



UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

RAPPORT DE STAGE MASTER 1

Etude et évaluation de la structure de donnée SVDAG et ses variantes pour le RayTracing en visualisation scientifique

Auteur : ANTOINE ROCHE

Tuteur de stage : JÉRÔME DUBOIS

Tuteur enseignant : MICKAEL KRAJECKI

08/04/2019 — 30/08/2019

CEA, DAM, DIF, F-91297, Arpajon, France

Remerciements

Je tiens à remercier Jérôme Dubois et Jacques-Bernard Lekien pour leur aide précieuse, ainsi que Michael Krajecki pour m'avoir donné l'opportunité de travailler au sein du CEA.

Sommaire

1	Introduction	4
2	Présentation CEA	5
3	Contexte du stage	12
4	Objet du stage	13
4.1	Structures de données	13
4.1.1	Sparse Voxel Octree	13
4.1.2	Sparse Voxel Directed Acyclic Graph	14
4.1.3	Adaptive Mesh Refinement	15
5	Travail réalisé	16
5.1	Prises de connaissances	16
5.2	Installation de l'environnement	16
5.3	Installation et fonctionnement du programme SVDAG	16
5.4	Analyse des performances du SVDAG	16
5.4.1	Résultats	17
5.4.2	Analyse	18
6	Expérience personnelle et professionnelle	18
6.1	Collaboration et contribution	18
6.2	Problèmes rencontrés	19
6.2.1	Installations	19
6.2.2	SVDAG	19
7	Travaux futurs	20
8	Conclusion	21
9	Glossaire	22
10	Bibliographie	23

1 Introduction

Le CEA, acteur majeur en matière de recherche et d'innovation, est considéré comme étant un expert du Calcul Haute Performance. Dans ce contexte, les équipes spécialisées en visualisation scientifique sont reconnues pour leur expertise en visualisation parallèle sur supercalculateur. Elles participent activement au développement de la bibliothèque open-source VTK, qui fournit un pipeline complet de visualisation associant différents types de données et d'algorithmes.

Pour répondre aux besoins de codes de simulation, le CEA a développé depuis quelques années des structures de données et des filtres VTK dédiés à la visualisation et la manipulation de maillages AMR. D'autres structures hiérarchiques utilisées dans le cadre de l'industrie cinématographique ou du jeux-vidéo permettent un rendu réaliste de type ray-tracing en temps réel en faisant appel à la programmation sur carte graphique (CUDA).

Ces structures sont basées sur les Smart Voxel Octrees (SVO), une représentation d'une scène voxelisée de type octree optimisée. Des techniques de compressions offrent par la suite des gains substantiels en tenant compte des motifs similaires : les Smart Voxel Directed Acyclic Graph (SVDAG) ou encore les Symmetric SVDAG (SSVDAg). L'objectif du stage est de permettre l'évaluation de ces structures hiérarchiques pour la visualisation scientifique.

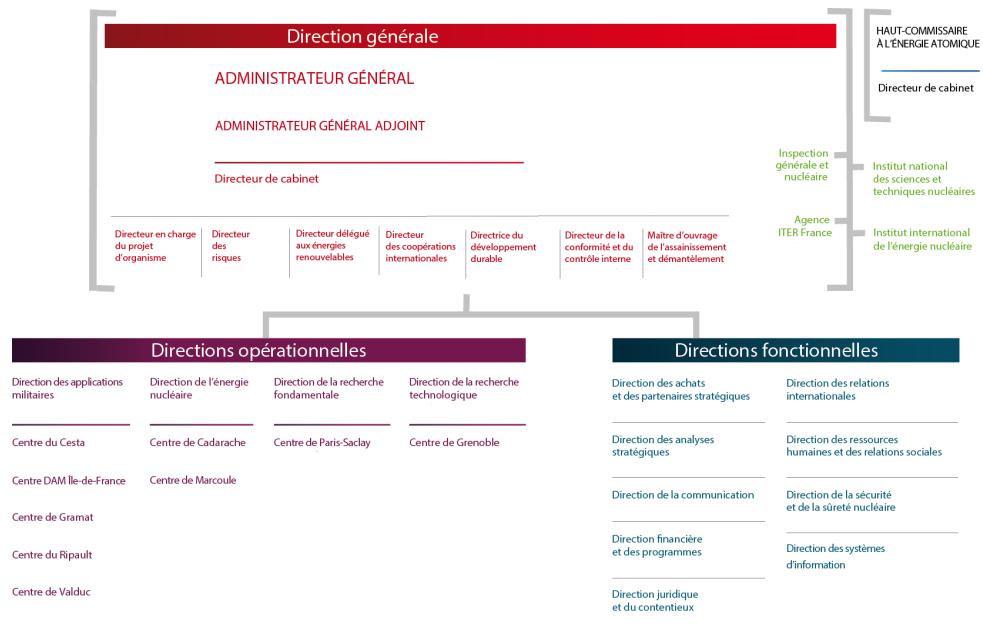
2 Présentation CEA

Acteur majeur de la recherche, du développement et de l'innovation, le Commissariat à l'énergie atomique et aux énergies alternatives intervient dans quatre domaines :

