



Projet pilote de connexion au nuage de l'Alliance

Description du PMV



Alliance de recherche
numérique du Canada

Digital Research
Alliance of Canada

1. Description du PMV pour plateforme de connexion au nuage

1.1. Présentation

Le présent document contient la description des caractéristiques requises pour l'élaboration du PMV de la plateforme de connexion au nuage de l'Alliance, ainsi qu'une liste de caractéristiques souhaitées. Il servira à orienter la création d'appels de participation pour le PMV du projet pilote.

1.1.1. Objet

Dans le cadre du projet pilote du nuage commercial, l'Alliance mettra au point une vitrine (ou portail) unique permettant aux chercheuses et chercheurs d'accéder sur demande à des services infonuagiques commerciaux essentiels, ce qui allégera les défis administratifs et techniques.

L'initiative traite des occasions suivantes, qui offrent toutes un meilleur accès aux calculs, aux logiciels et aux données pour la communauté de recherche du Canada.

Les exigences de l'initiative sont décrites par éléments. Bien que modulaires, ces éléments ont divers niveaux d'interdépendance.

Voici en quoi consiste le PMV de l'initiative :

1. Fournisseur – fournisseur de services/d'infrastructures infonuagiques (FSI) :

Les infrastructures seront principalement des IaaS (infrastructure-service), directement utilisables par les chercheuses et chercheurs et leur permettant de déployer des interfaces virtuelles libre-service de calcul, de stockage, de réseautage, etc. Les FSI offriront un service Kubernetes (K8s). Des PaaS (plateforme-service) et des SaaS (logiciels-services), comme des bases de données, des services d'intégration/de déploiement continu, des services de gestion du code source, des plateformes d'applications, des services de CHP, l'intelligence artificielle, l'apprentissage machine, etc., seront également nécessaires. Les chercheuses et chercheurs seront en mesure de déposer leurs documents de calculs, leurs applications Web, leur stockage, leurs réseaux virtuels, leurs conteneurs, leurs bases de données et plus encore sur les plateformes des FSI.

2. Plateforme

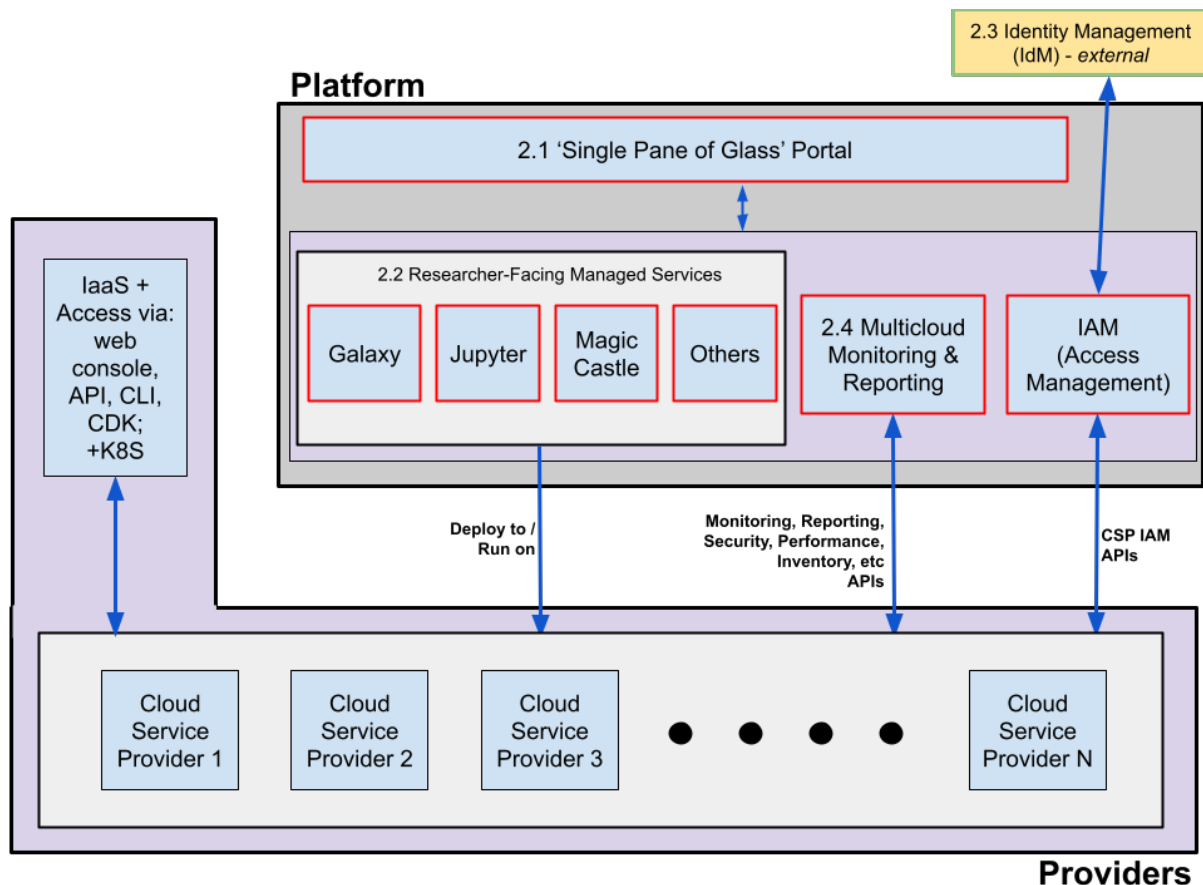
La plateforme se compose de six éléments :

