

**Contre-expertise
de l'évaluation socio-économique
du projet français Exascale**



Source : EuroHPC Joint Undertaking

AID René
FREY Pascal
LICHNEWSKY Alain
SERRIS Jacques

14 septembre 2020

Préambule

La loi du 31 décembre 2012 instaure l'obligation d'évaluation socio-économique préalable des projets d'investissements financés par l'Etat et ses établissements publics et une contre-expertise indépendante de cette évaluation lorsque le niveau de financement dépasse un seuil que le décret d'application de la loi a fixé à 100 M€.

C'est donc en respectant toutes les règles prévues dans le décret d'application (exigences du contenu du dossier, indépendance des contre experts, délais) que le SGPI a fait réaliser cette contre-expertise indépendante de l'évaluation de ce projet.

Ce rapport a été établi à partir des documents fournis par GENCI et des réponses apportées aux questions des contre-experts tout au long de la procédure. A l'été 2020, de nombreux éléments restent à décider, tant dans le projet européen que pour le projet de candidature française portée par GENCI. Le projet est encore dans une phase très amont, dans laquelle subsistent de nombreuses incertitudes sur l'architecture et la performance de la machine (inclusion du processeur européen), sur les partenariats internationaux, ainsi que sur le niveau de mobilisation des communautés scientifiques et des industriels utilisateurs potentiels.

Il ne saurait être reproché à ce rapport de ne pas tenir compte d'éléments qui n'auraient pas été communiqués à ses auteurs.

Sommaire

PREAMBULE	2
SOMMAIRE	3
1. INTRODUCTION	4
2. LA CONTRE-EXPERTISE DU RAPPORT DE GENCI	5
2.1. GENCI.....	5
2.2. CONTEXTE.....	6
2.2.1. <i>L'Exaflop: nouvelle frontière</i>	6
2.2.2. <i>Le périmètre du projet</i>	7
2.2.3. <i>L'accueil en France d'un super ordinateur Exascale</i>	8
2.2.4. <i>La configuration souhaitée</i>	9
2.2.5. <i>L'utilisation nationale du ordinateur</i>	10
2.3. PRESENTATION DE L'INVESTISSEMENT ET EXPLICITATION DES OBJECTIFS.....	12
2.4. LES DIFFERENTES OPTIONS D'INVESTISSEMENT.....	15
2.5. IDENTIFICATION DES EFFETS ATTENDUS DES OPTIONS D'INVESTISSEMENT.....	16
2.6. IDENTIFICATION DES EFFETS ATTENDUS DE L'OPTION DE REFERENCE.....	18
2.7. DESCRIPTION QUALITATIVE ET QUANTITATIVE DE TOUS LES EFFETS POUVANT ETRE PRODUITS.....	19
2.8. VALORISATION MONETAIRE DES EFFETS POUVANT ETRE VALORISES.....	20
2.9. ANALYSE DES RISQUES ET INCERTITUDES.....	21
2.9.1. <i>Les risques du programme</i>	21
2.9.2. <i>Les opportunités du programme</i>	26
2.9.3. <i>L'option « Calcul Quantique »</i>	27
2.9.4. <i>En résumé</i>	27
3. POUR ALLER PLUS LOIN DANS L'ANALYSE ECONOMIQUE	29
3.1. L'OPTION DE REFERENCE.....	29
3.2. LA DEMANDE.....	29
3.3. LES EFFETS ATTENDUS.....	30
4. CONCLUSIONS ET RECOMMANDATIONS	32

1. Introduction

Dans une première partie, le rapport reprend le plan de l'étude d'impact réalisée par GENCI et les éléments complémentaires apportés au cours de la contre-expertise pour la commenter. Dans une seconde partie sont présentées des pistes pour compléter l'étude d'impact. Le projet est en effet dans une phase amont : l'examen de son calendrier montre qu'un certain nombre de travaux complémentaires d'impact socio-économique seraient possibles et utiles pour éclairer la conduite du projet, optimiser les décisions d'investissement et développer un accompagnement du projet permettant d'en optimiser l'impact. Les conclusions et recommandations de la contre-expertise sont présentées en troisième partie.

2. La contre-expertise du rapport de Genci

2.1. Genci¹

Le dossier soumis à contre-expertise présente la très grande infrastructure de recherche GENCI : historique ; forme juridique ; mission – statuts ; instances, gouvernance ; les projets de recherche.

Cette description est bien sûr utile pour comprendre le rôle de GENCI. Elle a été complétée, lors des contacts pris dans le cadre de la contre-expertise par une présentation de l'ensemble des acteurs français responsables ou associés à la conduite du projet (les titulaires de « l'équipe de France ») et de leur interface avec l'organisation européenne.

- Au niveau européen, la *joint undertaking* (JU) EuroHPC est pilotée par un *Governing Board* composé de la Commission Européenne (50 % vote) et actuellement de 31 Etats participants. Le Directeur Exécutif de la *Joint Undertaking* est appuyé par deux instances de conseil : le RIAG (*Research & Innovation Advisory Group*) et l'INFRAG (*Infrastructure Advisory Group*). La représentation française au sein du *Governing board* est assurée par un représentant de la direction générale de la recherche et de l'innovation (DGRI) du Ministère de la recherche et de l'enseignement supérieur, dont le suppléant appartient à la direction générale des entreprises (DGE) du Ministère de l'économie et des finances.
- GENCI sera « l'architecte » du projet d'hébergement d'une machine européenne en France. Ce terme est choisi pour recouvrir un ensemble de rôles et de responsabilité qui sera précisé plus loin.
- La machine a vocation à être installée dans un centre de calcul associé de GENCI, le TGCC² du CEA.
- GENCI aura besoin de relais pour développer l'utilisation de cette nouvelle machine par les communautés de la recherche publique. Un appel à projets ANR dédié aux applications Exascale devrait être lancé dès 2020.
- Le financement des projets français (ou en participation des équipes françaises dans des consortia européens) sera apporté pour 50 % par des financements de l'ANR. L'ANR joue également le rôle d'Agence Nationale officielle de Financement garantissant la mise en œuvre des futurs accords qui seront passés entre la France et la JU (*Grant Agreement*).
- L'accès à la machine devrait être organisé dans la suite des procédures actuelles, au niveau national via GENCI, au niveau européen via le consortium PRACE (*Partnership for Advanced Computing in Europe*)³.
- La machine a vocation à être utilisée pour partie par des industriels. Certains grands industriels sont tout à fait armés pour cela. D'autres actions seront développées, qu'il s'agisse d'actions sectorielles (IA, santé ...) ou pour aider des entreprises à accéder à de

¹ GENCI le grand équipement national de calcul intensif créé en 2007 par les pouvoirs publics,

² Le TGCC (Très Grand Centre de Calcul) est une infrastructure pour le calcul scientifique très haute performance et le Bigdata, capable d'héberger des supercalculateurs d'échelle pétaflopique. Ce centre de calcul a été conçu pour accueillir la première machine pétaflopique française Curie.

³ PRACE a officiellement été créée au printemps 2010 sous la forme d'une association internationale sans but lucratif (AISBL) de droit belge. A cette date, PRACE regroupe 26 pays membres : l'Allemagne, l'Autriche, la Belgique, la Bulgarie, Chypre, le Danemark, l'Espagne, la Finlande, la France (représentée par GENCI), la Grande-Bretagne, la Grèce, la Hongrie, l'Irlande, Israël, l'Italie, le Luxembourg, la Norvège, les Pays-Bas, la Pologne, le Portugal, la Serbie, la Slovaquie, la Suède, la Suisse, la Tchéquie et la Turquie.

nouvelles technologies de calcul, les structures susceptibles de porter ces actions ne sont pas indiquées (ANR, Teratec, ...).

Les partenaires de premier rang du projet sont donc la DGRI du MESR (et la DGE du MEFR), GENCI, le TGCC du CEA, voire l'ANR. Un groupe miroir chargé de donner un avis sur la préparation de positions françaises a été constitué. Il comprend la DGRI, la DGE, Genci, le CEA, le CNRS, l'INRIA, Teratec⁴ et ATOS⁵. Au sein de GENCI, le projet Exascale est doté d'un comité de pilotage supervisé par un Comité Directeur.

Ce projet rassemble un grand nombre d'acteurs et il semble nécessaire d'en établir la liste afin de :

- Compléter l'analyse du coût du projet, en intégrant les coûts supportés par ces acteurs. En dehors du projet proprement dit, la disponibilité des ressources pour financer la montée en compétences est un facteur clé de succès.
- Compléter l'analyse des ressources humaines nécessaires, chez les différents acteurs (personnels en place ou recrutement à réaliser).
- Présenter l'organisation du projet et sa gouvernance. C'est évidemment important en termes d'évaluation des risques. Mais c'est aussi important pour évaluer dans quelle mesure on se donne les moyens de réaliser le potentiel du calcul haute performance. L'évaluation socio-économique devrait s'appuyer sur le potentiel du calcul haute performance (la performance théorique) d'une part, sur une évaluation du « rendement de l'écosystème », d'autre part, pour réaliser une estimation réaliste de l'impact socio-économique du projet.

2.2. Contexte

Nous reprenons ici le contexte du projet figurant dans l'étude d'impact objet de notre contre-expertise (abrégé « Dossier GENCI » ci-après) afin de faire apparaître plus nettement certaines interdépendances décisionnelles ou technologiques, la relation avec la stratégie nationale, européenne et celle des grands acteurs internationaux.

2.2.1. **L'Exaflop : nouvelle frontière**

La mesure de la puissance en Flops⁶ est utilisée depuis longtemps déjà pour décrire les calculateurs utilisés en simulation numérique scientifique, technologique ou industrielle. Le terme Exascale en descend, pour décrire une gamme capable de 10^{18} Flops. Depuis plus de 40 ans, le passage d'un ordre de grandeur au suivant intervient approximativement tous les 10 ans, et requiert la combinaison d'innovations scientifiques, technologiques, architecturales, logicielles, en mathématiques numériques et dans les domaines d'application.

S'il n'y a pas de « mur de l'Exaflop », le franchissement est néanmoins un repère pour la compétitivité des domaines concernés, l'orientation et la motivation des équipes scientifiques et industrielles. On en attend de nouvelles applications, à fort impact sur le paysage technologique et industriel, et le dossier GENCI comporte plusieurs documents, produits par la CE, pour l'illustrer.

⁴ Teratec : pôle européen de compétence en simulation numérique haute performance animé par le CEA et des partenaires industriels

⁵ ATOS : Atos est une entreprise de services du numérique (ESN) française, créée en 1997, dont BULL est une filiale

⁶ Opérations arithmétiques en virgule flottante 64 bit par seconde

Une convergence s'est effectuée entre le domaine du Calcul Intensif, celui de l'exploitation des données massives et le vaste domaine de l'intelligence artificielle. Sur la frontière de l'Exaflop, ces questions ne sont plus séparées mais innovent en se combinant.

La Commission Européenne (CE) a mis en place l'Entreprise Commune (JU EuroHPC) avec trois objectifs en synergie, au niveau de l'Exascale :

- Calcul Intensif pour la compétitivité scientifique ;
- Calcul Intensif pour son application aux politiques publiques (p.ex. climat, environnement, compétitivité économique) ;
- mise en place d'une filière de processeurs Européens.

Le dossier GENCI rappelle donc à juste titre les motivations, les planifications techniques et les engagements politiques des principaux acteurs mondiaux (USA, Chine, Japon).

Toutefois, il nous semble qu'il aurait pu être plus concret sur la question de la convergence HPC/IA/Données, en faisant référence aux résultats obtenus soit sur la machine Jean-Zay, soit au plan international (p.ex. Prix Gordon Bell de l'ACM) ou dans la collaboration PRACE.

