

Ordonnancement Des Workflows Scientifiques Sur Le Cloud Avec Optimisation De L'énergie

C.Houcine Abdelkader T.Abdelghani Chabane

Département d'Informatique
Université des sciences et de la technologie d'Oran - Mohamed Boudiaf

10 juin 2023

Sommaire

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Définition Du Cloud Computing

Définition

Cloud Computing est un modèle pour permettre un omniprésent, commode, accès a la demande a un parc partage de ressources informatiques configurables (Réseaux, serveur, stockage, applications et services) qui peuvent être mis rapidement a disposition et libère avec une intervention et interaction minimale du fournisseur de services[5, 3].

Modèles De Déploiement

Il existe 5 modèles de déploiement de Cloud Computing[5][1] :

- Cloud privé.
- Cloud communautaire.
- Cloud public.
- Cloud hybride.
- Multi-Cloud Computing.

Modèles De Déploiement

Il existe 5 modèles de déploiement de Cloud Computing[5][1] :

- Cloud privé.
- Cloud communautaire.
- Cloud public.
- Cloud hybride.
- Multi-Cloud Computing.

Modèles De Déploiement

Il existe 5 modèles de déploiement de Cloud Computing[5][1] :

- Cloud privé.
- Cloud communautaire.
- Cloud public.
- Cloud hybride.
- Multi-Cloud Computing.

Modèles De Déploiement

Il existe 5 modèles de déploiement de Cloud Computing[5][1] :

- Cloud privé.
- Cloud communautaire.
- Cloud public.
- Cloud hybride.
- Multi-Cloud Computing.

Modèles De Déploiement

Il existe 5 modèles de déploiement de Cloud Computing[5][1] :

- Cloud privé.
- Cloud communautaire.
- Cloud public.
- Cloud hybride.
- Multi-Cloud Computing.

Modèles De Services

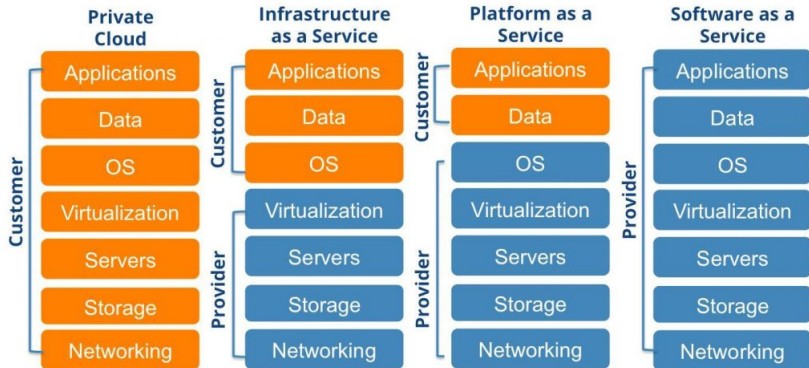


Figure – Les modèles des services Cloud quisted.net

Définition Du Workflow

Définition

Le workflow est la séquence de tâches, d'étapes et de décisions qui doivent être suivies pour mener à bien un processus spécifique. On peut le considérer comme un ensemble d'instructions qui décrivent comment un processus doit être effectué, y compris l'ordre dans lequel les tâches doivent être effectuées, qui est responsable de l'exécution de chaque tâche et ce qui doit se passer ensuite en fonction du résultat de chaque tâche[6].

Définition Des Workflow Scientifiques

Définition

Les workflows scientifiques sont des applications gourmandes en données représentant des sources de données distribuées et des calculs complexes dans divers domaines, à savoir l'astronomie, les sciences de l'ingénieur et la bioinformatique. Dans les environnements distribués, divers capteurs et processus expérimentaux génèrent un grand volume de données qui doivent être collectées et traitées dans des délais spécifiques[4].

