



Bulletin technologique de Calcul Canada

Novembre 2017

1. Vue d'ensemble

Le présent bulletin technologique s'adresse aux partenaires et aux fournisseurs de Calcul Canada. Il fournit un aperçu de l'état actuel du programme de renouvellement technologique, entrepris grâce à l'initiative de cyberinfrastructure de la Fondation canadienne pour l'innovation (FCI), dont la mise en place commencée depuis 2015 s'étendra jusqu'en 2019. Il décrit également des plans de croissance à venir.

Le présent document s'ajoute aux versions de 2015 et 2016¹. Il se veut utile et exact au moment de sa rédaction, mais ne garantit aucune action ni aucun résultat en particulier. Toute l'activité de Calcul Canada fait l'objet d'une surveillance de la part de la FCI ainsi que des établissements membre, de ses partenaires régionaux et d'autres intervenants.

2. Vue d'ensemble de Calcul Canada

Calcul Canada (CC) existe afin de permettre l'excellence en recherche et l'innovation dans l'intérêt du Canada, grâce au déploiement d'une infrastructure de calcul informatique de pointe (CIP) ultramoderne, efficace et durable, soutenue par une expertise de premier ordre couvrant de multiples disciplines. Calcul Canada se concentre sur l'intégration harmonieuse de la technologie et des services, s'efforçant de les rendre aisément accessibles et utiles aux chercheurs de tout le pays. CC est une composante essentielle de l'infrastructure de recherche numérique multicouche du Canada.

Calcul Canada concentre ses activités dans quatre secteurs stratégiques² :

- la fourniture d'excellents services de soutien à la recherche;

¹ [Bulletin technologique 2016 de Calcul Canada](#)

² [Plan stratégique de Compute Canada](#) 2014-2019

- l'édification d'une capacité de recherche;
- la création de retombées positives pour les Canadiens grâce à la recherche qui est soutenue;
- l'aspiration à l'excellence organisationnelle et à la durabilité.

Plusieurs éléments essentiels sont requis afin d'atteindre ces objectifs stratégiques. Notamment : une infrastructure de calcul informatique de pointe (CIP) ultramoderne; un personnel hautement qualifié centré sur le client; un environnement sécurisé et un engagement à l'excellence professionnelle et organisationnelle.

Calcul Canada soutient des projets de toutes tailles, que ce soit l'affaire d'un seul membre du corps professoral ou bien le plus grand des « projets scientifiques d'envergure » au pays. L'« Équipement » de CC comprend l'ensemble des infrastructures et des services de CIP qui sont gérés par la fédération et répartis dans divers emplacements au Canada. Cet Équipement permet la poursuite d'une recherche concurrentielle sur le plan international dans beaucoup de disciplines. Il soutient également les chercheurs de petits et grands établissements de recherche d'un océan à l'autre, dans de nombreux secteurs. Les services fournis par Calcul Canada sont destinés à combler les besoins variés du milieu de la recherche comme le calcul haute performance (CHP) classique en grappes étroitement liées, le calcul à haut débit (CHD), les mégadonnées et l'analytique des données massives, le traitement en série, l'infonuagique et le stockage en nuage, la visualisation et d'autres solutions technologiques.

La fédération de Calcul Canada fonctionne comme un établissement membre. Une entreprise sans but lucratif sert de gestionnaire de projet et la majorité des universités de recherche du Canada ainsi que d'autres établissements de recherche en sont membres. À la fin de 2017, on compte 37 établissements membres, répartis dans quatre régions.

3. Renouvellement technologique

La Fondation canadienne pour l'innovation (FCI) est la principale agence de financement du projet Calcul Canada. Un apport équivalent de la part des provinces, des établissements membres et d'autres sources s'ajoute à ce soutien. La FCI, elle-même financée par le ministère de l'Innovation, des Sciences et du Développement économique (ISED) du Canada, subventionne 17 projets dans le cadre du programme des initiatives scientifiques majeures (ISM), dont Calcul Canada. CC a commencé en 2017 son deuxième cycle de financement de 5 ans accordé par le programme des ISM qui fournit ainsi un soutien majeur pour l'exploitation et la maintenance de l'Équipement.