- La défense et la sécurité. ;
- Les énergies bas carbone (nucléaires et renouvelables). ;
- La recherche technologique pour l'industrie. ;
- La recherche fondamentale (sciences de la matière et sciences de la vie).

S'appuyant sur une capacité d'expertise reconnue, le CEA participe à la mise en place de projets de collaboration avec de nombreux partenaires académiques et industriels. Le CEA est implanté sur 9 centres répartis dans toute la France. Il développe de nombreux partenariats avec les autres organismes de recherche, les collectivités locales et les universités. A ce titre, le CEA est partie prenante des alliances nationales coordonnant la recherche française dans les domaines de l'énergie (ANCRE), des sciences de la vie et de la santé (AVIESAN), des sciences et technologies du numérique (ALLISTENE), des sciences de l'environnement (ALIEnvi) et des sciences humaines et sociales (ATHENA).

Reconnu comme un expert dans ses domaines de compétence, le CEA est pleinement inséré dans l'espace européen de la recherche et exerce une présence croissante au niveau international. Le CEA compte 15 942 techniciens, ingénieurs, chercheurs et collaborateurs pour un budget de 5 milliards d'euros (chiffres publiés fin 2017).



La direction des applications militaires

Une direction au service de la dissuasion

La Direction des applications militaires (DAM) du CEA, a pour mission de concevoir, fabriquer, maintenir en condition opérationnelle, puis démanteler les têtes nucléaires qui équipent les forces nucléaires aéroportée et océanique françaises.

La DAM est chargée de la conception et de la réalisation des réacteurs et de coeurs nucléaires équipant les bâtiments de la Marine nationale, sous-marins et porte-avions. Elle apporte son soutien à la Marine nationale pour le suivi en service et le maintien en condition opérationnelle de ces réacteurs.

La DAM est également responsable de l'approvisionnement des matières nucléaires stratégiques pour les besoins de la dissuasion.

Dans un monde en profonde mutation, la DAM contribue aussi à la sécurité nationale et

internationale à travers l'appui technique qu'elle apporte aux autorités, pour les questions de lutte contre la prolifération nucléaire et le terrorisme et de désarmement.

Depuis le transfert du centre de Gramat en 2010 de la Direction générale de l'armement au CEA, la DAM apporte son expertise à la Défense dans le domaine de l'armement conventionnel.



Une direction ouverte à la recherche

Le partage national et international des connaissances (lorsqu'il est possible), la confrontation à l'évaluation scientifique extérieure, l'intégration à des réseaux de compétences constituent des gages de crédibilité scientifique.

Les équipes de la DAM réalisent chaque année environ 2000 publications et communications scientifiques. Cette ouverture de la DAM passe également par la mise à la disposition de la communauté des chercheurs de ses moyens expérimentaux et par la contribution de ses équipes à d'autres programmes de recherche.

Une direction actrice de la politique industrielle française

La DAM partage très largement son activité avec l'industrie française : c'est ainsi que le montant des achats, auprès de celle-ci, représente plus des deux tiers de son budget ; le dernier tiers se répartit entre les salaires des personnels (un cinquième) et les taxes.

La politique industrielle de la DAM est originale à plus d'un titre :

- d'abord parce que la DAM conserve la maîtrise d'œuvre d'ensemble de la grande majorité des systèmes dont elle a la responsabilité : elle veille ainsi au juste équilibre entre les grands groupes industriels de la Défense et les PME souvent innovantes, en contractualisant directement avec ces dernières, leur permettant ainsi de recevoir la juste rémunération de leur production ;
- ensuite, parce que la répartition de son budget est sous-tendue par une répartition des travaux : la DAM conduit la recherche dans ses laboratoires grâce à son personnel de haut niveau scientifique et technologique. Une fois la définition d'un produit acquise, la DAM transfère la définition et les procédés vers les industriels qui en réalisent le développement, puis la production.