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Avantages Des Workflow Pour La Science[2]

- Le partage des connaissances en étant disponibles en tant que services pour les scientifiques collaborateurs.
- La prise en charge de l'analyse communautaire des résultats.
- Le traitement d'énormes quantités de données.
- Les workflow sont capables de s'exécuter dans des environnements distribués.
- L'automatisation des étapes lors de la conception et de l'exécution du workflow permet aux scientifiques de se concentrer sur la résolution de leurs principaux problèmes scientifiques.
- Le pouvoir d'effectuer des simulations scientifiques de manière parallèle et automatisée.

Cycle De Vie Des Workflow Scientifiques[2]

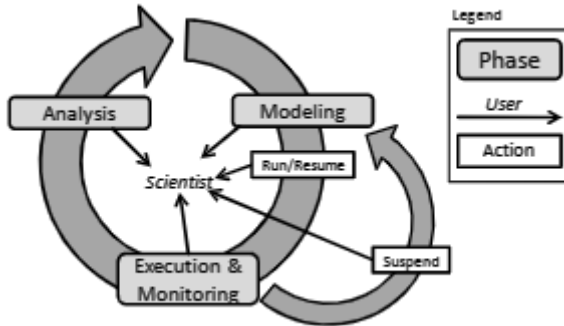


Figure – Cycle de vie des workflow scientifiques[2]

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 **Ordonnancement Des Workflow Scientifiques Sur Cloud**
 - **Ordonnancement Des Workflow Scientifiques**
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud**
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation**
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Objectives

- Notre algorithme considère les métriques suivantes :
 - Consommation de l'énergie
 - Coût
 - Maskespan
- Notre algorithme considère les contraintes suivantes :
 - Deadline
 - Les dépendances entre les tâches

Objectives

- Notre algorithme considère les métriques suivantes :
 - Consommation de l'énergie
 - Coût
 - Maskespan
- Notre algorithme considère les contraintes suivantes :
 - Deadline
 - Les dépendances entre les tâches

Description

Notre approche consiste à proposer un nouvel algorithme MOCS-OViC (Multi-Objective Cloud Scheduler with Optimized Virtual Machines Consolidation for Scientific Workflows) qui considère deux étapes principales pour la planification des workflows scientifiques sur le cloud :

- 1 L'étape de mappage pour affecter des tâches aux machines virtuelles (VM) appropriées. Pour cette étape, nous appliquons l'algorithme EViMA.
- 2 Nous abordons les questions de consommation d'énergie. Nous pensons que si la machine physique (PM) est sous-chargée ou surchargée, elle consomme plus d'énergie. Afin d'éviter ce problème, nous appliquons certaines stratégies de migration de VM au cours de l'étape d'exécution.

Description

Notre approche consiste à proposer un nouvel algorithme MOCS-OViC (Multi-Objective Cloud Scheduler with Optimized Virtual Machines Consolidation for Scientific Workflows) qui considère deux étapes principales pour la planification des workflows scientifiques sur le cloud :

- 1 L'étape de mappage pour affecter des tâches aux machines virtuelles (VM) appropriées. Pour cette étape, nous appliquons l'algorithme EViMA.
- 2 Nous abordons les questions de consommation d'énergie. Nous pensons que si la machine physique (PM) est sous-chargée ou surchargée, elle consomme plus d'énergie. Afin d'éviter ce problème, nous appliquons certaines stratégies de migration de VM au cours de l'étape d'exécution.

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

Lignes directrices

- 1 Cloud Computing
- 2 Workflow & Workflow Scientifique
- 3 Ordonnancement Des Workflow Scientifiques Sur Cloud
 - Ordonnancement Des Workflow Scientifiques
 - Objectives Ordonnancement
 - Modélisation
 - Représentation De Workflow
 - Datacenter
 - Allocation Des Machines Virtuelles (VM)
 - Formulation Du Problème D'optimisation Multi-objectives
 - État de l'art
- 4 Conception De Notre Approche
 - Mappage Des Tâches Sur Les Machines Virtuelles
 - L'emplacement Et La Consolidation Des VMs Sur Les PMs
- 5 Implémentation, Simulation & Discussion