Un projet de renouvellement technologique d'envergure est en cours grâce au programme Défi 2 de l'Initiative sur la cyberinfrastructure de la FCI. Le volet 1 du Défi 2 est presque terminé et la mise en œuvre de volet 2 est amorcée. Ensemble, ces deux volets représentent un investissement évalué à environ 125 M\$ sur trois ans et remplaceront la presque totalité de la cyberinfrastructure composant l'Équipement.

Dans le cadre du Défi 2, la FCI a guidé la consolidation de nombreux centres de données. Auparavant, 27 établissements membres exploitaient la cyberinfrastructure, qui comptait 58 systèmes dans l'ensemble du Canada. La nouvelle cyberinfrastructure, acquise grâce aux volets 1 et 2 du Défi 2, est en cours d'installation dans 5 nouveaux sites hôtes nationaux.

À l'issue d'un concours tenu entre les établissements membres de Calcul Canada, quatre sites ont été sélectionnés pour accueillir les nouveaux systèmes et leurs services connexes : l'Université de Victoria (UVic), l'Université Simon-Fraser (SFU), l'Université de Waterloo (Waterloo) et l'Université de Toronto (UofT). Le tableau 1 présente un sommaire des investissements du volet 1.

Tableau 1 : Sommaire des investissements du volet 1 (comprend les montants dépensés; les contributions en nature comptent pour un 20 % supplémentaire). Valeur totale : 75 M\$.

UVic	Calcul (GP1)	4 390 \$
	Stockage et infra	1 610 \$
SFU	Calcul (GP2)	11 550 \$
	Stockage et infra	5 150 \$
Waterloo	Calcul (GP3)	10 615 \$
	Stockage et infra	4 985 \$
UofT	Calcul (massivement parallèle)	16 850 \$
	Stockage et infra	2 850 \$
Logiciels et services	Pour les 4 sites	2 000 \$
	Total général	60 000 \$

Pour le volet 2, un autre concours a été tenu afin de déterminer un site national supplémentaire, ou plus. Calcul Canada a soumis sa proposition³ pour le volet 2 en mai 2016 et a été informé des résultats en septembre 2016, par contre la FCI n’a pas encore annoncé les résultats du concours pour le volet 2, c’est pourquoi le présent document ne nomme pas les sites sélectionnés. Toutefois, la planification est en cours pour le renouvellement technologique du volet 2, qui sera en grande partie terminé en septembre 2018.

Tableau 2 : Sommaire des investissements du volet 2 (comprend les montants dépensés; les contributions en nature comptent pour un 20 % supplémentaire). Valeur totale : 50 M\$.

Extension de GPx	Calcul	9 283 \$
	Stockage et infra	3 468 \$
Extension de GPx	Calcul	0
	Stockage et infra	1 000 \$
GP4	Calcul	16 788 \$
	Stockage et infra	7 463 \$
Systèmes expérimentaux	Pour les 3 sites	1 500 \$
Logiciels et services	Pour les 3 sites	500 \$
	Total général	40 000 \$

Le volet 1 a donné naissance à quatre nouveaux systèmes informatiques comportant de l’espace de stockage. Deux sites, Waterloo et SFU, ont été choisis comme sites de « stockage permanent » et recevront les sauvegardes et les autres données stockées à long terme. Les quatre sites ont travaillé de concert pour favoriser le développement d’un logiciel et d’un intergiciel de gestion des données de recherche et des services connexes afin d’améliorer la capacité de Calcul Canada à répondre aux besoins connus de ses utilisateurs. Le volet 1 mettait aussi l’accent sur le stockage, consacrant de 20 à 24 % de toutes les dépenses au stockage persistant (sur disque ou bande magnétique). Divers services d’infrastructure, comme un LDAP national et des systèmes de surveillance centralisés ont également été déployés dans le cadre du volet 1.

Avec le volet 2, un cinquième site national sera ajouté. Celui-ci sera l’hôte d’un autre gros système universel, comportant de nombreux types de nœuds différents afin de répondre à une

³ [Rapport complet du Défi 2 de l’initiative de cyberinfrastructure](#)

large gamme de besoins informatiques, comprenant des tâches pouvant demander jusqu'à 1024 cœurs de processeur. Les systèmes universels comprennent tous des partitions infonuagiques, de plus le système GP1 (à UVic) est entièrement centré sur les services infonuagiques OpenStack.

