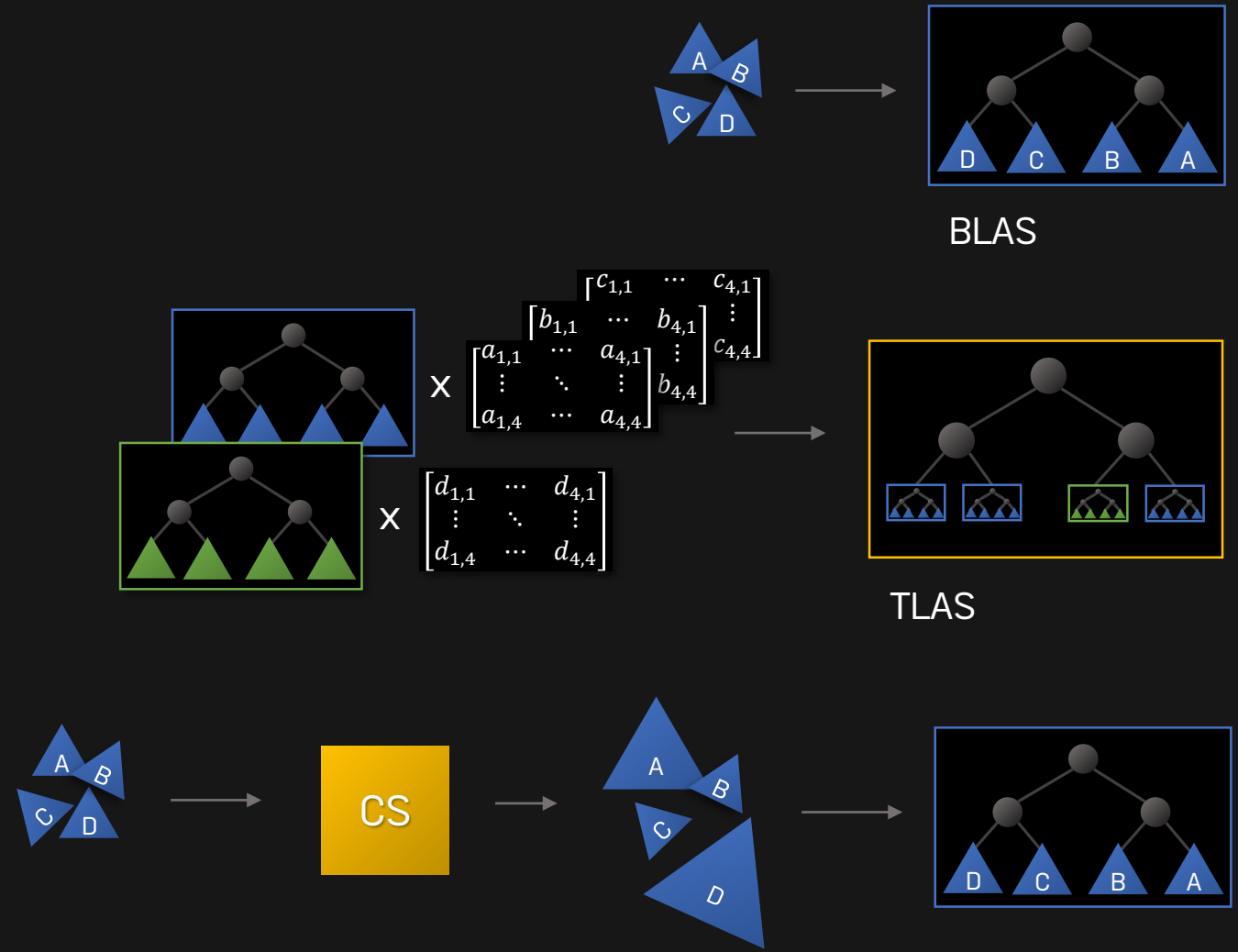
Intersection



Shading

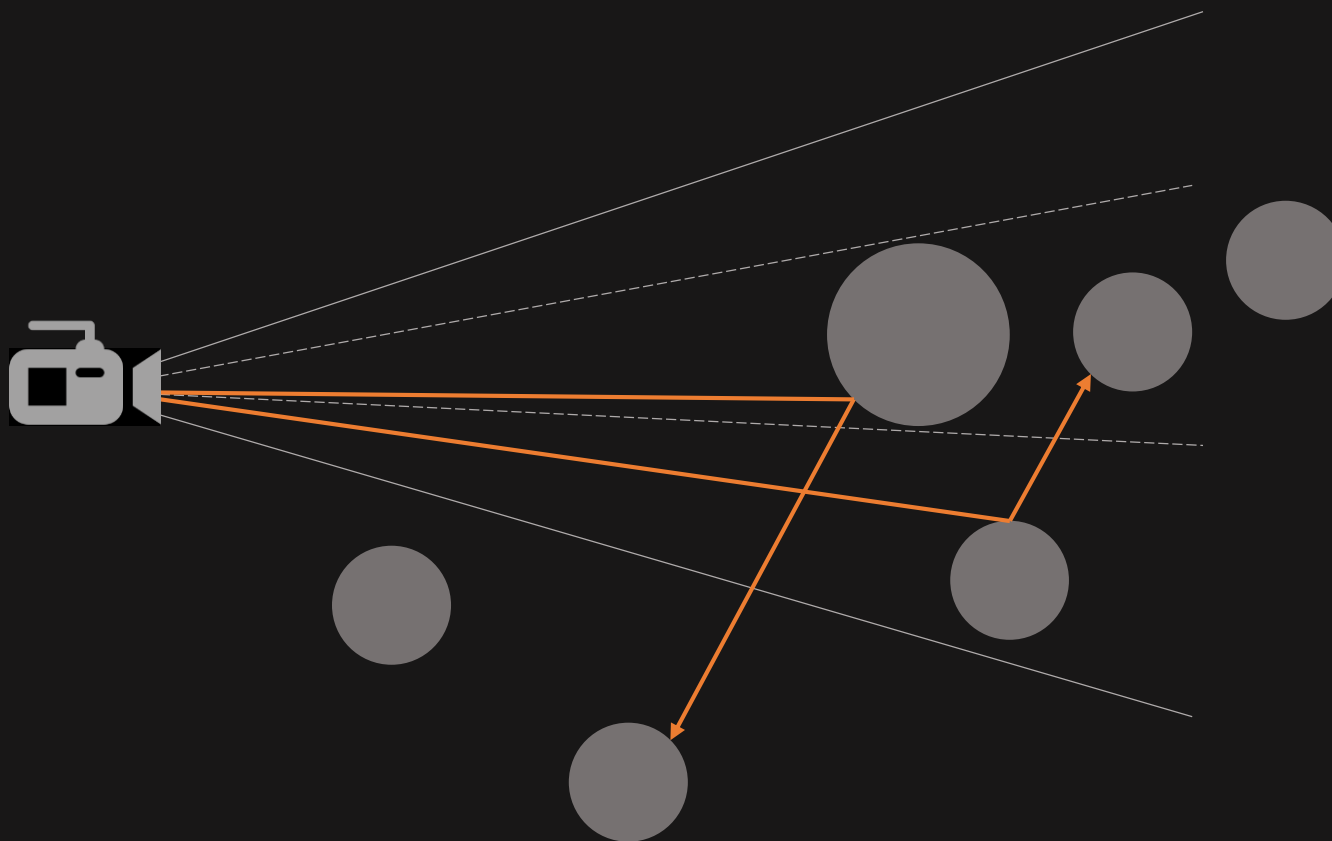
DXR BASICS

- **BLAS** - Bottom Level Acceleration Structure
- **TLAS** - Top Level Acceleration Structure
- **CS**
 - Skinning, Destruction
