



Simulating massive amounts of AI agents in video games



Supervisor: Vanden Abeele Alex

Coach: Verspecht Marijn

Graduation Work 2023-2024

Digital Arts and Entertainment

Howest.be



CONTENTS

Abstr	act & k	(ey words	3
Prefa	ce		2
List o	f Figure	25	5
Introd	ductior	1	7
Litera	iture Si	tudy / Theoretical Framework	8
1.	Data	a oriented programming	8
2.	Unr	eal Engine Mass	9
3.	Spa	tial partitioning	11
4.	Mul	tithreading	12
	4.1.	Multithreading in Unreal Engine	12
5.	Ren	dering/animations	12
	5.1.	Niagara	12
	5.2.	Animation Budget Allocator	13
	5.3.	Animation Sharing	13
	5.4.	Vertex Animations	14
6.	GPL	J programming	15
	6.1.	Compute Shaders	15
	6.2.	CUDA	16
	6.3.	Opencl	16
case s	study		17
1.	Ехр	eriment setup	17
	1.1.	Introduction	17
	1.2.	Requirements	17
	1.3.	Measurements	18
2.	Ехр	eriment Execution	18
	2.1.	Simple battle Simulator	18
	2.2.	Mass battle Simulator	24
	2.3.	Rendering/LODs	34
	2.4.	Multithreading	36
	2.5.	Spatial Partitioning	37
	2.6.	Animation Sharing	38
3.	Res	ults	41
Discu	ssion		42

Schaep Simon

1.	Frame Time Over Unit Count	43
2.	Game Thread Frame Time Composition	48
3.	Summary	51
Conclusion		52
Future work		53
Critical Reflection		54
Refere	nces	55
Acknowledgements		
Appen	dices	59

ABSTRACT & KEY WORDS

This paper covers the topic of optimizing the 3D simulation of massive amounts of independent AI agents in a video game made with a modern game engine. It includes a study on different aspects of optimization, of which a significant part is dedicated to the topic of data oriented programming. It is also accompanied by a case study on how to specifically optimize a battle simulator game in Unreal Engine 5, by utilizing the Mass system. We also explain how to combine techniques such as multithreading and spatial partitioning with Mass to achieve better performance.

PREFACE

I have always been amazed by big battles in games, I would often just look at the battle In games like Planetside 2 even when I'm supposed to play in them as well. I am also a big fan of the Total War games, because they can also have huge battles.

Ultimate Epic Battle Simulator was also a big inspiration for this graduation work. In that game, you can have millions of units in a single battle, so even achieving a small fraction of that amount would be amazing.

This paper was made during over the course of a single semester, while also working on another big project. Most of the case study was made in three weeks at the end of the semester.

The scope is more limited than I would have wanted, which also means this study allows for many different branches of future work.

This paper is meant for people like me 1 year ago, with basic knowledge of C++ programming and game development, and an interest in how to simulate big battles in games.

LIST OF FIGURES

FIGURES	
Figure 1: MassEntity archetype graph - MassEntity documentation (Epic Games)	9
Figure 2: MassEntity processor graph - MassEntity documentation (Epic Games)	9
Figure 3: Screenshot of an army rendered with Niagara - Simulating Large Crowds In Niagara Unreal Engine Games)	
Figure 4: Animation sharing graph - Animation Sharing Manager documentation (Epic Games)	14
Figure 5: Screenshot of a battle with 10 million units, almost the entire battlefield is covered - UEBS 2 The Ma Of Pt1 (BrilliantGameStudios)	_
Figure 6: Simple battle simulator animation state graph	23
Figure 7: PLugins required for using Mass in Unreal Engine 5	24
Figure 8: Example of a Mass spawner configuration	25
Figure 9: Example of an EntityConfig data asset configuration	26
Figure 10: Visualization trait configuration	33
Figure 11: Draw call amount with old model Figure 12: Draw call amount with new model	35
Figure 13: LOD Settings after generating an LODSettings asset	35
Figure 14: Visualization of the octree, the blue boxes represent the nodes of the tree	37
Figure 15: Finding the closest target in the octree	38
Figure 16: Screenshot from within Unreal Insights to show the amount of frame time the animations take	39
Figure 17: Creation of the animation sharing manager	39
Figure 18: Configuration of the AnimationSharingSetup asset	40
Figure 19: Registering and unregistering the actor with the animation sharing manager	40
Figure 20: Screenshot from within Unreal Insights showing how component ticks are not grouped	41
CODE SNIPPETS	
Code snippet 1: Units variable declaration	19
Code snippet 2: Unit arrays added to the Units variable	19
Code snippet 3: Target acquisition component header file	20
Code snippet 4: Move component header file	20
Code snippet 5: Health component header file	21
Code snippet 6: Attack component header file	22
Code spinnet 7: Example of a processor header file	24

Code snippet 8: FArmyIdFragment definition	26
Code snippet 9: Assigning the correct initial values to newly spawned entities in the initializer processor	27
Code snippet 10: PossibleTargetEntities variable declaration	27
Code snippet 11: TargetAcquisitionProcessor Execute functionality	28
Code snippet 12: Navigation mesh pathfinding in the navigation processor Execute function	29
Code snippet 13: MoveProcessor functionality in the Execute function	30
Code snippet 14: Attack parameters shared fragment definition	30
Code snippet 15: Part of the attack processor functionality	31
Code snippet 16: Part of the dying processor functionality in the Execute function	32
Code snippet 17: DestroyEntitiesProcessor Execute function	32
Code snippet 18: Transform translator execute function	33
Code snippet 19: Animation instance update functionality	34
Code snippet 20: Replacing the normal for loop with a ParallelFor	36
Code snippet 21: Using a mutex to prevent simultaneous writing to the deferred command buffer	36
CHARTS	
Chart 1: Game thread frame time over unit count	43
Chart 2: Render thread frame time over unit count	44
Chart 3: GPU frame time over unit count	44
Chart 4: Target acquisition frame time over unit count	45
Chart 5: Navigation frame time over unit count	46
Chart 6: Movement and avoidance frame time over unit count	46
Chart 7: Mass to actor frame time over unit count	47
Chart 8: Animation frame time over unit count	48
Chart 9: Game thread frame time distribution in the simple battle simulator with 800 vs 800 units	48
Chart 10: Game thread frame time distribution in the mass battle simulator with 800 vs 800 units	49
Chart 11: Game thread frame time distribution in the mass battle simulator with multithreading with 800 vs 8 units	
Chart 12: Game thread frame time distribution in the mass battle simulator with spatial partitioning with 800 units	
Chart 13: Game thread frame time distribution in the mass battle simulator with animation sharing with 800 units	

INTRODUCTION

How often do you encounter huge battles with thousands of units in a video game? Not very often is the answer. This is because it is not simply a matter of increasing the amount of units, it is mostly about having the battle still run at a playable framerate.

With this research, we try to figure out how to optimize big simulations of agents in video games. We find the most used techniques and implement some in our own battle simulator game, which is then compared to a standard battle simulator without optimizations, to see which techniques had which impact on performance.

The research assumes we are optimizing a battle simulator, but most optimization techniques can be applied to any game that has a large amount of independent agents in a simulation.

The main technique that will be used is data oriented programming and how to use other techniques together with DOP to achieve better performance.

The research question that we will try to answer in this research paper is the following:

How to simulate massive amounts of independent AI agents in real-time, in a 3D game made with a modern game engine?

We will answer the research question by testing the following hypotheses:

By utilizing Unreal Engine's Mass system, combined with other techniques, we can simulate more than 20 000 agents fighting each other at 30 fps.

- Using Unreal Engine's Mass system will improve overall performance.
- Multithreading certain processes within Unreal Engine's Mass system will improve overall performance.
- Using octree spatial partitioning will reduce the frame time needed for target acquisition.
- Using animation sharing will reduce animation frame time.

LITERATURE STUDY / THEORETICAL FRAMEWORK

The optimization of the simulation of large amounts of agents in a video game is a complex topic since there are many factors relating to performance. There are many different ways that you can optimize depending on the needs of your game. Some are more complicated and difficult to implement than others.

We will do our best to explain to most commonly applied methods and provide sources if you want to delve deeper into a topic.

1. DATA ORIENTED PROGRAMMING

Data oriented programming (sometimes called Data oriented design) is a programming approach where data is separated from code, with the main goal of making more efficient use of the cpu cache, by aligning data to avoid cache misses.

In a game engine, data oriented programming is usually applied by using ECS (Entity Component System). This is an approach where there are entities, represented by a unique ID, that are associated with components. Components mark an entity to have certain functionality, and hold the data needed for that functionality. Systems operate on all entities that have certain components, and manipulate data accordingly.

An example of this is a movement system, which evaluates all entities that have a position and movement component. The position component just holds a position, and the movement component holds a movespeed and direction. The system then goes over all position components and updates their position based on the direction and movespeed of the movement components with the same entity ID.

[1]–[5]

2. UNREAL ENGINE MASS

Unreal Engine has implemented their own ECS system, called Mass, with the introduction of Unreal Engine 5.0. This system uses the same concepts as ECS, but has named things a bit different. Entities are still called entities, components are called fragments, and systems are called processors.

There are some other concepts that they added, like traits, which are used to define what fragments an entity has. For example a movement trait might add a transform fragment and a movement fragment. They allow for easier configuration of entities.

Another concept they added are archetypes, which group entities with the same fragments together which are then evaluated in processors as separate chunks.

They also added tags, that are fragments without data, used to identify entities easier. An entity might have the "dead" tag to indicate it is dead. The movement processor will then not operate on archetypes with the "dead" tag.

Mass also includes lots of in-built fragments and tags that can help you set up basic functionality very quickly. It includes things like steering and avoidance, and also visualization. It is also made to work well with zonegraphs, which can be used to direct crowds. This was used extensively in the City Sample demo project, released with Unreal Engine 5.0.

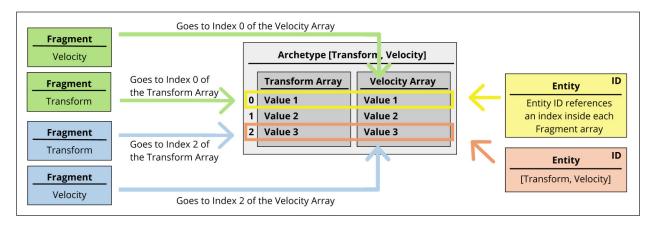


Figure 1: MassEntity archetype graph - MassEntity documentation (Epic Games)[6]

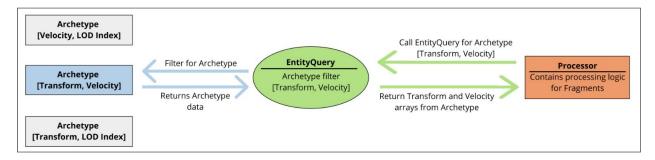


Figure 2: MassEntity processor graph - MassEntity documentation (Epic Games)[6]

Schaep Simon

The in-built visualization for Mass is quite complex and will work well enough for most projects. There are two different representation methods: instanced static meshes, and actors.