MOCS-OViC

Algorithm 4.6 MOCS-OViC

Require: Workflow, set of VMs and set of VM types ($VM_{HEC}, VM_{MEC}, VM_{LEC}$), PMs

- 1: $wt_{ReadyPool} = clustered(wt_1, wt_2, wt_3, \dots, wt_n)$
- 2: Apply algorithm EViMA
- 3: Sort PMs based on capacity /availability in descending order
- 4: $VM_{candidateList} = \phi$
- 5: Add all VMs that are in under-loaded PMs to $VM_{candidateList}$
- 6: Add all VMs that are in over-loaded PMs to $VM_{candidateList}$
- 7: $j=0$
- 8: **while** $VM_{candidateList} \neq \phi$ and $j < \text{size}(PMs)$ **do**
- 9: pick PM_j
- 10: $VM_{selected}$ which results minimum energy is selected
- 11: **if** $\text{requirements}(VM_{selected}) \leq \text{availableCapacity}(PM_j)$ **then**
- 12: place $VM_{selected}$ on PM_j
- 13: Update $VM_{candidateList}$
- 14: **end if**
- 15: $j = j+1$
- 16: **end while** Until no more VM or no more PM
- 17: Migrate all VMs
- 18: **foreach** pm in PMs **do**
- 19: **if** $\text{size}(VM_{candidateList}) > 0$ **then**

Applications Workflow

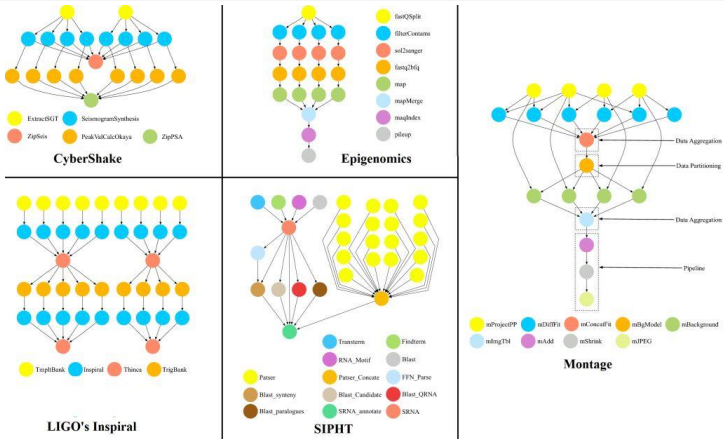


Figure – Les Workflows utilisés [confluence.pegasus.isi.edu]