2.2.2. Le périmètre du projet

Ainsi que le rappelle le dossier GENCI, le projet se place dans le contexte de politiques publiques de soutien au Calcul Intensif pour son importance stratégique et économique.

La Commission Européenne (CE), dans le cadre de sa Stratégie du Calcul Haute Performance, a fait adopter le Règlement du Conseil (de l'UE) établissant l'Entreprise Commune Européenne (JU-EuroHPC) pour le Calcul Haute Performance (abrégé RC-UE_JU dans la suite), et précisé ses statuts et missions⁷.

Dans cette décision, la composition de la JU-EuroHPC est précisée, les membres publics étant l'Union Européenne, représentée par la Commission (avec 50 % des parts), et des Etats participants. Les membres privés comprennent les associations ETP4HPC⁸ et BDVA. (Voir schéma d'organisation *infra*). Elle opère comme une agence de financement, dans le contexte de ses missions, et doit être objet d'une évaluation intermédiaire par la CE avant la fin de l'année 2022.

A l'appui de cette démarche, la CE a produit une étude d'impact qui précise ses attendus socio-économiques.

Dans ses statuts, il est précisé que la JU procèdera à des appels d'offres concernant notamment des supercalculateurs pré-Exascale (en cours, 3 consortia désignés) et à un appel d'offres concernant deux supercalculateurs Exascale. La France et l'Allemagne ont exprimé leur volonté de concourir.

Par le moyen de l'appel d'offres Horizon 2020 / ICT-42-2017, la Commission a sélectionné le consortium EPI (European Processor Initiative), en requérant que le candidat contribue à la Stratégie du Calcul Haute Performance. Juridiquement EPI est un projet « European Processor Initiative » financé par le Grant 826647 de la CE (AO : H2020-ICT-2016-2017/H2020-ICT-2017-2) sur la période allant du 01/12/2018 au 30/11/2021, coordonné par BULL SAS.

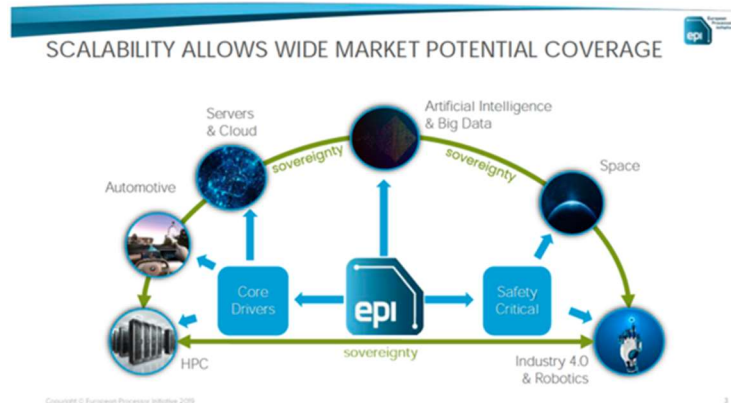
Ce consortium vise à couvrir les besoins d'une pluralité de marchés (voir la figure 1 suivante) allant du Calcul Haute Performance aux systèmes enfouis requis pour la conduite automobile autonome. Il doit notamment produire deux générations de processeurs ARM couvrant les besoins

⁷ Voir COM(2018) 8 final, du 11/1/2018

⁸ respectivement <http://www.etp4hpc.eu/> et <http://www.bdva.eu/>

« pre-Exascale » (Rhea en 2021) et « Exascale » (Cronos en 2022/2023). Il s'intéresse aussi à une seconde architecture dite RISC-V, notamment pour s'affranchir de possibles contraintes sur ARM maintenant propriété d'une société nipponne, mais aussi pour la performance induite par l'architecture.

Figure 1 : Flexibilité et couverture du marché potentiel



Source : © European Processor Initiative, 2019, figure extraite du courrier GENCI du 19/6/20

La Commission a exprimé son intérêt pour l'utilisation de la technologie européenne dans les programmes soutenus par EuroHPC. En réponse à notre demande d'information, la DGRI nous a informé qu'il est prévu qu'EuroHPC finance le consortium EPI par un contrat cadre de partenariat (FPA⁹) de 80 M€, couvert à 50 % par la Commission Européenne. Ceci prévoit également une ouverture de ce consortium, qui comprend déjà une masse critique d'acteurs industriels et académiques.

Dans cette contre-expertise, nous avons considéré EPI, et les entités qu'elle soutient comme externe au périmètre à évaluer, mais nous nous sommes intéressés à documenter les interactions avec le projet.

2.2.3. L'accueil en France d'un super ordinateur Exascale

Comme nous le verrons dans la partie consacrée aux risques, beaucoup reste encore à décider. Dans le calendrier ci-après nous distinguons les actions de la CE ou de la JU par une *typographie italique*.

Fin 2020-Début 2021:

- Mise en place d'un consortium national et Européen dirigé par GENCI (abrégié G&A¹⁰ dans la suite) pour concrétiser le projet d'une machine Exascale hébergée au TGCC (CEA/Bruyères le Chatel). Pour être en mesure de déposer sa candidature auprès de la JU, le G&A doit satisfaire aux contraintes imposées par la CE et/ou la JU, dont certaines sont connues, en raison de la formation en cours des consortia pré-Exascale. Notamment la doctrine d'inclusivité de la CE impose l'intégration de partenaires Européens.

⁹ Framework Partnership Agreement

¹⁰ G&A: GENCI et associés; la réelle désignation n'est pas connue des contre-experts

- Les partenaires français pourraient inclure, sur décision gouvernementale, des organismes de recherche, dont le CEA, aussi bien que des industriels en raison de leurs besoins Exascale ou pré-Exascale.
- Des discussions sont en cours avec l'Allemagne autour d'un scénario prévoyant un consortium commun se portant candidat pour l'hébergement des deux machines Exaflopiques, une dans chacun des deux pays.

Mi 2021- Fin 2021:

- La JU fait appel à candidatures pour les entités d'hébergement destinées à l'installation et l'opération des deux machines Exaflopiques. Le RC-UE-JU précise certains critères de la Convention d'Hébergement, les autres éléments étant décidés par le Comité Directeur de la JU.

Début 2022:

- *Sélection des entités d'hébergement par la JU, contractualisation*. Immédiatement après: *lancement des appels d'offres pour les machines Exascale*. Les modalités¹¹ prévoient que :
 - la JU soit seule responsable pour la mise en œuvre du processus d'acquisition ;
 - que le lancement de l'appel d'offres soit fait en étroite collaboration et avec le support de l'Entité d'Hébergement ;
 - que l'Entité d'Hébergement soit associée au processus p.ex. pour la vérification des spécifications techniques.

Fin 2022 - Debut 2023:

- *Choix des systèmes Exascale*. En pratique, pour chaque système la JU conduira une procédure de « dialogue compétitif » en étroite collaboration avec l'Entité d'Hébergement. C'est la procédure de choix pour les super calculateurs les plus puissants, qui sont *in fine* des systèmes définis sur mesure. Des benchmarks représentant le besoin de l'acquéreur, annexés au contrat, servent à la recette du matériel.

Mi 2023 :

- Mise en service du premier système, celle du second intervenant un an après.

2.2.4. La configuration souhaitée

Le dossier de GENCI décrit une configuration dans ses grandes lignes et caractérise l'utilisation qu'il est prévu d'en faire :

- une partition principale donnera l'essentiel de la puissance de calcul, avec des noeuds « accélérés » comportant des circuits « GPU », des partitions annexes sont prévues pour le calcul sans accélération, les pré- et post- traitement, un stockage de données sur circuits « mémoire flash » pour un débit de données très important ;
- est également prévue une partition comportant un dispositif de calcul quantique ;

¹¹ En supposant qu'elles soient semblables au cas pré-Exascale pour lequel nous disposons du document « Model Hosting Agreement to be annexed to the Call for Expression of Interest »

- il est prévu que l'équilibrage architectural de la partition principale en fera une machine « convergée » efficace aussi bien en simulation numérique classique que pour l'IA et le traitement de données massives.

Les moyens de stockage sont décrits séparément ainsi que leur financement (20 M€ CAPEX + 20 M€ OPEX). A l'été 2020, GENCI ne sait pas avec certitude si ce niveau d'environnement du calculateur sera ou non pris en charge par la JU.

Le dossier de GENCI précise que l'utilisation prévue de ce matériel est un mélange de mode de « capabilité¹² » et de « capacité », destiné à obtenir un bon rendement d'exploitation. Toutefois, le dossier manque de précision sur la description de la charge d'exploitation du calculateur, ce qui serait utile :

- pour mieux analyser les types d'utilisation considérés, et donc l'apport de la configuration Exaflopique¹³;
- pour comprendre l'équilibrage souhaitable entre les diverses partitions décrites ci-dessus ;
- pour permettre une meilleure comparaison avec les configurations concurrentes au niveau international, pour lesquelles des éléments sont disponibles, parfois avec des modes d'exploitation innovants.

Ce type d'information doit nourrir la constitution du cahier des charges, qui se fera en étroite collaboration avec la JU. Ce point est repris dans la discussion des risques.

Dans le contexte actuel de développements technologiques attendus (architectures quantiques, processeur européen EPI), mais comportant une part d'incertitude notamment quant à leur disponibilité et leur coût d'acquisition, un phasage de l'investissement pourrait s'envisager, en particulier quant à la partition exploratoire ou de rupture. Cette possibilité permettrait éventuellement de laisser du temps aux communautés scientifiques concernées ou intéressées pour migrer leurs codes et éviter une double rupture technologique.

2.2.5. L'utilisation nationale du calculateur

Les supercalculateurs d'une certaine importance ne se conçoivent que dans le contexte d'un écosystème de moyens matériels (dont calculateurs, réseau et gestion de données), logiciels et humains. Au niveau technique il s'agit de les alimenter, d'en suivre l'opération et d'exploiter les résultats ; aujourd'hui ceci demande aussi l'interopérabilité avec un grand nombre de ressources numériques. Cet écosystème doit aussi donner la liberté aux opérateurs de recherche, et aux industriels concernés, pour déployer des projets significatifs ou stratégiques dans la durée, bien au-delà de l'obsolescence d'une technologie ou d'une installation particulière.

En pratique, les écosystèmes du Calcul Intensif national et européen sont organisés sous la forme d'une pyramide de ressources interconnectées, qui tend de plus en plus à adresser les besoins en Calcul Intensif, Données Massives et réseaux « Cloud » de services distribués.

¹² La **capabilité** fait référence aux cas d'exploitation demandant un calculateur de cette configuration (nombre de processeurs, taille mémoire, organisation, etc.) et ne pouvant être exécutés raisonnablement sur de moindres configurations; la **capacité** fait référence au seul volume de calculs pouvant être exécutés.

¹³ Complémentarité avec les autres installations GENCI, avec les installations accessibles via PRACE, capacité de « workflow » avec des grands instruments (lignes de lumière, radiotélescope interférométriques, grandes bases de données génomiques, imagerie avancée)

Les niveaux les plus élevés de la pyramide nous concernent directement ici, mais leur pertinence demande un équilibre d'ensemble, aussi bien concernant les moyens techniques que de l'organisation qui permet de les affecter aux besoins.

- Le niveau le plus élevé était constitué des Tier-0 Européens de l'infrastructure PRACE, dont le but était de permettre des calculs ou simulations situés au-delà des capacités d'un pays seul. Au fur et à mesure du déploiement de l'investissement de la JU EuroHPC ce niveau sera peuplé des calculateurs pré-Exaflopiques et Exaflopiques.
- Le second niveau (dit Tier-1) est celui des calculateurs d'intérêt national, gérés par GENCI pour la recherche publique en France. Nous renvoyons au Rapport de la Cour des Comptes pour les exceptions à cette règle d'organisation telles le CC-IN2P3 traitant des données du CERN.
- Le troisième niveau est dit « méso-informatique » est spécialisé par localisation ou par destination.