Using instanced static meshes is the most optimized as it will create an instance of a specified static mesh at the location of every entity.

Using actors is heavier on performance since there will be an actor spawned for every entity and there is an extra step in communicating from Mass to the actor. However, actors allow for much more options in how you can render. The most obvious way of using an actor is to add a skeletal mesh to the actor to allow bone animations to be played. But you could also add logic to the actor to for example change material when you are hit.

You can even take it a step further and put gameplay logic on the actor and communicate back to Mass to inform the entity of any changes that have to be made.

[6]–[23]

3. SPATIAL PARTITIONING

In almost any type of simulation with agents, you will at some point need to query for other nearby agents. A good example is if you want agents to avoid each other. Using the simplest approach, you could go over all agents, and compare their distances. Doing this for one agent would be fine, but doing this for every agent will result in an algorithm of $O(n^2)$. With 10 agents, that would be 100 distance calculations, but with 20 agents, it would already be 400 calculations. This can result in huge performance issues when there are large numbers of agents in the simulation.

Luckily, there are spatial partitioning data structures that can help to optimize this process, since they organize their data based on positions, which in turn allows us to for example find the closest position, without calculating all distances. There are many different types of spatial partitioning structures, but they are usually either flat or hierarchical.

Flat structures are the easiest to understand and implement. They just divide the world in an array of cells and every position in the data structure is put in the according cell. If you then want to find other nearby positions, you just have to look through that cell and the neighboring ones. If the positions are moving, you will of course have a certain cost for updating which cell each position is in.

Hierarchical structures are more complex since they also take the density of the positions into account. There will for example be more cells in areas where there are more positions. This is usually more optimal when the positions aren't distributed uniformly and there are lots of empty spaces. The structure is more complicated to construct and update, but allows for faster evaluation of the data.

Spatial partitioning is a requirement for any type of simulation with many agents, and therefore is an important aspect in a battle simulator game.

[24]-[27]

4. MULTITHREADING

One of the more obvious ways of optimization is utilizing multiple threads to divide the work that one thread would otherwise have to do. Usually it makes the most sense to only multithread calculations that take the most time, since multithreading can be the cause of bugs and other issues if implemented poorly.

In the case of a battle simulator, most likely it will be the navigation, avoidance and target finding that will benefit the most from being multithreaded.

4.1. MULTITHREADING IN UNREAL ENGINE

Multithreading in Unreal Engine is very similar to standard C++, with Unreal having their own implementations for things like parallel for and mutex.

The engine runs of few processes on a different thread by default, such as rendering. But most game logic such as the game tick, runs on a single thread, so there is a lot of room for optimization there.

An important thing to be aware of is that some functionality is required to run on the game thread, such as changing the location of actors.

[28]–[30]

5. RENDERING AND ANIMATIONS

When working in a 3D game engine, rendering can be a potential bottleneck, especially when rendering a large amount of separate objects.

Along with rendering, animations can also take a lot of CPU frame time. Especially since skeletal animations are most commonly used, and require the CPU to evaluate the bones every frame.

The research presented here was mostly focused on how to optimize rendering and animations in Unreal Engine 5 specifically.

5.1. NIAGARA

Niagara can be used to render more than just visual effects. Using smart particles, you can render particles as if they were actors, and even add basic gameplay functionality to them. [31]



Figure 3: Screenshot of an army rendered with Niagara - Simulating Large Crowds In Niagara | Unreal Engine (Epic Games) [31]

5.2. ANIMATION BUDGET ALLOCATOR

Unreal Engine has a multiple systems to help optimize animations, one of which is the animation budget allocator. This system will dynamically lower tick rate of skeletal meshes that are further away from the camera, depending on the total budget it has. This means you can easily hard-limit the CPU time that is spent on updating animations. It is very easy to implement as it only involves switching skeletal meshes with budgeted skeletal meshes, and setting up the budget you want to give.

An important negative is that if you would exceed the budget too much, you might notice lower frame rates even on meshes that are closer.

This system is perfect for when you have more than average, but still relatively small amounts of skeletal meshes.

[32]

5.3. ANIMATION SHARING

Animation sharing is another system Unreal Engine has available to help optimize skeletal meshes.

Instead of each skeletal mesh evaluating their own animation individually, they can share it with each other. This works by evaluating one animation instance for every state required, and then each skeletal mesh component requests animation data from the animation sharing manager, which is also based on which state the mesh owner is in.

This system requires a bit of work to set up correctly, especially if you want it to seem as if each mesh has their own animation or you want to use animations that trigger on certain events, like attacking.

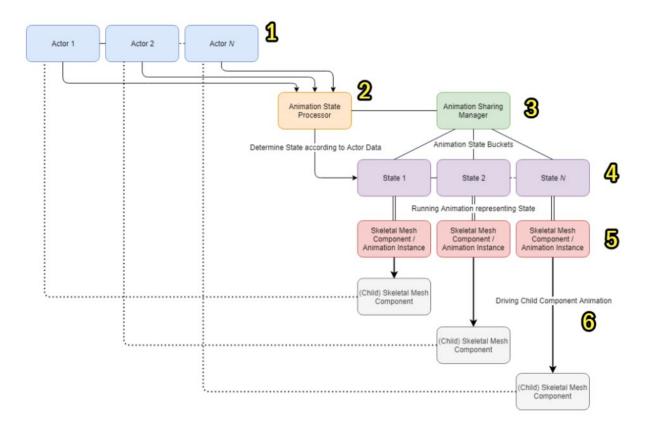


Figure 4: Animation sharing graph - Animation Sharing Plugin documentation (Epic Games)[33]

[33]

5.4. VERTEX ANIMATIONS

Vertex animations are an alternative to skeletal animations, the biggest advantage they have is that they run almost completely on the gpu.

Vertex animations store animation data in a texture, so the shader reads the texture to know where each vertex of the mesh needs to be at what frame in the animation.

A big disadvantage is that you lose a lot of cool things you can do with skeletal animations, like blending. It is still possible to blend between different animations, but you need separate textures for every possible blending scenario.

There is also extra labor required to create vertex animations. You usually first create a skeletal animation, and then afterwards use a tool to bake the animation to a texture.

Unreal Engine has a plugin called AnimToTexture, that allows you to bake animation sequences to textures.

[31], [34]

6. GPU PROGRAMMING

Whenever the CPU is your biggest bottleneck, while the GPU is still relatively free, it can be a good idea to delegate workload from the CPU to the GPU. The GPU is also able to do certain operations much more efficiently than the CPU, which can be another reason to use one of the following techniques.

However doing calculations on the GPU can add a lot of complexity to your code and should only be done if you need insane performance optimizations. For example in Ultimate Epic Battle Simulator 2, where there are millions of units in a battle, they achieved this by running most of their calculations on the GPU.



Figure 5: Screenshot of a battle with 10 million units, almost the entire battlefield is covered - UEBS 2 The Making Of Pt1 (BrilliantGameStudios)[35]

6.1. COMPUTE SHADERS

Compute shaders are shaders that instead of rendering, are used for computations. They work very similar to normal shaders, except that their output is not used for rendering.

Because they run on the GPU, they can do some operations much more efficient than the CPU. They also run on a separate thread, so they won't block the thread from which you call them, like with multithreading.

[36]-[39]

6.2. CUDA

CUDA is an API developed by Nvidia, to allow for programming on the GPU. It has the same purpose as compute shaders, but instead of writing a shader, you can use a programming language like C, C++ or Fortran. This often makes it easier or more intuitive to use than compute shaders.

Since this is developed by Nvidia, it will only run on Nvidia GPUs.

[40]-[45]

6.3. OPENCL

OpenCL is very similar to CUDA, but can run on both Nvidia and AMD GPUs since it is developed by Khronos Group. It is a more general parallel programming API, but its main use is GPU programming.

[36], [37], [39], [46]–[49]

CASE STUDY

1. EXPERIMENT SETUP

1.1. INTRODUCTION

In this case study, we create a battle simulator game, using Unreal Engine's Mass system, and implement different techniques to try to increase performance to allow for the simulation of massive amounts of agents at a playable fps. Playable fps is defined as 30 fps for the purpose of this paper.

We will then compare the performance of the battle simulator that uses Mass, with a simple battle simulator game that is created using usual methods. This is to confirm if the performance is actually a result of using Mass and the various techniques we added.

We will also compare the different additions to the Mass battle simulator to see which ones improved what aspects.

Some snippets of code will be included here, for the full source code, please refer to the Github repository: https://github.com/SimonSchaep/Research-UE5-Mass/tree/main/UnrealProjects

Code will be written with the standard Unreal Engine coding standards in mind: https://docs.unrealengine.com/5.3/en-US/epic-cplusplus-coding-standard-for-unreal-engine/

1.2. REQUIREMENTS

The simple and Mass battle simulator both have the following requirements:

- Two teams of units face each other.
- When the battle starts, the units run towards the closest enemy unit.
- When they are within a small distance (attack range), they will stop moving and start attacking.
- Attacking means that after a repeating delay, the unit deals damage to its target.
- When the target reaches zero hitpoints, it dies.
- The units have to avoid other units.
- The movement utilizes pathfinding to avoid big obstacles and navigate through the world.
- 3D models are used to represent the units, they are of standard quality.
- The following animations play at the correct time:
 - o Idle
 - Moving
 - Attacking
 - o Dying

To ensure the quality of the assets that are used we will be using models from the game Paragon[50]. This was a AAA game developed by Epic Games, which means the quality should be very good and it should be easy to use the assets in Unreal Engine. All assets are available on the Unreal Engine Marketplace, we will be using the paragon minion models specifically

For environmental assets we will be using the free Landscape Pro 2.0 package from the Unreal Engine Marketplace: https://www.unrealengine.com/marketplace/en-US/product/landscape-pro-auto-generated-material

1.3. MEASUREMENTS

We will measure performance of the simple battle simulator, the battle simulator made with Mass, and also every time we add a major optimization feature.

Eight different unit counts will be measured, starting at 100 vs 100, then doubling until we reach 12800 vs 12800. With these multiple tests we will be able to see how well everything scales with agent count.

Performance will be measured by frame time, which will be recorded using Unreal Insights, in a development build. This is not a shipping build, so there will still be a certain performance cost because debugging data is included, however, this is required to be able to properly profile.

We will record from the start of the battle (after spawning in the units), until one side has lost all their units. Meanwhile the camera will stay static, to eliminate any variance that could be caused by culling, LODs...

We will use the frame at 10 seconds into the battle to compare with others. Unless that frame has some irregularity, like a big flush, then we take the next frame.

The goal is not to have super precise measurements with exact numbers. Instead, we want to get a general idea of where we should optimize and which techniques have a noticeable impact on which aspect of the game.