Les Workflows De Petite Taille

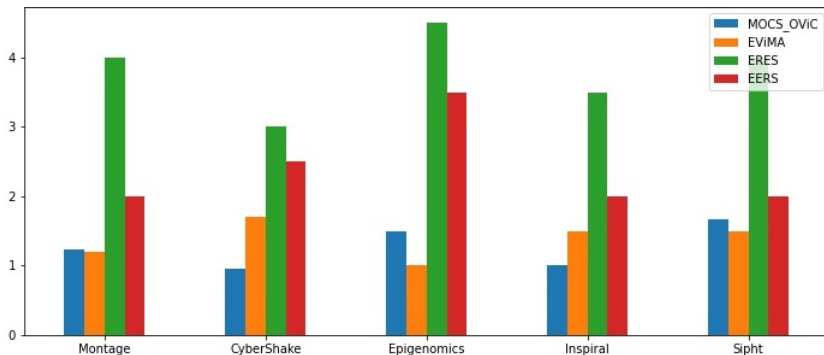


Figure – La consommation d'énergie dans les workflow de petite taille

Les Workflows De Taille Moyenne

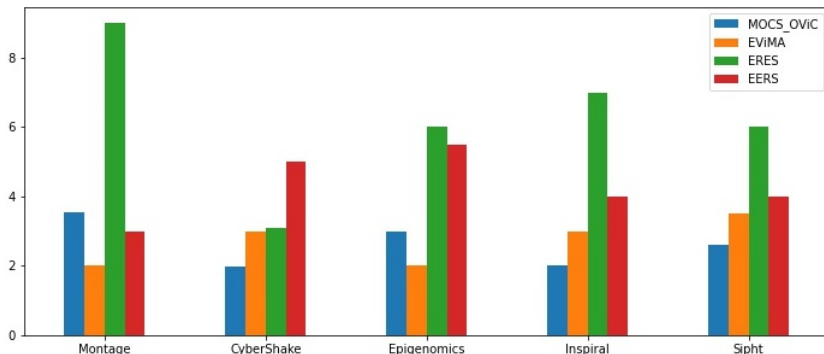


Figure – La consommation d'énergie dans les workflow de taille moyenne

Les Workflows De Grande Taille

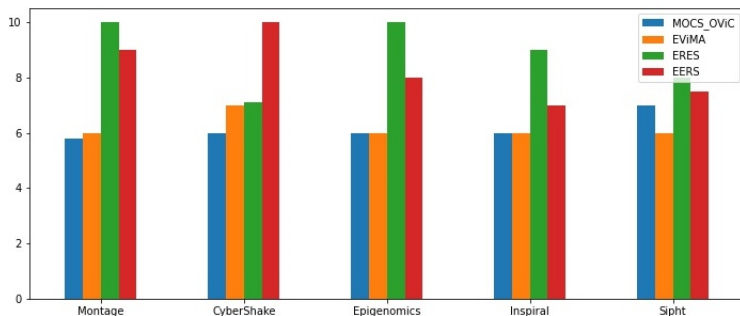


Figure – La consommation d'énergie dans les workflow de taille large

Les Workflows De Très Grande Taille

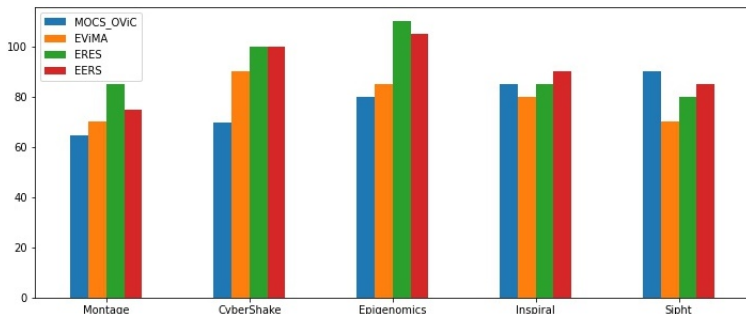


Figure – La consommation dans les workflows de très grande taille

Comparaison

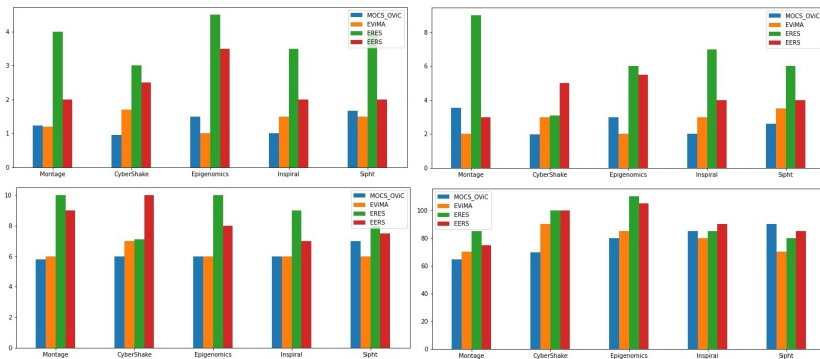


Figure – La comparaison de la consommation de l'énergie

Discussion

Avantage

Notre approche MOCS-OViC consomme moins d'énergie que les algorithmes ERES et EERS à tout moment sauf en 2 cas.