1. **Vitrine ou portail (élément 1)** : Interface Web (portail) qui fournit aux chercheuses et chercheurs une vitrine de l'ensemble des services de la plateforme (2, 3 et 4 ci-dessous) et des liens vers les consoles Web des FSI.

L'élaboration de cet élément requiert l'intégration des éléments ci-dessous et de liens vers les consoles Web des FSI. Le portail offre aussi à l'Alliance et aux administratrices et administrateurs de la communauté un affichage adapté de tous les éléments.

2. **Services de gestion à l'intention des chercheuses et chercheurs :**
Ensemble de services de calcul gérés par l'Alliance ou la communauté et hébergés par les FSI à l'intention des chercheuses et chercheurs, qui leur permet de déployer leur propre version de ces services. Les services (éléments) du projet pilote sont [Jupyter \(élément 2\)](#), [Galaxy \(élément 3\)](#) et [Magic Castle \(élément 4\)](#). Les interfaces utilisateur pour ces services seront intégrées au portail. Il est possible que d'autres services (*Autres – élément 5*) soient ajoutés au projet pilote, si leur pertinence, leur coût et le soutien de la communauté le justifient.
3. **Gestion de l'identité et de l'accès (IAM) (élément 6) :** Qui a) délègue la gestion de l'*identité* à une autorité centrale et b) représente l'autorité pour la gestion de l'*accès* et la transmission des politiques de gestion de l'accès aux FSI. Dans ce contexte, la gestion de l'accès s'entend principalement de la capacité à appeler, à gérer et à visualiser les services ou l'infrastructure des FSI. Les chercheuses et chercheurs pourront ouvrir une session avec leur authentifiant et accéderont aux ressources infonuagiques en fonction de leur rôle dans le projet. Le volet d'affichage administratif permettra à l'Alliance et aux administratrices et administrateurs de la communauté de gérer les comptes, les accès des FSI, entre autres.
4. **Un tableau de bord de gestion et de rapports multilingue (élément 7) :**
Cet élément permet aux chercheuses et chercheurs, à l'Alliance et aux administratrices et administrateurs de la communauté de surveiller, de contrôler et d'optimiser les coûts, les infrastructures et les services liés aux nuages. Le tableau de bord offrira une vue unifiée des différents fournisseurs de services, ce qui permettra la synthèse par les chercheuses et chercheurs, les gestionnaires de projet, les FSI et autres. Il comportera aussi d'autres fonctions : rapports et signalements sur la gestion des coûts, détection des intrusions et des anomalies, et alertes sur les problèmes de conformité, de réseau, de rendement ou liés à la politique de sécurité ou d'accès.

Ce [diagramme](#) donne une vue d'ensemble du projet pilote de l'Alliance.



1.1.2 Principes généraux

1. Chaque utilisatrice et utilisateur aura un compte d'utilisateur de l'Alliance (compte de facturation) pour chaque FSI du projet pilote qu'elles ou ils choisissent d'utiliser. Dans ce contexte, les utilisatrices et les utilisateurs *instancient* des ressources et *déploient* des logiciels à partir de ces ressources. Les chercheuses et chercheurs *utilisant* les ressources déployées dans le cadre du projet pilote n'ont pas nécessairement de compte d'utilisateur.
2. Les utilisatrices et utilisateurs devront exécuter tous les services et les infrastructures de FSI dans leur compte d'utilisateur de l'Alliance.
3. L'accès des chercheuses et chercheurs à tous les services et infrastructures infonuagiques sera géré par l'élément de gestion de l'identité et de l'accès du projet pilote.

