



L'infonuagique pour les chercheurs

Décembre 2016

Calcul Canada est souvent interrogé au sujet des possibilités d'externalisation vers des nuages commerciaux. Nous avons considéré cette voie comme solution de rechange à l'achat et à l'exploitation de ressources de calcul informatique de pointe (systèmes de CHP, systèmes axés sur l'infonuagique et stockage) ou comme complément à ces ressources.

Calcul Canada communique régulièrement avec les principaux fournisseurs de services infonuagiques canadiens et a établi un dialogue ouvert avec ces derniers en ce qui a trait aux tarifs, aux caractéristiques et aux services proposés. Ces discussions ont permis d'établir de nombreux éléments de notre stratégie infonuagique, y compris l'infonuagique hybride commerciale et les services de nuages sécurisés et évolutifs (Elastic Secure Cloud ou ESC, voir la description ci-dessous). Elles ont également influencé une partie de la division du travail ainsi que les attentes des utilisateurs quant au soutien de l'infonuagique.

Il est actuellement beaucoup plus rentable de faire l'acquisition de cyberinfrastructures et d'en assurer l'exploitation interne que de faire appel à des fournisseurs de services infonuagiques commerciaux.

Calcul Canada a procédé à une analyse détaillée de l'acquisition récente de deux grosses grappes d'ordinateurs et d'équipement de stockage pour une valeur totale de plus de 15 M\$ CA en dépenses d'investissement. Les tarifs proposés par un important fournisseur de services infonuagiques commerciaux ont été utilisés pour effectuer les comparaisons. Les configurations sélectionnées se rapprochent le plus possible des caractéristiques des nouveaux systèmes de Calcul Canada : unités centrales, stockage, etc. Les ristournes applicables ont été calculées et les montants en devises états-uniennes convertis en devises canadiennes.

Le résultat principal est que les coûts totaux, sur cinq ans, chez un fournisseur commercial sont considérablement plus élevés que le coût total de possession, sur la même période de cinq ans, des systèmes nouvellement acquis. De plus, le fournisseur de services infonuagiques commerciaux n'est pas en mesure d'offrir la gamme de services requis par un grand nombre de chercheurs canadiens.

Les coûts de l'infonuagique sont de 4 à 10 fois plus élevés que le coût que représentent la possession et l'exploitation de nos propres grappes. Certains éléments sont vraiment beaucoup plus chers, notamment le stockage persistant qui coûte 40 fois plus que celui de Calcul Canada.

Qui plus est, le fournisseur de services infonuagiques n'offre pas les caractéristiques les plus importantes de deux des nouveaux systèmes nationaux de Calcul Canada, soit ARBUTUS, à l'Université de Victoria, et CÈDRE, à l'Université Simon-Fraser. Comparé à ARBUTUS, le nuage commercial n'offre pas les espaces de stockage local et dédié adéquats. Les chercheurs qui ont besoin de stockage sur nœud de haute performance seraient extrêmement mal servis par un fournisseur commercial. On parle ici de champ de recherche comme la physique et les sciences environnementales, entre autres.

Comparé à CÈDRE, le nuage commercial n'offre pas d'ensembles de nœuds en grappes étroitement liées avec interconnexion de haute performance à faible latence. Cela constituerait d'énormes défis pour une grande partie de la charge de travail de CÈDRE, qui est dominée par le calcul parallèle multinœud. Ce type de configuration est habituel pour la recherche exigeant de grands volumes de données et se trouve à la base d'une bonne partie du calcul informatique de pointe dans des secteurs importants comme la science de l'environnement, la modélisation de systèmes complexes, la physique, etc.

Le résultat général de la comparaison des coûts peut se résumer comme suit :