 - Compute shader
 - Update each frame
 - Blas can update incrementally



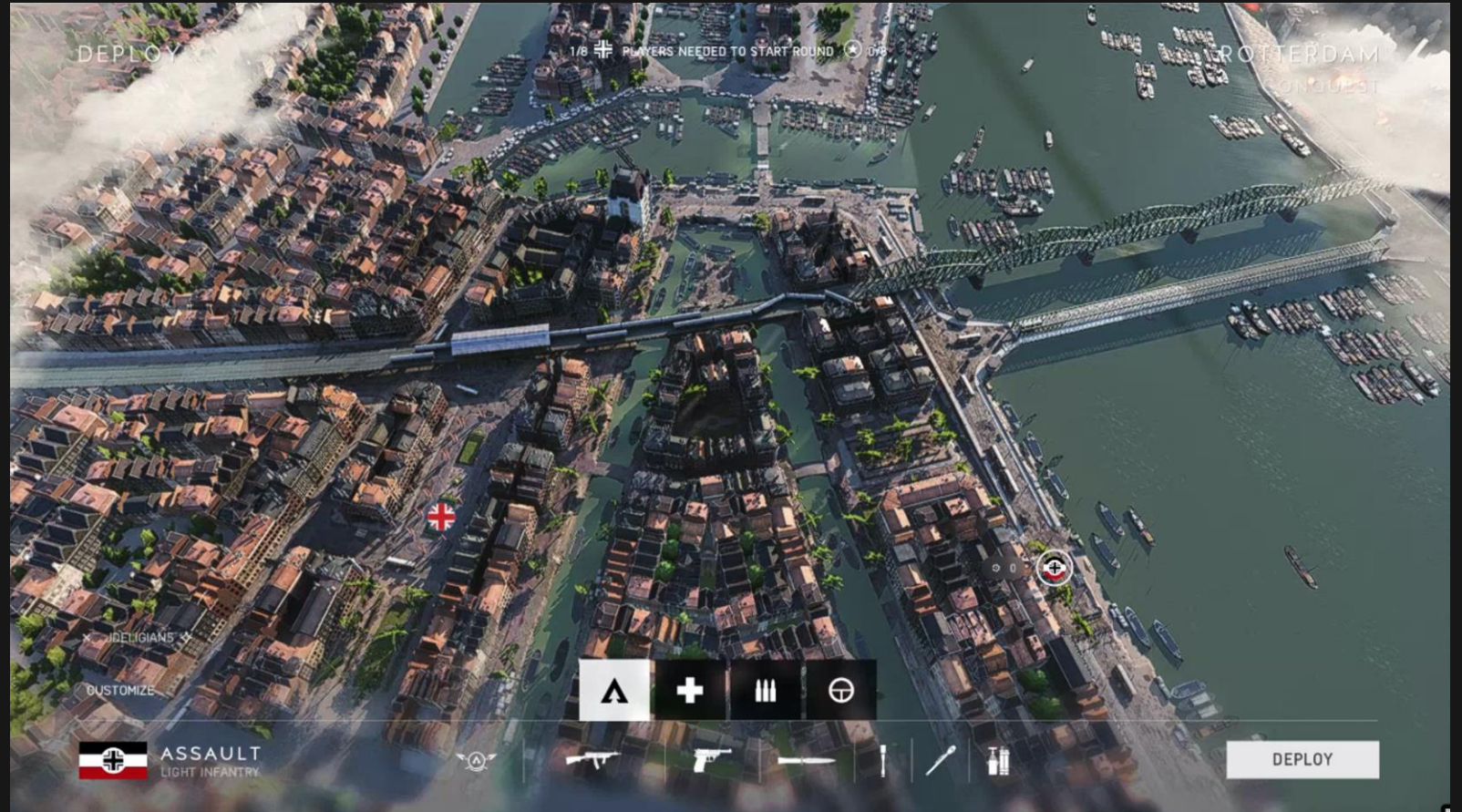
ACCELERATION STRUCTURE

- Which objects?
- **Frustum** Culling
- **Occlusion** Culling
- Easy... no culling!