1.2 Éléments

1.2.1 Infrastructure et services

1. **Fournisseur**
 - a) Le FSI doit proposer des services infonuagiques qui permettent de concevoir, de sécuriser, de mettre à l'essai, d'utiliser, de surveiller, de gérer et d'établir

et gérer le coût d'infrastructures, de services et d'applications infonuagiques, ainsi que d'héberger des applications Web publiques.

- b) Le FSI doit proposer des IaaS (infrastructures-services), des PaaS (plateformes-services), des FaaS (fonctions-services) et *quelques* SaaS (logiciels-services).
- c) Le FSI doit permettre l'accès à ses services et infrastructures infonuagiques par l'intermédiaire de sa console Web, de son API, de son CLI et de sa trousse de mise en réseau infonuagique.
- d) Le FSI *devrait* proposer des services automatisés (IA ou autre) pour analyser l'utilisation des utilisatrices et utilisateurs, et recommander des optimisations et des pratiques exemplaires dans divers domaines (performance, budget, gestion et protection des coûts, efficacité, sécurité, configuration, durabilité [empreinte carbone]).
- e) Le FSI doit pouvoir proposer des services de consultation à l'Alliance et aux chercheuses et chercheurs notamment sur les sujets suivants : architecture, sécurité, développement d'application, utilisation, et gestion et optimisation des coûts.
- f) Par ses services, le FSI doit appuyer l'application des politiques relatives à l'étiquetage (pour mettre en œuvre la stratégie d'étiquetage du projet pilote).
- g) Le FSI doit obligatoirement proposer les services spécialisés suivants :
 - i. Service Kubernetes (K8s) géré : gestion complète d'un service Kubernetes (K8s) [ex. : Google Kubernetes Engine (GKE), Amazon Elastic Kubernetes Service (EKS), Azure Kubernetes Service (AKS)]
 - ii. Informatique quantique : gestion d'un service d'informatique quantique infonuagique (ex. : Amazon Braket, plateforme IBM Quantum, Azure Quantum)
 - iii. Services d'intelligence artificielle (IA) et d'apprentissage machine (AM) : machines virtuelles dotées de GPU et services gérés d'IA et d'AM

2. Plateforme

a) IAM (gestion de l'identité et de l'accès)

- i. Gestion de l'identité : Sera initialement utilisé [CILogon](#), qui prend en charge l'inscription au portail par [EduGain](#). Si cela réussit, le projet pilote pourra également inclure les bases de données pour disques compacts, ainsi que d'autres fournisseurs d'identité comme AWS, Google, Azure et ORCID.
- ii. Gestion de l'autorité : Un fournisseur d'autorité sera déployé pour gérer l'accès des chercheuses et chercheurs aux fournisseurs de services ou de ressources infonuagiques. Ce fournisseur doit :
 1. être la source d'autorité pour la gestion de l'accès aux FSI;
 2. assurer la gestion des accès des fournisseurs de services infonuagiques (transmission des politiques en matière d'autorité aux FSI);
 3. entreprendre la création et la gestion des comptes de FSI;
 4. centraliser l'accès et l'autorisation des utilisatrices et des utilisateurs dans le cadre du projet pilote pour le déploiement des FSI par les chercheuses et chercheurs;

5. déléguer et cartographier le contrôle de l'accès à la gestion de l'identité et de l'accès native.
- iii. Centrale
- iv. Signature unique
- v. Doit prendre en charge le langage de balisage des assertions de sécurité
- vi. Facultatif
 1. Doit créer et gérer des équipes, des couches PLS, des projets
 2. A2F

b) Services gérés destinés aux chercheuses et chercheurs

Principes

Deux types de déploiement de ces services sont nécessaires :

i. Utilisatrices et utilisateurs généraux

Les services gérés destinés aux chercheuses et chercheurs doivent permettre un accès simple en ligne par le portail pour la personnalisation, la configuration, le déploiement, la gestion et la surveillance de leurs services. Les chercheuses et chercheurs, ainsi que leurs équipes, doivent instancier ces services (libre-service) en utilisant exclusivement l'interface Web du portail.

ii. Utilisatrices et utilisateurs avancés

Les services gérés destinés aux chercheuses et chercheurs ayant des compétences techniques plus avancées doivent également leur donner la possibilité de déployer des infrastructures pour leurs besoins et exigences qui ne sont pas aisément remplis par la seule utilisation du portail.