Inconvénient

MOCS-OViC peut économiser plus d'énergie que EViMA lorsque CyberShake et Inspiral sont exécutés pour toutes tailles et pour Epigenomics de très grande taille.

Les Workflows De Petite Taille

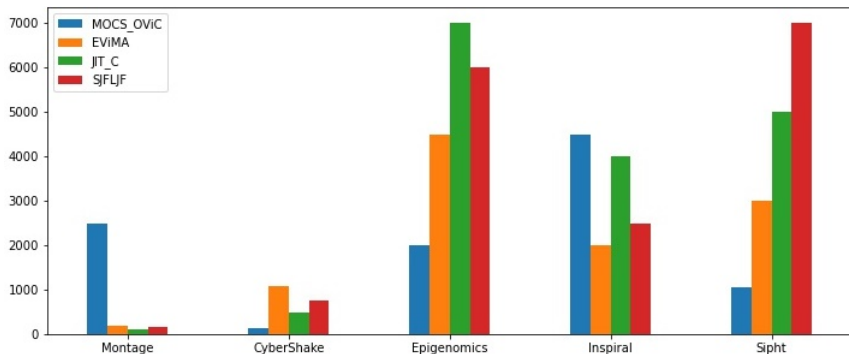


Figure – Le Makespan dans les workflow de petite taille

Les Workflows De Taille Moyenne

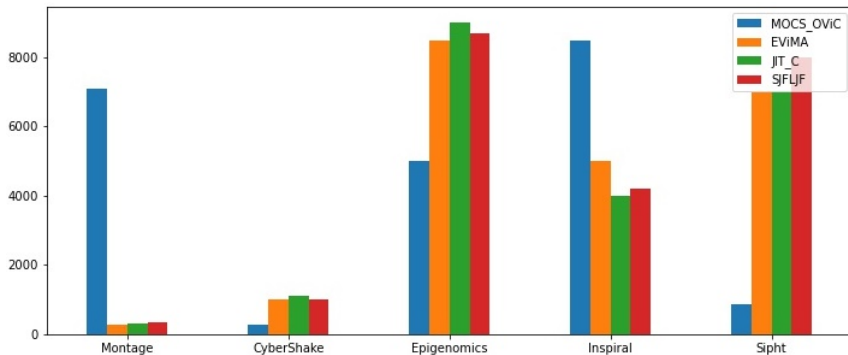


Figure – Le Makespan dans les workflow de taille moyenne

Les Workflows De Grande Taille

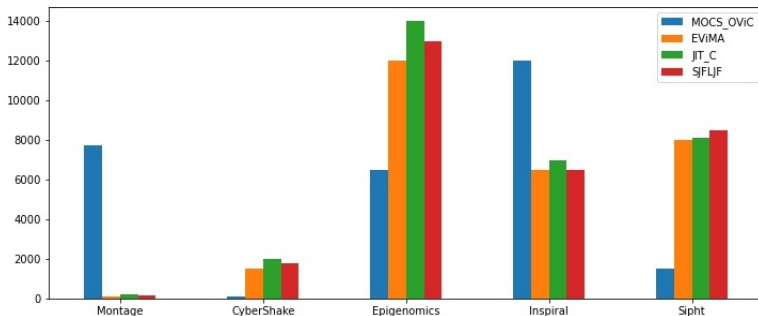


Figure – Le Makespan dans les workflow de taille large

Les Workflows De Très Grande Taille

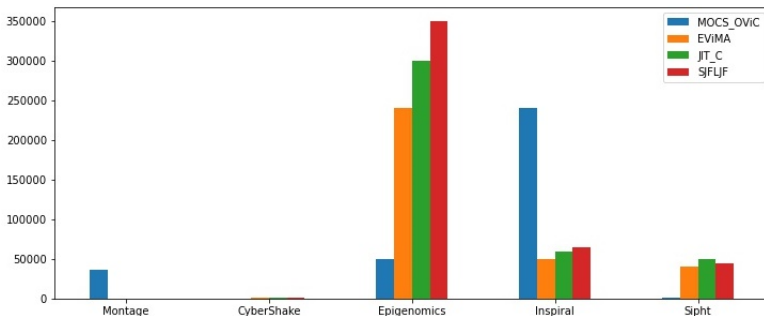


Figure – Le Makespan dans les workflows de très grande taille

Comparaison