There will only be one device used for measurements, an Omen Gaming Laptop with the following specs:

CPU: AMD Ryzen 7 5800H 3.20 GHz

RAM: 16 GB

GPU: NVIDIA GeForce RTX 3060 Laptop GPU

[51]-[53]

2. EXPERIMENT EXECUTION

2.1. SIMPLE BATTLE SIMULATOR

For the simple battle simulator, we will follow the normal pipeline for creating an Unreal Engine game.

We create a unitmanager as a component of the gamemode. This manager spawns a specified number of actors within a specified area. It also holds two arrays of AActor pointers to those spawned actors, so these can later be

queried when looking for nearby actors. There are two arrays, so that the actors can easily be filtered based on which army they belong to.

```
TArray<TArray<AActor*>> Units;
```

Code snippet 1: Units variable declaration

```
Units.Add(TArray<AActor*>());
Units.Add(TArray<AActor*>());
Units[0].Reserve(SpawnCount);
Units[1].Reserve(SpawnCount);
```

Code snippet 2: Unit arrays added to the Units variable

The agent/unit is an actor that inherits from ACharacter, since that will allow us to use UCharacterMovementComponent, which has in-built RVO avoidance.

Most functionality will be implemented in separate components, since that allows reusability in case we would implement different types of units. It will also be easier to translate to Mass, since the components are somewhat similar to the components in an ECS. (fragments in Unreal's case)

2.1.1. COMPONENT OVERVIEW

TARGET ACQUISITION COMPONENT

Holds the ID of the army this unit belongs to, so it can filter the arrays of units from the unit manager Every Tick, get all alive units from the unit manager.

Find and store the one that is the closest to this unit

```
UCLASS( classGroup=(Custom), meta=(BlueprintSpawnableComponent) )
class BATTLESIMULATOR_API UTargetAcquisitionComponent : public UActorComponent
{
    GENERATED_BODY()

public:
    UTargetAcquisitionComponent();
    virtual void BeginPlay() override;
    virtual void TickComponent(float DeltaTime, ELevelTick TickType, FActorComponentTickFunction* ThisTickFunction) override;

    AActor* GetClosestTarget()const;
    float GetClosestTargetDistanceSqr()const;

UFUNCTION(BlueprintCallable)
    int GetTeamId()const;

private:
    void FindClosestTarget();
    class UUnitManager* UnitManager;
    AActor* ClosestTargetDistanceSqr = FLT_MAX;
    int TeamId;
};
```

Code snippet 3: Target acquisition component header file

MOVE COMPONENT

Every Tick, read the current target from UTargetAcquisitionComponent.

If we are outside StopRange, move towards the target, following a path on the navmesh.

Otherwise, stop all movement.

```
UCLASS( classGroup=(Custom), meta=(BlueprintSpawnableComponent) )
class BATTLESIMULATOR_API UMoveComponent : public UActorComponent
{
    GENERATED_BODY()

public:
    UMoveComponent();

    virtual void BeginPlay() override;
    virtual void TickComponent(float DeltaTime, ELevelTick TickType, FActorComponentTickFunction* ThisTickFunction) override;

UFUNCTION(BlueprintCallable)
    bool IsMoving()const;

private:
    class UTargetAcquisitionComponent* TargetAcquisitionComponent;
    class UTargetAcquisitionComponent* TargetAcquisitionComponent;

    float StopRangeSqr;
    bool bIsMoving = false;

UPROPERTY(EditAnywhere)
        float StopRange = 150.f;
};
```

Code snippet 4: Move component header file

HEALTH COMPONENT

Holds the current health amount.

Has a function that subtracts health, and broadcasts a delegate event when health reaches zero or below. After death, we leave the actor alive, since it looks cool and won't bring our performance lower than it was at the start of the battle (and it looks cool). If you are in scenario where you spawn in more units as the battle progresses, it is probably better to destroy the actor after death.

```
UCLASS( ClassGroup=(Custom), meta=(BlueprintSpawnableComponent))
class BATTLESIMULATOR_API UHealthComponent : public UActorComponent
{
    GENERATED_BODY()

public:
    UHealthComponent();

    virtual void BeginPlay() override;

    DECLARE_DYNAMIC_MULTICAST_DELEGATE(FOnDeath);
    UPROPERTY(BlueprintAssignable)
        FOnDeath OnDeath;

    void DoDamage(float Amount);

    UFUNCTION(BlueprintCallable)
    bool HasDied()const;

private:
    class UUnitManager* UnitManager;

    UPROPERTY(EditAnywhere)
        float CurrentHealth = 100;

};
```

Code snippet 5: Health component header file

ATTACK COMPONENT

Holds data regarding attacking (damage, attack interval, attack interval timer)

Every Tick, read the current target from UTargetAcquisitionComponent.

If the target is within attack range, count down the timer.

If the timer reaches zero or below, deal damage to the target and reset the timer.

```
UCLASS( ClassGroup=(Custom), meta=(BlueprintSpawnableComponent) ) class BATTLESIMULATOR_API UAttackComponent : public UActorComponent
    GENERATED_BODY()
    UAttackComponent();
    virtual void BeginPlay() override;
    virtual void TickComponent(float DeltaTime, ELevelTick TickType, FActorComponentTickFunction* ThisTickFunction) override;
    bool IsAttacking()const;
    void AttackTarget();
    bool bIsAttacking = false;
    class UTargetAcquisitionComponent* TargetAcquisitionComponent;
    float AttackDelayTimer = 0.f;
    float RangeSqr;
    UPROPERTY(EditAnywhere)
        float AttackDelay = 1.f;
    UPROPERTY(EditAnywhere)
        float AnimationAttackDelay = .4f;
    UPROPERTY(EditAnywhere)
        float Range = 200;
    UPROPERTY(EditAnywhere)
         float Damage = 20;
```

Code snippet 6: Attack component header file

2.1.2. ANIMATIONS

We still have to set up animations. For that, we add our model to the skeletal mesh component in the unit blueprint and we create a simple animation blueprint with four different states: Idle, Moving, Attacking and Dead.

We assign the correct animations to each state and we get our current state from a UAnimationInstance class that we create. This class will evaluate the actor that the skeletal mesh is a part of, and by using the components, will determine the state that the unit is in.

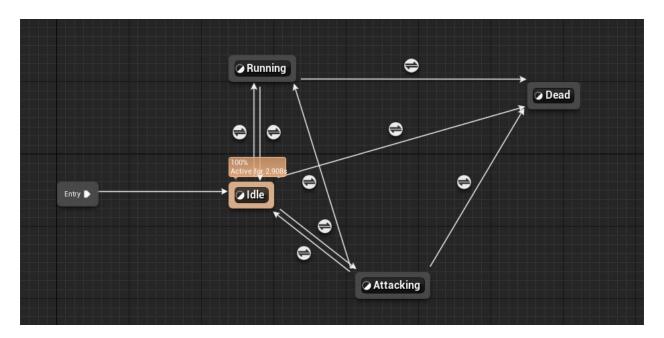


Figure 6: Simple battle simulator animation state graph

2.1.3. PROJECT SETTINGS

We don't change many project settings, except for changing the global illumination method from lumen to none

We leave the anti-aliasing method to the default Temporal Super Resolution, because other methods look noticeably worse in this project.

We leave the shadow map method to virtual shadow maps.

2.2. MASS BATTLE SIMULATOR

For the battle simulator made with Mass we have to first include the required plugins, these include MassEntity (for all base Mass functionality) and MassAI, MassGameplay and MassCrowd (which include lots of pre-built functionality)

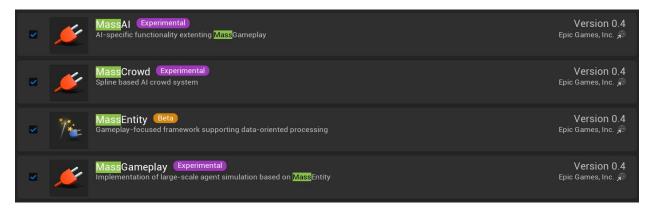


Figure 7: PLugins required for using Mass in Unreal Engine 5

Next, we will create our fragments and processors. Fragments are very simple, they just hold data. Setting up a processor usually requires the following three functions: the contructor, ConfigureQueries and Execute.

In the constructor, we define the initial value of some variables, like whether the processor will automatically register itself, or if we want to do that at another time through our own code. We can also configure which tick phase we want our processor to execute during, and there are many more options than that of course. ConfigureQueries is used for configuring which archetypes the processor will operate on, you can define which fragments and/or tags the archetype needs to have or not have.

In the Execute function, you can program the operations that the processor needs to perform on each entity chunk. Chunks are entities grouped by archetype.

```
UCLASS()
class BATTLESIMULATORMASS_API UDestroyEntitiesProcessor : public UMassProcessor
{
    GENERATED_BODY()

public:
    UDestroyEntitiesProcessor();

protected:
    virtual void ConfigureQueries() override;
    virtual void Execute(FMassEntityManager& EntityManager, FMassExecutionContext& Context)override;

private:
    FMassEntityQuery EntityQuery;
};
```

Code snippet 7: Example of a processor header file

2.2.1. FUNCTIONALITY

SPAWNING

Spawning Mass entities can be done in two ways by default:

Adding a MassAgent component to an actor, and spawning in the actor.

Or using the MassSpawner to spawn in entities based on an EntityConfig data asset.

We could create our own spawner aswell if we want custom spawing behavior, but for our purposes, the MassSpawner will do.

The EntityConfig asset can be used to define which traits, so in turn, which fragments will be part of the spawned entities.

