

Projet présenté par le Conseil d'Etat

Date de dépôt : 21 juin 2017

Projet de loi

ouvrant un crédit au titre de subvention cantonale d'investissement de 15 546 000 F pour financer l'évolution des infrastructures et services associés de calcul à haute performance et de stockage long terme afin d'optimiser et faciliter l'utilisation des données issues de la recherche dans les hautes écoles universitaires genevoises

Le GRAND CONSEIL de la République et canton de Genève
décrète ce qui suit :

Art. 1 Crédit d'investissement

Un crédit global fixe de 15 546 000 F (y compris TVA et renchérissement) est ouvert au Conseil d'Etat au titre de subvention cantonale d'investissement pour l'équipement nécessaire à l'évolution des infrastructures et services de calcul à haute performance et de stockage long terme afin d'optimiser et faciliter l'utilisation des données issues de la recherche.

Art. 2 Planification financière

¹ Ce crédit d'investissement est ouvert dès 2018. Il est inscrit sous la politique publique A – Formation (rubrique 0326 5640).

² L'exécution de ce crédit est suivie au travers d'un numéro de projet correspondant au numéro de la présente loi.

Art. 3 Subventions d'investissement accordées

Les subventions d'investissement accordées dans le cadre de ce crédit d'investissement s'élèvent à 15 546 000 F.

Art. 4 Amortissement

L'amortissement de l'investissement est calculé chaque année sur la valeur d'acquisition (ou initiale) selon la méthode linéaire et est porté au compte de fonctionnement.

Art. 5 But

Cette subvention doit permettre d'acquérir et de mettre en œuvre les équipements nécessaires à l'évolution des infrastructures et services de calcul à haute performance et de stockage long terme afin d'optimiser et faciliter l'utilisation et la conservation des données primaires issues de la recherche.

Art. 6 Durée

La disponibilité du crédit d'investissement s'éteint à fin 2023.

Art. 7 Aliénation du bien

En cas d'aliénation du bien avant l'amortissement complet de celui-ci, le montant correspondant à la valeur résiduelle non encore amortie est à rétrocéder à l'Etat.

Art. 8 Loi sur la gestion administrative et financière de l'Etat

La présente loi est soumise aux dispositions de la loi sur la gestion administrative et financière de l'Etat, du 4 octobre 2013.

Certifié conforme

La chancelière d'Etat : Anja WYDEN GUELPA

EXPOSÉ DES MOTIFS

Mesdames et
Messieurs les députés,

De nos jours, les chercheurs produisent plus de données scientifiques en une année que la contribution cumulée au cours de toute l'histoire¹. Cette abondance de données constitue la base d'une nouvelle science et l'on parle déjà de 4^e paradigme pour la désigner². Ce 4^e paradigme vient compléter les méthodes classiques que sont l'expérimentation, la théorie et la simulation, et représente donc un élément déterminant pour l'exploration et le progrès de la connaissance en général, ainsi que pour promouvoir une recherche interdisciplinaire à fort potentiel pour résoudre des problèmes scientifiques et sociétaux complexes.

Ce phénomène ne concerne pas uniquement les sciences dites « dures » (physique, astronomie, génomique, informatique, neurosciences, sciences de l'environnement, etc.), mais également les sciences sociales et humaines. En effet, par le biais de ce que l'on nomme les « digital humanities » (humanités numériques), la puissance de calcul des ordinateurs rend possible l'examen de larges corpus³ (textes ou autres types de média), y compris des millions de livres numérisés. Bien qu'émergente, cette approche transdisciplinaire, qui bénéficie de l'initiative du libre accès⁴, conduit déjà à de nouvelles connaissances en linguistique, en histoire, en archéologie, dans l'interprétation d'anciens textes, etc.⁵.

¹ A. Bly (2012) *Science Re-Imagined*, *The Serials Librarian* 62(1-4) pp. 33-46.

² *The Fourth Paradigm. Data-Intensive Scientific Discovery*. Microsoft Research, T. Hey, S. Tansley, and K. Tolle (eds) 2009.

³ Par exemple, des dictionnaires en ligne, Wikipédia, Google books, le corpus « year of speech », etc.