Le volet 2 verra également une extension importante de l'un des systèmes universels du volet 1, dont l'augmentation de la capacité de stockage, y compris pour les sauvegardes et le stockage de proximité. Les logiciels et intergiciels pour la gestion des données de recherche et d'autres besoins seront compris, cette fois encore. Une nouvelle catégorie d'investissement s'ajoute pour les systèmes expérimentaux. Elle permettra des investissements relativement petits dans des technologies qui intéressent un ou plusieurs utilisateurs actuels et qui fourniront de l'information utile pour les choix technologiques à venir.

Tableau 3 : Principaux systèmes de calcul des volets 1 et 2

Système	Description	Date de mise en production (prévue)
(GP1) Arbutus – Université de Victoria	Système infonuagique OpenStack. Plus de 7000 cœurs de processeur. MV persistantes et sur demandes.	8 septembre 2016
(GP2) Cedar – Université Simon-Fraser	Grappe universelle. Plus de 20 000 cœurs de processeur, plus de 580 000 dispositifs d'UTG. De nombreux types de nœuds.	25 juillet 2017
(GP3) Graham – Université de Waterloo	Grappe universelle. Plus de 33 000 cœurs de processeur, plus de 360 000 dispositifs d'UTG. De nombreux types de nœuds.	28 juillet 2017
(LP1) Niagara – Université de Toronto	Système massivement parallèle. 60 000 cœurs, Skylake.	(Début 2018)
GP4 – Site 5	Grappe universelle.	(Milieu 2018)

4. Composantes principales de la plateforme nationale

Les sections qui suivent donnent des détails sur les principales composantes du renouvellement technologique de Calcul Canada, qui comprend la mise à jour de la plateforme nationale de calcul informatique de pointe. Ce bulletin ne contient pas tout l'historique des besoins des utilisateurs et des choix d'infrastructures précédents qui ont mené à ces plans. Cet historique est inclus dans les propositions de cyberinfrastructure et d'ISM qui sont disponibles à www.computecanada.ca/publications.

4.1 Grappes universelles

Avant de présenter à la soumission du volet 1, Calcul Canada a procédé à des consultations à la grandeur du pays pour comprendre les besoins actuels et à venir des utilisateurs. Ces consultations visaient à déterminer les points faibles et les déficiences de la plateforme nationale telle qu'elle était alors ainsi qu'à prévoir les futurs besoins en infrastructure. L'une des principales conclusions a été qu'il existe un besoin considérable pour des systèmes informatiques conçus pour des charges de travail variées, allant d'un seul cœur à environ mille cœurs. Différents types de nœuds étaient nécessaires, y compris des nœuds d'UTG, de grands nœuds de mémoire et des nœuds offrant du stockage local (sur le nœud). Les systèmes de ce type ont été nommés grappes universelles (en anglais *general purpose*, ou GP) et le volet 1 a mis en place deux grands systèmes universels et un autre plus petit centré principalement sur l'exploitation d'un nuage OpenStack.

Arbutus (Université de Victoria) : Arbutus (anciennement « GP1 ») est un nuage OpenStack dont la fonction première est d'exploiter des machines virtuelles et d'exécuter d'autres charges de travail infonuagiques. Ce système possède 6944 cœurs de processeur répartis sur 248 nœuds, possédant chacun de l'espace de stockage sur le nœud et une capacité réseau de 10 Gb. Il accède à 1,6 Po de stockage persistant, principalement par Ceph, avec une configuration à triple redondance. Ce système est entré en fonction en septembre 2016 en tant qu'extension du système « Cloud West » de Calcul Canada. En 2018, une extension d'environ 1/3 de la capacité de calcul et de stockage d'Arbutus aura lieu.