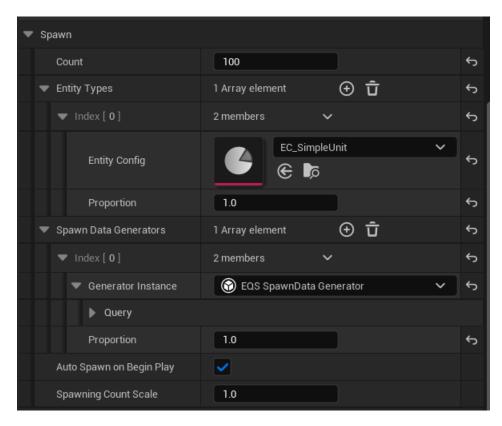


Figure 8: Example of a Mass spawner configuration

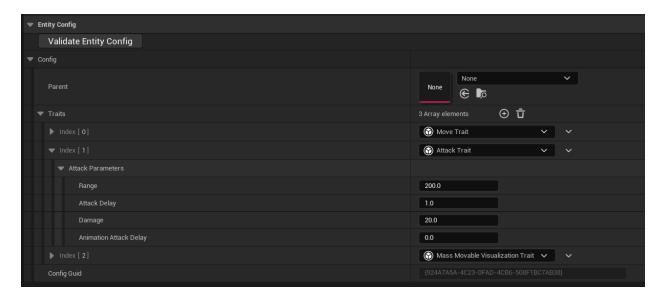


Figure 9: Example of an EntityConfig data asset configuration

TARGET ACQUISITION

For the target acquisition, we first need to be able to associate every entity with a certain army, for that, we add an ArmyldFragment, this just holds an integer value that represents the army this entity belongs to.

```
USTRUCT()
struct BATTLESIMULATORMASS_API FArmyIdFragment : public FMassFragment {
    GENERATED_BODY()
    int ArmyId = 0;
};
```

Code snippet 8: FArmyIdFragment definition

Ideally, we should be able to configure a spawner to spawn units with a specific army ID. This is not possible with the default spawner – data generator setup. To solve this issue, we create our own spawn data generator, based on the default EQS SpawnPoints generator, we just add an extra value, Armyld, to the spawndata. Then we also create a copy of the default UMassSpawnLocationProcessor that sets up all spawned entities with the correct spawn transforms, we again just add the army ID.

```
const int32 NumEntities = Context.GetNumEntities();
for (int32 EntityIndex{}; EntityIndex < NumEntities; ++EntityIndex)</pre>
   //Spawn Transform
   const int32 AuxIndex = FMath::RandRange(0, Transforms.Num() - 1);
   TransformList[EntityIndex].GetMutableTransform() = Transforms[AuxIndex];
   if (bShouldOverrideSpawnRotation)
       TransformList[EntityIndex].GetMutableTransform().SetRotation(SpawnRotation);
   Transforms.RemoveAtSwap(AuxIndex, 1, /*bAllowShrinking=*/false);
   RepresentationList[EntityIndex].StaticMeshDescIndex = RepresentationSubsystem->FindOrAddStaticMeshDesc(StaticMeshInstanceDesc);
   RepresentationList[EntityIndex].LowResTemplateActorIndex
       LowResTemplateActor.Get() ? RepresentationSubsystem->FindOrAddTemplateActor(LowResTemplateActor.Get()) : INDEX_NONE;
   RepresentationList[EntityIndex].HighResTemplateActorIndex =
       \label{thm:highResTemplateActor.Get()} \ensuremath{\text{RepresentationSubsystem-}} Find Or Add Template Actor (High ResTemplate Actor.Get())} : INDEX_NONE;
   ArmyIdList[EntityIndex].ArmyId = UnitArmyId;
   TargetAcquisitionSubsystem->AddPossibleTargetEntity(
       Context.GetEntity(EntityIndex),
       UnitArmyId
```

Code snippet 9: Assigning the correct initial values to newly spawned entities in the initializer processor

Next, we need to be able to get a list of all possible targets. For that, we will be using a World Subsystem, which is similar to a singleton in that it can easily be accessed from any class. It is created right before the world initializes, and destroyed after the world cleans up.

This subsystem will hold an array of arrays of FMassEntityHandle, similar to the simple battle simulator. FMassEntityHandle is a struct that can be used together with the EntityManager to operate on a specific entity, for example getting a specific fragment or tag.

```
TArray<TArray<FMassEntityHandle>> PossibleTargetEntities;
```

Code snippet 10: PossibleTargetEntities variable declaration

We then create a TargetAcquisitionProcessor, where we go over every entity and for each one, we go over all entity arrays that are not at the index of our own army ID, to find the one that is closest. We then update our TargetAcquisitionFragment with the target we found.

Code snippet 11: TargetAcquisitionProcessor Execute functionality

MOVEMENT

For movement, we will use the Steering, Avoidance, Movement and SmoothRotation traits that are included in the Mass plugin.

We still have to set the steering target and do pathfinding ourselves. To do so, we will create a navigation processor and a movement processor.

The navigation processor will go over every entity, and find a path on the navmesh towards the target from the TargetAcquisitionFragment. It uses UNavigationSystemV1::FindPathSync to do so.

```
FMassMoveTargetFragment& MoveTarget = MoveTargetList[EntityIndex];
const FTransform& Transform = TransformList[EntityIndex].GetTransform();
const FMassEntityHandle& TargetEntity = TargetAcquisitionList[EntityIndex].CurrentTarget;
if (!EntityManager.IsEntityValid(TargetEntity)) continue;
FMassEntityView EntityView{ EntityManager, TargetEntity };
const FVector& TargetEntityLocation = EntityView.GetFragmentData<FTransformFragment>().GetTransform().GetLocation();
//Configure path query params
FNavAgentProperties NavAgentProperties{};
FPathFindingQuery NavParams{};
NavParams.NavData = NavigationSystem->MainNavData;
NavParams.QueryFilter = NavParams.NavData->GetDefaultQueryFilter();
NavParams.StartLocation = Transform.GetLocation();
NavParams.EndLocation = TargetEntityLocation;
FPathFindingResult Result = NavigationSystem->FindPathSync(NavAgentProperties, NavParams, EPathFindingMode::Regular);
if (Result.IsSuccessful() && Result.Path->GetPathPoints().Num() >= 2)
    const FVector& Target = Result.Path->GetPathPoints()[1]; //Take point after starting point of path
    MoveTarget.Center = Target;
else //We are outside navmesh, try to get back
    FNavLocation NavLocation{};
    NavigationSystem->ProjectPointToNavigation(Transform.GetLocation(), NavLocation, FVector(1000, 1000, 1000));
    MoveTarget.Center = NavLocation.Location;
```

Code snippet 12: Navigation mesh pathfinding in the navigation processor Execute function

The movement processor will check if the unit is within stop range, and if so, will set the steering target to its own position. With this, there is still one issue, the avoidance will still be active. To solve this, we set the velocity and force values of the according fragments to zero every frame. It would be more ideal to have our own avoidance system, but that is out of scope of this research.

```
const FTransform& Transform = TransformList[EntityIndex].GetTransform();
FMassMoveTargetFragment& MoveTarget = MoveTargetList[EntityIndex];
//If in stop range -> stop moving
if (TargetAcquisitionList[EntityIndex].ClosestTargetDistanceSqr <= FMath::Square(UnitMoveParams.StopDistance))
    MoveTarget.Center = Transform.GetLocation();
    MoveTarget.DistanceToGoal = 0;
    VelocityList[EntityIndex].Value = FVector::Zero();
    ForceList[EntityIndex].Value = FVector::Zero();
    //Set anim state
    if (AnimStateList[EntityIndex].UnitAnimState == EUnitAnimState::Moving)
        AnimStateList[EntityIndex].UnitAnimState = EUnitAnimState::Idle;
    continue;
MoveTarget.Center.Z = 0;
MoveTarget.DistanceToGoal = TargetAcquisitionList[EntityIndex].ClosestTargetDistanceSqr;
MoveTarget.Forward = (MoveTarget.Center - Transform.GetLocation()).GetSafeNormal();
MoveTarget.DesiredSpeed = FMassInt16Real(MovementParams.DefaultDesiredSpeed);
AnimStateList[EntityIndex].UnitAnimState = EUnitAnimState::Moving;
```

Code snippet 13: MoveProcessor functionality in the Execute function

ATTACKING

For our attacking functionality, we need two fragments to hold data, and Attack and a Health fragment. The attack fragment will hold the attack delay timer, while the health fragment will keep track of the current health. For our constant data, like damage and attack delay, we will use a shared fragment. This fragment will be shared between all entities of the same archetype, allowing for more efficient memory usage. It is a good practice to make use of shared fragments wherever possible.

```
USTRUCT()
struct BATTLESIMULATORMASS_API FUnitAttackParameters : public FMassSharedFragment
{
    GENERATED_BODY()

    UPROPERTY(EditAnywhere)
        float Range = 200.f;

    UPROPERTY(EditAnywhere)
        float AttackDelay = 1.f;

    UPROPERTY(EditAnywhere)
        float Damage = 20.f;

    UPROPERTY(EditAnywhere)
        float AnimationAttackDelay;
};
```

Code snippet 14: Attack parameters shared fragment definition

The attack processor will simply go over all entities, and count down their attack delay timer.

Once this timer is below or equal to zero, the health fragment of the target entity will lose health equal to the damage of the shared fragment.

If the health falls below 0, it will simply add a DyingTag to the entity, which will exclude it from most processors and allow a DeathProcessor to take over.

Code snippet 15: Part of the attack processor functionality

DYING

Dying will be handled in two separate stages: dying and dead, which will be indicated by the appropriate tag.

Dying entities are still in the process of dying, they need to play their animation, but no longer be doing any functionality like movement. They also need to be excluded from the target acquisition subsystem.

The dying fragment will use a timer to indicate how long it takes until it is dead. The processor will go over every entity and count down the timer, once done, it will spawn a new entity, according to a DeadEntity EntityConfig asse. This entity will just have a transform and visualization.

```
//Timer
if (DeathList[EntityIndex].DeathDelayTimer > 0)
{
    DeathList[EntityIndex].DeathDelayTimer -= WorldDeltaTime;
    if (DeathList[EntityIndex].DeathDelayTimer > 0) continue;
}

//Mark for destroy
Context.Defer().AddTag<FDeadTag>(Context.GetEntity(EntityIndex));

FTransform Transform = TransformList[EntityIndex].GetTransform();
int ArmyId = ArmyIdList[EntityIndex].ArmyId;

//Create new dead entity
Context.Defer().PushCommand<FMassDeferredCreateCommand>(
```

Code snippet 16: Part of the dying processor functionality in the Execute function

We will not immediately destroy the now dead entity, as a good practice is to instead mark it as dead by using a tag, and then using DestroyEntities() to destroy multiple entities at once.

```
void UDestroyEntitiesProcessor::Execute(FMassEntityManager& EntityManager, FMassExecutionContext& Context)
{
    EntityQuery.ForEachEntityChunk(EntityManager, Context, ([&](FMassExecutionContext& Context)
    {
        EntityManager.Defer().DestroyEntities(Context.GetEntities());
    }));
}
```

Code snippet 17: DestroyEntitiesProcessor Execute function

2.2.2. VISUALIZATION

We will visualize the units by adding the MassMovableVisualizationTrait to the EntityConfig asset. This trait requires us to configure which actors or static meshes will represent the entity, along with at what distance actors or instanced static meshes will be shown. The static meshes will be just one extracted frame from the idle animation, and the actor will be a newly created actor blueprint.

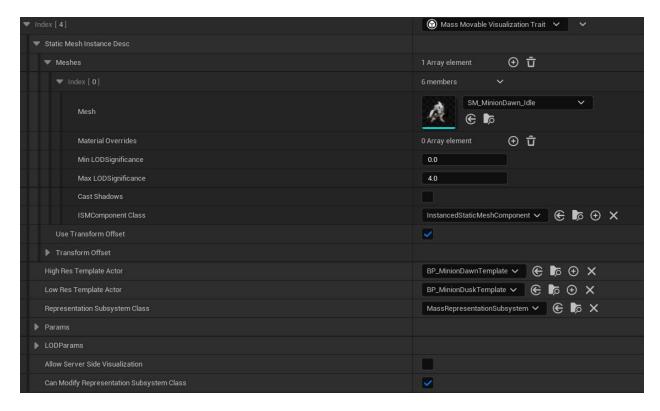


Figure 10: Visualization trait configuration

This actor needs to have a MassAgentComponent to be able to work properly. In this component, we can then assign a sync trait. There are some sync traits that come with the plugin, but these assume that our actor inherits from ACharacter and have a CharacterMovementComponent. This is not the case for us, so we will create our own sync trait and translator. The translator is actually just a processor that sets the transform of the actors associated with an entity.