La situation actuelle

Aujourd'hui, l'allocation de moyens (temps de calcul) au niveau Tier-0 se fait par les instances de la TGIR-PRACE, sur dossier scientifique; au niveau Tier-1 GENCI a mis en place plusieurs possibilités d'allocation, toujours sur critère scientifique. Même sur des calculateurs identiques les politiques d'attribution de PRACE et GENCI ne sont pas identiques:

- la part PRACE est très compétitive, vise à des allocations importantes pour des projets d'excellence, et qui ont besoin de ressources exceptionnelles. Il n'y a pas de garantie de pouvoir déployer une recherche sur plusieurs années par des allocations successives.
- la part GENCI est allouée aux projets via les comités d'évaluation et d'attribution, sur la base de dossiers scientifique. Le comité scientifique d'évaluation permet de suivre l'évolution des capacités scientifiques dans le temps, par exemple avec des continuations pluriannuelles de projets pour les équipes.

La combinaison des deux modes contribue à l'efficacité de la politique de recherche. Il en va de même de l'équilibre de l'écosystème en pyramide évoqué plus haut.

Dans les deux cas, des programmes particuliers permettent : l'allocation aux projets industriels, l'accès préparatoire, la satisfaction de besoins de communautés identifiées, par exemple la communauté d'Intelligence Artificielle sur la machine Jean-Zay de GENCI ou la communauté « climat ».

Le cas du calculateur GENCI Joliot-Curie, installé au CEA/TGCC, préfigure la situation à venir au niveau Exaflopique. En effet cette machine est partagée entre l'utilisation nationale (47 % en 2019) et l'utilisation affectée à PRACE (53 % en 2019). Tous les utilisateurs peuvent utiliser l'intégralité des ressources si leur code le justifie, le partage s'effectuant au niveau temporel. Ce mécanisme permet l'évolution des méthodes et des codes pour en accroître le parallélisme, ce qui reste un enjeu méthodologique et prépare l'Exascale. Suivant les exigences ou les phases de leurs projets, les utilisateurs peuvent évoluer entre les niveaux méso-, Tier-1 et Tier-0 sans obstacle.

Dans le cas de l'Exascale :

- 50 % sera réservé à l'allocation prévue au niveau Européen par la JU-EuroHPC ;
- 50 % sera réservé à l'allocation dans le cadre du consortium national (G&A ci-dessus).

Au total, les utilisateurs français pourront utiliser :

1. la machine Exascale hébergée en France, en candidatant selon 2 voies :

- la part qui leur sera attribuée dans le cas des appels d'offres compétitifs effectués par la JU, probablement dans l'esprit du mode Tier-0 décrit ci-dessus ;
 - la part des 50 % disponible une fois pris en compte le partenariat G&A.
2. les autres machines de la JU, par les appels d'offres compétitifs de la JU ;
 3. le cas échéant, des ressources PRACE non associées à la JU ;
 4. les autres ressources de GENCI dans les conditions actuelles. Ce point est mentionné ici, bien qu'il ne relève pas directement du projet Exascale, en raison du possible impact de ce projet sur ces ressources, notamment des difficultés budgétaires soulevées dans le dossier de GENCI.

L'expérience passée de PRACE-1 et PRACE-2 montre que les utilisateurs français ont pu obtenir respectivement 20 % puis 15 % de la ressource, la diminution lors du passage à PRACE-2 étant due à l'entrée de la Suisse comme membre hôte et aussi comme utilisateur.

Des inconnues subsistent, mais nous notons que, selon son résumé exécutif publié, le Conseil de la JU-EuroHPC, pour la séance des 9 et 10 Décembre 2019, a débattu de politique de distribution des temps d'accès sur la base de travaux de PRACE et du Comité INFRAG de la JU. PRACE prépare une offre de service pour la JU, dans la phase PRACE-3 (2021) pour mettre en œuvre l'allocation du temps de calcul sur les machines de la JU et apporter des services de formation et d'accompagnement aux utilisateurs.

2.3. Présentation de l'investissement et explicitation des objectifs

La Commission Européenne investit conjointement avec les États membres pour mettre en place une infrastructure de supercalculateurs de classe mondiale à l'échelle européenne, via l'entreprise commune européenne pour le calcul haute performance EuroHPC. Comme le Conseil des Ministres l'a souligné dès 2018, cette collaboration est vitale pour assurer la compétitivité et l'indépendance de l'UE dans une économie fondée sur les données. En effet, lors de la création d'EuroHPC, un tiers des ressources mondiales de calcul haute performance étaient consommées par des entreprises de l'EU, alors qu'elles ne fournissaient que 5 % de ces ressources. L'objectif affiché est de se doter d'une infrastructure de calcul haute performance européenne et de soutenir le développement d'un écosystème européen de la recherche et de l'innovation destiné à stimuler le secteur de l'équipement technologique¹⁴.

L'investissement dans une machine Exascale française doit permettre de lever deux problèmes conjoncturels : d'une part la nécessité de pouvoir traiter des applications scientifiques extrêmement gourmandes en ressources de calcul et d'analyse de données (simulations multi-échelles, résolution accrue des modèles et des domaines de calcul, modélisation multi-physique, ou encore convergence d'usages entre la simulation numérique et l'intelligence artificielle) actuellement hors de portée, et d'autre part de soulager la pression sur les moyens de calcul actuels de GENCI. Sans oublier que de tels systèmes assurent un accès souverain à des moyens de stockage de données aux chercheurs scientifiques et industriels européens. Dans le document GENCI, quelques enjeux scientifiques majeurs (calcul et stockage dans les domaines du climat, de la mécanique des fluides, de la fusion magnétique) sont rappelés afin de mesurer l'importance de ce type d'investissement auprès des communautés scientifiques concernées.

L'investissement prévu comprend deux volets spécifiques : le supercalculateur Exascale et les travaux d'aménagement du centre. L'étude réalisée par GENCI est basée sur l'hypothèse d'un coût d'investissement de l'ordre de 160 M€¹⁵ avec un taux d'incertitudes de +/- 7 %.

¹⁴ EU Budget for the Future, Investing in the Future, Digital Transformation 2021-27 (Programme pour une Europe Numérique), juin 2018.

¹⁵ Montant communiqué par la Ministre de l'enseignement supérieur et de la recherche en janvier 2019.

Les spécifications et la composition précise du supercalculateur Exascale que JU EuroHPC et GENCI souhaitent co-acquérir sont encore inconnues à ce stade de l'étude ; elles feront l'objet d'un appel d'offres ouvert, coordonné par la JU. Néanmoins, l'architecture technologique modulaire suggérée devrait s'articuler autour d'une partition principale, basée sur des nœuds de calculs à base de processeurs scalaires couplés à des accélérateurs de calcul (de type GPU) fournissant environ 85 % de la puissance nominale de la machine, couplée à une partition secondaire basée sur des nœuds scalaires à grande mémoire. Deux partitions viennent compléter cette architecture classique, une partition de pré et de post-traitement de données et une partition exploratoire ou de rupture, qui pourrait intégrer des technologies émergentes ou de rupture (à base d'accélérateurs quantiques), plus spécifiquement destinée à des communautés et des défis scientifiques ciblés (dynamique moléculaire, santé, science des matériaux ou intelligence artificielle). Un espace de stockage privé de premier niveau (à très haut débit) sera intégré dans le dispositif envisagé. Ce supercalculateur est conçu pour un usage dit convergé, c'est-à-dire capable d'adresser les besoins en modélisation/simulation et ceux de l'intelligence artificielle, dont beaucoup se recouvrent désormais. Une partie du budget (de l'ordre de 40 M€) sera dédiée à la mise en place de moyens de stockage de données massives à des fins de traitement (niveau secondaire) ou d'archivage (troisième niveau).

Le centre d'hébergement prévu (centre de calcul du TGCC) doit être aménagé pour accueillir une machine dont les spécifications techniques diffèrent de celles des machines actuellement installées. Le refroidissement par eau chaude (eau à température ambiante) devrait permettre d'obtenir une efficacité accrue, avec un PUE¹⁶ cible de l'ordre de 1.1. La maîtrise de la consommation énergétique et l'efficacité des dispositifs d'alimentation électrique et de refroidissement sont une priorité qui sera affichée dans la mise en place d'appels d'offres.

Les évaluations des coûts prévisionnels, regroupant les aspects relatifs aux travaux d'infrastructures, au coût d'acquisition du calculateur et aux coûts de fonctionnement, sont basées sur des hypothèses relatives au dimensionnement du système et à ses spécifications techniques. En dépit d'un probable appel d'offre européen euroHPC attendu fin 2020 qui devrait fixer les orientations et donner des spécifications techniques pour devenir « Hosting Entity », les estimations sont données sans marge pour aléa. Le coût total d'exploitation du projet (TCO) est estimé à 405.9 M€. Le document fournit deux hypothèses de co-financements des coûts par la JU (182,8 M€ et 203 M€), en complément du budget GENCI (82.4 M€) et de l'apport du PIA3 (80 M€). L'hypothèse la plus favorable table sur un co-financement à hauteur de 50 % du TCO de JU, à l'instar de la contribution aux projets pré-Exascale.

Quelle que soit l'hypothèse de budget retenue, un financement complémentaire (estimé entre 40,6 M€ et 60,7 M€) est indispensable à la candidature française pour une machine Exascale. Faute d'un tel apport, la France ne pourra pas concourir à l'hébergement d'une telle machine. Les pistes évoquées par GENCI incluent la constitution d'un consortium élargi à d'autres pays et la recapitalisation de GENCI. La mise en place d'une cotisation financière annuelle aux organismes de recherche non associés à GENCI est également évoquée. Il convient de noter que si ce dernier apport financier ne comble pas le déficit de budget présenté par GENCI, il permettrait néanmoins de financer des recrutements en support et accompagnement des communautés utilisatrices des moyens de calcul. L'hypothèse d'une contribution financière des organismes non partenaires, pour un retour de l'ordre de 1 M€ annuel, semble possible. Elle demande toutefois l'adhésion de ces organismes au projet Exascale et l'articulation avec leurs politiques technologiques, car ils disposent pour la plupart de centres de calcul et de compétences.

Le dimensionnement de la machine apparaît à ce stade de l'étude comme une variable d'ajustement du projet, avec en filigrane la possibilité d'un retrait de la candidature française en cas de sous-financement. Pour autant, les recommandations européennes en matière de priorités scientifiques et le recul dont dispose GENCI dans la gestion de centres de calcul haute performance

¹⁶ Indicateur d'efficacité énergétique

devraient permettre de détailler dès aujourd'hui les partitions de la future machine Exascale. GENCI indique, sans toutefois les détailler, les coûts prévisionnels de la machine Exascale, en se référant au coût des dernières machines installées (i.e., Joliot Curie et Jean Zay). Le coût d'investissement intègre trois éléments : le calculateur, la capacité de stockage de premier niveau (mais pas les niveaux deux (stockage de masse) et trois (archivage)) et les travaux d'infrastructure. Les coûts d'exploitation et d'environnement prévisionnels des machines à l'IDRIS et au CINES ont été reconduits à l'identique.

Dans l'hypothèse d'un co-financement européen, la France disposera de 50 % des 10 milliards d'heures de calcul de la machine Exascale, représentant un gain d'un facteur 2,5 par rapport au nombre d'heures de calcul actuelles disponibles pour les chercheurs français.

Une réflexion engagée dans d'autres centres européens de calcul haute performance (Jülich par exemple¹⁷) propose d'estimer la meilleure configuration technologique pour un TCO donné. Le document cité propose par ailleurs un modèle de calcul de coûts complet par nœuds sur la durée de vie de l'investissement.