Cedar (Université Simon-Fraser) : Le système Cedar (anciennement GP2) est une grappe hétérogène convenant à diverses charges de travail. Ce système, refroidi par liquide, est installé dans le centre de données de l'édifice Water Tower de l'Université Simon-Fraser. Il comprend plus de 27 000 cœurs de processeur et 576 dispositifs d'UTG, des nœuds de mémoire de base (128 Go) et des grands nœuds (256 Go), ainsi que des nœuds « bigmem »

Bulletin technologique de Calcul Canada – novembre 2017 : Page 6



de 512 Go, 1,5 To et 3 To de mémoire. Pour plus d'information : <https://docs.computecanada.ca/wiki/Cedar>

Graham (Université de Waterloo) : Le système Graham (anciennement GP3) a une architecture semblable à celle de Cedar, ainsi qu'une pile de logiciels partagés et des caractéristiques uniformisées de configuration qui permettent la portabilité des charges de travail et facilitent l'utilisation. Tout comme Cedar, Graham accueille une petite partition OpenStack et possède divers types de nœuds. Il comprend plus de 33 000 cœurs de processeur et 320 dispositifs d'UTG. Dans la topologie, les îlots ont 1280 cœurs et l'un d'eux en a 4096. Pour plus d'information : <https://docs.computecanada.ca/wiki/Graham>

GP4 (le site n'est pas encore annoncé) : Le système GP4 aura un profil similaire à Cedar et Graham et fournira aux utilisateurs une expérience visuelle et intuitive cohérente avec ces systèmes. Il comportera divers types de nœuds, dont des nœuds avec accélérateurs. Il fournira une petite partition OpenStack fédérée à d'autres systèmes universels. La mise en production complète est prévue pour la deuxième moitié de 2018.

4.2 Calcul massivement parallèle

Beaucoup d'utilisateurs exécutent des tâches qui excèdent la conception cible de Cedar, Graham ou GP4. Le système de type massivement parallèle est conçu pour supporter ces plus grosses tâches ainsi que des tâches plus petites traitant des mégadonnées, et d'autres utilisations. Le volet 1 a financé le système Niagara, qui est le seul système de ce type soutenu jusqu'à maintenant par la FCI dans le cadre du Défi 2.

Niagara (Université de Toronto) : le système NIAGARA (anciennement LP) sera déployé à la fin de 2017 et entrera en production d'ici avril 2018. Il comportera 60 000 cœurs de processeur, une couche de stockage de mémoire tampon de paquets et une interconnexion de haute performance avec routage adaptatif. Pour plus d'information : <https://docs.computecanada.ca/wiki/Niagara>

4.3 Cyberinfrastructure nationale de données

La cyberinfrastructure nationale de données (CND) fournit aux utilisateurs divers types de stockage en ligne. Elle permet également le réseautage entre les sites afin de rendre possible le transport de données pour une vaste gamme de besoins. Un logiciel/intergiciel comme [Globus](#), est utilisé pour faciliter l'accès aux données et les mouvements de données.

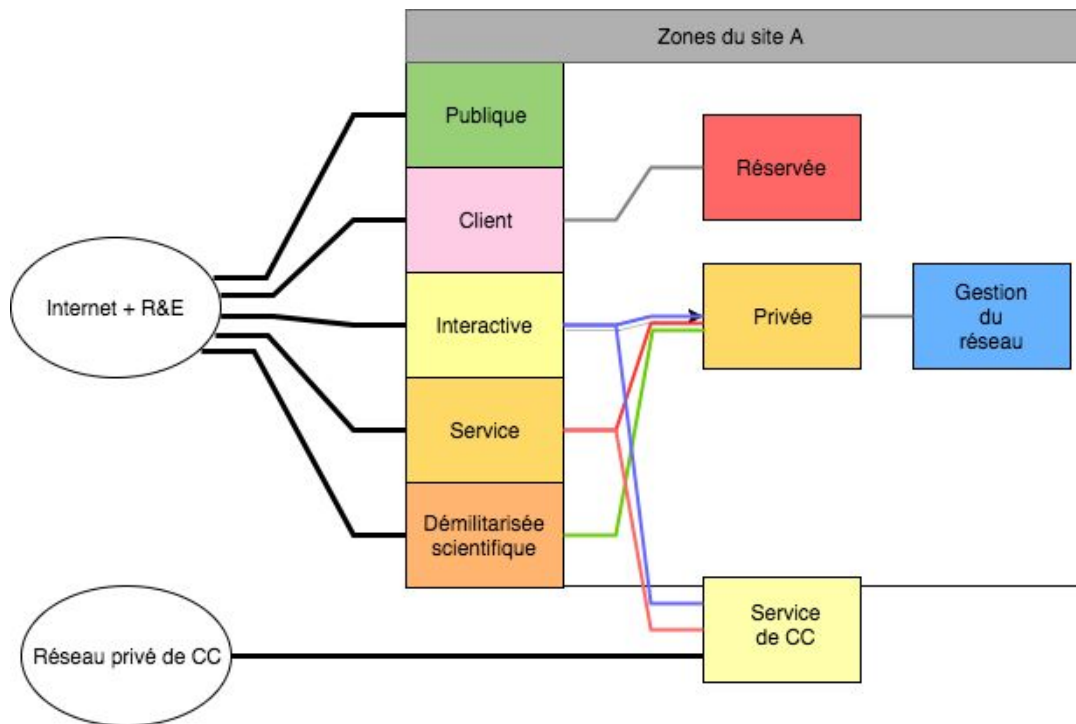
Chacun des cinq sites hôtes nationaux déploie des éléments de la CND ayant des configurations et des capacités spécifiques, correspondant à la charge de travail du site. La capacité totale de stockage en ligne du volet 1 atteint approximativement 62 Po et le volet 2 l'augmentera jusqu'à environ 100 Po. La combinaison du stockage par blocs (SBB) et du stockage défini par logiciel permet une gestion efficace et rentable de l'espace ainsi que la souplesse requise pour offrir une variété de types et de configurations de stockage.