Code snippet 18: Transform translator execute function

We also need to add the LODCollectorTrait, to make sure that the correct LOD tags are added to the entity, so that the visualization processor knows when to show what type of visualization.

2.2.3. ANIMATION

We will not implement vertex animations, so whenever the instanced static meshes are shown, they will not be animated. However, we will use skeletal meshes in the actors that are used for visualization up close. This means we add a skeletal mesh component to our actor, assign the correct model and create an animation blueprint. The animation blueprint is exactly the same as the one used in the simple battle simulator, but the animation instance class that is used to gather the state of the actor will be different.

We now need to access the MassAgentComponent, get the associated EntityHandle, and get a fragment that holds our animation state. This fragment's state will be set by various processors, like the movement and attack processors.

```
void UUnitAnimInstance::NativeUpdateAnimation(float DeltaSeconds)
{
    if (EntityManager == nullptr) return;
    if (!EntityManager->IsEntityValid(EntityHandle))
    {
        if (!TryGetEntityHandle())
        {
            return;
        }
    }

//Get animation state from fragment
    auto DataStruct = EntityManager->GetFragmentDataStruct(EntityHandle, FUnitAnimStateFragment::StaticStruct());
    if (!DataStruct.IsValid()) return;
    const FUnitAnimStateFragment& UnitAnimStateFragment = DataStruct.Get<FUnitAnimStateFragment>();
    UnitAnimState = UnitAnimStateFragment.UnitAnimState;
}
```

Code snippet 19: Animation instance update functionality

[6], [10]–[18], [20], [21]

2.3. RENDERING/LODS

We soon noticed that our draw calls exceeded insane amounts, and the draw thread got overloaded way sooner than expected. This is because we originally used models from the Paragon Greystone pack, which were intended to be models for a main player character. This meant it had 18 different materials applied resulting in an insane amount of draw calls per unit. After replacing this with the paragon minions models, the draw calls went back to a more normal amount.



Figure 11: Draw call amount with old model



Figure 12: Draw call amount with new model

However, the primitive count was still too high and could become a bottleneck. To solve/remedy this, we generated new LODs for the models. Going from 5000 vertices to only 250 at far distances. Generating LODs in Unreal Engine is luckily very easy, you just configure an LOD data asset by specifying the percentage of triangles that you wish to keep at which LOD levels, and click regenerate.

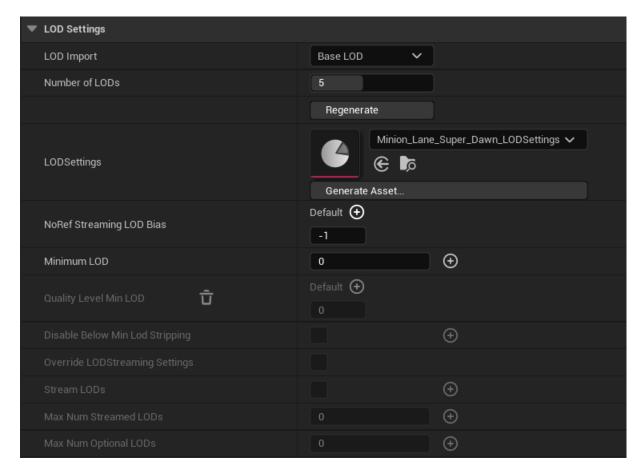


Figure 13: LOD Settings after generating an LODSettings asset

2.4. MULTITHREADING

As of 5.1, Mass already does some multithreading by default. Every processor will execute on a separate thread, unless bRequiresGameThreadExecution is true. This might happen in parallel with other processors, but the execution order often won't allow it.

However, a single processor can still take a lot of frame time, and it might benefit from dividing tasks between multiple threads.

The processors that currently take the most frame time, are the target acquisition, avoidance and navigation processors. We could multithread the avoidance processor by copying it and making our own, but since that would cost too much time, we will skip this one. The target acquisition processor can easily be multithreaded by just replacing the normal for loop with a ParallelFor.

Code snippet 20: Replacing the normal for loop with a ParallelFor

We can do the exact same in our navigation processor, but we have to replace FindPathToLocationSynchronously with FindPathSync, since the first one requires to be executed on the game thread. There is also the option to use a normal for loop and use FindPathASync, but from what we tried, this was slower than FindPathSync in a ParallelFor.

We also multithreaded the movement and attacking processor, which in hindsight was not needed since they already barely take any frame time at all.

Also make sure to use a mutex (FCriticalSection in Unreal Engine) whenever you don't want two threads accessing the same section of your code at the same time.

```
//Check if dead
if (TargetEntityHealth.CurrentHealth <= 0)
{
    Mutex.Lock();
    Context.Defer().AddTag<FDyingTag>(TargetEntity);
    Mutex.Unlock();
}
```

Code snippet 21: Using a mutex to prevent simultaneous writing to the deferred command buffer

[28]-[30]

2.5. SPATIAL PARTITIONING

There are many types of spatial partitioning structures, ideally, we would test out multiple and compare which give the best result. However, we don't have enough time for that since we would like to focus on all aspects of optimizing a battle simulator.

The avoidance processor that is included in the plugin, uses a hashgrid for its spatial partitioning, however, we will use an octree. We will likely have a high variance in density of our units, at the start of the battle, they will be evenly divided, but after they meet each other, potentially less than 10% of the space will be used. Therefore, we expect better results for our case if we use an octree, since it is supposed to have fast querying of data when there is variance in density.

We could have also used a quadtree (2D version of an octree), since our movement is limited to a 2D plane right now, but that might change in the future so we use an octree instead.

Luckily, Unreal Engine has some spatial partitioning structures built into the engine, like the hashgrid that was used for avoidance. They also have multiple octree implementations, the one we will use is TOctree2 (the original TOctree is deprecated), since it is the most similar to a classic octree. There is also FSparseDynamicOctree3, which is a grid of octrees, which could be interesting but we will not experiment with it during this research.

Implementing the octree is simply a matter of replacing the arrays in the TargetAcquisitionSubsystem with an octree. The documentation of TOctree2 is very limited so it is somewhat difficult to figure out how to properly use it, so we will just look at other places in the engine code where it is used, and base our implementation on that.

After defining the octree, we will create functions that add, remove, and update our octree elements. Updating an element is done by removing and reinserting it in the octree. This will get done through a processor that updates all entities in the octree every frame. This can become an expensive operation to do for every entity, so it would be a good idea to use a tag to mark an entity to require an update, that way we avoid having to update the octree for every entity every frame.

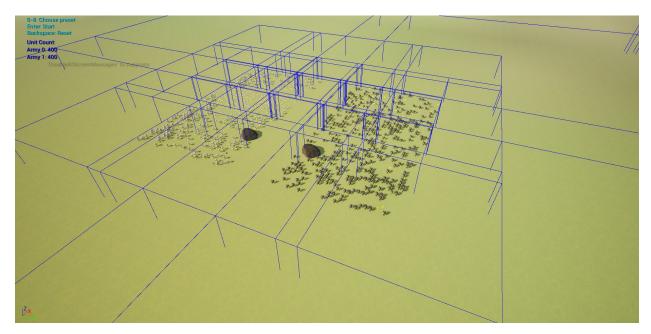


Figure 14: Visualization of the octree, the blue boxes represent the nodes of the tree

Next, for finding the closest target in the octree, we don't have access to all nodes directly, but we have some functions that will help us operate on elements in the nodes. First, we will use FindNearbyElements, to go over all elements in the same node as our entity. This will give us a close entity, but might be inaccurate, since there could be a neighboring node, that has an entity very close to the edge that is closer to us. So to get an accurate result, we will use FindElementsWithBoundsTest afterwards, with the radius of our bounds equal to the closest distance that we found so far.

```
(int32 OctreeIndex{}; OctreeIndex < Octrees.Num(); ++OctreeIndex)</pre>
//Skip those with the same army id if (OctreeIndex == ArmyIdList[EntityIndex].ArmyId)
     continue;
Octrees[OctreeIndex].FindNearbyElements(TransformList[EntityIndex].GetTransform().GetLocation(), [&](const FUnitOctreeElement& OctreeElement)
           const FMassEntityHandle& Handle = OctreeElement.EntityHandle;
           if (!EntityManager.IsEntityValid(Handle)) return;
           auto HandleTransform = EntityManager.GetFragmentDataStruct(Handle, FTransformFragment::StaticStruct()).Get<FTransformFragment>();
float DistanceSqr = FVector::DistSquared(TransformList[EntityIndex].GetTransform().GetLocation(), HandleTransform.GetTransform().GetLocation());
           float DistanceSqr = FVector::DistSqua
if (DistanceSqr < ClosestDistanceSqr)</pre>
                TargetAcquisitionList[EntityIndex].CurrentTarget = Handle;
ClosestDistanceSqr = DistanceSqr;
//Do bounds check to see if there's any elements in other nodes that are closer
float ClosestDistance = FMath::Sqrt(ClosestDistanceSqr);
FBoxCenterAndExtent Bounds{
    TransformList[EntityIndex].GetTransform().GetLocation(),
     FVector{ClosestDistance,ClosestDistance,ClosestDistance}
Octrees[OctreeIndex].FindElementsWithBoundsTest(Bounds, [&](const FUnitOctreeElement& OctreeElement)
           const FMassEntityHandle& Handle = OctreeElement.EntityHandle;
           if (!EntityManager.IsEntityValid(Handle)) return;
           auto HandleTransform = EntityManager.GetFragmentDataStruct(Handle, FTransformFragment::StaticStruct()).Get<FTransformFragment>();
float DistanceSqr = FVector::DistSquared(TransformList[EntityIndex].GetTransform().GetLocation(), HandleTransform.GetTransform().GetLocation());
if (DistanceSqr < ClosestDistanceSqr)</pre>
                TargetAcquisitionList[EntityIndex].CurrentTarget = Handle;
ClosestDistanceSqr = DistanceSqr;
```

Figure 15: Finding the closest target in the octree

[24]-[27]

2.6. ANIMATION SHARING

After optimizing the game logic, we notice that our newest bottleneck is the animations. Skeletal mesh updates and animation blueprint ticks seem to take too much frame time.



Figure 16: Screenshot from within Unreal Insights to show the amount of frame time the animations take

There are multiple ways to optimize animations, the most optimal solution would be to use vertex animations. This would take a long time to implement and might potentially be challenging to combine with the static mesh instance rendering from Mass that we are currently using.

Another option would be to use the animation budget allocator, however, this could cause noticeable animation lag with our large amount of units.