La DAM a également pour objectif que ses centres participent à la vie économique locale par leur implication dans les pôles de compétitivité. Hors de son propre champ d'utilisation, elle valorise ses recherches par le transfert de technologies vers l'industrie et le dépôt de nombreux brevets.

Le format

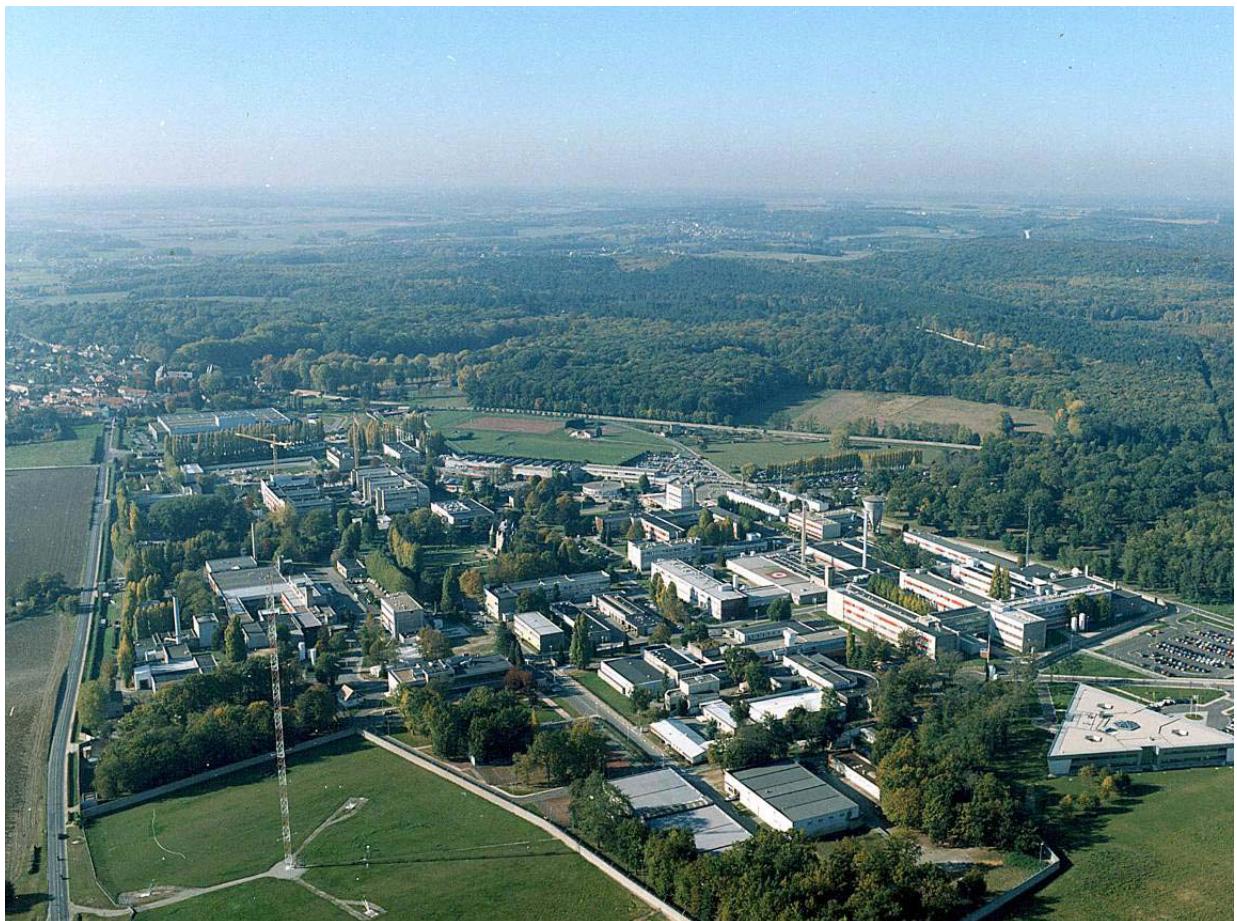
La DAM comprend cinq centres aux missions homogènes, dont les activités se répartissent entre la recherche de base, le développement et la fabrication :

- **DAM Ile-de-France (DIF)**, à Bruyères-le-Châtel, où sont menés les travaux de physique des armes, les activités de simulation numérique et de lutte contre la prolifération nucléaire ; DIF est aussi le centre responsable de l'ingénierie à la DAM ; enfin, au centre DIF est rattachée l'INBS-Propulsion Nucléaire du centre CEA/Cadarache, en région Provence Alpes-Côte d'Azur, où sont implantées les installations d'essais à terre et une partie des fabrications de la propulsion nucléaire ;