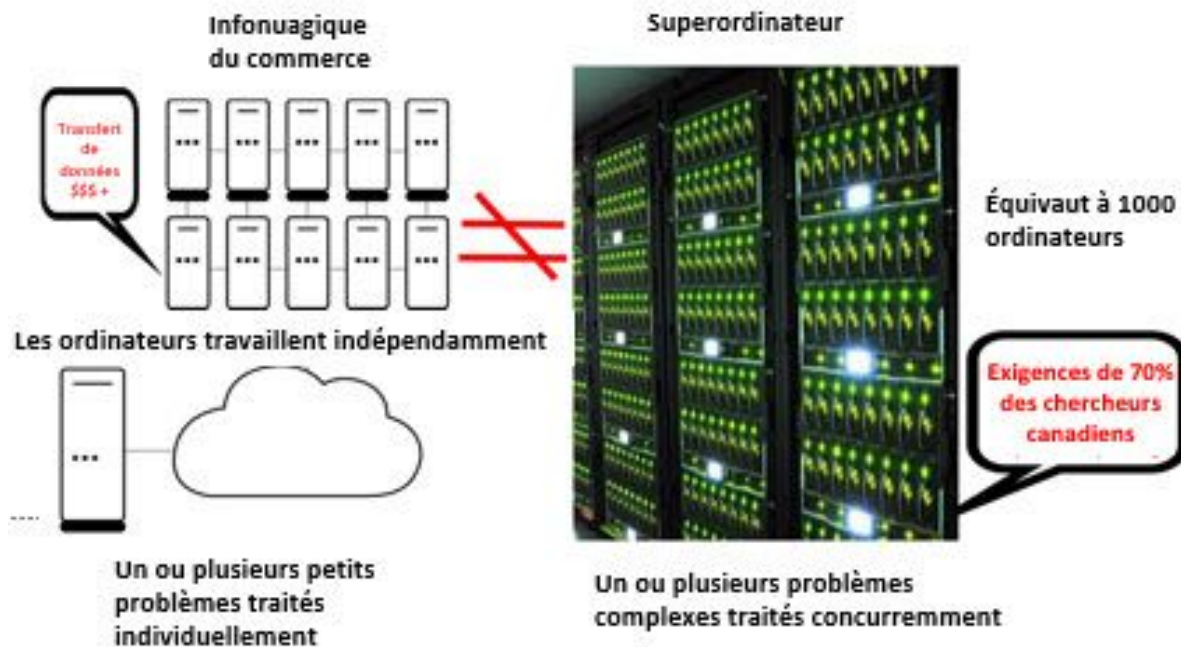
- Pour ARBUTUS, les coûts directs (en capital) de cinq ans d'une configuration externe comparable représentent approximativement 7,4 fois le coût de l'achat de GP1. Cela, sans compter que cette solution externe ne serait pas en mesure de répondre aux exigences de la charge de travail d'ARBUTUS en raison d'un manque de stockage local sur nœud.
- Pour ARBUTUS, lorsqu'on ajoute cinq ans de coûts d'exploitation, les coûts de la solution externe demeurent environ 4 fois plus élevés que ceux de l'achat.
- Pour CÈDRE, les coûts directs (en capital) de cinq ans d'une configuration externe comparable représentent environ 15,4 fois le coût de l'achat de GP1 : ils sont 6,6 fois plus élevés pour le calcul et 40 fois pour le stockage. Cela, sans compter que cette solution externe ne serait pas en mesure de répondre aux exigences de la charge de travail de CÈDRE en raison d'un manque de stockage local sur nœud (pour la partition OpenStack) et de l'absence d'interconnexion de haute performance à faible latence.
- Pour CÈDRE, lorsqu'on ajoute cinq ans de coûts d'exploitation, les coûts de la solution externe demeurent environ 10,9 fois plus élevés que ceux de l'achat.

Les fournisseurs de services infonuagiques commerciaux offrent souvent des calculateurs d'analyse de coût ou de rendement du capital investi (RCI) pour aider les utilisateurs potentiels à prendre une décision relativement à l'impartition vers des services infonuagiques commerciaux. Ces analyses ne sont pas adaptées à la situation de Calcul Canada, pour plusieurs raisons :

- Le modèle de financement de la cyberinfrastructure canadienne, y compris le financement jumelé, favorise les biens en équipements;
- Les sites membres possèdent et exploitent déjà des centres de données et d'autres infrastructures physiques, il n'est donc pas nécessaire de les construire, de les acheter ou de les louer;
- Le calcul de recherche accapare généralement la totalité des ressources des systèmes et les garde occupées en tout temps. Les calculs de RCI supposent souvent que les systèmes seront inutilisés la majeure partie du temps, alors que les systèmes de calcul de recherche fonctionnent normalement à plein régime en tout temps;
- Calcul Canada est doté de réseaux robustes et efficaces, depuis les réseaux locaux jusqu'aux réseaux régionaux et nationaux; ces réseaux entraînent peu de coûts directs pour la plateforme nationale et aucun coût additionnel par unité d'utilisation. Ces réseaux sont nécessaires, peu importe l'endroit où se trouve la cyberinfrastructure;
- L'expertise de Calcul Canada est déjà en place et appuie tous les aspects des besoins des utilisateurs, par conséquent, aucun coût différentiel ne doit être pris en compte pour le soutien aux utilisateurs, lequel est souvent prévu dans les analyses de RCI.
- Calcul Canada fournit une cyberinfrastructure entièrement canadienne, tandis que celles de la majorité des fournisseurs de services infonuagiques commerciaux et le soutien qui les accompagne ne le sont pas.