So we arrive at our third option, the animation sharing manager. This seems to be an ideal solution, since we can keep using our skeletal meshes. And because all of our units share the same skeleton, we should be able to improve performance quite a bit by sharing animations.

To set up the animation sharing manager, we have to first create an instance of the manager in our gamemode.

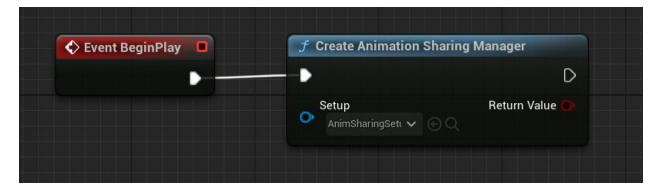


Figure 17: Creation of the animation sharing manager

Then, we have to define an AnimationSharingSetup asset. In here, we define which skeletons are used and which animations they have. We also have to enable blending and set an amount of different instances per animation, to try and make it less noticeable that animations are shared.

For our attack and death animations, we need to set them to be on demand, since we want an individual instance to play every time a unit attacks or dies.

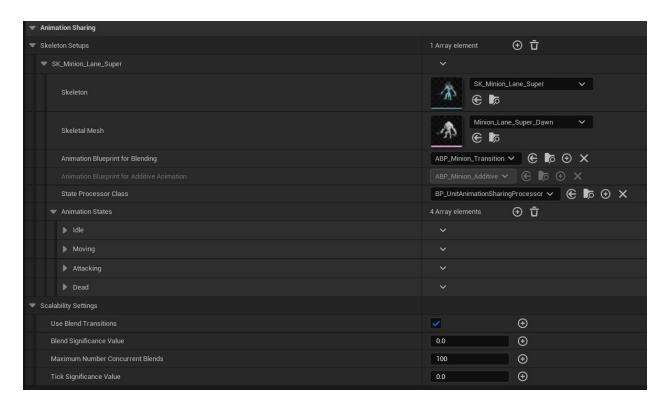


Figure 18: Configuration of the AnimationSharingSetup asset

Lastly, we have to register every actor with the animation sharing manager on BeginPlay, and unregister on EndPlay. Unregister is weirdly not available to blueprints, so we expose the function ourselves through C++;

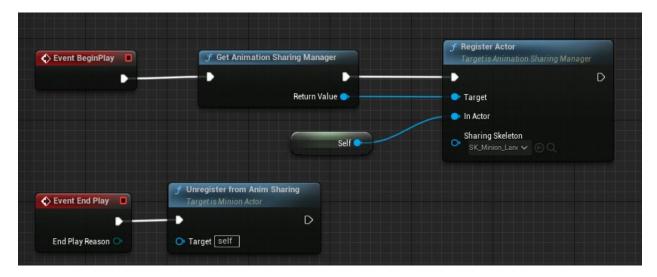


Figure 19: Registering and unregistering the actor with the animation sharing manager

[33]

3. RESULTS

Results are only measured on one device, and each trace was only recorded once. This can lead to inaccurate results, but they will still present a general idea that can be insightful. We don't think more accurate measurements will heavily influence the conclusion.

You can find an excel sheet with all the recorded data here: https://github.com/SimonSchaep/Research-UE5-Mass/blob/main/Results.xlsx

Use the drop-down list in the top right of the master sheet to determine what amount of units you want to see data for.

You can see the most impactful and interesting frame times in the table.

The navigation and target acquisition costs in the simple battle simulator are difficult to distinguish, since they don't happen in order. We will present these frame costs combined under the target acquisition cost, please be aware that this also includes navigation in the case of the simple battle simulator.

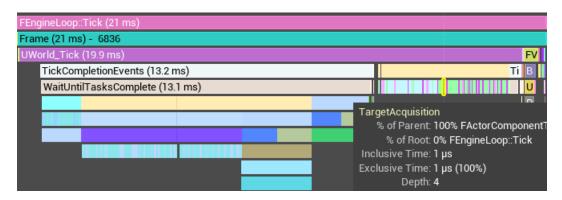


Figure 20: Screenshot from within Unreal Insights showing how component ticks are not grouped

For target acquisition with spatial partitioning, the time required to update the octree is also included in the target acquisition time.

[51]-[53]

If you want to look at the traces recorded with Unreal Insights, that were used to extract this data, they are available here: https://drive.google.com/file/d/1dy27FoWe-HDYQR6C2MbmFpC7WUoz FB9/view?usp=sharing

DISCUSSION

Let's first address the many inconsistencies in some results when going to higher frame times. We took the frame snapshot after exactly 10 seconds of the trace start, but that means that differences in frame times can affect the amount of units that are still alive at 10 seconds, which means that frame time measurements could potentially be misleading.

We could solve that by taking a frame after only one second, but this might also make results inaccurate, since there will be different animations playing, density might be different for spatial partitioning, etc...

This also means that results with frame times above 100 might be almost useless.

We will be using graphs to discuss the data, if you want to see the exact numbers we measured, please refer to the excel file with all the data here: https://github.com/SimonSchaep/Research-UE5-Mass/blob/main/Results.xlsx

1. FRAME TIME OVER UNIT COUNT

The game thread takes more frame time than the render thread and the GPU in all cases, once we go past 800 vs 800 units. This means that our game functionality scales the worst over unit count, it is also our bottleneck and should be the focus of optimizations.

An unexpected result is that the battle simulator with Mass seems to only very barely improve the frame time compared to the simple battle simulator.

We can also see that the addition of spatial partitioning seems to greatly improve the scaling of the frame time.

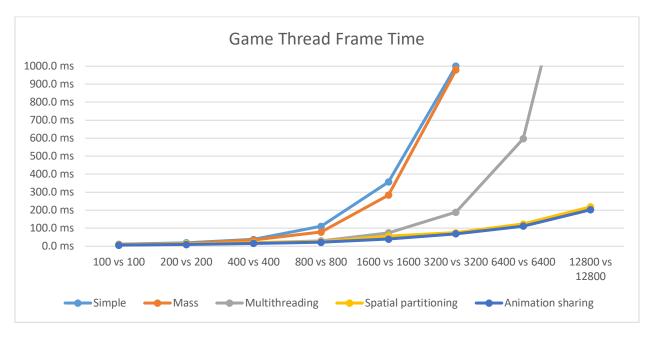


Chart 1: Game thread frame time over unit count

In case of the render thread, it seems like mass greatly reduces the frame time. Probably as a result of the instanced static meshes that are used for rendering at longer distances instead of actors.

Something noticeable is that spatial partitioning seemingly improves rendering frame time. However, it is possible that because the frame time gets so high without spatial partitioning, there is a different amount of units alive at the time of the frame, which will of course influence rendering speed.

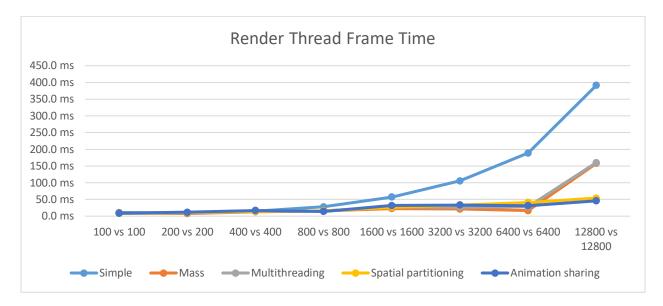


Chart 2: Render thread frame time over unit count

The GPU frame time seems to vary a lot, seemingly randomly. We can also deduce that the GPU time goes up when there are more units on screen, and that Mass, or any other technique doesn't have a clear obvious impact on GPU performance.

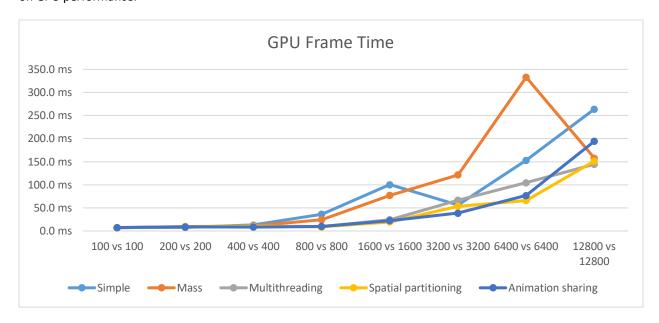


Chart 3: GPU frame time over unit count

Schaep Simon

For the simple battle simulator, we couldn't gather data for target acquisition and navigation separately, so both costs are included together in this one. This makes it much more difficult to compare the simple battle simulator with the others.

Target acquisition seems to be scaling very badly, and is at least one of the reasons of why the game thread in general scales badly over unit count. The addition of multithreading seems to help a lot with the frame time, but doesn't improve the scaling. However, the addition of spatial partitioning seems to fully solve that issue.

It is noticeable that the Mass battle simulator has a worse frame time than the simple one, even with the simple one also including navigation. This is likely caused by a mistake in our implementation or measurements, and would need to be further analyzed to find the reason.

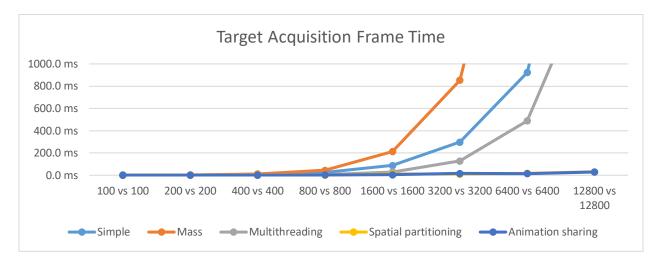


Chart 4: Target acquisition frame time over unit count

The navigation frame time of the simple battle simulator was included in target acquisition time, since we couldn't distinguish those measurements. As a result it is excluded from this graph.

Navigation seems to scale at roughly O(n) time. Multithreading improved navigation frame time by quite a lot, but didn't affect the scaling.

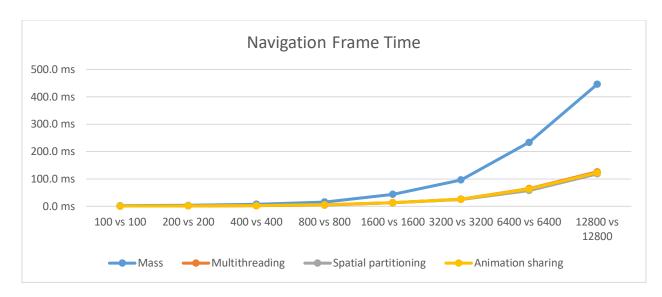


Chart 5: Navigation frame time over unit count

In this graph, we can see that the character movement component that was used in the simple battle simulator, seems to be very inefficient in general, even with only a few hundred units.

The movement in the Mass battle simulator is very fast, and the avoidance processor seems to be very optimized. It is not multithreaded however, so there is no noticeable difference with the multithreading addition.

Movement and avoidance take over 25 milliseconds in the final test, which means they will still need more optimizations if we want to run the battle simulator at 30 fps.