- **Le Cesta**, en Aquitaine, consacré à l'architecture des armes, aux tests de tenue à l'environnement. Il met en oeuvre le Laser Mégajoule, équipement majeur de la Simulation ;
- **Valduc**, en Bourgogne, dédié aux matériaux nucléaires et à l'installation expérimentale Epure du programme Simulation ;
- **Le Ripault**, en région Centre, dédié aux matériaux non nucléaires (explosifs chimiques...);
- **Gramat**, (ex-DGA) en Midi-Pyrénées, qui conduit au profit de la Défense des activités en vulnérabilité des systèmes et efficacité des armements. ;

Le centre DAM Ile-de-France



Centre DAM Île-de-France

Le CEA/DAM - Île de France (DIF) est l'une des directions opérationnelles de la DAM. Le site de la DIF compte environ 2000 salariés CEA et accueille quotidiennement environ 600 salariés d'entreprises extérieures. Il est situé à Bruyères-le-Châtel à environ 40 km au sud de Paris, dans l'Essonne.

Les missions de la DIF comprennent :

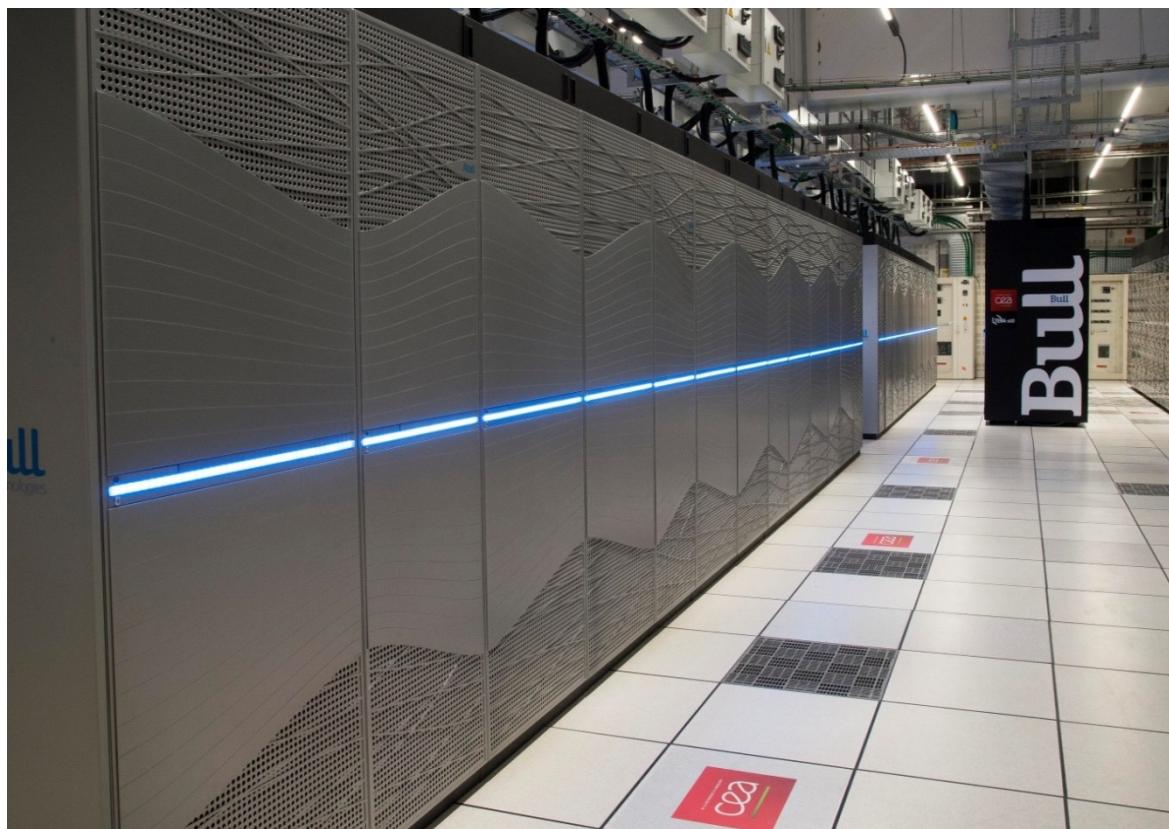
- La conception et garantie des armes nucléaires, grâce au programme Simulation. L'enjeu consiste à reproduire par le calcul les différentes phases du fonctionnement d'une arme nucléaire et à confronter ces résultats aux mesures des tirs nucléaires passés et aux résultats expérimentaux obtenus sur les installations actuelles (machine radiographique, lasers de puissance, accélérateurs de particules).;
- La lutte contre la prolifération et le terrorisme, en contribuant notamment au programme de garantie du Traité de Non-Prolifération et en assurant l'expertise tech-

nique française pour la mise en œuvre du Traité d’Interdiction Complète des Essais Nucléaires (TICE).