⁴ La recherche encouragée par des fonds publics devrait – en premier lieu dans l'intérêt de la science elle-même – être accessible autant que possible au public. C'est à cet effet et après concertation avec la Conférence des recteurs des universités suisses (CRUS) que le FNS a émis des dispositions sur le libre accès en vigueur depuis septembre 2007. Pour plus d'information sur ce sujet, voir le site <http://www.snf.ch/fr/leFNS/points-de-vue-politique-de-recherche/open-access/Pages/default.aspx>

⁵ Pour des exemples concrets de projets en « digital humanities », voir par exemple les sites <http://mith.umd.edu/> et <http://www.iath.virginia.edu/>

Dès lors que les politiques et les instances qui financent la recherche œuvrent à promouvoir l'accès libre aux données scientifiques (sous une forme intelligible), la culture de la science telle que nous la connaissons en sera fortement impactée : des chercheurs de pays émergents, des étudiants, des citoyens, etc., qui reconnaissent la capacité de la science à transformer la société, mais qui n'ont jusqu'à présent pas (ou peu) accès aux résultats de la recherche, auront de nouvelles sources d'information à disposition leur permettant de mieux contribuer au développement économique d'une région, ou tout simplement, d'élever leur niveau d'éducation.

Pour que ce 4^e paradigme puisse se réaliser, il y a à ce jour une nécessité de développer de nouveaux services et les infrastructures informatiques associées permettant d'une part d'optimiser la gestion et le traitement des données scientifiques (dont le volume croît de manière exponentielle), et d'autre part de rendre accessibles sur le long terme aux chercheurs, aux étudiants des hautes écoles, et dans certains cas au citoyen. Ces nouveaux services contribueront fortement à améliorer la qualité de la formation délivrée par l'Université et ses partenaires, en offrant aux enseignants l'accès à des outils plus performants leur permettant d'élargir leur propre expérience et aux étudiants de se former dans des environnements modernes leur permettant d'acquérir des compétences qu'ils pourront faire valoir dans leur vie professionnelle.

Pour répondre à ces nouveaux défis des solutions concrètes sont proposées, dont le déploiement va nécessiter des efforts coordonnés et des financements adéquats. Ces solutions, qui s'adressent à tous les chercheurs des hautes écoles genevoises, avec l'intention de faciliter les collaborations intra- et inter-institutionnelles (au niveau national et international), se déclinent en quatre objectifs prioritaires :

1. Mettre en place une infrastructure centralisée et partagée de calcul haute performance (« High Performance Computing » ou « HPC ») pouvant répondre de manière optimisée aux besoins en matière de calcul et de stockage, selon une démarche visant à réduire les empreintes écologiques;
2. Mettre en place une architecture de stockage sécurisée construite sur les standards internationaux permettant la conservation et l'accès à long terme des données scientifiques;
3. Développer des interfaces logicielles qui répondent aux besoins des chercheurs et facilitent l'utilisation de ces infrastructures aussi bien pour le calcul que pour le dépôt, la gestion, le traitement, et l'accès aux données;

4. Concevoir des environnements informatiques favorisant la collaboration entre chercheurs, facilitant l'échange des données scientifiques et leur visualisation⁶, ainsi que leur utilisation dans l'enseignement et l'apprentissage.

Cette thématique de calcul et de stockage des données scientifiques mobilise un nombre croissant d'institutions, sur le plan suisse et international, qui s'organisent pour apporter des solutions concrètes à cette problématique. Au niveau européen, cette thématique devient un axe prioritaire puisque dans la consultation préliminaire du nouveau programme stratégique pour la recherche, qui prendra la relève du FP7 actuel dès 2013, il ressort que les e-infrastructures au service des chercheurs, de même que l'accès aux données scientifiques et l'accès libre à la connaissance, sont jugés clés pour la recherche et l'innovation⁷. Sur le plan suisse, les écoles polytechniques fédérales et les universités de ZURICH et Berne sont déjà dotées de moyens de calcul conséquents⁸ pour la recherche. Se dessinent d'autre part dès 2013 des initiatives concrètes au niveau national visant à interconnecter, développer et partager les infrastructures et services soutenant la gestion et le traitement de l'information scientifique.

Rester en retrait de ces initiatives, ne pas réaliser les quatre objectifs précités dans un avenir proche, serait préjudiciable à nos institutions en termes de compétitivité et d'attractivité. Cela reviendrait à réduire notre capacité à traiter les données scientifiques, ce qui à moyen terme péjorerait notre potentiel à innover et donc à maintenir, voire à consolider, notre position d'excellence dans certains domaines scientifiques. A plus long terme, le patrimoine scientifique contenu dans les données primaires, voire secondaires⁹, produites par nos institutions serait en péril. Nous sommes par

⁶ Voir par exemple le projet Visualizing.org (<http://opendata-tools.org/en/visualization>).

⁷ Green Paper on a Common Strategic Framework for EU Research and Innovation Funding : Outcome of the Public Consultation, https://ec.europa.eu/research/horizon2020/pdf/consultation-conference/summary_analysis.pdf

⁸ Représentant une capacité de calcul annuelle comprise entre 6 et 40 millions d'heures suivant les institutions.