ACCELERATION STRUCTURE – FIRST PASS

- Rotterdam
- 20200 TLAS instances...
- 5000 BLAS rebuilds...
- GPU rebuild **64 ms (!)**

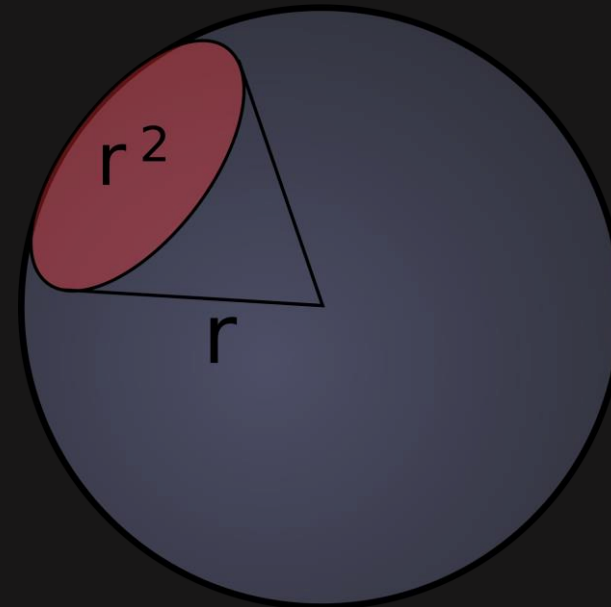


WHAT TO DO?

- Idea: Reduce instance count
- Use a culling **heuristic**
- Accept (some) minor artifacts

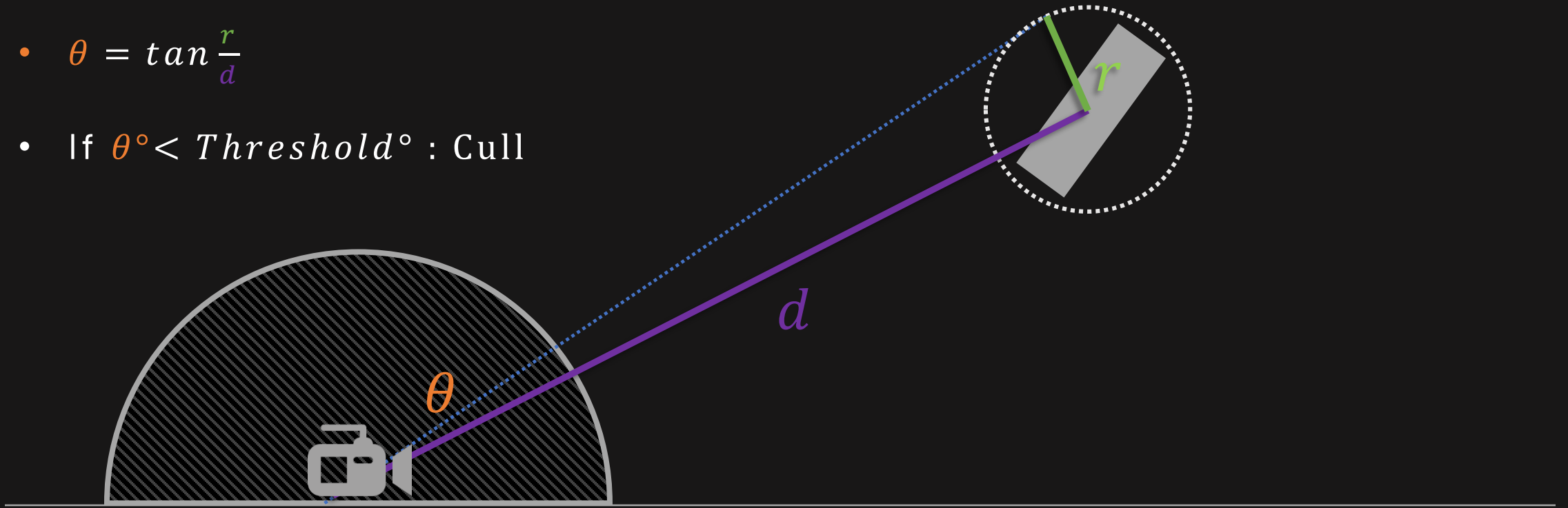
CULLING HEURISTIC

- Assumption:
 - **Far away** objects not important
 - Except for large objects
 - Bridge, building etc
- Need some kind of measurement...



CULLING

- Project bounding sphere
- $\theta = \tan^{-1} \frac{r}{d}$
- If $\theta^\circ < Threshold^\circ$: Cull