Nous recommandons que GENCI s'inspire de cette étude pour produire une analyse plus détaillée de ces coûts.

Le saut technologique sera tel avec ce type d'architecture émergente, qu'il imposera a minima une réécriture de la plupart des logiciels et des codes de calcul, voire au-delà, la conception de nouveaux algorithmes. GENCI et ses associés ont conscience que la réussite de cette opération réside dans la capacité d'accompagner les communautés dans le portage et la réécriture des applications scientifiques.

L'approche préconisée par GENCI consiste à faire appel à des financements issus de projets de recherche déposés en réponse à des appels à projets (ANR, européen).

Ce mécanisme nous semble comporter une grande part d'aléa pour répondre de manière concrète et pertinente à un tel enjeu. Il y a un risque significatif de décrochage scientifique de communautés qui n'auraient pas été capables (par exemple faute d'appels ciblés) de lever les fonds nécessaires qu'impliquent ces réécritures. Par ailleurs, il faut s'assurer de la mise en place de guichets adaptés au financement du portage et de la réécriture de logiciels applicatifs (qui par nature requièrent des ingénieurs experts), à côté des guichets plutôt réservés à des projets de recherche et au recrutement de chercheurs.

Une base nécessaire du renforcement du support est esquissée dans le document de GENCI qui envisage de mutualiser une équipe de support applicatif entre les centres nationaux et avec certains centres régionaux, avec un abondement budgétaire de l'ordre de 1.5 à 2 M€ annuel permettant de recruter des CDD experts. Ces investissements en ressources humaines sont fondamentaux pour permettre la migration et le développement des applications.

La relative faiblesse des budgets consentis pour la migration d'applications nous semble à ce stade à peine suffisante pour permettre de soutenir les domaines prioritaires des organismes de recherche. Il est en outre peu probable que des communautés émergentes (sciences humaines et sociales, santé/médecine), moins familières du calcul intensif et de l'analyse de données, puissent ainsi accéder à ce type de ressources. Confier aux centres régionaux la tâche d'accompagner ces communautés est certes vertueux, mais intègre un risque lié aux disparités de ressources et de politiques scientifiques de ces centres.

¹⁷ 8th European Workshop on HPC Centre Infrastructures 4-7 April 2017, Swiss National Supercomputing Centre - CSCS

En résumé, le saut technologique et le niveau de performances de cette machine demande une politique proactive d'accompagnement et de préparation à son utilisation, financée en marge du projet ou, mieux, dans une phase préparatoire. De telles préparations ont été développées sur des périodes de 4 à 7 ans aux USA ainsi qu'au Japon, et par les appels d'offres « Centres d'Excellence » de la CE¹⁸. Au vu des pratiques constatées à l'étranger, les actions entreprises au niveau national ne sont ni assez développées ni suffisamment proactives pour répondre au besoin. Faute d'anticipation, il y a un risque notable et significatif de sous-utilisation ou d'utilisation non optimale de l'architecture. La valeur nette de l'investissement pourrait décroître et la part relative française pourrait diminuer singulièrement au profit d'autres utilisateurs. Ce problème de phasage ne doit pas être négligé ou minoré dans la mise en valeur de ce projet.

2.4. Les différentes options d'investissement

Le document de GENCI présente comme « option de référence » le renoncement de la France à co-acquérir une configuration Exascale, soit a) en se retirant complètement du projet européen Exascale, soit b) en participant comme partenaire co-financeur et non hébergeur d'un consortium européen.

Avec l'option a), le décrochage technologique serait important, rapide et probablement irréversible, puisque la France ne disposerait avec ses trois supercalculateurs scientifiques que d'une puissance cumulée d'environ 140 Pflops, inférieure par exemple à la puissance disponible par un projet pré-Exascale européen déployé dès 2021 (Espagne Italie et Finlande), de l'ordre de 180- 200 Pflops.

Avec l'option b), la France intègre un consortium européen dans des conditions qui risquent d'impacter la souveraineté des données produites ou collectées et le niveau d'influence dans les décisions prises par ses partenaires et la JU concernant les caractéristiques et l'usage du système.

L'option Exascale vise à doter la France d'une machine européenne Exascale en coacquisition avec EuroHPC, dont elle bénéficie de 50 % des ressources. Les deux supercalculateurs des centres nationaux de l'IDRIS et du CINES venant compléter cette offre technologique, pour satisfaire l'ensemble des besoins nationaux.

D'autres options pourraient être envisagées.

La décision d'investir dans un supercalculateur européen est une priorité scientifique affirmée de l'Europe, dans un contexte de compétition économique et politique entre les nations développées. EuroHPC doit être présent dans le haut du classement TOP500 des superordinateurs. Cette décision a pour conséquence de « figer » la proposition de GENCI autour d'une architecture Exascale polyvalente. Elle exclut *de facto* toute option de renforcement des moyens de calcul par grand domaine scientifique (ex. météo, études aéronautiques, santé/médecine, géosciences, physique des hautes énergies, etc.).

Avec l'option b) de l'option de référence, une question reste en suspens qui mérite d'être adressée dans le document. Si on fait l'hypothèse que d'autres pays se porteraient candidats à l'accueil d'une machine Exascale, est-ce que la France n'obtiendrait pas un impact plus important en rejoignant un consortium et en consacrant moins d'argent à l'infrastructure et plus d'argent à l'accompagnement des communautés scientifiques françaises, pour gagner accès au calcul haute performance, dans le cadre des appels à projets compétitifs internationaux ? La question qui est

¹⁸ Premier appel d'offres: H2020-EINFRA-2014-2015 ouvert le 24/9/14, <https://ec.europa.eu/programmes/horizon2020/en/news/overview-eu-funded-centres-excellence-computing-applications>

soulevée ici est celle de l'optimisation du mix investissement dans l'infrastructure / investissement dans l'utilisation de l'infrastructure.

Il faut cependant souligner que la candidature française à l'hébergement d'un ordinateur Exascale, présente évidemment l'avantage d'être pleinement en phase avec la politique d'investissement européenne sur les infrastructures de calcul haute performance de classe mondiale. Vu le retard pris par l'Europe vis-à-vis des USA, de la Chine et du Japon, retard qui est à l'origine de l'initiative Exascale, une « stratégie du coucou » serait contre-productive de la part d'un grand pays européen.

2.5. Identification des effets attendus des options d'investissement

Les effets identifiés par GENCI de l'option de référence sont ceux d'un décrochage technologique de la communauté scientifique française (perte de développement dans des nouveaux algorithmes, codes qui ne sont plus à l'état de l'art, perte de compétences du capital humain). Ce décrochage induit une fuite des cerveaux vers les pays disposant des moyens d'effectuer des recherches de pointe (baisse d'attractivité pour les jeunes et fuite des chercheurs), qui induit à son tour une incapacité de la France à répondre aux grands enjeux scientifique qui sous-tendent la simulation du climat, de l'énergie, de la santé et du transport et une perte de souveraineté (données hébergées à l'étranger, dépendance à un tiers pour ses propres simulations et perte de compétences)).

Les effets identifiés de l'option Exascale sont l'opposé des effets précédents. Dans les 2 cas, l'analyse reste qualitative.

Cette identification des effets attendus ne correspond pas aux recommandations inscrites au Guide de l'Evaluation Socioéconomique des Investissements Publics (chap. 3, sect. 1). En effet, il est demandé d'identifier en premier lieu les acteurs affectés et en second lieu, la demande qui sera adressée à l'investissement envisagé, puis d'identifier les effets marchands et non-marchands pesant sur les différents acteurs.¹⁹

Dans la liste des effets attendus identifiés par GENCI, on perçoit les effets attendus sur la communauté scientifique (baisse de compétitivité, fuite des cerveaux) sans pourtant que soit fourni un indicateur quantitatif permettant d'apprécier cette baisse prévisionnelle. En ce qui concerne la demande de calcul dans le cas de l'option de référence, il est seulement précisé qu'elle explosera. Enfin, on ne trouve aucune trace d'effet marchand ou non-marchand.

Le document de GENCI ainsi que les entretiens conduits au cours de la contre-expertise laissent apparaître les acteurs suivants :

- l'Etat français qui va fournir la moitié du financement de l'investissement ;
- GENCI qui est l'opérateur de l'investissement ;
- la JU qui effectuera l'appel d'offre et sera propriétaire de l'investissement ;
- la communauté scientifique française qui aura un accès réservé à 50% de la machine Exascale ;
- la communauté scientifique européenne (incluant également la France) qui aura un accès au 50 % restant de la machine Exascale ;
- la communauté scientifique internationale qui aura accès à la machine Exascale à travers des partenariats avec les laboratoires européens ou bien en séjournant dans un laboratoire européen ;

¹⁹ Voir <https://www.strategie.gouv.fr/publications/guide-de-levaluation-socioeconomique-investissements-publics>

- les tissus industriels français et européens qui auront accès à la machine Exascale dans des conditions encore à déterminer ;
- les mêmes industriels, non utilisateurs directs, qui seront concernés par les externalités: progression de l'état de l'art, amélioration des capacités des méthodes et logiciels résultant de l'utilisation de la machine Exascale par les concepteurs desdits logiciels, au-delà des seuls industriels.

En ce qui concerne la demande en calcul, bien que le rapport d'évaluation ne fasse état d'aucune statistique, GENCI dispose de dix années d'historique — de la demande et de l'utilisation — par champ scientifique mais également par type de technologies, processeur scalaire ou GPU, et par niveau de parallélisation (taux d'applications utilisant plus de 1000 cœurs de calcul, par exemple). Ces historiques associés aux différents enjeux stratégiques européens et français en matière de priorités de recherche scientifique devraient être utiles pour dimensionner la taille et la configuration de la machine Exascale.

En ce qui concerne les effets marchands, la JU envisage d'offrir un accès tarifé à la machine Exascale. Dans la présentation du projet EuroHP fournie par Genci, on trouve la mention d'un accès jusqu'à 20 % de la machine à des usages commerciaux (*'up to 20 % of access time for pay-per-use commercial services'*). Si cette possibilité se confirmait, elle aurait un impact significatif sur la valeur créée par l'investissement, qui doit être pris en compte dans l'évaluation socioéconomique.

Enfin, bien qu'ils ne soient pas cités explicitement, on peut citer plusieurs effets non-marchands parmi lesquels :

- le gain de productivité scientifique obtenu par une réduction du temps de production d'une publication scientifique ;
- le coût évité par la communauté scientifique de l'usage de solutions de calcul alternatives ou le recours à des pratiques et méthodes moins efficaces ou plus éloignées de l'état de l'art international ;
- les subventions implicites induites par les heures de calcul accordées aux entreprises ;
- les coûts d'adaptation ou de renouvellement des codes de simulation à la nouvelle architecture de la machine Exascale ;
- le coût des émissions de carbone liées à l'alimentation électrique de la machine Exascale et de son environnement de refroidissement.

Cette liste d'effets, que nous appellerons globaux, ne prétend pas être exhaustive et ne saurait justifier que GENCI limite son analyse à ce périmètre. Toutefois, il est requis qu'a minima, tous ces effets non-marchands fassent l'objet d'une estimation précise dans le cas des différentes configurations possibles d'une machine Exascale.

Par ailleurs, à la liste précédente qui concerne les effets non-marchands globaux lié à une infrastructure de calcul, s'ajoute un second niveau d'effets attendus qui concerne les futures applications scientifiques elles-mêmes qui feront usage de la machine Exascale. Cette analyse plus fine est requise pour permettre une évaluation de son intérêt. Elle peut être conduite en usant de la méthodologie fondée par le groupement JASPERS²⁰ pour l'évaluation des projets de R&D&I pour la Commission Européenne. Préconisée également par le rapport France Stratégie de Février 2019 portant sur l'évaluation des projets immobiliers dans l'enseignement supérieur, cette méthode classe les bénéfices sociaux-économiques d'un projet public d'infrastructure de R&D en :

- bénéfiques industriels et commerciaux :

²⁰ Swerdlow R., Teichmann D. et Young T. (2016), « Economic analysis of research infrastructure projects in the programming period 2014-2020 », Staff Working Papers, Jaspers Smart Development Division.