⁹ Par données dites « primaires » il faut comprendre l'ensemble des données produites par un projet de recherche, principalement sous forme numérique. Il s'agit donc des travaux préparatoires qui définissent le projet de recherche et en défendent l'existence future (dossier de requête de subsides, par exemple), les données brutes, ces mêmes données après traitement et analyse, la documentation qui accompagne ces données, etc. Par données « secondaires », il faut entendre les publications et communications scientifiques qui reposent sur ces données primaires.

ailleurs obligés d'avancer dans cette voie dans la mesure où, d'une part, de plus en plus d'éditeurs de journaux prestigieux (par exemple, Nature, Science, etc.) demandent à ce que les données de recherche ayant servi à la publication soient accessibles aux communautés de chercheurs et, d'autre part, des bailleurs de fonds, particulièrement dans le domaine public, imposent que les données issues des recherches soient, dans un souci d'éthique, accessibles au citoyen et à la société civile. Cependant, les données sous forme numérique ont une durée de vie courte, en raison de l'obsolescence des formats et des environnements¹⁰, et nécessitent pour assurer leur conservation à long terme des architectures adaptées et des standards stricts reconnus par la communauté internationale¹¹, sans quoi l'effort de conservation sera vain.

Un facteur clé de succès du projet sera la possibilité d'intégrer ces technologies au métier du chercheur d'une manière harmonieuse. Cela va nécessiter une sensibilisation du chercheur qui aujourd'hui, dans 80% des cas, utilise son ordinateur personnel pour stocker ses données de recherche¹², avec tous les risques que ceci engendre. Le chercheur est de plus en plus conscient de ces risques et cite volontiers les avantages à utiliser des infrastructures de stockage plus adaptées qui, au-delà d'une pérennité accrue des données, lui permettent de réanalyser ses données à des fins de validation des résultats, de promouvoir des collaborations interdisciplinaires, de justifier les fonds publics, et d'obtenir une plus grande citabilité de ses travaux de recherche, ce qui par ailleurs a un impact direct sur le rang des universités dans les différents classements internationaux¹³.

Besoins

L'Université de Genève a pu acquérir durant ces dernières années des compétences certaines et reconnues dans les domaines du calcul à haute performance et du stockage des données – en particulier au travers des

¹⁰ Voir par exemple l'article récent de P.-Y. Burgi (2011) *Préservation des connaissances*, Market.ch no 94 (décembre), p. 64.

¹¹ Par exemple le standard OAIS, voir <http://public.ccsds.org/publications/BlueBooks.aspx>

¹² <http://libereurope.eu/wp-content/uploads/2010/01/PARSE.Insight.-Deliverable-D3.4-Survey-Report.-of-research-output-Europe-Title-of-Deliverable-Survey-Report.pdf>

¹³ Dans les classements internationaux, l'Université de Genève se situe régulièrement parmi les 100 meilleures universités sur 20'000 institutions universitaires dans le monde.

centres de calcul romands CADMOS¹⁴ et Vital-IT¹⁵ – qui sont issus de collaborations entre l'Université de Lausanne, l'EPFL, et l'Université de Genève. D'autres projets en HPC ont également été réalisés avec l'HEPIA (Haute Ecole du Paysage, d'Ingénierie et d'Architecture) dans le contexte de projets nationaux. Mais ces collaborations restent spécifiques à certains champs d'étude, et il s'est avéré nécessaire de dresser un paysage plus complet des besoins en HPC et en gestion des données scientifiques à l'Université de Genève. Par conséquent, les huit facultés ont fait l'objet, en 2011, d'une consultation au travers d'enquêtes et de réunions impliquant les responsables des commissions informatiques des facultés, et les chercheurs utilisateurs de ces ressources. Au terme de cette consultation, il a été recensé un besoin annuel de l'ordre de 65 millions d'heures de calcul (correspondant à un parc de machines d'environ 10'000 cœurs), et de l'ordre de 10 Peta Bytes (PB) de capacité de stockage sécurisé¹⁶ d'ici 2016-17. Ces chiffres, bien qu'ils soient en augmentation continue, nous fournissent une bonne indication sur les investissements à prévoir ces cinq prochaines années, et c'est sur cette base que nous établissons le plan financier de ce projet de loi.