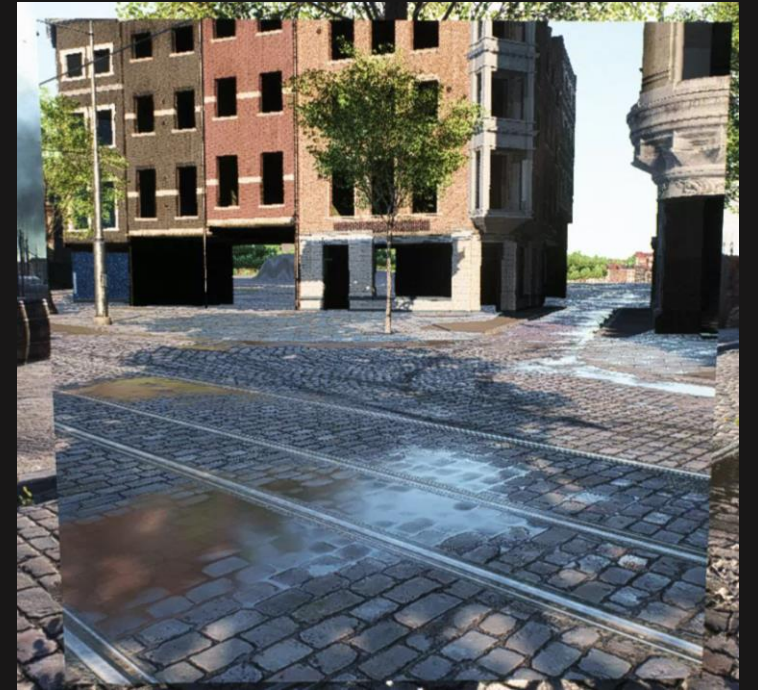
CULLING



reference – no culling



$\theta = 4^\circ$



$\theta = 15^\circ$

CULLING

Culled Objects



reference – no culling



$\theta = 4^\circ$

CULLING - RESULTS

- 4 deg culling
- 5000 -> 400 BLAS rebuilds each frame
- 20000 -> 2800 TLAS instances
- TLAS + BLAS build (GPU): 64 ms -> 14.5 ms
- Pros
 - Faster
- Cons
 - Occasional popping
 - Missing objects

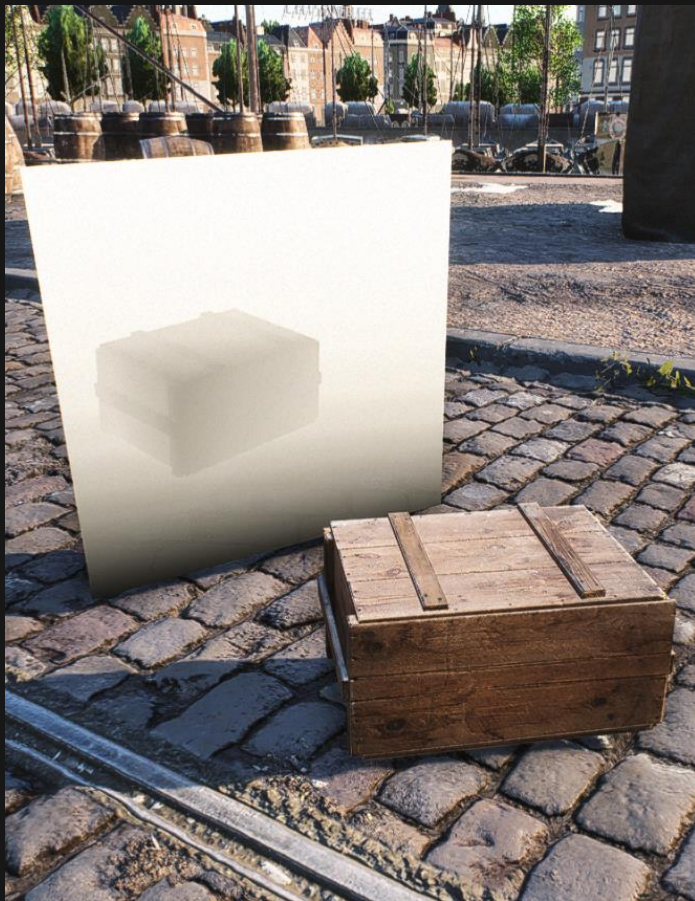
BLAS UPDATE OPTIMIZATIONS

- Still expensive! More ideas:
 1. **Stagger** full and incremental BLAS rebuild
 - N frames incremental before full rebuild
 2. D3D12_RAYTRACING_ACCELERATION_STRUCTURE_BUILD_FLAG_PREFER_FAST_BUILD
 3. Avoid **redundant** rebuilds
 - Check CS input (bone matrix)
 - 400 -> 50
- **Overlap** BLAS update with GFX
 - Gbuffer, shadowmaps

RESULTS

- TLAS + BLAS build (GPU): 14.5 ms -> 1.15 ms
- RayGen (GPU): 0.71 ms -> 0.81 ms (staggered refit + flags)
- Much better 😊

SHADING (OPAQUE)