Calcul Canada surveille continuellement les tarifs et les services, et interagit avec les autres fournisseurs de services de calcul de recherche de même qu'avec les fournisseurs de services infonuagiques commerciaux afin d'évaluer les avantages possibles de variations dans l'équilibre des services. D'ailleurs, les perspectives d'approvisionnement pour les systèmes du défi 2 sont ouvertes à tous les fournisseurs qui sauront répondre aux exigences concernant les systèmes et les services, y compris les fournisseurs de services infonuagiques.

Calcul Canada a amorcé une collaboration avec les fournisseurs de services infonuagiques afin d'offrir les avantages de services infonuagiques hybrides aux utilisateurs qui le désirent. L'objectif est de réduire, pour les utilisateurs, le temps d'accès à une solution. En ayant recours aux nuages commerciaux, ceux-ci pourraient disposer de plus de ressources ou y accéder plus rapidement que sur les systèmes de Calcul Canada. Les services infonuagiques commerciaux peuvent également proposer des services ou des fonctionnalités que Calcul Canada n'offre pas. Les coûts d'utilisation de ces nuages commerciaux seraient cependant assumés par les utilisateurs et non par la plateforme nationale.



Aperçu des services infonuagiques de Calcul Canada

Depuis 2015, Calcul Canada offre des services infonuagiques au milieu canadien de la recherche universitaire. Les services fournis à ce jour visaient principalement les utilisateurs ayant déjà des compétences en administration des systèmes et en soutien. Le nuage de Calcul Canada se classe dans la catégorie des nuages privés partagés, occupant plusieurs sites et fondés sur la plateforme OpenStack. Cette plateforme constitue une norme en matière de services infonuagiques scientifiques.

Le concept de **fédération** se trouve au cœur de la plateforme de services infonuagiques de Calcul Canada. Le nuage de Calcul Canada est constitué d'une fédération de plusieurs sites. La fédération, dans ce contexte, est ainsi définie :

1. Système d'ouverture de session unifié. Les utilisateurs accèdent aux ressources infonuagiques en utilisant les mêmes nom d'utilisateur, mot de passe, UID, GID et autres caractéristiques d'identification à la grandeur du nuage fédéré de Calcul Canada.

2. Portabilité des charges de travail. Grâce à la plateforme OpenStack, une même machine virtuelle fonctionne de façon identique sur n'importe quel système infonuagique.
3. Expérience utilisateur semblable. En plus de la plateforme OpenStack, les autres aspects des systèmes infonuagiques sont essentiellement les mêmes, y compris les systèmes de fichiers et les logiciels.
4. Service d'assistance centralisé. Tous les services de soutien infonuagique sont accessibles par l'intermédiaire du système national d'assistance (à cloud@computecanada.ca), qui fournit des services locaux et régionaux à l'intérieur du modèle de service de Calcul Canada.
5. Service d'assistance local ou régional, le cas échéant. Le service d'assistance, sur place et régional, joue un rôle important dans la plateforme nationale et l'assistance infonuagique fait partie intégrante de ce service.

Stratégie infonuagique de Calcul Canada pour 2016–2018

Au cours de la prochaine année, Calcul Canada terminera sa plateforme infonuagique multirégionale, qui offrira un meilleur accès aux sites fédérés. La portabilité des charges de travail des utilisateurs sera encore plus grande et les utilisateurs pourront même automatiser le processus d'équilibrage des charges de travail entre les différents sites. D'ici là, le nuage fédéré de Calcul Canada sera réalisé par la mise en œuvre de pratiques communes aux différents sites infonuagiques. Cette mesure permettra d'assurer le déploiement des caractéristiques définies ci-dessus dans l'ensemble de la plateforme infonuagique fédérée, qui ne sera toutefois pas exploitée comme une plateforme infonuagique multirégionale unique.