Ce projet de loi vise par conséquent en premier lieu le développement des infrastructures et services informatiques nécessaires à l'évolution de la recherche pour l'ensemble des domaines scientifiques de la place universitaire genevoise, y compris les besoins des projets particuliers tels que :

- le centre de compétences bioinformatiques Vital-IT, qui dispose déjà d'une expertise dans le domaine; ce projet de loi lui permettrait de renforcer les infrastructures déjà installées à Genève, tout en favorisant la collaboration avec les partenaires académiques romands;
- le développement des départements de mathématiques, de physique et d'astronomie;
- les projets interdisciplinaires (par exemple, dans le domaine de l'environnement et du climat, en neurosciences, en médecine, etc.);
- les « digital humanities » en sciences humaines et sociales.

Ces nouvelles infrastructures viendront compléter celles de CADMOS, créé par l'UNIGE, l'UNIL et l'EPFL, qui s'inscrit dans la stratégie HPC

¹⁴ <http://www.cadmos.org/index.php/fr/>

¹⁵ <http://www.vital-it.ch>

¹⁶ Un Peta Byte (PB) correspond à 1'000 Tera Bytes (TB). Il s'agit de stockage principalement en termes de « spinning disks », technologiquement plus contraignant qu'un stockage sur bande magnétique.

suisse, et dont les super calculateurs sont plus spécialement adaptés aux calculs dits « massivement parallèles ».

Cette première étape favorisera une utilisation optimale des capacités de calcul et de stockage selon le concept moderne de « cloud computing »¹⁷. Elle sera réalisée en collaboration avec les services de l'Etat qui, tout comme l'Université, étudie les possibilités d'hébergement des futures infrastructures informatiques.

En complément à ce travail d'envergure visant à optimiser l'emploi des infrastructures de calcul et de stockage des données, il sera question de réaliser et d'intégrer au système d'information des développements logiciels ciblés sur des besoins récurrents des chercheurs en matière de traitement des données scientifiques. Pour cela, l'Université de Genève a besoin de renforcer son équipe technique pendant la durée du projet afin de poursuivre le développement de son système d'information dans ce secteur en très forte évolution, de capitaliser son expertise, et d'étendre ce service plus largement au sein de l'ensemble de la communauté scientifique genevoise. Les services associés à ces infrastructures vont inclure concrètement :

- Un « cloud computing » permettant aux chercheurs d'utiliser les ressources d'une manière flexible, selon leurs besoins, à la demande. A terme cette infrastructure devrait s'intégrer dans une logique suisse de cloud computing académique.
- L'archivage des données selon des standards internationaux en vigueur dans le domaine archivistique, permettant la conservation à moyen terme (moins de 10 ans) et long terme (au-delà de 10 ans) des données scientifiques.
- Du conseil et de l'expertise en calcul haute performance afin de garantir aux chercheurs une meilleure adéquation entre leurs algorithmes et l'architecture des machines disponibles.
- L'accompagnement de chercheurs dans leurs projets d'exploitation de leurs données selon de nouveaux algorithmes, particulièrement dans les sciences sociales et humaines (digital humanities), domaine qui est amené à se développer et dont le potentiel, largement sous-exploité à ce jour, intéresse de plus en plus les chercheurs de l'Université de Genève (par exemple, en facultés des lettres, de théologie, de traduction et d'interprétation, et des SES) et plus largement de l'arc lémanique.

¹⁷ Pour une définition concise de ce concept, cf. http://fr.wikipedia.org/wiki/Cloud_computing

- L'optimisation des applications couramment utilisées par les chercheurs afin qu'elles exploitent au mieux les différentes architectures de calcul en place.
- Le développement d'interfaces permettant aux étudiants et plus largement au citoyen d'accéder et de visualiser des données issues de la recherche et rendues spécifiquement compréhensibles pour cette population d'utilisateurs.

Pour des raisons de pérennité des solutions informatiques, les logiciels sous-jacents à ces infrastructures seront pour la plupart des logiciels libres soutenus par des communautés internationales, pour la majorité liées au milieu académique et actives dans le domaine de la conservation de l'information et des connaissances. Il serait en effet illusoire de compter sur des solutions propriétaires, sujettes aux aléas des marchés économiques mondiaux. Par conséquent, les frais de licences seront faibles, voire inexistantes (ou se limiteront à du service autour de certains logiciels libres). Afin de faciliter l'usage de ces technologies par les communautés scientifiques et optimiser les compétences nécessaires à leur maîtrise et leur évolution, les choix seront opérés en étroite collaboration avec les réseaux de compétences existant au sein de la communauté universitaire suisse.