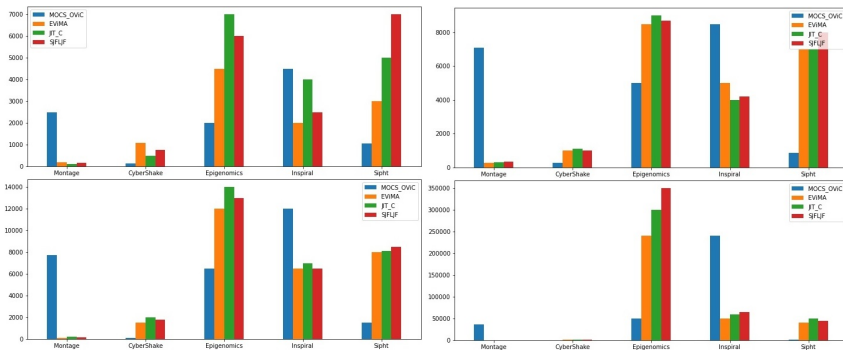


Figure – La comparaison du Makespan

Discussion

Avantage

Notre approche produit des plans d'exécution les plus courts pour CyberShake, Epigenomics et Sipht dans presque toutes les tailles parce que ces 3 workflows sont des applications gourmandes en CPU et notre algorithme MOCS-OViC équilibre les charges entre les hôtes mieux que les autres algorithmes.

Inconvénient

MOCS-OViC génère les pires plannings lorsque Montage et Inspiral sont exécutés. En effet, Inspiral nécessite beaucoup de mémoire, donc les migrations de machines virtuelles prennent plus de temps. Le montage est gourmand en transfert des données, notre algorithme ne considère pas le temps de transfert dans la génération des plans d'exécutions.