Multithreading could potentially be a good enough solution.

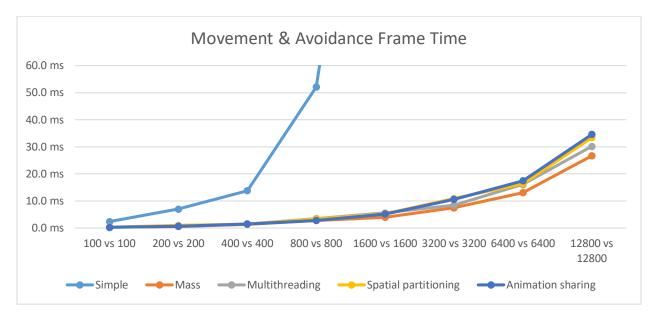


Chart 6: Movement and avoidance frame time over unit count

The time it takes to translate from Mass to the actors is noticeable enough to be worth analyzing. In the first three tests, the frame time seems to steadily grow, after which it stabilizes. This is because at that point, most new units will be spawned outside of the range where actors get used for visualization.

The frame time in case of the animation sharing implementation deviates from the others. This might be because there is an extra step involved with the skeletal meshes copying data from the animation sharing manager. It could also be a measuring mistake.

Multithreading did not improve translation frame time, because we are not able to manipulate actor positions from a thread other than the game thread. Which means we are not able to multithread this process at all.

Using vertex animations could potentially eliminate this entire cost, since we could use instanced static meshes to render everything in that case.

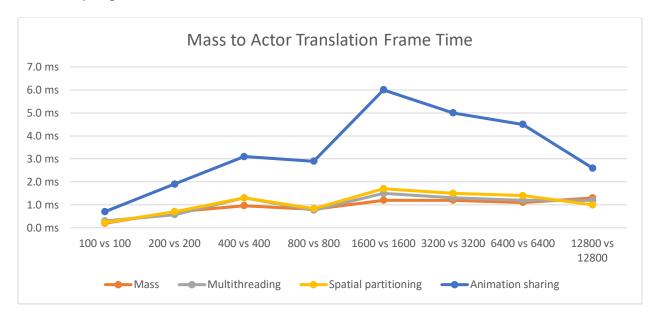


Chart 7: Mass to actor frame time over unit count

Animation frame times vary a lot, this might be because of the kind of animations that are playing at that specific frame, for example an attacking animations might be more costly than an idle one. Or perhaps it is more costly while the animations are blending between each other.

We can see that the simple battle simulator needs a lot more frame time for animations compared to the Mass implementations. This is because in Mass, the furthest away units are instanced static meshes and don't play any animations. This is definitely somewhat noticeable though, so the improved animation performance comes at a cost of quality in this case.

The animation sharing implementation greatly reduces animation frame time in all tests, which means that it works quite well.

Vertex animations might be a better solution than animation sharing, since it would solve the issue of far away units not having animations and also delegate the animation work to the GPU.

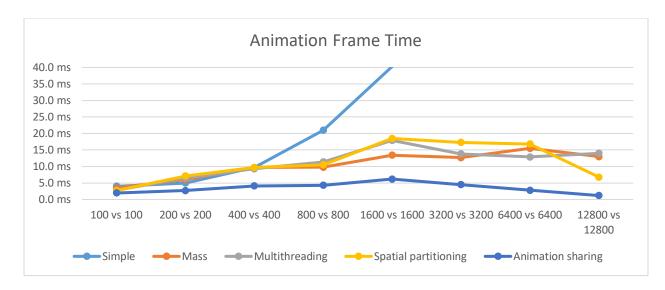


Chart 8: Animation frame time over unit count

2. GAME THREAD FRAME TIME COMPOSITION

Through the following graphs, we will take a look at which elements take the most time in the game thread, since that is our biggest bottleneck. These measurements are from the 800 vs 800 test, because that is about where the fps drops below 60 in all cases.

In the simple battle simulator, it is obvious that the character movement component is the biggest performance issue. Creating your own movement component and avoidance system might help with that. After that, target acquisition takes the most frame time.

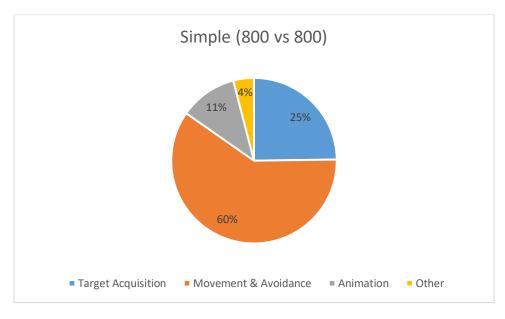


Chart 9: Game thread frame time distribution in the simple battle simulator with 800 vs 800 units

In the Mass battle simulator, target acquisition takes most of the frame time. This could be the result of a mistake in the implementation.

It is also noticeable that the navigation starts taking a big chunk of frame time as well.

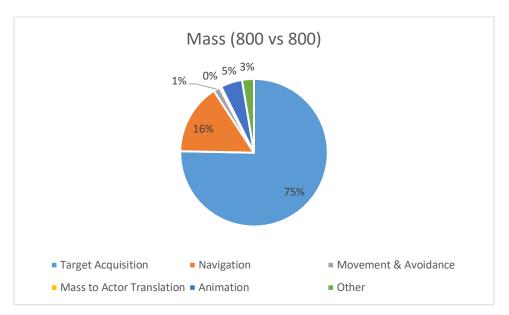


Chart 10: Game thread frame time distribution in the mass battle simulator with 800 vs 800 units

With the addition of multithreading, we can see that the distribution of frame time is much more even than before. However, target acquisition is still clearly the biggest chunk, while animation and navigation start to become very noticeable as well.

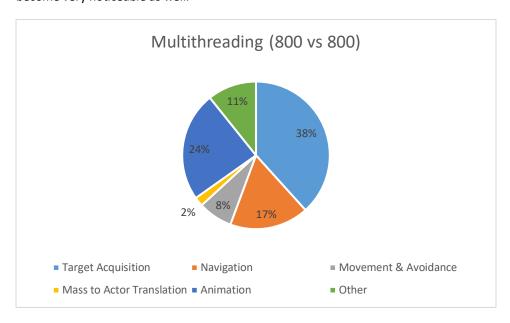


Chart 11: Game thread frame time distribution in the mass battle simulator with multithreading with 800 vs 800 units

Spatial partitioning resolves the scaling issue that target acquisition had. As a result, the biggest consumers of frame time are now animation and navigation.

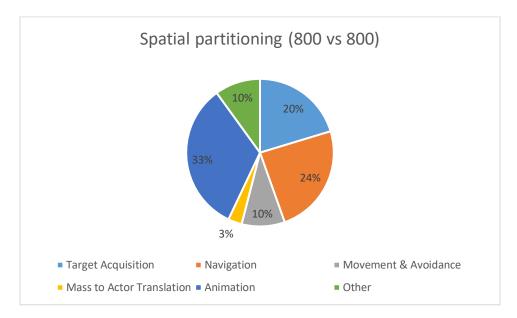


Chart 12: Game thread frame time distribution in the mass battle simulator with spatial partitioning with 800 vs 800 units

With our final addition of animation sharing, we drastically reduce the frame time that animation takes. With the newest game thread bottleneck being the navigation.

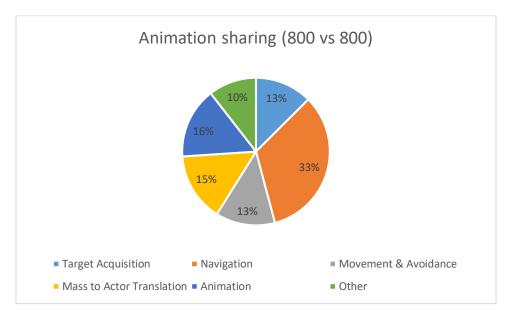


Chart 13: Game thread frame time distribution in the mass battle simulator with animation sharing with 800 vs 800 units

3. SUMMARY

With all these results we can now verify if the hypotheses were wrong or right.

We can conclude that with our implementations, we cannot simulate 20 000 units at 30 fps.

Using Unreal Engine's Mass system, does increase overall performance if you take away the target acquisition aspect of the simulation, however, our Mass implementation seems to decrease the performance of that target acquisition.

Multithreading does improve overall performance when combined with Mass.

Octree spatial partitioning does reduce frame time needed for target acquisition by improving how well it scales with unit count.

Animation sharing does reduce animation frame time.

CONCLUSION

While we didn't reach the target we expected when forming our hypotheses, we still gained some interesting insights from our case study on how to optimize huge battles in a video game:

- First of all, the game thread is the primary bottleneck in a battle simulator game.
- Secondly, spatial partitioning is extremely important in simulations with thousands of units.
- A data oriented system has the potential of increasing overall performance, and makes it easier to
 optimize other aspects, like multithreading especially.

We also have an example of the usage of Mass available now, which other people can use to learn how the system works.

If your game has different requirements than this case study, you will be able to do more specific optimizations. For example, if your battlefield does not have any obstacles, you could completely remove navigation. If it fits the battle, you could also group your units together in squads, which could also give a huge boost to game thread performance. [54]

Reducing tick rate of certain operations like target acquisition and navigation, could also make an immense impact on performance at the cost of simulation accuracy of course. In many cases however, it will be worth it.

FUTURE WORK

There is a lot of potential future work that can be done on the basis of this study.

You could delve deeper into most aspects that were applied in the case study. For example, you could try out many different spatial partitioning structures and measure which ones have the biggest impact on performance in which situations. Or you could go deeper into Mass or other data oriented systems, to try and make the most optimal use of it.

This case study could also be expanded by continuing to optimize here. The two things that would probably have the biggest impact would be to optimize navigation, since that seems to be a big consumer of frame time. Another optimization would be to use vertex animations and only instanced static meshes to render the Mass entities. This could potentially even be combined with Niagara's rendering system to try and get the best possible rendering performance.

There could also be a study on how to add projectiles to the simulation. This is quite a big addition, since you might need hit detection, efficient spawning of entities, etc...

Another big topic that could be its own research project, is GPU programming. This is how the biggest battle simulator on the market, Ultimate Epic Battle Simulator, achieves millions of units in a single battle. It is a very complex topic, but could definitely pay off.

CRITICAL REFLECTION

Overall I am happy with what I accomplished during this research project. However, I really wish I had more time to fully optimize the battle simulator and reach my goal of 20 000 agents fighting each other.

I also would have liked to take more time to explore Mass further since I only scraped the surface.

I underestimated how much work the CPU needs to do for animations and rendering as a whole.

I learned a lot about Mass and how a data oriented system can be implemented in a game engine. I would like to take the time to also look into Unity DOTS, and compare the two systems.

I also learned how to use the Unreal Engine source code to understand how their systems work, since that was my main source of information about Mass, because there are very few tutorials and documentation about more advanced Mass aspects.