Critères d'évaluation de l'investissement proposé

Les objectifs poursuivis par ce projet de loi s'inscrivent dans le cadre du plan directeur d'évolution du système d'information (PDSI) de l'Université, lui-même inscrit dans le cadre du plan stratégique de l'institution. L'évaluation des projets de système d'information au sein de l'Université suit une méthodologie en tout point comparable à celle des 7 axes recommandés par la Cour des comptes. Le processus d'analyse de ces projets est décrit systématiquement dans un dossier de cadrage, qui sert de base de décision pour la gouvernance. En outre, l'outil de gestion des projets utilisé, PSNext, garantit le suivi du projet tout au long de son cycle de vie, en fournissant les données pour tout type d'analyse et de contrôle.

Coûts de l'infrastructure

Le fait de pouvoir proposer tous ces services dans le cadre d'un cloud communautaire « académique » interconnecté aux hautes écoles suisses représente un avantage en termes aussi bien de coûts que de sécurité, par rapport à des services commerciaux de type cloud Amazon EC2¹⁸. En effet, en considérant uniquement l'aspect HPC, tout en prenant en compte

¹⁸ Amazon Elastic Compute Cloud (Amazon EC2) (<http://aws.amazon.com/ec2>).

l'amortissement des machines sur 5 ans, le prix de l'électricité et du refroidissement, un cloud académique permettrait de réduire les coûts par un facteur d'environ 3. Pour le volume de 10 000 cœurs et 10 PB de stockage demandés dans ce projet de loi, le tableau suivant résume les coûts annuels en millions de francs :

	Cloud académique ¹⁹	Cloud Amazon EC2 ²⁰
HPC	1,7	4,9
Stockage	1,6	8,5
Salaires	0,6	0
Total	3,9	13,4

Pour répondre aux besoins des chercheurs, le coût annuel d'un cloud académique pour les hautes écoles genevoises s'élève donc à environ 3,9 millions de francs. En comparaison, l'utilisation d'un cloud commercial tel que celui d'Amazon engendrerait un surcoût annuel de l'ordre d'environ 10 millions de francs. Sur le plan économique, il est donc très nettement préférable de construire un cloud académique.

D'autre part, avec une solution locale de cloud telle que proposée, la confidentialité des données (un aspect particulièrement sensible par exemple en médecine) est assurée, contrairement aux clouds proposés par ces entreprises dont les serveurs peuvent se trouver a priori dans n'importe quel pays, ce qui amène des questions d'ordre juridique (cf. la LIPAD²¹). Finalement, l'exploitation en local permet de développer des compétences dans un domaine qui reste clé pour l'avenir de la recherche.

Efficacité métier

Aujourd'hui, les chercheurs ont constamment besoin d'ordinateurs plus puissants, de disques aux capacités de stockage plus importantes, ceci afin de s'adapter à l'augmentation continue du volume de données à traiter. Cette

¹⁹ Estimation basée sur deux offres obtenues en novembre 2011 pour des machines à 1000 cœurs (50 W de consommation par cœur, y inclus le refroidissement), un coût d'électricité de 20 cts le kWh et un taux d'utilisation de 80%.

²⁰ Basé sur une utilisation « lourde » (correspondant à l'usage prévu), un taux d'utilisation de 80%, et une remise de 25% pour des volumes importants.

²¹ http://www.ge.ch/legislation/rsg/f/s/rsg_a2_08.html

course effrénée engendre des coûts financiers et des risques mal maîtrisés du fait que chaque laboratoire s'équipe comme il peut pour faire face au plus urgent. En conséquence, le retour sur investissement (ROI) des deniers publics, utilisés pour financer cette recherche, se réduit d'autant plus que des données scientifiques clés se retrouvent fréquemment dans des infrastructures locales de stockage non pérennes, et généralement non accessibles à la communauté scientifique.

Au lieu d'acheter individuellement des ordinateurs et disques de stockage, ou de souscrire à des services de type Amazon EC2 qui restent onéreux et risqués (voir ci-dessus), ce projet de loi vise au contraire à mutualiser des infrastructures de calcul et de stockage à long terme qui permettront aux chercheurs une meilleure gestion de leurs données de recherche en réduisant les coûts de leur traitement informatique, ce qui permettra de mieux utiliser leurs budgets dans des domaines plus spécifiques à leur recherche, de mieux valoriser leurs données primaires et secondaires dans le long terme auprès d'une communauté plus large et interdisciplinaire, de faire bénéficier les résultats de leur recherche au domaine de l'éducation, et plus généralement, à la société civile. Les moyens mis en œuvre pour poursuivre cette voie vont s'appuyer sur :

1. La modernisation et consolidation des infrastructures de calcul et de stockage qui sont aujourd'hui lacunaires;
2. L'interconnexion de ces infrastructures au système d'information de l'Université (et plus largement des hautes écoles suisses) afin de bénéficier de services de base, par exemple, backups sur sites distants, sécurisation des données, outils de recherche, etc.;
3. La mise à disposition de services à valeur ajoutée et d'outils informatiques spécifiquement conçus pour faciliter la gestion des calculs et des données scientifiques associées (par exemple outils de visualisation).