Date cible pour la mise en œuvre de la plateforme infonuagique multirégionale : la date coïncide avec le déploiement des ressources infonuagiques du volet 2 du défi 2, soit vers le milieu de 2017.

L'amélioration des capacités de déploiement de nouvelles ressources en CHP pourrait également réduire la nécessité de ressources infonuagiques dédiées. Les charges de travail actuelles d'intégration de données expérimentales ou d'observation sont sensibles au temps, mais pourraient tolérer le posttraitement en lots. Dans des cas semblables, il peut être plus efficace d'utiliser des ressources de CHP dans le cadre du flux de travail, plutôt que d'affecter la totalité du flux de travail à des ressources infonuagiques dédiées.

Date cible pour que les flux de travail s'étendent à la fois à la plateforme infonuagique et aux ressources de CHP, y compris pour les mouvements de données : à être développé selon les besoins dans le cadre de la mise en place des services du volet 1 du défi 1, à partir de la fin de 2016.

Intégration des données et des capacités de calcul

L'intention de Calcul Canada est d'assurer l'intégration de l'ensemble des données et des capacités de calcul qui sont sur place à la plateforme fédérée et, par conséquent, de les conserver dans les limites des frontières nationales. Cela comprend les données d'authentification, les copies de sécurité, le stockage, ainsi que tout autre trafic ou échange de données.

Infonuagique hybride commerciale

Ayant conscience que ses ressources, tous types de calcul confondus, ne sont pas adéquates pour répondre à la demande actuelle des utilisateurs, Calcul Canada prévoit faciliter le déplacement de leur charge de travail entre son service infonuagique et les nuages commerciaux.

Certains utilisateurs de la plateforme de Calcul Canada peuvent souhaiter utiliser principalement des ressources infonuagiques commerciales et faire appel de temps à autre aux systèmes de calcul et de stockage de Calcul Canada. D'autres utilisateurs peuvent souhaiter utiliser principalement les ressources de Calcul Canada et faire occasionnellement appel aux ressources infonuagiques commerciales.

Calcul Canada n'a pas l'intention de payer pour les ressources infonuagiques commerciales dans ces situations. Toutefois, il est dans l'intérêt de toutes les parties concernées de simplifier l'accès bidirectionnel aux ressources de calcul et aux données. Par conséquent, Calcul Canada a proposé à la FCI qu'une portion des efforts de déploiement du volet 2 du défi 2 soit consacrée à l'intégration de services infonuagiques hybrides bidirectionnels aux ressources infonuagiques et de CHP existantes. L'analyse initiale de ces capacités se poursuit, mais les efforts de développement et de soutien à grande échelle dépendront de l'évaluation que fera la FCI de la proposition relative au volet 2 du défi 2.

Date cible pour la mise en œuvre des ressources infonuagiques hybrides bidirectionnelles : dans le cadre des activités du volet 2 du défi 2, à partir du milieu de 2017. Analyse initiale en cours, du milieu à la fin de 2016.

Plateforme infonuagique sécurisée et évolutive

Le concept de nuage sécurisé et évolutif (NSE) a été proposé à la FCI dans le cadre du volet 2 du défi 2. Le NSE renforcera le modèle infonuagique fédéré en améliorant ses capacités d'isolation de données et de calcul dans des partitions infonuagiques. Un ou plusieurs nœuds ou machines virtuelles pourraient être alloués à un projet infonuagique pour le NSE. Ces nœuds auraient accès à des partitions de données qui seraient inaccessibles aux autres locataires de la plateforme infonuagique. S'ils le désirent, les locataires du NSE pourraient utiliser leurs propres services d'authentification et d'autorisation plutôt que ceux de Calcul Canada. D'autres processus de vérification et procédures seraient également développés, comme la certification des administrateurs de système, les sauvegardes chiffrées ou hors site, ou l'isolation de portions du trafic du réseau. La polyvalence et l'évolutivité d'une plateforme infonuagique sécurisée peuvent être réalisées par la modification, avec le temps, des paramètres quantitatifs et de la configuration des ressources affectées aux divers projets. Les concepts de nuage sécurisé sont déjà en place dans plusieurs des établissements membres de Calcul Canada et serviront de bases au NSE dans la fédération infonuagique.