- a) développement de nouvelles entreprises innovantes,
- b) développement de nouveaux produits,
- c) retombées collectives d'une connaissance commune
- bénéfiques pour les chercheurs et les étudiants :
 - a) découverte d'un nouveau champ de recherche,
 - b) formation du capital humain,
 - c) développement du capital social (c'est-à-dire développement de nouvelles connaissances communes)
- bénéfique pour la population dans son ensemble :
 - a) réduction des risques environnementaux,
 - b) réduction des risques sanitaires,
 - c) effet culturel sur les visites

Nous renvoyons au document de JASPERS décrivant le contenu des différents bénéfices identifiés bénéfiques ou effets que nous appellerons *particuliers*.

En outre, relativement aux deux références citées ci-dessus, nous recommandons de prendre en compte dans l'analyse socio-économique les éléments caractéristiques du domaine numérique, et notamment la propriété intellectuelle représentée dans les codes (applicatifs, système, de conception, ...) et données. La méthodologie sur ces points pourrait être affinée en tenant compte des pratiques distinguant « Open source » et propriétaire. Le projet Exascale peut nécessiter, soutenir ou susciter le développement de cette propriété intellectuelle.

Nous recommandons que GENCI reprenne l'analyse socioéconomique de l'option Exascale en suivant deux grilles d'évaluation, la grille des effets globaux attendus et la grille des effets liés aux applications scientifiques qui utiliseront la machine.

2.6. Identification des effets attendus de l'option de référence

GENCI détaille trois grands effets attendus de l'option de référence :

1. remise en cause de la France dans l'Europe du numérique ;
2. effet négatif sur la compétitivité scientifique et économique ;
3. perte de compétence dans un domaine stratégique et de souveraineté.

Pour le premier point, il s'agirait d'un signal négatif envoyé par la France à ses partenaires européens.

Les effets négatifs sur la compétitivité scientifique et économique seraient de trois ordres : perte d'accès à des heures de calcul sur des machines performantes ; perte de capacité de simulation des modèles actuels à des mailles plus fines ; perte de souveraineté dans des domaines stratégiques.

Enfin, l'option de référence conduirait à une perte de compétence dans les domaines de la cyber-sécurité et de la défense nationale.

Le second point peut faire l'objet d'une prise en compte dans le cadre d'une évaluation socioéconomique d'un investissement public, selon la méthodologie recommandée par France Stratégie.

En effet, l'option de référence conduirait à une réduction du temps de calcul alloué aux communautés françaises et à une délocalisation des simulations numériques pour des applications de grande ampleur, comme les simulations du climat nécessaires au GIEC, vers des pays disposant d'une machine Exascale. La quantification, même grossière, de cette perte de productivité de la communauté scientifique française permettrait d'estimer une partie de la perte économique liée à l'option de référence.

En revanche, la question de la souveraineté ne fait pas l'objet d'une description suffisamment précise pour entrer dans le cadre de l'évaluation socioéconomique demandée. Les organismes ayant besoin de calcul haute performance pour des usages de Défense Nationale disposent de leur propre moyen de calcul dédiés. L'articulation de l'hébergement d'une machine Exascale dont GENCI ne sera pas propriétaire avec les questions de souveraineté n'est pas claire.

La question de la souveraineté nationale doit entrer au titre des effets attendus de l'option Exascale. Il est nécessaire de donner à cette notion un contenu suffisamment précis pour lui donner une place dans l'une de ces deux grilles d'analyse. On peut penser, par exemple et sans valeur d'exhaustivité, que le fait de disposer de l'infrastructure sur le sol français permet dans certaines situations de conflits d'intérêts avec d'autres états de réaliser sans contrainte des études stratégiques ou de vérifier celles présentées par nos partenaires. Ce sont ces situations qui nécessitent d'être précisées.

Nous recommandons donc que GENCI et la DGRI reprennent l'évaluation des effets de l'option de référence en suivant les grilles d'analyse préconisées dans la section 2.5.

2.7. Description qualitative et quantitative de tous les effets pouvant être produits

La liste des effets pouvant être produits se fonde sur le rapport d'étude d'impact de l'infrastructure GENCI réalisée en 2017 à l'occasion de son dixième anniversaire. Cette étude d'impact comporte deux volets : un volet d'impact scientifique mesuré à l'aide d'indicateurs de type bibliométrique et un volet d'impact économique mesuré à l'aide d'indicateurs économétriques.

L'étude d'impact porte sur les effets globaux des services que l'infrastructure de calcul apporte à la communauté scientifique. En résumé, elle montre qu'environ 700 publications par an en France font usage des ressources de calcul de GENCI et que ces articles ont en moyenne un facteur d'impact supérieur à la moyenne nationale.

En ce qui concerne l'impact économique, une analyse économétrique montre qu'un euro de dépense de GENCI induit 1,93 euro de production et 0,87 euro de valeur ajoutée à l'économie française et que 12 emplois sont soutenus par million d'euro. Enfin, environ 15% par an des projets bénéficiant d'heures de calcul sur les machines de GENCI sont réalisés en partenariat avec des industriels.

La description qualitative et quantitative des effets attendus ne correspond pas aux attendus de la doctrine de France Stratégie sur l'évaluation des investissements publics.

Tout d'abord, les éléments d'impact économique présentés n'apportent pas d'information sur l'intérêt de l'investissement dans la machine Exascale. La méthodologie préconisée par le Guide d'évaluation s'intéresse aux effets spécifiques au projet.

Par ailleurs, les informations bibliométriques fournies donnent bien un premier éclairage quantitatif sur les retombées scientifiques et économiques de l'investissement public en machine haute performance. Toutefois, ces données concernent une estimation de la productivité scientifique liée à l'usage du calcul haute performance en général, alors que serait nécessaire une estimation relative pour permettre une estimation du gain de productivité apportée par

l'infrastructure GENCI. Enfin, l'évaluation de l'option Exascale nécessite une estimation du gain de productivité qu'apporterait spécifiquement cette machine.

A cet effet attendu sur la productivité scientifique, nous recommandons que GENCI complète son analyse par le détail des effets cités dans la section 2.5, effets globaux et effets liés aux futures applications elles-mêmes utilisant l'option Exascale.

2.8. Valorisation monétaire des effets pouvant être valorisés

Le rapport ne fait pas état de la valorisation monétaire des produits de GENCI, en général, et de l'investissement envisagé dans l'option Exascale en particulier. Les raisons invoquées pour cette absence sont que GENCI est une société civile ne générant pas de chiffre d'affaire et qu'elle fournit un accès gratuit à son infrastructure.

L'ensemble des effets attendus, globaux et particuliers, cités au paragraphe 2.5, peuvent cependant faire l'objet d'une valorisation monétaire.

En ce qui concerne les effets particuliers liés aux applications scientifiques utilisant l'option Exascale, nous renvoyons au document descriptif de la méthode d'évaluation de JASPERS cité plus haut.

Pour les effets globaux attendus, nous soulignons les points suivants.

En ce qui concerne les publications scientifiques, le rapport de France Stratégie intitulé « L'évaluation socioéconomique des projets immobiliers de l'enseignement supérieur et de la recherche » de Février 2019 produit par le groupe de travail présidé par M. Emile Quinet fournit une méthode d'estimation de la valeur d'une publication accompagnée d'un exemple. Cette valeur est identifiée au coût marginal de production d'un article scientifique. Il est en première approximation le temps consacré par un chercheur à la production d'un article que multiplie son salaire annuel brut. En comptant une publication par an et un salaire brut annuel moyen de 80 k€, il est possible de valoriser une publication à ce montant. L'analyse d'impact réalisée par GENCI en 2017 indique une production de l'ordre de 700 publications par an s'appuyant sur ses infrastructures de calcul. Cela correspond à une valeur de production annuelle de l'ordre de 56 M€.

Dans ce contexte, comme indiqué au paragraphe précédent, cette estimation absolue ne reflète pas le gain de productivité obtenue par la recherche française grâce à l'accès à l'infrastructure GENCI, ni même le gain que procurerait une machine Exascale bien précise. La section 3.1 propose des pistes pour réaliser cette estimation.

En ce qui concerne les solutions alternatives accessibles aux chercheurs pour obtenir du temps de calcul, rien dans le document d'évaluation ne permet d'estimer que l'acquisition de la machine Exascale est la solution qui procure le plus grand coût évité pour un objectif de calcul fixé.

En revanche, il est possible de fournir une estimation de la valeur de production du temps de calcul. En effet, de nombreux opérateurs offrent désormais de l'espace de stockage numérique et du temps de calcul (Amazon, Google, Microsoft, Apple...). Si l'on se réfère à la grille tarifaire d'Amazon pour ses services de calcul en ligne sur des processeurs scalaires, on trouve un prix de l'heure de calcul de 13 cEuro pour un usage spot, de 7,8 cEuro pour un usage avec engagement d'un an et de 5 cEuro pour un usage avec engagement sur trois ans. En 2019, GENCI a alloué 2 milliards d'heure de calcul. Le cycle d'attribution des projets étant annuel, en considérant une valeur de 7,8 cEuro, une estimation grossière de la valeur de la production de GENCI est de l'ordre de 160 millions d'euros.

De plus, le prix de marché de l'heure de calcul GPU est bien plus élevé que son équivalent scalaire. Amazon affiche un prix de 230 cEuro pour une heure de calcul GPU avec engagement d'un an. Une

décomposition plus précise entre les heures GPU et scalaires allouées conduiraient à une estimation supérieure de la valeur de la production de GENCI.

Par ailleurs, l'infrastructure GENCI étant ouverte aux entreprises à hauteur de 15%, la valeur précédente des heures de calcul donne un ordre de grandeur du montant de subvention implicite accordée aux entreprises françaises.

Dans le contexte précis de l'option Exascale, la possibilité d'une commercialisation de 20 % du temps de calcul par la JU peut changer radicalement l'appréciation de la valeur de l'option Exascale.

En ce qui concerne les coûts d'adaptation des codes de simulation, GENCI dispose d'un retour d'expérience du temps d'adaptation couvrant en particulier les temps ingénieur et chercheur nécessaires à l'adaptation des codes pour les différents portages au cours des dix dernières années. Ces coûts permettraient d'estimer les coûts à prévoir pour l'adaptation à la machine Exascale.

Enfin, le contenu carbone du kWh français étant connu (de l'ordre de 60 gCO₂/kWh), une valeur tutélaire du carbone étant préconisée par le rapport Quinet « La valeur de l'action pour le climat » (de l'ordre de 100 euro/ton sur la période concernée par l'usage de la machine Exascale), et la consommation de la machine pouvant être estimée, il en ressort qu'il est possible d'estimer le coût environnemental de l'installation de cette machine en France.

Nous recommandons donc que GENCI développe ces valorisations monétaires en les adaptant au cas de la machine Exascale.

2.9. Analyse des risques et incertitudes

Ci-dessous sont passés en revue les principaux risques et opportunités du programme.

2.9.1. **Les risques du programme**

Les principales variables à suivre dans l'analyse de risque nous paraissent être:

- les modalités restant à décider par la JU EuroHPC ;
- la capacité de mise en œuvre du projet, technique, scientifique ou humaine ;
- la pratique du pilotage national et européen ;
- la capacité et le coût d'accès aux technologies requises, notamment Européennes ;
- la situation concurrentielle, internationale et intra-Européenne ;
- les capacités financière, et les modalités de mise en œuvre (CAPEX, OPEX, partenariats, participation de la CE,...).