Impact

Ce projet va impacter dans le travail quotidien la communauté scientifique genevoise, à savoir environ 4 500 enseignants/chercheurs²² (Université et HES). Dans la mesure où les étudiants et, dans certains cas, le citoyen, devraient également bénéficier des résultats de ce projet, particulièrement par une accessibilité facilitée aux données, un impact

²² Pour des statistiques récentes, cf. <http://www.unige.ch/dadm/stat/dernierepublication-1.html> et <https://www.hesge.ch/geneve/hes-so-geneve/qui-sommes-nous/en-chiffres>

sociétal, bien que plus difficile à évaluer, est néanmoins avéré et toucherait une population bien plus large. Par ailleurs, la mise en œuvre de ce « 4^e paradigme » grâce à une infrastructure moderne contribuera à une attractivité accrue de nos hautes écoles genevoises pour de futurs chercheurs, ce qui permettra de maintenir notre compétitivité dans les domaines de la recherche.

Efficacité financière

L'efficacité financière se décline en 3 axes principaux :

1. En optimisant les dépenses et en évitant que les investissements soient réalisés sans coordination dans les laboratoires de recherche, ou que des crédits soient investis dans des clouds commerciaux, comme par exemple « Amazon EC2 », qui restent très onéreux (voir ci-dessus);
2. En pérennisant et mutualisant les données, ce qui évite des coûts certains de collecte d'information et/ou de calculs à refaire dus à des données dispersées ou érodées par le temps, voire perdues;
3. Par une approche réfléchie en termes d'une part d'espace et de gestion d'énergie des ressources de calcul (synergies avec l'étude d'un centre de calcul de l'Etat) et d'autre part de mutualisation des compétences entre et au sein des institutions.

Degré de contrainte

Au-delà de la démarche visant à optimiser les ressources, il est également nécessaire de réduire les empreintes écologiques dues à une utilisation effrénée de ressources de calculs non contrôlées. Sur le plan éthique, le citoyen qui contribue au financement de la recherche par ses impôts est en droit d'accéder aux résultats, voire même aux données primaires qui ont conduit à ces résultats. Un aspect sociétal est à prendre en compte également dans la mesure où garantir la pérennité des données scientifiques c'est aussi assurer un patrimoine des connaissances pour les futures générations.

Descriptif du développement technique et informatique

Pour réaliser tous les objectifs inscrits dans ce projet de loi, il va falloir :

- A. Concevoir une infrastructure répondant aux besoins des chercheurs et qui puisse s'adapter aux évolutions technologiques. Plus précisément, il faudra s'assurer de la mise à niveau de l'infrastructure de calcul haute performance et de l'architecture de conservation long terme des données scientifiques selon une approche de type « cloud computing » permettant de s'affranchir de certaines contraintes physiques et d'optimiser l'utilisation des ressources en les mutualisant.

- B. Développer les logiciels permettant de piloter ces infrastructures, d'automatiser les processus sous-jacents aux calculs/données, et de valoriser les résultats (visualisation, agrégation des données, outils de recherche, etc.). Ces développements vont tenir compte aussi bien des besoins des chercheurs, que de ceux qui seront amenés à utiliser les résultats de la recherche (enseignants, étudiants, éditeurs, bailleurs de fonds et le public au sens large dans certains cas).
- C. Intégrer ces environnements dans le système d'information de l'Université et plus largement des hautes écoles suisses.

Ces 3 volets seront réalisés en parallèle et, pour chacun d'eux, des ressources humaines seront prévues sur une durée de 5 ans, à savoir :

- 2 ingénieurs système, spécialistes du HPC et du stockage qui concevront l'architecture technique, déploieront les infrastructures HPC et de stockage, en assureront le contrôle et le suivi et offriront des conseils et de l'expertise afin de garantir aux chercheurs une meilleure adéquation entre leurs besoins et l'architecture des machines disponibles, ainsi que pour les aider à formaliser et qualifier leurs besoins en ressources de calcul et de stockage;
- 2 informaticiens de développement pour concevoir et réaliser les couches logiques d'interfaçage du cloud avec les utilisateurs, ainsi que pour concevoir et réaliser l'architecture logicielle de préservation des données scientifiques;
- 2 informaticiens de développement, qui accompagneront les chercheurs dans leurs projets de mise en place de nouveaux algorithmes et leur intégration dans le système d'information, ceci particulièrement dans le domaine émergent des « digital humanities »;
- adjoint-e scientifique, spécialiste de la conservation des données, qui viendra compléter cette équipe afin de garantir une préservation optimale des données selon une expertise métier avérée.