Les Workflows De Petite Taille

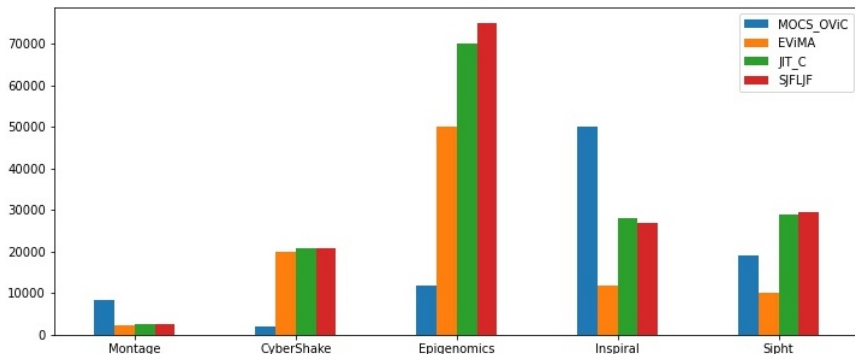


Figure – Le coût de l'exécution des workflow de petite taille

Les Workflows De Taille Moyenne

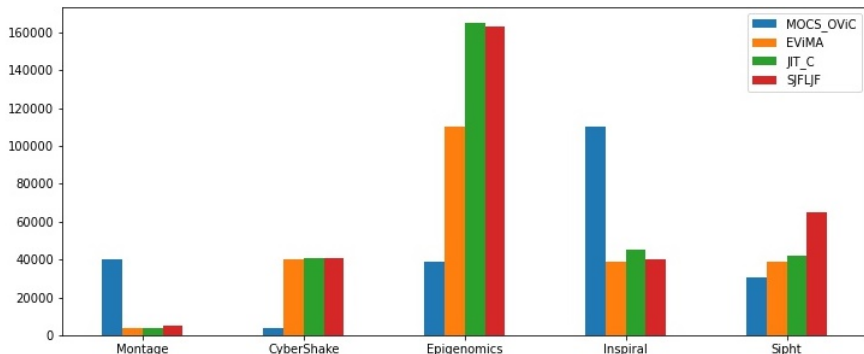


Figure – Le coût de l'exécution des workflow de taille moyenne

Les Workflows De Grande Taille

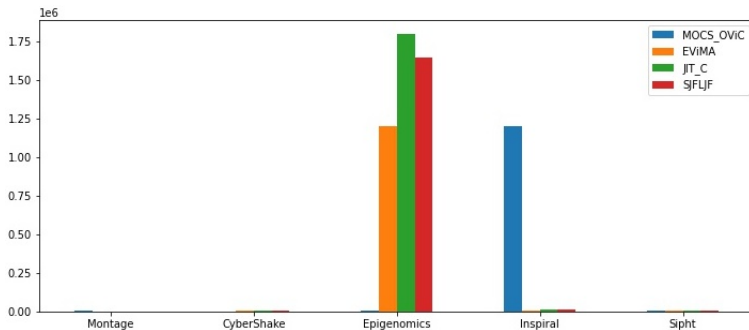


Figure – Le coût de l'exécution des workflow de taille large

Les Workflows De Très Grande Taille

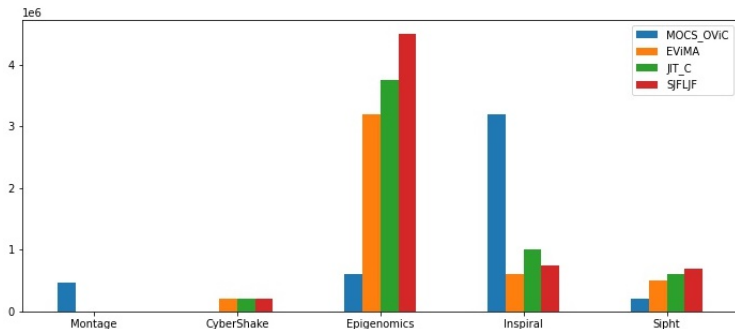


Figure – Le coût de l'exécution des workflows de très grande taille

Comparaison

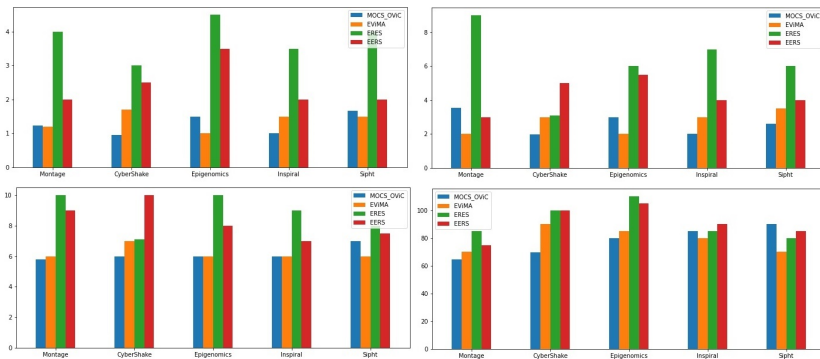


Figure – La comparaison des coûts de tous les algorithmes

Discussion

Avantages

MOCS-OViC produit des plus courts pour CyberShake, Epigenomics et SiphT dans presque toutes les tailles. En effet, notre algorithme génère une meilleure durée de fabrication, de sorte que l'utilisation des machines virtuelles est moindre.

Inconvénient

MOCS-OViC génère les pires plannings lorsque Montage et Inspiral sont exécutés. MOCS-OViC est un algorithme qui ne prend pas en charge les données. Comme on le sait, Inspiral est un flux de travail gourmand en mémoire et Montage est un algorithme gourmand en données.