A notre avis, l'étude soumise à la contre-expertise ne développe pas suffisamment ces points.

Les modalités restant à décider :

Il ressort du dossier et des interviews que si le montant budgétaire de la contribution de la Commission Européenne est fixé, restent à déterminer par la JU :

- le cahier des charges de l'appel d'offre pour devenir structure d'accueil d'une machine Exascale ;

- le cahier des charges technique sur les caractéristiques de l'équipement et les modalités de recette (« benchmark »).

Le risque peut être technique²¹, tel les délais et les caractéristiques de la machine, concerner l'adéquation aux besoins et ambitions nationales, ou financier (par ex. qualité de service exigée induisant plus de charge qu'escompté (chez les utilisateurs ou dans le centre prestataire)).

La composition et l'organisation interne du consortium dirigé par GENCI (G&A ci-avant) restent à finaliser. Ceci aura notamment des conséquences sur les moyens financiers et humains et corrélativement sur la répartition de l'utilisation des ressources.

Une fois identifié, ce type de risque devrait être contenu par une cohérence d'action dans les structures techniques et de gouvernance, par l'expertise disponible qui aura pu être développée au plan national ou au travers des programmes coopératifs « PRACE Implementation Phase » financés par la CE depuis 2011.

Dans ce dernier cas, la valorisation de l'investissement matériel et humain, contribuant au projet, pourrait partir des données budgétaires desdits programmes.

La capacité de mise en œuvre du projet, technique, scientifique ou humaine:

Le dossier GENCI aborde ce sujet, en page 37, sans vraiment aller au fond des choses, ni développer les atouts :

- Les partenaires de GENCI, et notamment le CEA qui doit héberger, disposent de compétences sur de nombreux points. Si besoin il est donc possible de former des personnels supplémentaires pour les fonctions techniques et d'assistance aux utilisateurs, à condition de disposer du financement et du temps nécessaire. L'Etat pourrait investir par avance, pour ses propres besoins et pour compenser le fait que le marché ne suscite peut-être pas suffisamment les vocations dont une part importante est accaparée par les grandes entreprises internationales (GAFAM).
- Face à ces risques, GENCI pourrait faire état des acquis des programmes PRACE (Implementation Phase) qui ont permis de développer des compétences et de former des personnels (prototypes, définition et mise en place de standards techniques, utilisation d'un protocole d'acquisition de technologie précompétitive²², etc.). Il doit également être possible de profiter des compétences développées chez les partenaires européens qui n'ont pas vocation à installer cette catégorie de matériels.
- La formation des chercheurs à l'usage de l'Exascale est évoquée sommairement, et il semblerait qu'on soit en attente d'une programmation adéquate à l'ANR.
- La contribution provenant des programmes Européens améliorant les perspectives nationales du projet pourrait être documentée dans le dossier GENCI, distinguant la part de subvention et la contrepartie d'investissement national²³. Il en va de même des soutiens ANR²⁴.

²¹ La question des risques et challenges technologiques est l'objet du §2.2

²² PCP (Pre-Commercial Procurement) et PPI (Public Procurement of Innovative Solutions) selon la terminologie de la CE cf. <https://ec.europa.eu/digital-single-market/en/news/frequently-asked-questions-about-pcp-and-ppi>. Concernant le PPI PPI4HPC, voir <https://ppi4hpc.eu/> et https://ted.europa.eu/udl?uri=TED:NOTICE:219297-2020:TEXT:EN:HTML&WT.mc_id=RSS-Feed&WT.rss_f=Computer+and+Related+Services&WT.rss_a=219297-2020&WT.rss_ev=a

²³ Programmes PRACE-Implementation Phase, Centre d'Excellence

²⁴ Alors que de nombreux projets soutenus par l'ANR mentionnent le calcul intensif ou l'IA, nous ne disposons pas d'une synthèse sur ce point.

- Le dossier GENCI ne documente pas de risques concernant le délai de préparation des installations techniques²⁵ du centre pour accueillir le calculateur Exascale ; étant donné le calendrier serré on peut se demander si l'Entité d'Hébergement n'est pas conduite à débiter certains travaux par anticipation, d'où risque financier ou de reprise si les spécifications du constructeur diffèrent des attentes.

Il est également à noter que les installations destinées à accueillir des supercalculateurs de la classe visée font l'objet d'innovations, par exemple pour limiter la consommation énergétique. Sur ce point également, CEA, GENCI et leurs partenaires ont une expertise qui pourrait être valorisée.

En outre, la CE a contribué à faire émerger une série d'outils pour faire de l'achat public un levier de développement technologique, notamment les modalités PPI et PCP ; il en va de même pour la JU-EuroHPC. Les partenaires de PRACE, dont GENCI et CEA, participent déjà à des PCP et PPI, dont l'évaluation n'entre pas dans le périmètre de la contre-expertise. L'expérience acquise vient conforter les capacités du consortium en cours de constitution.

La contre-expertise souligne le bon *track record* et l'expérience de GENCI et de ses partenaires, mais note la pression qui s'exercera sur les ressources humaines.

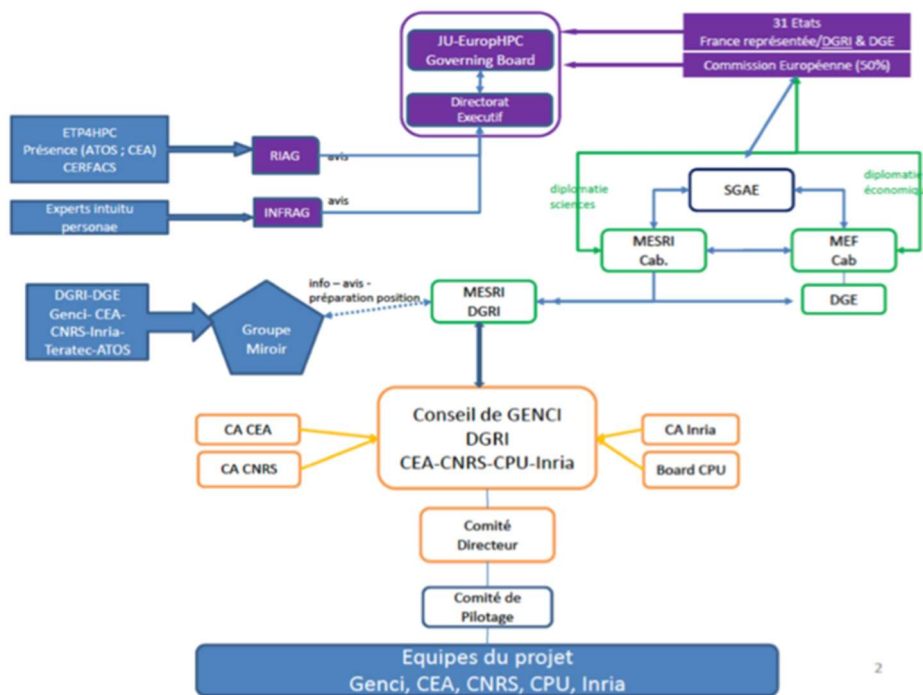
La pratique du pilotage national et européen :

L'organisation en Entreprise Commune (JU) est nouvelle et son fonctionnement interne en cours de formalisation. Cette variable décrit pour une part la déclinaison organisationnelle et managériale du point précédent ; d'autre part la valorisation des coopérations permises par l'exploitation scientifique de calculateurs au niveau Exascale.

Nous avons inclus ci-après une illustration du réseau explicitant les moyens offerts à la partie française pour faire valoir ses intérêts auprès de la JU-EuroHPC; ceci est plus détaillé que les éléments fournis dans le dossier GENCI initial.

²⁵ Fluides (refroidissement), distribution électrique, organisation de la salle

Figure 2 : Réseau explicitant les moyens offerts à la partie française



Source : GENCI, courrier du 19/6/2020

Il convient de considérer les risques associés aux difficultés de la coordination du suivi de ce dispositif :

- besoin d'arbitrage entre des critères possiblement antagonistes à un instant donné (par exemple besoin scientifique / retard dans la mise à disposition d'une technologie) ;
- nécessité pour la partie française d'être prête pour les appels d'offres de la JU, qui contiendront vraisemblablement des contraintes temporelles, par exemple pour caler les dates de mise en œuvre des supercalculateurs avec le calendrier d'introduction des composants électroniques critiques, notamment EPI.
- décisions à prendre concernant l'évolution des coopérations entre certains Etats, par exemple pour mutualiser les risques ou optimiser par adoption de phasage ;
- décisions à prendre concernant les aides aux utilisateurs, dont appels d'offres scientifiques, programmes ANR et PCRDT, l'impact sur d'autres programmes (climat, IA, etc.).

Nous avons reçu de la DGRI des précisions satisfaisantes sur l'existence de démarches visant à la mitigation de ce type de risque (programmes ANR, PIA, discussions bilatérales Franco-Allemandes). Ces démarches devraient être consolidées avec l'appui des acteurs concernés.

Cependant, si on compare le programme européen aux programmes américains ou japonais, le lien avec les utilisateurs potentiels semble moins fort, avec un programme d'accompagnement qui peine à prendre de l'ampleur en France.

La capacité et le coût d'accès aux technologies requises, notamment Européennes:

Nous retenons trois lignes principales pour documenter ce type de risque.

La capacité d'accéder aux technologies est liée au risque d'embargo » ou « mise en place de tarifs onéreux » concernant certains composants matériels ou logiciels. A ce titre on distingue des éléments: a) à acquérir dans un grand volume (mémoires); b) pour lequel le nombre de fournisseurs est limité au plan mondial (GPU, interconnexion rapides). Dans le cas d'éléments « commodity », le volume du marché Européen nous protège car les fournisseurs ont intérêt à le fournir; il paraît néanmoins utile de documenter ce point.

Les coûts d'accès dépendent d'une part des effets de marché, d'autre part d'effets pouvant limiter l'accès aux dispositifs les plus récents, ou basé sur la décroissance rapide de la valeur d'usage en raison de l'évolution technologique. A l'inverse, les projets de supercalculateurs sont optimisés en interceptant la technologie au meilleur moment, voire par l'utilisation de composants de meilleur rapport performance/prix que le tout haut de gamme. Documenter ce point: a) évite des surprises lors des procédures d'appel d'offre, de dialogue compétitif et d'acquisition; b) peut permettre d'enrichir le lot de « benchmark » pour représenter fidèlement les besoins; c) informe les décisions à prendre lorsque des compromis sont nécessaires.

La France souhaite utiliser des technologies européennes notamment l'EPI. Il y a lieu d'accorder, dans la gestion du projet, une attention particulière au fait que les porteurs d'enjeux (utilisateurs, fournisseurs) peuvent avoir des intérêts divergents en cas de retard ou de surcoût. Nous estimons qu'une attention particulière aux cahiers des charges et « benchmarks » est le moyen d'établir l'objectivité dans l'intérêt de toutes les parties.

La situation concurrentielle, internationale et intra-Européenne:

Ce point est couvert par les documents de la Commission Européenne joints au dossier déposé par GENCI. Certains aspects mériteraient d'être repris, caractérisant la position nationale, les atouts et les risques :

- ratios entre les subventions accordées à l'étranger et en France ;
- ratios concernant les moyens humains formés et les financements ;
- comparaison sur les durées des programmes de soutien ou de préparation par anticipation proactive, les domaines couverts (recherche de nouvelles méthodes, adaptation aux architectures Exascale, développement et qualification de logiciels,...).

Ces comparaisons internationales peuvent permettre d'évaluer le niveau d'investissement nécessaire pour traiter les risques tels que les retards, les reports budgétaires, l'inadéquation du programme de développement RH et ces politiques d'accompagnement.