Ces personnes viendront, pendant la durée du projet, compléter les équipes de la Division du Système et des Technologies de l'Information et de la Communication (DiSTIC). Elles constitueront la base d'un noyau de compétences qui permettra ensuite à l'Université de poursuivre l'exploitation et la maintenance des infrastructures et des services mis en œuvre dans le cadre de ce projet de loi.

Coût des ressources matérielles et logicielles requises

Les coûts des ressources matérielles et logicielles ont été évalués sur la base des besoins recensés (voir Besoins, page 6). Il s'agit d'une part de

l'infrastructure HPC, principalement constituée de serveurs dont la puissance globale doit permettre de répondre au besoin exprimé d'environ 65 millions d'heures de calcul par année et, d'autre part, de l'architecture de stockage et de conservation à long terme et sécurisée des données évaluée à 10 Peta Bytes (PB) :

	Budget (CHF)
Infrastructure HPC (serveurs)	3'800'000
Stockage sécurisé (10 PB)	4'500'000
Stockage répliqué à distance	1'900'000
Total HT	10'200'000
TVA (8%)	816'000
Total TTC	11'016'000

Coût des ressources humaines requises

Les coûts annuels des ressources humaines financées par cette loi sont indiqués dans le tableau suivant :

	Nb	Coût annuel moyen, charges comprises	Total (CHF)
Informaticiens (ingénieurs, développeurs)	6	150'000	900'000
Adjoint scientifique	1	120'000	120'000
Total (annuel)			1'020'000

Récapitulatif des coûts financiers et planning des dépenses (K CHF)

	2018	2019	2020	2021	2022	Total
Investissements	1'016	2'000	3'000	3'000	2'000	11'016
RH (salaires)	450	1'020	1'020	1'020	1'020	4'530
Total	1'466	3'020	4'020	4'020	3'020	15'546

Charges d'exploitation

L'exploitation de ces ressources mutualisées dans le cadre d'une architecture de type cloud permettra de disposer de mesures relatives à leur utilisation et d'optimiser leur mise à disposition et leur évolution en fonction

des besoins. La mise en œuvre de ce cloud local s'inscrivant aussi dans la stratégie de développement d'un cloud académique suisse permettant de renforcer encore son élasticité, ces mesures contribueront également à l'établissement d'un modèle économique permettant de répartir équitablement les charges d'utilisation et d'évolution de ces services au sein des hautes écoles genevoises et suisses.

Au bénéfice de ces explications, nous vous remercions, Mesdames et Messieurs les députés, de réserver un bon accueil au présent projet de loi.

Annexes :

- 1) *Préavis financier (art. 30 RPF CB – D 1 05.04)*
- 2) *Planification des dépenses et recettes d'investissement du projet (art. 31 RPF CB – D 1 05.04)*
- 3) *Planification des charges et revenus de fonctionnement du projet (art. 31 RPF CB – D 1 05.04)*



REPUBLIQUE ET
CANTON DE GENEVE

PRAVIS FINANCIER

Ce préavis financier ne préjuge en rien des décisions qui seront prises en matière de politique budgétaire.

1. Attestation de contrôle par le département présentant le projet de loi

- ♦ Projet de loi présenté par le département de l'instruction publique, de la culture et du sport.
- ♦ Objet : Projet de loi ouvrant un crédit au titre de subvention cantonale d'investissement de 15 546 000 F pour financer l'évolution des infrastructures et services associés de calcul à haute performance et de stockage long terme afin d'optimiser et faciliter l'utilisation des données issues de la recherche dans les hautes écoles universitaires genevoises
- ♦ Rubrique budgétaire concernée : 0326 5640
- ♦ Politique publique concernée : A - Formation
- ♦ Coût total du projet d'investissement :

Dépenses d'investissement	15.546
- Recettes d'investissement	0
= Investissements nets	15.546

- ♦ Coût total du fonctionnement lié :

Charges liées de fonctionnement	0
- Revenus liés de fonctionnement	0
= Impacts nets sur les résultats annuels	0

- ♦ Planification pluriannuelle de l'investissement :

(en mios de F)	2018	2019	2020	2021	2022	2023	2024	2025	Total
Dépense brute	1.5	3.0	4.0	4.0	3.0	0.0	0.0	0.0	15.5
Recette brute	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Invest. net	1.5	3.0	4.0	4.0	3.0	0.0	0.0	0.0	15.5

♦ Planification des charges et revenus de fonctionnement liés et induits :

oui non Les tableaux financiers annexés au projet de loi intègrent la totalité des impacts financiers découlant du projet.