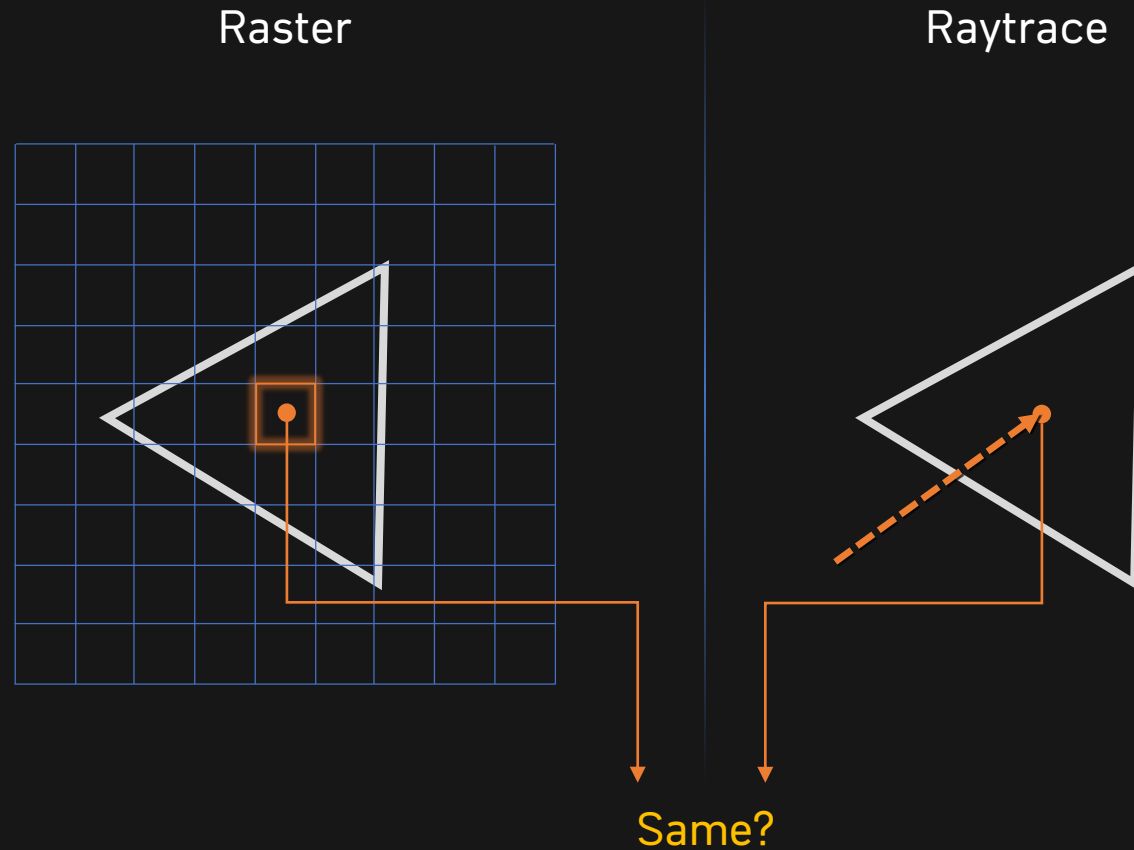
RT ON | SHADING OFF



RT ON | SHADING ON

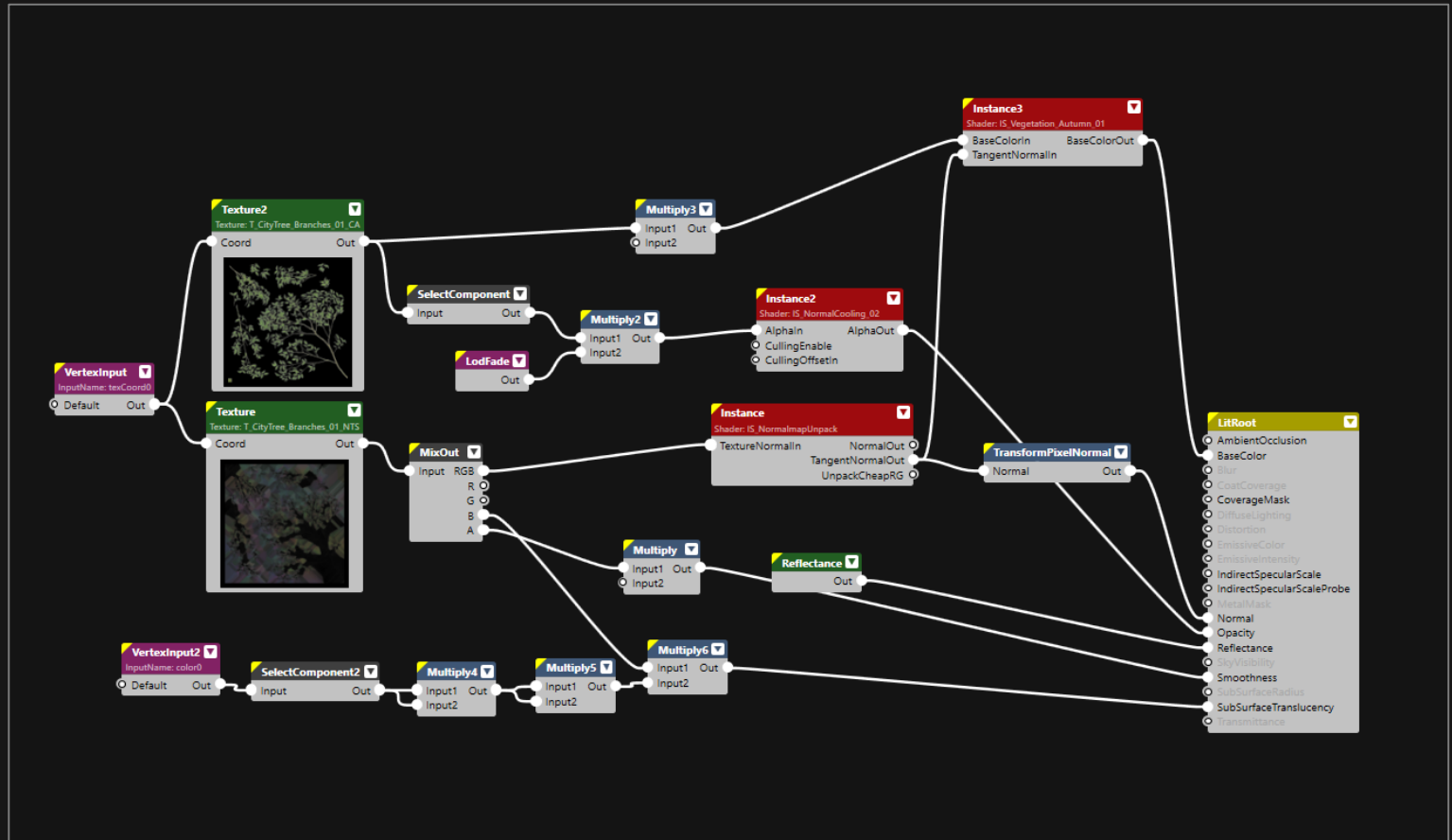
RAYTRACING REQUIREMENTS

- Shader output must **match!**
- ClosestHit Shader
- AnyHit Shader

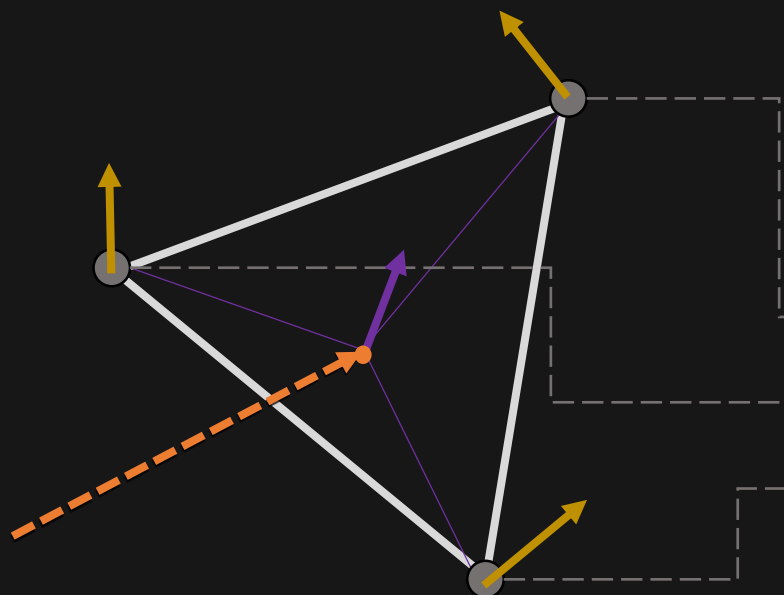


SHADERS IN FROSTBITE

- VS - Handwritten
- PS - Shader Graphs
 - Graph -> .hlsl
- Manual conversion... no
 - 1000s of shaders
- **Auto** VS + PS to HitGroup



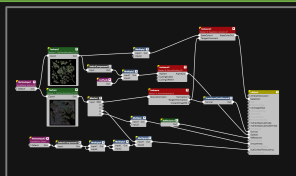
HIT SHADER TEMPLATE



VS – VertexFragment
World Space Normal

#define Sample(s, uv

PS –



```
[shader("closesthit")]
void chMain()
{
```

IA0 = iaMain(id + 0)

IA1 = iaMain(id + 1)

IA2 = iaMain(id + 2)

V0 = vsMain(IA0)

...clip?

Texture MIP?

```
#define ddx(x) x
#define ddy(x) x
... ^\_(\ツ)\_/
```

... ddx/ddy?

Vertex buffer
UV, Normal

ALPHA TESTING

- AnyHit Shader:
- If (AlphaTest(alphaValue))
 IgnoreHit();



ALPHA TEST



ANY HIT OFF



ANY HIT ON

ALPHA TEST



```
Conns: Periodic - Rotate sun around Y-axis (current speed: 0)
Camera [lock] [x]:219.66 [y]:138.82 [z]:42.80, use the mouse scroll to change sun rotation speed
Memory CPU/GPU: 35.25/2.00 Take screenshot (and optionally add it to performe
```

[illegible]

Descriptor Heap PlaceResourceHeap ()

SUMMARY OPAQUE

- **Closest Hit** Shader
 - Always
- **Any Hit** Shader (Optional)
 - Alpha tested
- **Compute** Shader (Optional)
 - Skinning, destruction etc

RAY PAYLOAD

- Payload: returned on ray intersection
- Same format as Gbuffer RTV
- Contains Material Data
 - Normal
 - Base Color
 - Smoothness
 - ...etc

```
struct GbufferPayloadPacked
{
    uint data0; // R10G10B10A2_UNORM
    uint data1; // R8G8B8A8_SRGB
    uint data2; // R8G8B8A8_UNORM
    uint data3; // R11G11B10_FLOAT
    float hitT; // Ray length
};
```

VERIFYING CORRECTNESS

1. Rasterizer output
2. Shoot primary rays in to scene
3. Compare Payload with Gbuffer
4. Non zero output? Bug!
5. Fix bug