Date cible pour le déploiement d'un nuage sécurisé et évolutif dans l'ensemble de Calcul Canada : dans le cadre des activités du volet 2 du défi 2, à partir du milieu de 2017.

Vision à long terme de l'intégration nuage-CHP et du CHP-service

Calcul Canada suit de près les développements de l'industrie de l'informatique en recherche ainsi que dans d'autres secteurs d'activité. À long terme, l'industrie prévoit que les distinctions actuelles entre le calcul de haute performance (CHP) par traitement en lots et les plateformes infonuagiques du genre OpenStack s'estomperont, pour éventuellement fusionner en un tout.

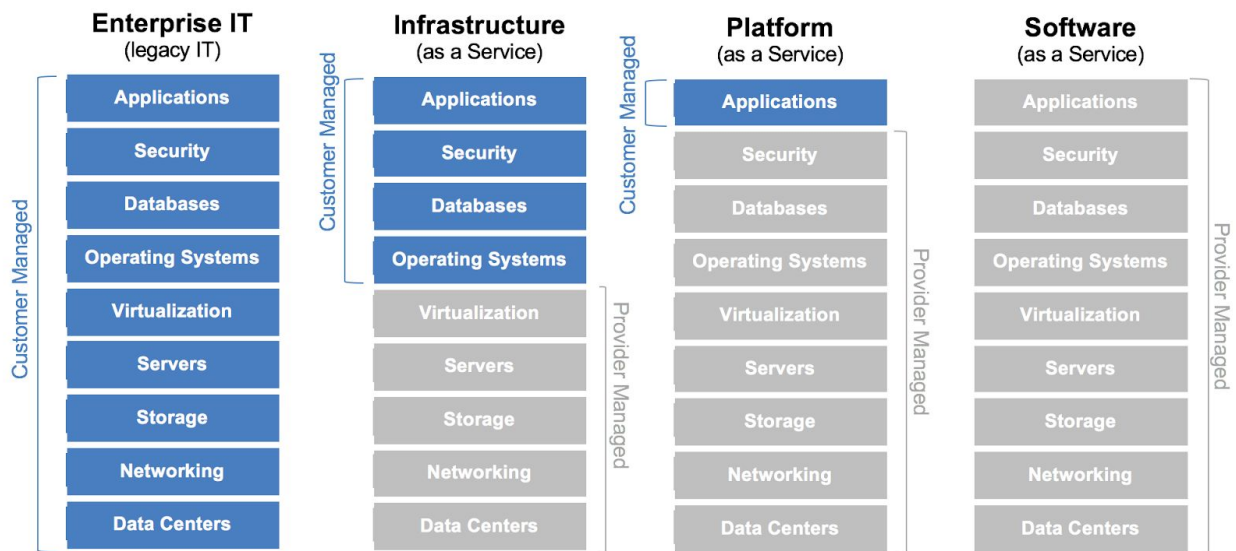
L'une des étapes vers une telle fusion sera l'allocation de *grappes* virtuelles et non seulement de machines virtuelles (MV) sur les plateformes infonuagiques. Une grappe virtuelle, qu'elle soit créée dans un système OpenStack ou par le programme de commande d'un système de CHP par lots, permettra à l'utilisateur de développer un environnement logiciel personnalisé d'une façon similaire à l'infrastructure-service (IaaS) sur une MV. La différence sera que le flux de

travail englobera les calculs parallèles, comprenant l’utilisation efficace de réseaux avec interconnexion de haute performance, et les systèmes de fichiers parallèles.

Calcul Canada prévoit qu’un nombre grandissant d’utilisateurs ou de projets d’utilisateurs souhaiteront utiliser des environnements personnalisés de ce type. Tout comme c’est le cas avec les MV d’aujourd’hui, les choix de grappes virtuelles seront persistents pour un usager donné et pourront être appliqués sur différents systèmes physiques.

De même, comme c’est le cas sur le nuage fédéré actuel, Calcul Canada offrira l’environnement logiciel sous-jacent pour le soutien des grappes virtuelles.

Infrastructure-service (IaaS)



Mention d’auteur : The enterprise Cloud Blog

<https://mycloudblog7.wordpress.com/2013/06/19/who-manages-cloud-iaas-paas-and-saas-services/>