- L’expertise scientifique et technique, dans le cadre de la construction et du démantèlement d’ouvrages complexes ainsi que pour la surveillance de l’environnement et les sciences de la terre.
- L’alerte des autorités, mission opérationnelle assurée 24h sur 24, 365 jours par an, en cas d’essai nucléaire, de séisme en France ou à l’étranger, et de tsunami dans la zone Euro-méditerranéenne. La DIF fournit aux autorités les analyses et synthèses techniques associées.

Depuis 2003, le centre DAM-Île-de-France héberge le complexe de calcul scientifique du CEA, qui regroupe l’ensemble des supercalculateurs du CEA, et qui comprend à ce jour :

- Le supercalculateur Tera1000-1 pour les besoins du programme Simulation du CEA/-DAM, mis en service en 2016, dispose d’une puissance de calcul de 2,5 petaflops, c’est à dire capable d’effectuer 2,5 millions de milliards d’opérations par seconde. Il est complété en 2018 par Tera1000-2, autre composante du projet Tera1000, qui préfigure les architectures et technologies du futur supercalculateur qui sera installé à l’horizon 2020. Sa puissance de calcul est de 12,5 petaflops.



- Le supercalculateur Cobalt du Centre de Calcul pour la Recherche et la Technologie (CCRT), ouvert à la communauté civile de la recherche et de l'industrie, pour une puissance globale de 1,5 petaflops.



- Le supercalculateur IRENE, d'une puissance de 9 petaflops, deuxième élément d'un réseau de supercalculateurs de classe petaflopique destiné aux chercheurs de la communauté scientifique européenne. Ce supercalculateur est hébergé au TGCC (Très Grand Centre de Calcul) et exploité par les équipes du CEA, qui apporte ainsi sa contribution à la participation de la France au projet PRACE (Partnership for Advanced Computing in Europe)

3 Contexte du stage

En contact avec les équipes de recherche suédoises à la base du SVDAG, le stagiaire devra :

1. Tout d'abord, étudier et comprendre la théorie et l'implémentation des structures SVO/SVDAG/SSVDAG pour en évaluer les performances de rendu ray-tracing dans un contexte maîtrisé.
2. Un premier convertisseur VTK-AMR vers SVDAG/SSVDAG sera à implémenter.
3. Le travail se poursuivra par l'application de rendu réaliste de type lancé de rayons à base de SVDAG sur ces données issues de VTK-AMR. Plusieurs développements seront possibles en vue de prendre en compte la notion de variables ou de compresser encore plus les informations dans le cadre de rendu de films interactifs. Les développements effectués seront évalués aussi bien sur GPU que sur les nombreux accélérateurs et processeurs disponibles au CEA.

4 Objet du stage

4.1 Structures de données

4.1.1 Sparse Voxel Octree

Le Sparse Voxel Octree (SVO) est une représentation de données en arbre. Sa structure permet d'accéder à une information rapidement. Il est construit à partir d'une scène 3D non structurée, qui est ensuite découpée en grille dont les cases contenant un élément sont à leur tour découpées jusqu'à atteindre le niveau défini. Chaque feuille du dernier niveau devient alors un voxel.

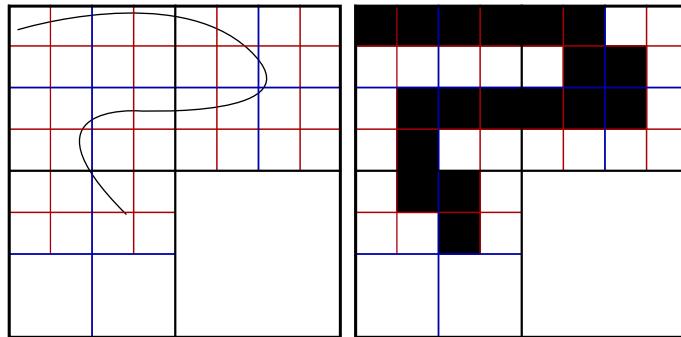


FIGURE 1: Découpage d'une scène 2D avec 3 niveaux et représentation des "voxels" obtenus