Les capacités financière, et les modalités de mise en oeuvre (CAPEX, OPEX, partenariats, participation de la CE,...).

Trois points critiques sont à souligner :

Le budget prévisionnel du projet présenté par GENCI présente encore une très grosse incertitude avec un besoin de financement complémentaire compris entre 40 et 60 M€.

Le dossier indique que GENCI souhaite le renouvellement de l'ensemble du parc des machines existantes (Jouvence à partir de 2025. En l'état actuel de sa subvention annuelle, le tableau prévisionnel communiqué par GENCI fait alors apparaître une trésorerie négative en 2028 (-59 M€).

Le dossier n'intègre pas par ailleurs tous les financements du programme d'accompagnement du projet Exascale, qui restent à préciser, et en l'absence desquels le projet Exascale ne pourra pas donner l'ensemble de son potentiel.

Nous recommandons de consolider le budget en amont des décisions d'engagement du projet d'accueil d'un ordinateur Exaflopique, avec la finalisation des accords du partenariat d'hébergement.

2.9.2. Les opportunités du programme

Le dossier de GENCI décrit succinctement les actions qui pourraient être entreprises pour la bonne utilisation de la machine Exascale par le développement de code et l'accompagnement des chercheurs. Ceci nous semble insuffisant par comparaison avec les préparations entreprises par les principaux partenaires Européens et internationaux.

Du fait de l'évolution technologique qui se poursuit, un matériel de cette classe verra son avantage compétitif se réduire au fil du temps, ce qui reviendrait à une dépréciation d'environ 20 % à 30 % par an. A côté de la traduction économique, il y a également la réduction de l'avantage offert par une préparation proactive efficace de pouvoir effectuer des « premières » mondiales à fort impact.

Le dossier de GENCI décrit en pages 13 et 14 les programmes mis en place par nos principaux concurrents, depuis 2016 pour les USA qui ont prévu d'atteindre 2Exaflops en 2023. Il s'agit de programmes proactifs ambitieux visant au développement scientifique et technique d'applications justifiant de calculateurs exaflopiques dans une pluralité de domaines clés.²⁶

Ceci nous paraît argumenter en faveur d'actions à entreprendre au plan national suivant une démarche proactive plus forte.

Il faut développer les actions entreprises soit au plan national (p.ex. expertise devant provenir de l'exploitation de la machine Jean-Zay ou des programmes spécifiques COVID), soit au plan Européen (expertises obtenues dans PRACE, dans les Centres d'Excellence et par l'accès de la communauté nationale aux machines « pré-Exascale » financées par la JU).

Il convient d'évaluer l'opportunité d'effectuer des recherches (algorithmes, méthodes de modélisation,...) en vue d'une utilisation innovante ou optimisée des systèmes Exascale par :

- la réalisation d'une analyse des approches et résultats obtenus par les principaux concurrents²⁷ ;
- la valorisation des capacités offertes par des initiatives en coopération, approche également suggérée dans la prospective scientifique PRACE (« The Scientific Case for Computing in Europe 2018-2016 ») ;
- la mobilisation des contributions françaises aux initiatives pertinentes, dont les Centres d'Excellence.

La responsabilité d'une action dans ce domaine relève, à notre avis, de la DGRI et/ou de l'ANR. Toutefois, le projet Exascale a pour but de placer la France et l'Europe en position de leader, il importe donc de sortir d'une stratégie de « rattrapage » et d'entreprendre plus de R&D proactive « *high risk/high reward* »; l'étude socio-économique est le cadre approprié pour argumenter cela.

²⁶ Exascale Computing in the United States, D.B.Kothe, I.Qualters , Computing in Science and Engineering · January 2019

²⁷ La liste des récipiendaires et nominés pour le Prix (annuel) Gordon Bell de l'ACM est un point de départ; d'autres éléments sont disponibles dans les publications des projets Européens PRACE-Implementation Phase et EESI(-2) dont GENCI est partenaire. Le programme américain ECP a publié son plan de travail de façon très détaillée.

Par ailleurs, nous souscrivons à ce que GENCI nous a répondu au sujet du taux d'utilisation de la machine :

« Développer des modes d'accès qui permettront d'accompagner nouvelles communautés et nouveaux usages : [...] Tout cela nécessitera un décloisonnement des décisions interministérielles pour couvrir tous les domaines des défis sociaux et une volonté des organismes et établissements de recherche de la MIRES de rationaliser/consolider leurs besoins dans une TGIR reconnue telle que GENCI ».

La gouvernance du projet Exascale doit monter en puissance, non seulement pour piloter le projet lui-même, mais également pour impulser les actions d'accompagnement nécessaires afin d'en développer les usages.

2.9.3. L'option « Calcul Quantique »

Sur la base des informations disponibles, la mission est perplexe concernant l'inclusion d'une partition « exploratoire » construite autour d'un « accélérateur quantique ». L'évaluation de l'état de l'art par l'Académie des Sciences des USA en 2019 (cf. ²⁸) indique (p.184):

« Key Finding 7: Although the feasibility of a large-scale quantum computer is not yet certain, the benefits of the effort to develop a practical QC are likely to be large, and they may continue to spill over to other nearer-term applications of quantum information technology, such as qubit-based sensing. »

Il nous semble qu'un tel projet devrait faire l'objet d'un projet d'infrastructure de recherche bien identifié et géré suivant les procédures recommandées par ESFRI (« design study », phase préparatoire, suivi scientifique à haut niveau, budget, accompagnement). Il devrait constituer un bassin d'attraction d'une large communauté scientifique et technologique, constituée pour en explorer tous les aspects.

A l'instar de ce qui est fait en Allemagne au centre de Jülich, des équipes de spécialistes capables de traduire les besoins algorithmiques des utilisateurs dans le paradigme du calcul quantique doivent être constituées. Ceci suggère de renforcer le contact avec une large communauté de spécialistes d'algorithmique.

L'interconnexion avec un supercalculateur Exascale pourrait devenir une option profitable à une étape à déterminer du déploiement, mais il est également possible que le besoin d'interface vers un système numérique soit différent (architecture système, équilibrage des fonctions).

2.9.4. En résumé

Au total, l'absence de quantification de l'impact socio-économique du projet ne permet pas une analyse quantitative de la plupart des risques identifiés. Seuls les risques ayant un impact direct sur les coûts sont inclus dans la marge de risque de +/- 7 % indiquée par GENCI.

Une valorisation monétaire des effets du programme aurait par exemple permis de quantifier l'impact d'un sous-investissement dans la mise à niveau de la communauté scientifique et de mettre cet impact en regard du coût du programme d'accompagnement.

Parmi les risques qui ont été évoqués précédemment, la plupart sont des risques inhérents à tout projet ambitieux et la contre-expertise n'a pas de doute sur le fait que GENCI et ses partenaires

²⁸ National Academies of Sciences, Engineering, and Medicine 2019. Quantum Computing: Progress and Prospects. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25196>

sont armés pour les traiter correctement. En revanche, la contre-expertise souhaite mettre en évidence :

- que la gouvernance complexe du projet aura du mal à s'assurer que les performances correspondent aux besoins français ;
- que le programme d'accompagnement des utilisateurs, indispensable compte tenu du saut technologique engagé avec ce projet, ne semble pas à la hauteur de ce que font d'autres pays.

3. Pour aller plus loin dans l'analyse économique

Le projet est encore dans une phase très amont, dans laquelle subsistent naturellement de nombreuses incertitudes sur l'architecture et la performance de la machine (inclusion du processeur européen), les partenariats internationaux, le niveau de mobilisation des communautés scientifiques et des industriels pour utiliser la machine. Ces incertitudes expliquent certaines lacunes de l'étude d'impact réalisée par GENCI.

Des exemples passés, ou l'examen de ce que font d'autres pays lancés dans la course vers l'exascale, donnent cependant des pistes.

Des travaux complémentaires d'étude d'impact socio-économique seraient possibles et utiles pour éclairer la conduite du projet, optimiser les décisions d'investissement et développer un accompagnement du projet permettant d'en optimiser l'impact in fine.

L'évaluation socioéconomique d'un projet d'infrastructure de recherche reste un exercice difficile. Cela dit, les opérateurs amenés à développer ce type de projet disposent actuellement de méthodologies qui permettent de cadrer l'analyse. Outre le guide général sur l'évaluation des investissements public de France Stratégie tenant lieu de doctrine, GENCI peut s'appuyer sur le document plus spécifique dédié à l'enseignement supérieur et à la recherche, mais également sur le document produit par le groupement JASPERS sur l'évaluation des bénéfices sociaux des projets de R&D.

Nous décrivons comment de telles méthodologies pourraient se mettre en œuvre dans le cas de l'option Exascale.

3.1. L'option de référence

En premier lieu, nous recommandons de porter la plus grande attention à la justification de l'option de référence. Dans l'état actuel de l'analyse, le décideur est laissé dans l'alternative ou bien d'exercer l'option Exascale ou bien de ne rien faire d'autre que de remplacer à l'identique les machines actuelles de GENCI. Un tel écart entre les options ne peut que survaloriser l'option Exascale, le seul maintien à l'identique des capacités de calcul entraînant de façon mécanique une dépréciation du capital.

Dans son analyse, GENCI doit expliquer pourquoi d'autres alternatives ne sont pas envisageables, comme par exemple d'investir dans le capital humain et l'exploitation d'une machine installée dans un autre pays d'accueil.

3.2. La demande

En second lieu, nous recommandons que soit réalisée une description quantifiée aussi précise que possible de la demande prévisionnelle à laquelle la machine Exascale devra répondre. Cette demande doit être décomposée par type d'applications, d'enjeux stratégiques de la politique française en matière de recherche mais également par type de technologies requises (processeurs scalaires, convergés...) et par besoin en puissance de crête pour une exécution en un temps raisonnable.

En effet, les bénéfices socioéconomiques dépendent en partie des retombées escomptées par les applications scientifiques qui utiliseront la machine Exascale. Ces applications vont de l'étude de l'explosion des étoiles à la modélisation des réactions de notre système immunitaire. Ces applications n'offrent pas les mêmes effets attendus au sens de la grille des bénéfices socioéconomique préconisée par le groupement JASPERS. Par ailleurs, elles ne demandent pas

non plus les mêmes capacités de calcul en termes de puissance de crête. En effet, les rapports annuels de GENCI jusqu'à 2016 indiquaient le taux de projets faisant un usage de plus de 1000 coeurs. Ce taux était de 27 % en 2016, ce qui correspond à un volume de 200 projets sur les 600 que traitent annuellement GENCI. Si ces 200 projets sont les projets présentant la plus grande espérance de bénéfices socioéconomiques et qu'ils sont limités dans leur progression par la puissance de crête, cela justifierait pleinement l'accès à l'Exascale. Dans le cas contraire, la question de la répartition des moyens de calcul selon une autre configuration se poserait.

De plus, dans un contexte de croissance des applications liées à l'intelligence artificielle mais également de non-alignement possible des priorités scientifiques de l'Union Européenne et de la France, il paraît crucial de disposer d'une évaluation prospective aussi solidement argumentée que possible sur la répartition des applications qui seront exécutées sur la machine Exascale pour définir sa configuration. La répartition entre part GPU et scalaire est un facteur déterminant à la fois du prix de la machine mais aussi des communautés auxquelles elle va pouvoir s'adresser.

Par ailleurs, l'attribution des heures de calcul étant gratuite, quel que soit le dimensionnement de la machine, elle sera saturée rapidement, pour peu que les codes aient été portés sur la nouvelle architecture. Le phénomène qui s'est produit à la mise en service de la machine Jean Zay en septembre 2019 est susceptible de se reproduire pour la machine Exascale. Les 600 projets annuels qui ont accès actuellement à l'infrastructure de calcul GENCI augmenteront la définition du maillage ou le nombre de simulations pour obtenir des résultats plus précis et satureront les capacités de la part française de la nouvelle machine. De ce fait, une analyse qualitative et quantitative des applications et des problèmes qui requièrent l'accès à une machine Exascale pour parvenir à des résultats nouveaux est nécessaire.