REFERENCES

- [1] *CppCon 2014: Mike Acton 'Data-Oriented Design and C++'*, (Sep. 30, 2014). [Online Video]. Available: https://www.youtube.com/watch?v=rX0ltVEVjHc
- [2] CppCon 2018: Stoyan Nikolov "OOP Is Dead, Long Live Data-oriented Design", (Oct. 26, 2018). [Online Video]. Available: https://www.youtube.com/watch?v=yy8jQgmhbAU
- [3] 'Data-oriented design', *Wikipedia*. Oct. 03, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Data-oriented_design&oldid=1178464448
- [4] 'Entity component system', Wikipedia. Dec. 20, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Entity_component_system&oldid=1190943102
- [5] The ECS Architecture Performance in UE4, (Feb. 06, 2021). [Online Video]. Available: https://www.youtube.com/watch?v=kXd0VDZDSks
- [6] 'MassEntity Overview'. Accessed: Oct. 21, 2023. [Online]. Available: https://docs.unrealengine.com/5.2/en-US/overview-of-mass-entity-in-unreal-engine/
- [7] 'Community Mass Sample', Github. [Online]. Available: https://github.com/Megafunk/MassSample
- [8] 'On All Fronts', Github. [Online]. Available: https://github.com/HaywireInteractive/OnAllFronts-Public
- [9] 'MassAlTesting', Github. [Online]. Available: https://github.com/Ji-Rath/MassAlTesting
- [10] Large Numbers of Entities with Mass in UE5 | Feature Highlight | State of Unreal 2022, (Apr. 26, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=f9q8A-9DvPo
- [11] 'Mass Avoidance Overview'. [Online]. Available: https://docs.unrealengine.com/5.2/en-US/mass-avoidance-overview-in-unreal-engine/
- [12] 'MassAl, Mass Crowd, State Tree with Custom Character | UE5 | Community tutorial', Epic Developer Community. [Online]. Available: https://dev.epicgames.com/community/learning/tutorials/0be9/unreal-engine-massai-mass-crowd-state-tree-with-custom-character-ue5
- [13] 'MassGameplay Overview'. [Online]. Available: https://docs.unrealengine.com/5.2/en-US/overview-of-mass-gameplay-in-unreal-engine/
- [14] The Matrix Awakens: Generating a World | Tech Talk | State of Unreal 2022, (Apr. 05, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=usJrcwN6T4I
- [15] J. Keeling, 'Tutorial: Your First 60 Minutes with Mass', Epic Developer Community Forums. [Online]. Available: https://forums.unrealengine.com/t/tutorial-your-first-60-minutes-with-mass/794314
- [16] *UE5 MassAl Crowd in your project step by step guide Part 1*, (May 08, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=bVvYn7dEqMA
- [17] UE5 MassAI Crowd in your project step by step guide Part2 AnimationBlueprint, (Jun. 29, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=4-qcSrAxGCc&list=PLKNsvWABWRceGFLyCWBz33f1aw4WjyYGs&index=3
- [18] *UE5 Tutorial MassAl Crowd in any project in less then 15 mins.*, (Jul. 03, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=2LvUB3 PAhl
- [19] *Unity vs. Unreal Which has the Best ECS in 2022?*, (May 06, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=fFVuzlTlye4
- [20] *Unreal Engine 5 Mass Crowd Ai Tutorial*, (May 03, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=w 6LGyACVz0
- [21] *Unreal Engine 5 MassEntity Plugin Demo*, (Jul. 11, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=0dwgnH3SoC0

- [22] 'Large Numbers of Entities with Mass in Unreal Engine 5 | Talks and demos', Epic Developer Community. [Online]. Available: https://dev.epicgames.com/community/learning/talks-and-demos/37Oz/large-numbers-of-entities-with-mass-in-unreal-engine-5
- [23] 'Unreal Source Discord Server', Discord. [Online]. Available: https://discord.com/invite/unrealsource
- [24] Symphonym, 'Choosing a spatial partitioning structure', r/gamedev. [Online]. Available: www.reddit.com/r/gamedev/comments/22lxg5/choosing a spatial partitioning structure/
- [25] M. KelleghanAugust 12 and 1997, 'Octree Partitioning Techniques', Game Developer. [Online]. Available: https://www.gamedeveloper.com/programming/octree-partitioning-techniques
- [26] 'Space partitioning', *Wikipedia*. Oct. 16, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Space partitioning&oldid=1116447126
- [27] Bob Nystrom, 'Spatial Partition · Optimization Patterns · Game Programming Patterns'. [Online]. Available: https://gameprogrammingpatterns.com/spatial-partition.html
- [28] 'Multithreading and Performance in Unreal', Epic Developer Community Forums. [Online]. Available: https://forums.unrealengine.com/t/multithreading-and-performance-in-unreal/1216417
- [29] *UE4 / UE5 Multithreading with ECS*, (Aug. 15, 2021). [Online Video]. Available: https://www.youtube.com/watch?v=HQ3jNg1z63A
- [30] 'UE5 Multithreading With FRunnable And Thread Workflow Unreal C++ API'. Available: https://store.algosyntax.com/tutorials/unreal-engine/ue5-multithreading-with-frunnable-and-thread-workflow/
- [31] Simulating Large Crowds In Niagara | Unreal Engine, (Jan. 15, 2021). [Online Video]. Available: https://www.youtube.com/watch?v=CqXKSyAPWZY
- [32] 'Animation Budget Allocator'. [Online]. Available: https://docs.unrealengine.com/5.3/en-US/animation-budget-allocator-in-unreal-engine/
- [33] 'Animation Sharing Plugin'. [Online]. Available: https://docs.unrealengine.com/5.3/en-US/animation-sharing-plugin-in-unreal-engine/
- [34] Stephen Phillips, 'Baking out vertex animation in editor with AnimToTexture | Community tutorial', Epic Developer Community. [Online]. Available:

 https://dev.epicgames.com/community/learning/tutorials/daE9/unreal-engine-baking-out-vertex-animation-in-editor-with-animtotexture
- [35] 10 MILLION Characters On Screen?! Ultimate Epic Battle Simulator 2 The Making Of Pt1, (Feb. 12, 2021). [Online Video]. Available: https://www.youtube.com/watch?v=kpojDPlIjdQ
- [36] R. Banerjee, 'Answer to "What is the difference between OpenCL and OpenGL's compute shader?", Stack Overflow. [Online]. Available: https://stackoverflow.com/a/15868564
- [37] C. Rau, 'Answer to "What is the difference between OpenCL and OpenGL's compute shader?"', Stack Overflow. [Online]. Available: https://stackoverflow.com/a/15874988
- [38] Jeremy Frank, 'Simple compute shader with CPU readback | Community tutorial', Epic Developer Community. [Online]. Available: https://dev.epicgames.com/community/learning/tutorials/WkwJ/unreal-engine-simple-compute-shader-with-cpu-readback
- [39] Maiss, 'What is the difference between OpenCL and OpenGL's compute shader?', Stack Overflow. [Online]. Available: https://stackoverflow.com/q/15868498
- [40] 'CUDA', Wikipedia. Jan. 05, 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=CUDA&oldid=1193733902
- [41] SCIEMENT, '[C++][CUDA][UE4]CUDA 関数を Unreal Engine 4 で用いる サイアメント技術メモ'. [Online]. Available: https://www.sciement.com/tech-blog/c/cuda in ue4/

- [42] 'CUDA'. [Online]. Available: https://docs.unrealengine.com/5.1/en-US/API/Runtime/CUDA/
- [43] Brad Nemire, 'CUDA Spotlight: GPU-Accelerated Agent-Based Simulation of Complex Systems', NVIDIA Technical Blog. [Online]. Available: https://developer.nvidia.com/blog/cuda-spotlight-gpu-accelerated-agent-based-simulation-complex-systems/
- [44] 'Fast Large-Scale Agent-based Simulations on NVIDIA GPUs with FLAME GPU', NVIDIA Technical Blog. [Online]. Available: https://developer.nvidia.com/blog/fast-large-scale-agent-based-simulations-on-nvidia-gpus-with-flame-gpu/
- [45] anonymous_user_526a5b67, 'How to use CUDA file in UE4?', Epic Developer Community Forums. [Online]. Available: https://forums.unrealengine.com/t/how-to-use-cuda-file-in-ue4/90965
- [46] 'OpenCL', Wikipedia. Oct. 20, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=OpenCL&oldid=1181029989
- [47] 'OpenCL The Open Standard for Parallel Programming of Heterogeneous Systems', The Khronos Group. [Online]. Available: https://www.khronos.org//
- [48] 'OpenCL Guide', Github. [Online]. Available: https://github.com/KhronosGroup/OpenCL-Guide
- [49] anonymous_user_968689a2, 'OpenCL Plugin', Epic Developer Community Forums. [Online]. Available: https://forums.unrealengine.com/t/opencl-plugin/18426
- [50] 'Paragon (video game)', Wikipedia. Dec. 25, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Paragon_(video_game)
- [51] New Features of Insights: Unreal Engine's Built In Profiling Tools | Inside Unreal, (Mar. 30, 2023). [Online Video]. Available: https://www.youtube.com/watch?v=af_M38Z325I&list=PLGAgBU0oz5KZXHvaHAgas90IGBkJqsgzV
- [52] [UE5] Understanding Render Thread and Animation Thread in Unreal Engine, (Mar. 06, 2023). [Online Video]. Available: https://www.youtube.com/watch?v=RRwNIntV10I
- [53] Maximizing Your Game's Performance in Unreal Engine | Unreal Fest 2022, (Nov. 03, 2022). [Online Video]. Available: https://www.youtube.com/watch?v=Gulav71867E&list=PLGAgBU0oz5KZXHvaHAgas90IGBkJqsgzV
- [54] Phantomx1024, 'Optimizations For High Unit Count Al', r/unrealengine. [Online]. Available: www.reddit.com/r/unrealengine/comments/dmockg/optimizations for high unit count ai/

ACKNOWLEDGEMENTS

I want to thank the lecturer of the Gradwork course, Alexander Deweppe, for giving insightful lectures on the topic of creating a research project.

I also want to thank my supervisor Alex Vanden Abeele, who organized several feedback sessions where we would discuss the state of the project.

I want to thank my coach Marijn Verspecht. For assisting with the general outline of the Gradwork course.

And I want to thank everyone at Howest and Digital Arts and Entertainment for giving me the opportunity to develop the skills that were required to complete this project.

I also want to thank my family and friends for their support during this project.

APPENDICES

Github repository:

https://github.com/SimonSchaep/Research-UE5-Mass

Source code:

https://github.com/SimonSchaep/Research-UE5-Mass/tree/main/UnrealProjects

Excel file with results from case study:

https://github.com/SimonSchaep/Research-UE5-Mass/blob/main/Results.xlsx

Trace files used to extract the data for the results:

https://drive.google.com/file/d/1dy27FoWe-HDYQR6C2MbmFpC7WUoz_FB9/view?usp=sharing