Conclusion

- De nombreux travaux de recherche étudient le problème de la planification des workflow scientifiques sur le Cloud.
- On a étudié dans ce travail le problème d'optimisation du coût, du makespan et de l'énergie.
- Nous proposons notre approche MOCS-OViC qui :
 - se base sur EViMA pour mapper les tâches aux VMs.
 - utilise certaines politiques de placement et de consolidation des VMs pour optimiser la consommation d'énergie.
 - détermine ensuite quelles VMs à migrer depuis la source PM et quand.
 - le PM de destination est sélectionné pour placer la VM sur celui-ci.

Conclusion

- De nombreux travaux de recherche étudient le problème de la planification des workflow scientifiques sur le Cloud.
- On a étudié dans ce travail le problème d'optimisation du coût, du makespan et de l'énergie.
- Nous proposons notre approche MOCS-OViC qui :
 - se base sur EViMA pour mapper les tâches aux VMs.
 - utilise certaines politiques de placement et de consolidation des VMs pour optimiser la consommation d'énergie.
 - détermine ensuite quelles VMs à migrer depuis la source PM et quand.
 - le PM de destination est sélectionné pour placer la VM sur celui-ci.

Conclusion

- De nombreux travaux de recherche étudient le problème de la planification des workflow scientifiques sur le Cloud.
- On a étudié dans ce travail le problème d'optimisation du coût, du makespan et de l'énergie.
- Nous proposons notre approche MOCS-OViC qui :
 - se base sur EViMA pour mapper les tâches aux VMs.
 - utilise certaines politiques de placement et de consolidation des VMs pour optimiser la consommation d'énergie.
 - détermine ensuite quelles VMs à migrer depuis la source PM et quand.
 - le PM de destination est sélectionné pour placer la VM sur celui-ci.

Conclusion

- Nous avons évalué notre algorithme :
 - en exécutant des expériences et en comparant les résultats avec certains algorithmes existants.
 - en utilisant 5 workflow scientifiques de différentes tailles.
- Les résultats sont satisfaisants lorsque des workflow intensifs en CPU sont exécutés.
- Le délai d'acquisition et de terminaison d'instance ne sont pas pris en compte.

Conclusion

- Nous avons évalué notre algorithme :
 - en exécutant des expériences et en comparant les résultats avec certains algorithmes existants.
 - en utilisant 5 workflow scientifiques de différentes tailles.
- Les résultats sont satisfaisants lorsque des workflow intensifs en CPU sont exécutés.
- Le délai d'acquisition et de terminaison d'instance ne sont pas pris en compte.





Conclusion

- Nous avons évalué notre algorithme :
 - en exécutant des expériences et en comparant les résultats avec certains algorithmes existants.
 - en utilisant 5 workflow scientifiques de différentes tailles.
- Les résultats sont satisfaisants lorsque des workflow intensifs en CPU sont exécutés.
- Le délai d'acquisition et de terminaison d'instance ne sont pas pris en compte.

Conclusion

- Nous avons l'intention d'appliquer ce travail dans un environnement Fog Computing :
 - réduire la consommation d'énergie et les coûts opérationnels et l'impact environnemental.
 - réduire du makespan pour améliorer QoS et l'expérience utilisateur des applications Fog.
 - améliorer la fiabilité et la scalabilité et éviter la congestion du réseau.

Références

-  So many clouds-what's the difference ?, May 2020.
-  Katharina Görlach, Mirko Sonntag, Dimka Karastoyanova, Frank Leymann, and Michael Reiter.
Conventional workflow technology for scientific simulation.
Guide to e-Science : Next Generation Scientific Research and Discovery, pages 323–352, 2011.
-  Hachem Guerid.
Introduction au cloud computing, 3 2022.
-  Mandeep Kaur and Rajni Aron.
An energy-efficient load balancing approach for scientific workflows in fog computing.