Arbutus consacre la majeure partie de sa capacité de stockage à Ceph, pour soutenir OpenStack. Graham, Cedar et GP4 ont (ou auront) de 20 à 30 Po d'espace « projet » (« /project » de Lustre) et divers espaces de stockage dédiés, dont les serveurs de base de données, Ceph, dCache, etc.

4.3.1 Architecture de réseau et zone démilitarisée scientifique

La CND comprend également, pour chaque site, de nouveaux routeurs de périphérie leur permettant de se connecter, à 100 Gb, au réseau régional et, de là, à CANARIE, le réseau national de recherche et d'éducation. Une architecture de réseau est en cours d'implémentation pour fournir le routage, le rendement et l'isolation du service qui sont requis. La zone démilitarisée scientifique est appliquée pour isoler le trafic de certains services à l'intérieur d'une zone de réseau qui sera dirigée vers le réseau étendu, mais non à l'extérieur.

Figure 1 : Zones de l'architecture du réseau



4.3.2 Stockage permanent

Le stockage permanent est un espace qui est optimisé pour les coûts et non pour la performance. Il comprend le stockage de proximité (disque à bande automatisé ou semi-automatisé) et les sauvegardes.

Trois sites de stockage permanent sont compris dans le Défi 2 : SFU, Waterloo et le nouveau GP4. Ils disposeront chacun de plus de 80 Po de capacité de stockage sur bande pour les sauvegardes, les copies d'archives et d'autres utilisations semblables. Habituellement, les sauvegardes comprennent une deuxième copie hors site dans un autre site de stockage permanent.

4.3.3 Gestion de la capacité

Tous les sites du volet 2 participeront à un nouveau consortium, par appel d'offres ou un autre processus du même genre, de solutions de gestion de la capacité. Ces solutions devaient faire partie de l'appel d'offres de stockage 2016 (comme la gestion hiérarchique du stockage), mais n'ont pas été acquises à ce moment-là. Maintenant que d'autres éléments de la CND et du stockage permanent sont connus et forment la base de la capacité de stockage pour les sites des volets 1 et 2, les spécifications pour les solutions de gestion de la capacité seront plus conformes au mélange de technologies existant.

Le principe de la gestion de la capacité est de décharger la capacité des supports en lignes qui coûtent très cher vers des supports de proximité qui sont moins dispendieux. En pratique, cela comprend le transfert automatisé des données de disque à bande, cependant il existe de nombreuses approches pour la mise en place d'un tel système. Il faudra peut-être plus d'une solution afin de répondre à toute la gamme des besoins.