(en mios de F)	2018	2019	2020	2021	2022	2023	2024	Dès 2025
NET LIE et INDUIT	-0.03	-0.09	-0.17	-0.25	-1.87	-3.42	-3.42	-3.42

♦ Planification financière (modifier et cocher ce qui convient) :

oui non Le crédit d'investissement est ouvert dès 2018, conformément aux données des tableaux financiers.

oui non Les charges et revenus de fonctionnement liés et induits de ce projet sont inscrits au budget de fonctionnement dès 2018.

oui non Le crédit d'investissement et les charges et revenus de fonctionnement liés et induits de ce projet sont inscrits au plan financier quadriennal 2018-2021.

oui non Ce projet génère des charges de fonctionnement liées nécessaires à sa réalisation (ces charges n'étant pas comprises dans la demande de crédit du présent projet de loi, elles doivent faire l'objet d'une inscription annuelle au budget de fonctionnement).

oui non Autre(s) remarque(s) : ./.

Le département atteste que le présent projet de loi est conforme à la loi sur la gestion administrative et financière de l'Etat (LGAF), à la loi sur les indemnités et les aides financières (LIAF), au modèle comptable harmonisé pour les cantons et les communes (MCH2) et aux dispositions d'exécution adoptées par le Conseil d'Etat.

Genève, le : 11/05/2017 Signature du responsable financier du département investisseur :

-Tissot

2. Approbation / Avis du département des finances

oui non Remarque complémentaire du département des finances : ce projet de loi est conforme au PDI 2018-2027

Genève, le : 03/05/2017 Visa du département des finances :

S. Jandeau

N.B. : Le présent préavis financier est basé sur le PL, son exposé des motifs, les tableaux financiers et ses annexes transmis le 04 MAI 2017.

1. PLANIFICATION DES DEPENSES ET RECETTES D'INVESTISSEMENT DU PROJET
Projet de loi ouvrant un crédit au titre de subvention cantonale d'investissement de
15 546 000 F à l'Université de Genève pour des infrastructures de calcul à haute
performance et de stockage

Projet présenté par département de l'instruction publique, de la culture et du sport

<i>(montants annuels, en millions de F)</i>		2018	2019	2020	2021	2022	2023	2024	TOTAL
Dépenses d'investissement		1.5	3.0	4.0	4.0	3.0	0.0	0.0	15.5
Recettes d'investissement		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Investissement net	Durée	1.5	3.0	4.0	4.0	3.0	0.0	0.0	15.5
Informatique - Subv. Invest.	5 ans	1.5	3.0	4.0	4.0	3.0	0.0	0.0	15.5
Recettes		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Aucun		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Recettes		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Aucun		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Recettes		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Aucun		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Recettes		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Remarques :

Date et signature direction financière (investisseur) :

P. Tissot

le 11/5/2017



2. PLANIFICATION DES CHARGES ET REVENUS DE FONCTIONNEMENT DU PROJET

Projet de loi ouvrant un crédit au titre de subvention cantonale d'investissement de 15 546 000 F à l'Université de Genève pour des infrastructures de calcul à haute performance et de stockage

Projet présenté par département de l'instruction publique, de la culture et du sport

(montants annuels, en mio de F)	2018	2019	2020	2021	2022	2023	2024	dès 2025
TOTAL charges liées et induites	0.03	0.09	0.17	0.25	1.87	3.42	3.42	3.42
Charges en personnel [30]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Biens et services et autres charges [31]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Charges financières	0.03	0.09	0.17	0.25	1.87	3.42	3.42	3.42
Intérêts [34] 2.000%	0.03	0.09	0.17	0.25	0.31	0.31	0.31	0.31
Amortissements [33 + 366 - 466]	0.00	0.00	0.00	0.00	1.55	3.11	3.11	3.11
Subventions [363 + 369]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Autres charges [30 à 36]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TOTAL revenus liés et induits	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Revenus [40 à 46]	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RESULTAT NET LIE ET INDUIT	-0.03	-0.09	-0.17	-0.25	-1.87	-3.42	-3.42	-3.42
RESULTAT NET LIE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RESULTAT NET INDUIT	-0.03	-0.09	-0.17	-0.25	-1.87	-3.42	-3.42	-3.42

Remarques :

Date et signature direction financière (investisseur) :

P. Tissot le 11/5/2017
S.T.