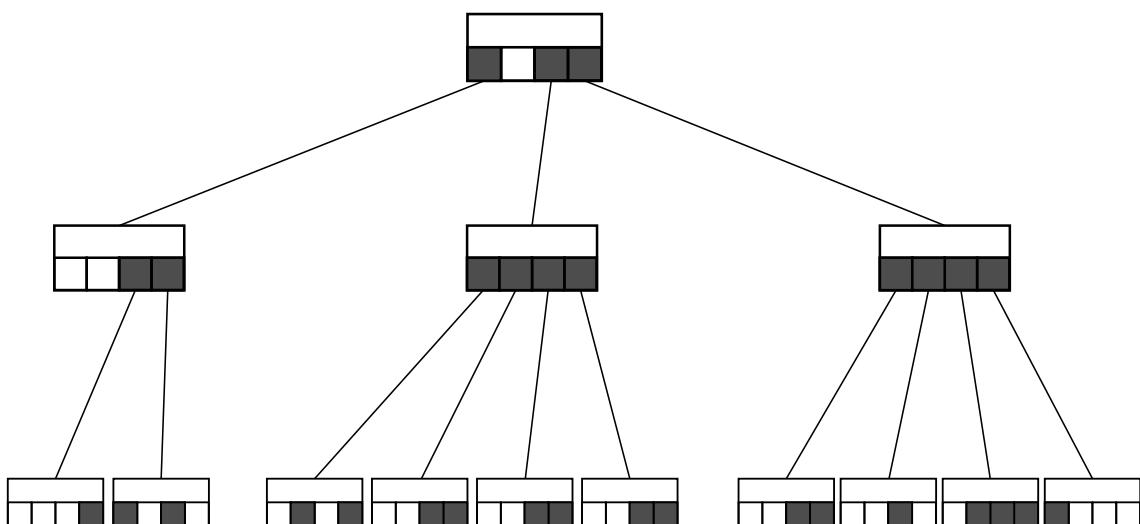


FIGURE 2: Représentation de l'arbre construit

4.1.2 Sparse Voxel Directed Acyclic Graph

Le Sparse Voxel Directed Acyclic Graph (SVDAG) est un arbre de voxel compressé. Il peut être construit à partir d'un SVO. Sa particularité est qu'un noeud peut être pointé par plusieurs parents, évitant ainsi de stocker plusieurs noeuds s'ils sont identiques. Il est également acyclique, nous pouvons descendre dans l'arbre mais pas remonter. Cela offre un gain de stockage supplémentaire en ne spécifiant pas les parents des noeuds. Tout cela permet de faire tenir en mémoire des structures trop volumineuses à la base et de rendre les accès mémoires plus rapides.

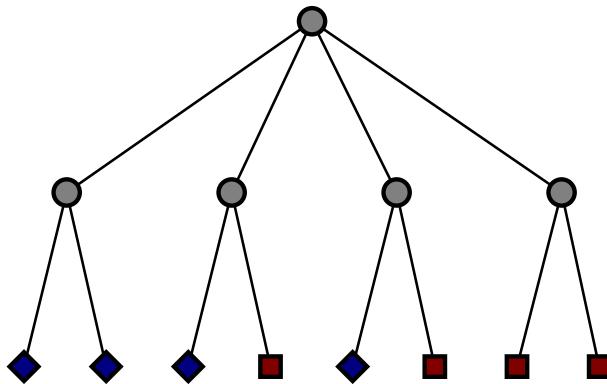


FIGURE 3: Représentation d'un arbre de données

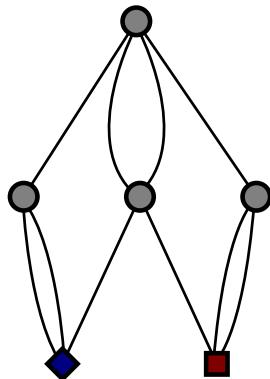


FIGURE 4: Représentation de l'arbre après compression via SVDAG

4.1.3 Adaptive Mesh Refinement

L'Adaptive Mesh Refinement (AMR) est un arbre semblable au SVO, à la grande différence que les voxels peuvent se situer à des niveaux différents dans l'arbre. Ceci implique qu'une scène peut contenir des voxels plus grossiers que d'autres si un affinage n'est pas nécessaire. En contrepartie, des zones peuvent être très détaillées sans allourdir la mémoire de façon excessive.

De plus, il est aisément de diminuer le niveau maximal des voxels pour obtenir une scène moins détaillée mais beaucoup plus fluide (voire en temps réel) afin de configurer rapidement les différents paramètres pour le rendu final.