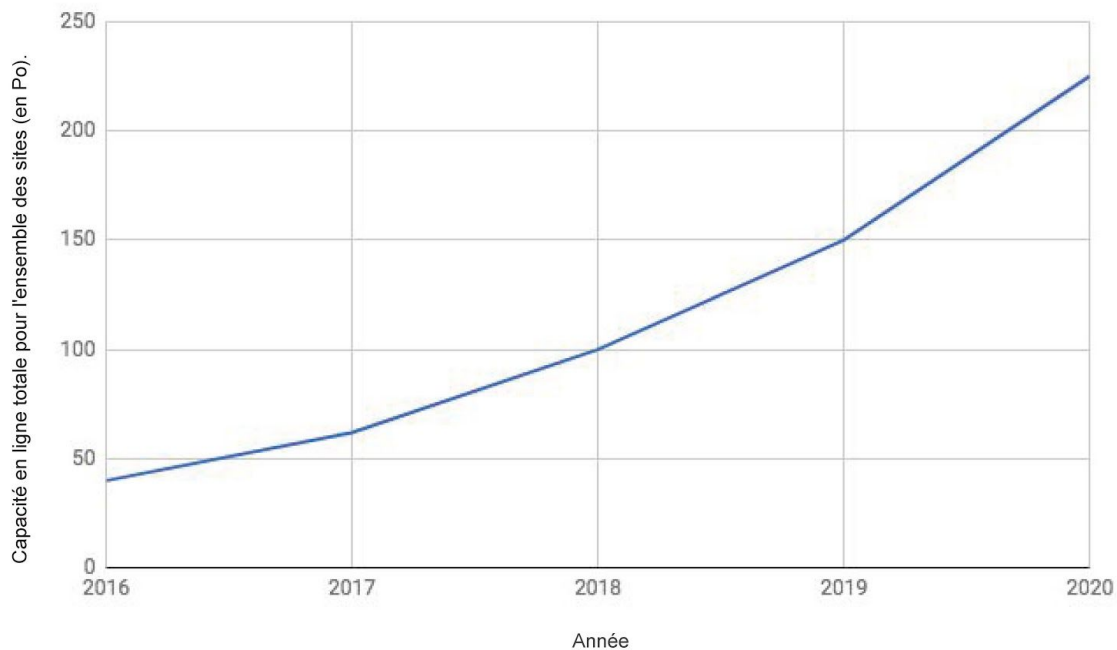
Les besoins de gestion de capacité de Calcul Canada comprennent :

- les gros systèmes de fichiers (principalement Lustre, mais aussi certains systèmes universels), dont les données doivent être automatiquement transférées de disque à bande, selon une politique qui sera déterminée en fonction de caractéristiques comme le temps écoulé depuis la dernière utilisation, la taille, les quotas des utilisateurs ou des groupes d'utilisateurs. C'est là le domaine classique des systèmes de gestion hiérarchique du stockage et il pourrait être appliqué de la même façon à d'autres types de systèmes de fichiers;
- le stockage objet (S3 ou des mécanismes semblables), qui doit fournir des mécanismes de sauvegarde ou des mécanismes de reproduction des données pour inclure des copies autonomes ou de proximité;

- la « mémoire froide » d'objets de données, comme lorsqu'un utilisateur désire entreposer des ensembles de données pour de l'archivage à long terme;
- la hiérarchisation basée sur la performance, qui déplace les objets de donnée d'un support à haute performance vers un support à performance moyenne ou vers le stockage de proximité;
- la reproduction des données et l'équilibre de la charge, permettant à de multiples copies (en ligne et de proximité) d'être conservées efficacement dans plusieurs sites;
- la disponibilité des données, les transferts en entrée et en sortie, par lesquels les objets de données seront rendus disponibles localement pour des calculs de haute performance, puis transférés de nouveau une fois le traitement terminé.

Compte tenu des grandes capacités que représentent l'ensemble des sites et du grand nombre d'utilisateurs, Calcul Canada préfère les modèles de licences en volume. CC étudie également des options gratuites ou à faibles coûts, même si elles requièrent un peu plus de temps de la part du personnel. Cette étude comprend la modélisation du coût total de propriété (CTP) pour différentes plages de capacité.

Tableau 4 : Projection de capacité de stockage en ligne



4.4 Systèmes expérimentaux

Les trois sites du volet 2 accueilleront des systèmes expérimentaux. Un processus d'appel de propositions peu élaboré sera utilisé dans chacun des sites de façon à équilibrer l'utilité immédiate, la dépense et le niveau d'effort demandé de la part de Calcul Canada, l'intention étant de trouver des technologies prometteuses, de les mettre rapidement en place et de les évaluer ainsi que de répondre aux demandes de certains utilisateurs pour des technologies nouvelles qui n'avaient pas été considérées dans le volet 2.