Gbuffer (BaseColor)
Reference



Raytraced (BaseColor)
Primary Rays

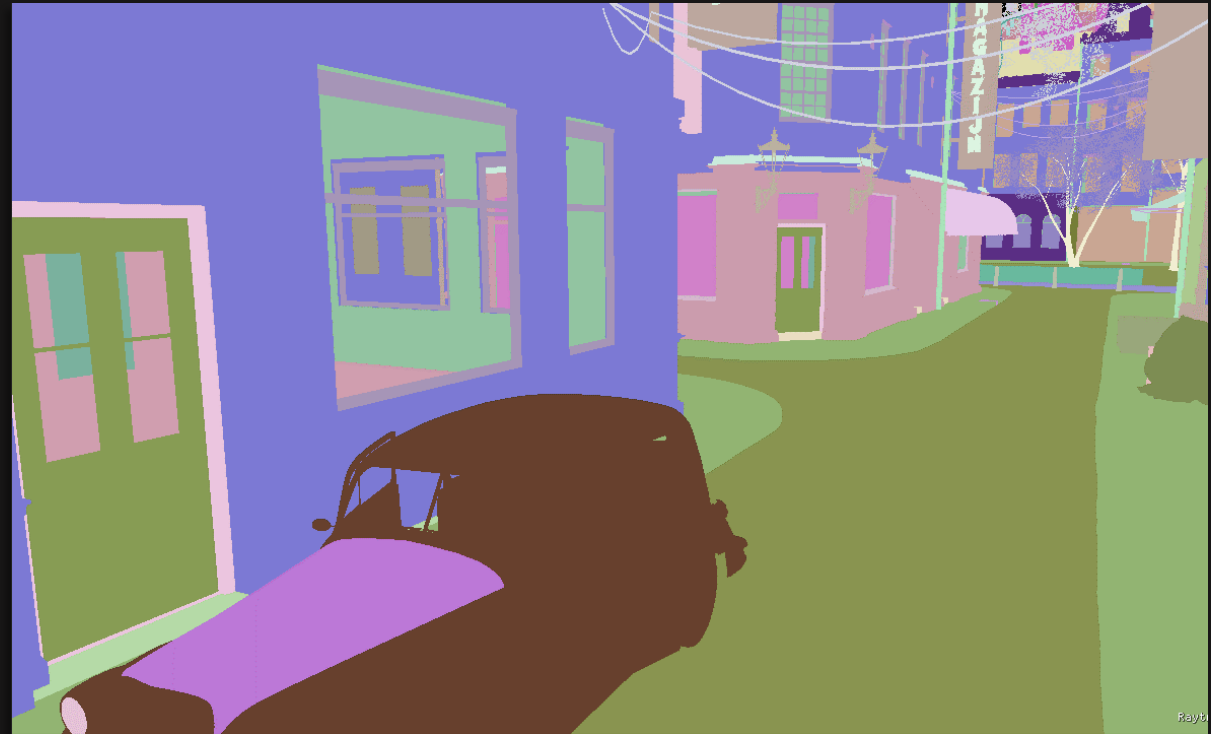
=



Delta

SHADER COMPILATION

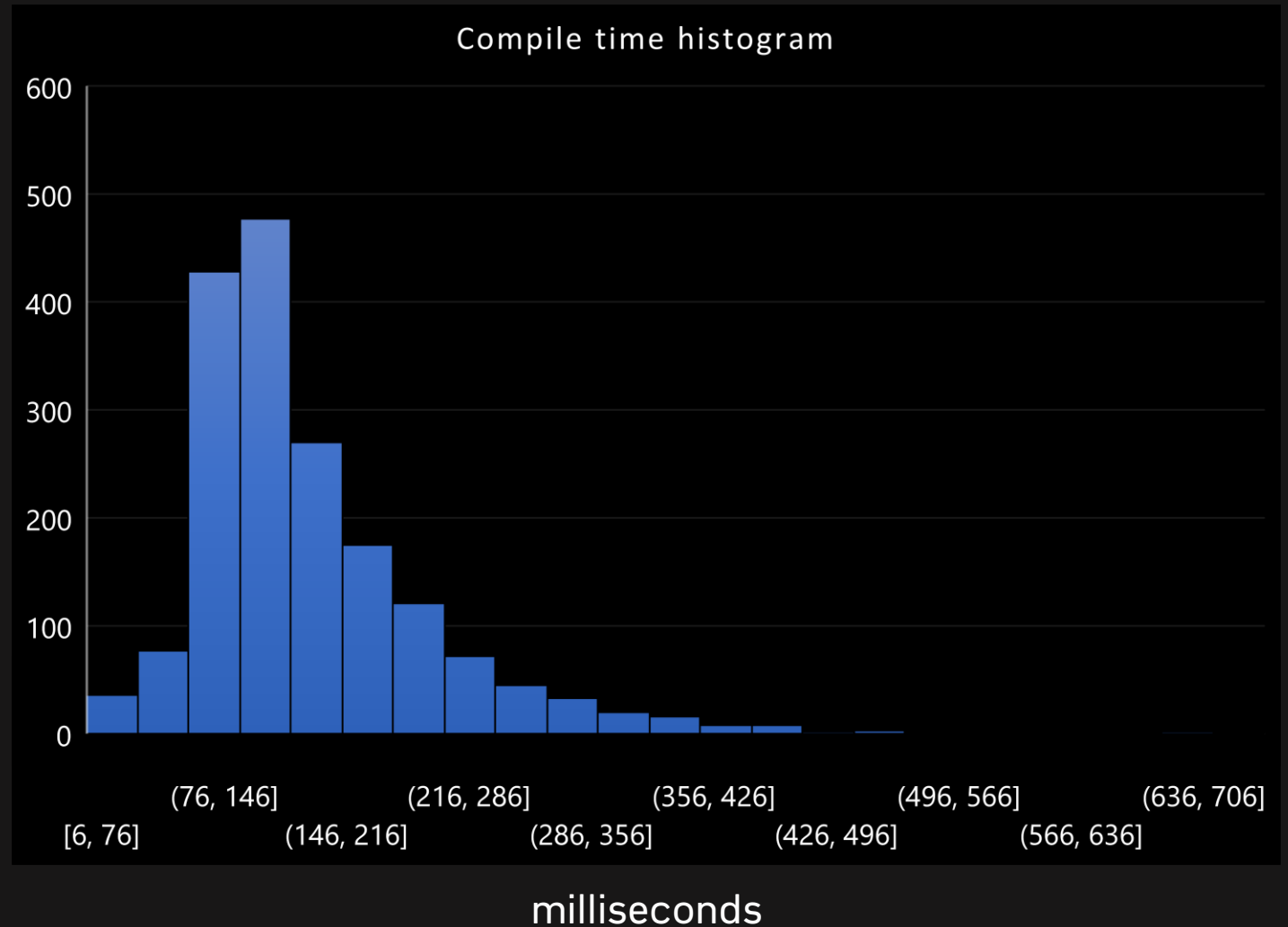
- All shaders generated 😊
- ~3000 per level
- ~250 per frame
- Single RT PSO
- Runtime compile times?



Color coded Closest Hit Shaders

PSO GENERATION

- Dx12 GFX PSO...
- ... DXR 3000 shaders ☹
- Compile times?
 - Majority > 100ms
- Cold cache
 - 7 min 30 sec thread time
 - 6 threads: 1 min 30 sec
- Warm cache
 - 1 min 30 sec thread time
 - 6 threads: 15 sec