3.3. Les effets attendus

Les effets attendus de l'investissement dans une machine Exascale proposés dans la section 2.5 peuvent être globaux (gain de productivité de la recherche française, coût évité par des solutions alternatives, subventions implicites à l'industrie, coût d'adaptation des codes, coût des émissions induites) ou liés aux applications scientifiques elles-mêmes usant de la machine à travers leurs retombées socioéconomiques.

En ce qui concerne les effets globaux non-marchands, nous avons indiqué au paragraphe 2.7 que l'estimation de la valeur des publications fournie était une estimation absolue. L'investissement dans la machine Exascale se justifie par l'amélioration du niveau de publication, en quantité (plus de publications) et en qualité (facteur d'impact). Les statistiques issues de l'étude d'impact réalisée par GENCI en 2017 ne permettent malheureusement pas d'estimer l'accroissement du niveau des publications lié à l'accès à l'infrastructure de calcul en général.

En revanche, il serait possible d'obtenir une estimation de ce gain de productivité en utilisant des méthodes bibliométriques comparant la productivité des laboratoires (ou des équipes de recherche) qui ont accès à GENCI avec celle qui n'ont pas accès et qui n'ont pas déposé de projet. Le rapport de France Stratégie sur l'évaluation socioéconomique des projets immobiliers de l'enseignement supérieur préconise également de tenir compte du facteur d'impact des publications et propose une méthode simple pour le faire.

Par ailleurs, cette même étude d'impact indique que la communauté scientifique usant de HPC a atteint un plateau de l'ordre de 700 publications annuelles. Dans ce contexte apparent de saturation des capacités de publications de cette communauté, il apparaît important de quantifier la croissance d'impact que va procurer l'accès à l'Exascale ou de quantifier l'augmentation prévisible des nouvelles communautés accédant aux infrastructures de GENCI.

En ce qui concerne les coûts évités par la communauté scientifique grâce à l'accès aux machines de GENCI, nous avons indiqué au paragraphe 2.7 comment il était possible de valoriser les heures

de calcul distribuées par GENCI en utilisant des prix de marché. Toutefois, la question se pose de façon différente pour les heures fournies spécifiquement par une nouvelle machine, ici la machine Exascale. En effet, comme nous l'indiquions plus haut au titre de la demande, seules sont concernées les applications qui ne pouvaient pas s'exécuter en un temps raisonnable sur les machines actuelles de GENCI. Il est donc crucial que l'identification exhaustive des applications et des problèmes requérant ce niveau de puissance soit effectuée.

En ce qui concerne les effets attendus marchands, la possibilité ouverte par la JU de commercialiser 20% du temps de calcul de la machine Exscale change drastiquement le paysage de sa valorisation. Pour donner un ordre de grandeur, si l'on considère la machine Jean Zay, machine d'une puissance de calcul de 16 Pflops disposant d'une architecture mixte composée d'une partie scalaire à 5 Pflops et d'une partie GPU à 11 Pflops et produisant environ 800 Mh/an (rapport activité GENCI 2019). Cette machine a coûté un montant de l'ordre de 25 M€ (source web), induit un coût de fonctionnement annuel de l'ordre de 2 M€ par an et a une espérance de vie de l'ordre de 7 ans et dispose d'une valeur résiduelle nulle. Considérant un taux d'actualisation tutélaire de 3,5 %, un prix de vente de l'heure GPU à 2,3 euro (grille Amazon EC2), il suffirait de commercialiser moins de 1 % des heures GPU pour que la VAN financière du projet soit positive ! Ces chiffres n'ont pas de valeur de business plan. Ils ont seulement pour but d'attirer l'attention sur le gisement de valeur que recèle l'activité de commercialisation du temps de calcul.

En ce qui concerne les effets attendus liés aux applications scientifiques exécutées sur la machine Exascale, l'estimation de leurs bénéfices socioéconomiques selon la grille d'analyse préconisée par JASPERS requiert des données sur les entreprises créées dans leur lignée, des brevets déposés, les produits nouveaux lancés, les réductions de risques environnementaux et sanitaires obtenus ou espérés et les gains liés à l'ouverture de l'infrastructure à des opérateurs privés (cf. Table 11 du document de JASPERS pour une liste exhaustive).

Les différents documents (celui de JASPERS comme le rapport France Stratégie de Février 2019) fournissent des valeurs tutélaire pour la valeur d'un brevet ou la valeur des emplois créés par création d'entreprise ou la réduction d'un risque climatique. Ces valeurs permettent des estimations quand aucune autre information n'est disponible. Toutefois, GENCI dispose d'un historique de 600 projets par an à partir duquel il est possible d'obtenir des estimations des brevets déposés, des emplois nouveaux créés dans des entreprises nouvelles, des produits nouveaux ou bien des estimations des risques environnementaux et sanitaires réduits.

Aux indicateurs identifiés par la méthode utilisée par la Commission Européenne, il est possible également d'ajouter des mesures des bénéfices industriels quand ceux-ci peuvent être quantifiés. Ainsi, si l'on prend l'exemple de la validation des méthodes numériques de simulation de test d'accident effectué par Renault sur la machine Curie de GENCI, ce calcul a permis au constructeur de gagner 5 ans sur la validation de leur méthode. Il s'agit d'un exemple où le coût évité est quantifiable (nombre d'ETP par an dévolu à la validation des méthodes fois le salaire moyen en coût complet des ingénieurs fois la durée de temps gagnée).

Nous recommandons donc la systématisation de cette approche sur l'ensemble des projets usant des ressources de calcul de GENCI. Cette étude permettra une appréciation retrospective plus précise des bénéfices socioéconomiques du calcul haute performance qui alimentera l'évaluation prospective des bénéfices liés à l'investissement dans la machine Exascale.

4. Conclusions et recommandations

Le dossier d'évaluation socio-économique a pour objet d'objectiver l'intérêt socio-économique de l'acquisition d'un ordinateur Exascale installée en France à l'usage de la communauté de recherche scientifique et industrielle. Des exemples nombreux fournis dans les rapports d'activités de GENCI de 2010 à 2019 attestent d'un ancrage des applications scientifiques dans les problèmes sociétaux. Aujourd'hui, la dynamique scientifique et industrielle conduit clairement l'Europe à se doter de calculateurs haute performance Exascale.

Toutefois, la conjonction d'une option de référence manquant de justifications, d'une identification lacunaire des effets attendus et d'un déficit d'indicateurs quantifiés des bénéfices socioéconomiques donne l'impression que l'acquisition d'un supercalculateur Exascale ne résulte pas d'une autre logique que de celle de l'acquisition de la machine la plus puissante possible dans l'enveloppe budgétaire affectée, dans le cadre d'un plan européen ambitieux.

Nous résumons nos recommandations autour des points suivants :

1. A court terme, et pour que l'acquisition de la machine réponde au mieux au besoin de la communauté scientifique française, et sachant que la configuration et la puissance de la machine Exascale sont encore largement incertaines, il nous paraît incontournable que GENCI et la DGRI se livrent à un exercice prospectif détaillé d'estimation de la demande en calcul sur la durée de vie estimée de la machine. Le dimensionnement de la machine doit se fonder sur une connaissance des applications et des problèmes nécessitant ce type de machine et la configuration qu'elles requièrent pour s'exécuter en un délai raisonnable. Cette analyse de la demande peut s'appuyer sur les Comités Thématiques en charge de la sélection des projets pour la demande des communautés existantes. Pour les nouvelles communautés, un processus *ad hoc* doit être constitué.

2. A moyen terme, compte tenu de la possibilité ouverte par la JU de commercialiser une partie du temps de calcul de la machine Exascale et compte tenu de la valeur de marché de cette ressource, la DGRI doit réinterroger le modèle économique de son opérateur d'infrastructure de calcul, GENCI. Si le modèle de la société civile fournissant un accès à la ressource a permis un rattrapage de la France en matière de Calcul Intensif sur les dix dernières années, ce modèle doit être adapté à la décennie qui se profile. L'ouverture à des partenaires privés pourrait permettre des économies d'échelle ou des gains de compétitivité. Le partage de l'investissement ou la commercialisation à des entreprises privées d'une fraction de la production de la machine Exascale dégagerait une valeur qui réduirait la charge de l'État. Par ailleurs, la question de l'accès gratuit à la ressource, même pour des acteurs publics, doit être posée. Étant données les différences de performance des machines, un système de péages dépendant de la machine permettrait *a minima* de faire s'exprimer les préférences des différents utilisateurs. Un système de tarification devrait être investigué par la DGRI, en ligne avec sa politique scientifique et indépendamment du mode de financement des laboratoires pour accéder à l'infrastructure.

3. A plus long terme, le processus de planification de l'investissement en moyens de Calcul Intensif doit être construit, accompagné d'une analyse de sa valorisation socioéconomique. Il devrait s'appuyer sur une réflexion de stratégie scientifique et d'innovation couvrant mieux les opportunités et les défis apportés par l'évolution du domaine numérique. A l'heure actuelle, on ne distingue pas d'autre logique économique que celle d'une course à l'armement numérique de l'Europe avec les USA, la Chine et le Japon. L'état actuel des documents fournis et des échanges avec GENCI et la DGRI, ne permettent pas de dire si le projet est le mieux adapté pour répondre aux enjeux sociétaux relevant de la communauté scientifique, ni si l'effort consenti par l'État est au bon niveau. Une analyse précise de la valeur créée par le Calcul Intensif pourrait révéler qu'un

effort financier supérieur aurait tout son sens. Le processus de planification requis plus haut pourrait interroger la question de la souveraineté, de la résilience, de l'indépendance de calcul en lui donnant un sens précis et quantifié.

4. S'agissant d'une infrastructure numérique avancée, il importe d'accorder une attention particulière aux aspects relatifs à la formation et au capital humain, à l'accompagnement par des actions préparatoires de développement applicatif et à la valorisation de la propriété intellectuelle notamment constituée par les codes de calcul.

5. Concernant les risques, nous observons que l'absence de valorisation de l'impact socio-économique du projet ne permet pas une analyse du coût potentiel de la plupart des risques identifiés. Seuls les risques ayant un impact direct sur les coûts sont inclus dans la marge de risque de +/- 7 % indiquée par GENCI.

Une valorisation monétaire des effets du programme aurait par exemple permis de quantifier l'impact d'un sous-investissement dans la mise à niveau de la communauté scientifique et de mettre cet impact en regard du coût du programme d'accompagnement.

Parmi les risques que nous avons identifiés, la plupart sont des risques inhérents à tout projet ambitieux et la contre-expertise n'a pas de doute sur le fait que GENCI et ses partenaires sont armés pour les traiter correctement. En revanche, la contre-expertise souhaite alerter GENCI et la DRI :

- la gouvernance complexe du projet aura du mal à s'assurer que les performances correspondent aux besoins français ;
- le programme d'accompagnement des utilisateurs, indispensable compte tenu du saut technologique engagé avec ce projet, ne semble pas à la hauteur de ce que font d'autres pays.

En conclusion, nous recommandons donc que GENCI et la DGRI consacrent un effort significatif à l'analyse rétrospective et prospective de la valeur socioéconomique créée par le Calcul Intensif et la future machine Exascale selon les préconisations décrites précédemment. Cet effort donnerait des éléments concrets pour améliorer le pilotage du projet Exascale en cours et des projets à venir.

Table des figures

Figure 1 : Flexibilité et couverture du marché potentiel	8
Figure 2 : Réseau explicitant les moyens offerts à la partie française	23