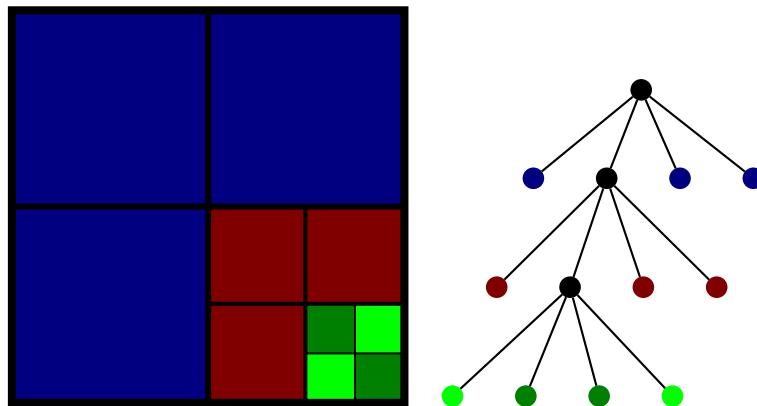


FIGURE 5: Grille de type AMR de niveau 3 avec son arbre associé

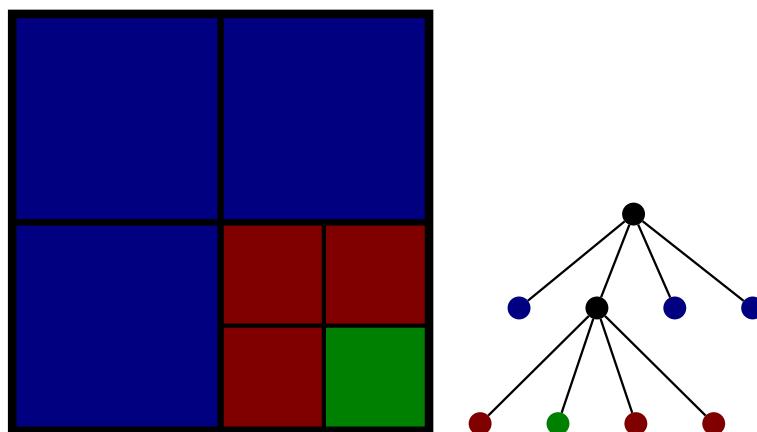


FIGURE 6: Grille de type AMR de niveau 2 avec son arbre associé

5 Travail réalisé

5.1 Prises de connaissances

La première semaine du stage fut dédiée à la connaissance et à la compréhension des structures de données citées plus haut. Pour cela, j'ai lu plusieurs papiers de recherche, de manière plus ou moins approfondie.

5.2 Installation de l'environnement

Plusieurs semaines ont ensuite été dédiées à l'installation de l'environnement nécessaire. N'ayant pas les droits d'administrateur, j'ai installé Spack en tant que gestionnaire de paquets.

J'ai ensuite installé VTK, contenant des modules pour l'AMR, et le programme SVDAG fourni, ainsi que toutes les dépendances.

5.3 Installation et fonctionnement du programme SVDAG

Contrairement à tous les autres code sources, celui du SVDAG a été fait sur Windows. J'ai donc du résoudre les problèmes de compatibilité, notamment causés par un paquet dépendant au système.

Le programme était maintenant fonctionnel, et le fichier d'entrée donné en exemple également. Afin d'utiliser d'autres fichiers en entrée, j'ai du chercher et installer un constructeur de SVO. Je l'ai également modifié de sorte à sortir les fichiers aux bons formats, pour qu'ils soient compatibles avec le SVDAG. De l'autre côté, j'ai ajusté l'entrée du SVDAG. J'ai également ajouté une option permettant d'écrire le SVDAG sur le disque pour ne pas avoir à le construire à chaque fois que l'on veut le visualiser.

5.4 Analyse des performances du SVDAG

Avec tout en place, j'ai pu tester le programme afin de comprendre comment il se comporte selon différents paramètres. J'ai également analysé ses performances avec quelques jeux de données connus.

5.4.1 Résultats



FIGURE 7: Rendu de modèles avec du raycasting sur SVDAG
De gauche à droite : Bunny (Stanford), Dragon (Stanford), Hairball (NVIDIA Research), Lucy (Stanford)

	Nombre Triangles	Taille non-structuré	Temps construction SVO	Temps construction SVDAG
Bunny	69 451	3.0 Mo	3.9s	4.3s
Dragon	871 414	33.8 Mo	5.7s	3.0s
Hairball	2 880 000	236.1 Mo	58.7s	47.7s
Lucy	28 055 742	533.1 Mo	77.5s	1.7s