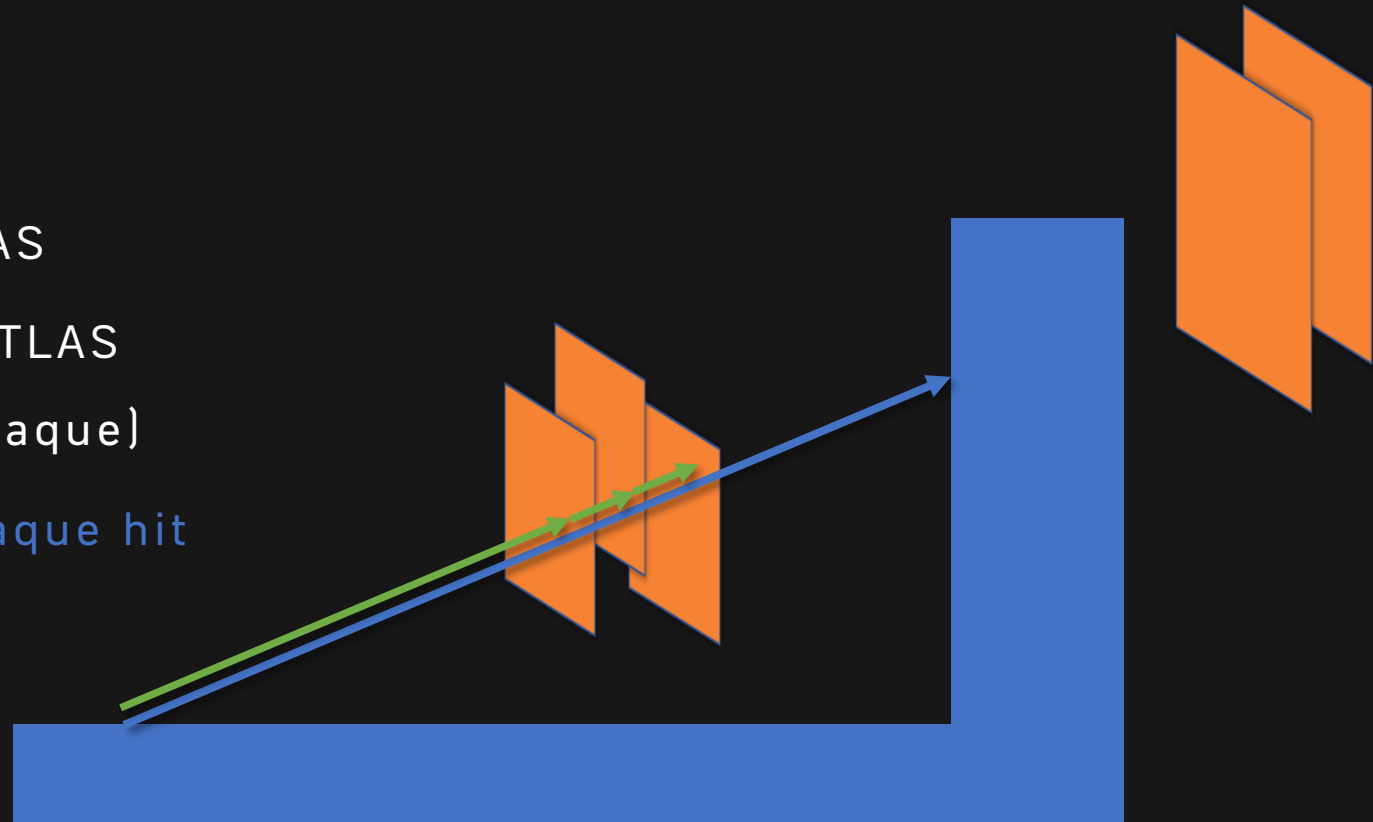
PARTICLES



Smoke, Fire and Exposions. Important elements in BFV!

PARTICLES

- Particle = Transparent+Billboard
- Basic algorithm
 1. Shoot ray in Opaque TLAS
 2. Shoot again in Particle TLAS
(Max ray length from Opaque)
 3. Blend particles with opaque hit



THE PROBLEM WITH PARTICLES

- Camera aligned billboards ☹️
- Rotate odd particles 90 deg around Y?



Billboards visible when viewed from the side.

ROTATED BILLBOARDS



Before: Billboards visible in reflection



After: Rotating odd quads produces a more volumetric look

PERFORMANCE

- Accumulate intersections along ray
- 1 rpp => N rpp 😞
- RayGen loop
- Sounds... expensive?

RayGen Shader

```
*... init ray using opaqueHitT and currT*
for (hitCount = 0; hitCount < MaxIntersectionCount; ++hitCount)
{
    ...

    ForwardPayloadPacked forwardPayloadPacked;
    initForwardPayloadPacked(forwardPayloadPacked);
    TraceRay(g_tlasPartices, 0, 0xFF, 0, 1, 0, ray, forwardPayloadPacked);
    ForwardPayload forwardPayload = unpackForwardPayload(forwardPayloadPacked);
    if (forwardPayload.hitT <= 0.0f) // Miss, tracing done
        break;

    * ... update ray using forwardPayload.hitT, accumulate color, alpha *
}
```


THE (SECOND) PROBLEM WITH PARTICLES



RayGen loop 0.96ms

OPTIMIZING PARTICLES

- Loop Idea: AnyHit shader?
- Same... but different
- Inspired by WBOIT*
 - $\text{Weight} = \max(\text{luminance}, r, g, b, \alpha)$
 - Emissive, fire

RayGen Shader

```
... init ray using maxT and currT
TraceRay(g_tlasPartices, 0, 0xFF, 0, 1, 0, ray, forwardPayloadPacked);
* ... process payload and calculate weighted average *
```

Any Hit Shader

```
struct Attributes { float2 barycentrics; };
[shader("anyhit")]
void main(inout ForwardPayloadPacked payloadPacked, in Attributes attributes)
{
    *... Calculate color, transparency *
    payloadPacked.alpha += alpha * weight;
    payloadPacked.color += color * weight;
    payloadPacked.weight += weight;
    IgnoreHit();
}
```

*Weighted Blended Order-Independent Transparency:

<http://jcgt.org/published/0002/02/09/>

PARTICLES - RESULTS



'Naive' Closest Hit **0.96ms**
Slow but accurate



Order Independent AnyHit **0.34ms**
Really fast, but slightly different look



THANK YOU!
...ANY QUESTIONS?