Les technologies présentant un intérêt pour l'appel de proposition du volet 2 comprennent :

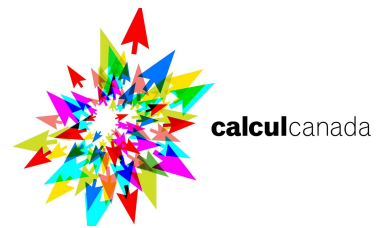
- l'intégration infonuagique commerciale et l'extension au nuage public;
- les technologies informatiques de la prochaine génération, y compris l'informatique quantique;
- des niveaux de stockage ultrarapides;
- des accélérateurs et des dispositifs reconfigurables pour le calcul, le stockage et le réseautage;
- la cybersécurité;
- les logiciels et les solutions pour l'analytique des données, l'apprentissage profond et la recherche disciplinaire.

Les systèmes expérimentaux peuvent comprendre d'acquisition d'équipement, de logiciels et de services. L'équipement peut être loué ou situé hors lieux, le cas échéant. Les dépenses peuvent aussi couvrir le personnel, la formation ou d'autres éléments nécessaires à l'amélioration de l'utilité du système expérimental.

4.5 Infrastructure de services

De nouveaux investissements dans le développement de l'infrastructure de service financée au volet 1 sont prévus. Au volet 1, on parlait de « gestion des données de recherche ». Les deux volets rendent compte du besoin de logiciels, d'intergiciels et de services pour répondre à des nécessités variées. Les activités de service sont principalement centrées sur le développement de logiciel à l'intérieur de Calcul Canada, souvent à partir de logiciels existants.

L'idée étant que si plusieurs utilisateurs ou groupes expriment le besoin d'un service, par exemple dans le cadre d'un sondage ou d'un livre blanc, alors Calcul Canada devrait considérer la possibilité d'offrir ce service à l'échelle nationale. Ces activités comprennent : l'implantation de technologies d'authentification à signature unique et de gestion d'identité; le déploiement de services de surveillance et de publication de ressources; la production de rapport; les services



de transfert de données; l'intégration de programmeur de lot et de nuage; le métaordonnement; la gestion des charges de travail pour le calcul et les données et la gestion des données de recherche.

5. Vision 2020

En qualité de premier fournisseur d'infrastructure de recherche numérique (IRN), Calcul Canada aborde l'infrastructure de données et de calcul de façon globale afin d'en faire bénéficier toutes les sphères de la société. Le renouvellement technologique et la modernisation des équipements, soutenus par la FCI dans le cadre des volets 1 et 2 du Défi 2, ont permis de mettre des ressources modernes et efficaces au service de projets de recherche canadiens exceptionnels reposant sur le calcul et le traitement de données.

Calcul Canada collabore avec des organismes de financement gouvernementaux et d'autres fournisseurs d'IRN afin de mettre à la disposition des chercheurs des systèmes, des services et du soutien perfectionnés, intégrés et efficaces, de calibre mondial. Les chercheurs de l'avenir auront un accès transparent à ces ressources, intégrées pour une efficacité et une performance optimales, sans égard à l'emplacement géographique ou au fournisseur.

D'ici 2020, Calcul Canada offrira une gamme entière de ressources couvrant toutes les étapes du cycle de recherche de données pour aider les chercheurs et leurs partenaires industriels et internationaux à se tailler une place sur l'échiquier mondial. De concert avec les autres fournisseurs d'IRN du pays, Calcul Canada offrira des systèmes et des services simplifiant les charges de travail et pouvant facilement s'étendre à différentes ressources comme les ressources de calcul informatique en laboratoire, en milieu universitaire ou à l'échelle nationale, l'équipement d'analyse, les archives de publication ou la collaboration. La proximité du soutien et de l'engagement demeurera une marque d'excellence du service pour tous les utilisateurs. L'avenir commence aujourd'hui, notamment par la modernisation de la cyberinfrastructure nationale de données de Calcul Canada, réalisée grâce aux investissements de la FCI dans le cadre du Second défi.