	Nombre voxels	Taille SVO Structure + grandeurs	Taille SVDAG Structure + grandeurs	Taux compression
Bunny	3 591 666	14.4 + 57.5 Mo	1.3 + 8.5 Mo	86.3%
Dragon	2 688 970	10.8 + 43.0 Mo	1.0 + 6.4 Mo	86.2%
Hairball	41 521 450	166.1 + 664.3 Mo	13.3 + 94.6 Mo	87.0%
Lucy	1 540 004	6.2 + 24.3 Mo	0.6 + 3.6 Mo	86.3%

TABLE 1: Résultats avec des scènes voxélisées de résolution $1K^3$

5.4.2 Analyse

Tout d'abord, les modèles sont ordonnés selon le nombre de triangles qui les composent. Nous pouvons en premier lieu constater que la taille des modèles non-structurés dépend du nombre de triangles. Cependant, le nombre de voxels obtenus, quant à lui, dépend du volume du modèle dans le cube. Par exemple, la boule de poil est sphérique et "pleine" tandis que Lucy est un modèle creux et haut, rendant une grosse partie du cube vide.

Concernant le SVO, sa taille dépend du nombre de voxels, tandis que son temps de construction dépend du nombre de triangles.

La taille et le temps de construction du SVDAG sont liés au nombre de voxels. De plus, sa taille dépend également de la régularité des modèles. Dans notre cas, les modèles sont tous autant réguliers car la couleur est distribuée avec le même procédé.

La structure du SVDAG est environ 10 fois moins volumineuse que le SVO, et les grandeurs 85% plus légères. Ce qui donne un taux de compression total d'environ 86% ;

6 Expérience personnelle et professionnelle

6.1 Collaboration et contribution

Afin de comprendre certaines choses sur le SVDAG, j'ai sollicité Dan Dolonius et Ulf Assarsson, deux chercheurs suédois à l'origine du SVDAG. Malheureusement, aucune de mes questions n'a abouti à une réponse de leur part. Seules les solutions que je pouvais apporter (le problème de compatibilité notamment) a suscité leur intérêt.

De plus, le programme fourni est une ancienne version de ce qu'ils possèdent actuellement, et en avoir une plus récente nous permettrait beaucoup plus de choses. Je leur ai donc envoyé un mail récemment, en espérant avoir une réponse positive dans les prochains jours.

Quant à VTK et à l'AMR, Jérôme Dubois et Jacques-Bernard Lekien, contributeurs à VTK, m'ont beaucoup expliqué le fonctionnement de la structure AMR et sa construction, me donnant de bonnes bases pour la suite.

6.2 Problèmes rencontrés

6.2.1 Installations

Tout d'abord, l'accès aux machines internes a pris du temps, ce qui a retardé les installations de quelques jours. De plus, les restrictions de l'entreprises nous empêche d'utiliser curl et wget comme bon nous semble, il a donc fallu créer des miroirs. Utiliser Roméo pour cela était la meilleure solution, malheureusement une coupure de courant au Moulin de la Housse survint au mauvais moment. La seule façon de faire était de télécharger les paquets un par un via un navigateur et de les placer dans un miroir.

Concernant VTK, l'installateur télécharge des programmes de test, ce qui bloquait tout. Et les télécharger à la main n'était pas envisageable étant donné leur grand nombre. Nous avons donc attendu le retour de Roméo pour les télécharger puis les transférer en ssh.

De plus, l'installation de VTK pour python est en option, ce que je ne savais pas, et les programmes d'exemples qui m'étaient donnés étaient en python. Après plusieurs heures de recherches sur l'importation de VTK (car je pensais toujours que VTK python était installé), j'ai fini par restranscrire les exemples en C++, ce qui m'a permis au passage de comprendre les différentes étapes.

6.2.2 SVDAG

Le problème de compatibilité expliqué plus haut m'a également pris beaucoup de temps. En effet, en plus du paquet dépendant au système, l'appel d'une fonction template externe ne passait pas à la compilation. Les solutions sur internet liées à l'erreur retournée ne parlant que de la fonction en elle-même et non de son appel, j'ai du en partie comprendre les types des paramètres donnés (qui étaient des vecteurs très complexes) afin de trouver la source du problème.

7 Travaux futurs

Durant le reste du stage, mon travail consistera dans un premier temps à maîtriser la structure VTK-AMR afin de faire un convertisseur AMR vers SVDAG. Ensuite, des rendus seront faits à partir de données AMR afin d'évaluer les performances et les améliorations possibles.

8 Conclusion

9 Glossaire

Obligatoire dans le rapport

10 Bibliographie

Obligatoire (il existe des solutions pour générer automatiquement les bib)