Infrastructure en tant que code

1. Des applications sont nécessaires pour utiliser l'infrastructure en tant que code (IaC), définie ici comme la distribution utilisant une ou plusieurs méthodes suivantes :
 - a) [Terraform/OpenTofu](#)
 - b) [Pulumi](#)
 - c) [CFEngine](#)
 - d) [Chef](#)
 - e) [Puppet](#)
 - f) [SaltStack](#)
 - g) [Ansible](#)
 - h) [Helm](#)
 - i) [Kustomize](#)

2. Pour un déploiement propre aux FSI, l'laC de base du FSI peut être acceptable. Exemples : AWS [CloudFormation](#), [modèles Azure Resource Manager](#).
3. Un déploiement utilisant l'interface de ligne de commande, l'API et la trousse de mise en réseau infonuagique du FSI peut être envisagé s'il est justifié de le faire.
4. L'laC servant au déploiement *devrait* comporter une estimation des coûts (et potentiellement de l'empreinte carbone) du déploiement aux utilisatrices et utilisateurs et aux administratrices et administrateurs. L'estimation devrait être effectuée avant le déploiement. Exemple : [estimation des coûts dans Terraform](#)
5. L'laC doit s'inscrire dans la stratégie d'étiquetage du projet pilote (conçue dans le cadre dudit projet).
6. Tout le code des laC, des scripts, des modèles et autres doit être stocké dans les dépôts de code et être accessible aux administratrices et administrateurs de l'Alliance.

Les services

1. Service Galaxy

Permet aux utilisatrices et utilisateurs de déployer et de gérer les instances [Galaxy](#), peu importe leur nombre, sur une infrastructure de FSI.

Deux méthodes de mise en œuvre :

1. Utilisatrices et utilisateurs généraux : instances Galaxy définissables pouvant se déployer sur une interface Web simple.
 - a) Doit pouvoir se déployer avec au moins deux FSI.
2. laC (utilisatrices et utilisateurs avancés)
 - a) Doit pouvoir se déployer avec au moins un FSI.

2. Service Jupyter Notebook

Permet aux utilisatrices et utilisateurs de déployer et de gérer n'importe quel nombre d'instances de l'application Jupyter Notebook sur une infrastructure de FSI.

Deux méthodes de mise en œuvre :

1. Utilisatrices et utilisateurs généraux : doit être définissable et pouvoir se déployer sur une interface Web simple.
 - a) Doit pouvoir se déployer avec au moins deux FSI.
2. laC (utilisatrices et utilisateurs avancés)
 - a) Doit pouvoir se déployer avec au moins un FSI.

3. Service Magic Castle

Permet aux utilisatrices et utilisateurs de déployer et de gérer n'importe quel nombre d'instances Magic Castle sur une infrastructure de FSI.

Deux méthodes de mise en œuvre :

1. Utilisatrices et utilisateurs généraux : doit être définissable et pouvoir se déployer sur une interface Web simple.
 - a) Doit pouvoir se déployer avec au moins deux FSI.
2. laC (utilisatrices et utilisateurs avancés)
 - a) Doit pouvoir se déployer avec au moins un FSI.

4. Autre

D'autres services destinés aux chercheuses et chercheurs qui peuvent fournir d'importants services à la communauté sont demandés. La demande vient principalement des besoins et des exigences des chercheuses et chercheurs, du coût, du temps et des ressources nécessaires pour la mise en œuvre et l'exploitation dans le cadre du projet pilote, et du fait que le service répond aux besoins des chercheuses et des chercheurs à l'échelle nationale ou internationale (c.-à-d. non restreint à l'utilisation régionale ou institutionnelle). Pour les groupes de recherche qui comptent des « utilisateurs avancés », et sont en mesure de mettre à disposition leur plateforme en tant qu'instance d'un flux de travail d'infrastructure en tant que code (p. ex. déploiement Terraform), il est souhaitable de rendre ce déploiement facilement accessible aux utilisatrices et utilisateurs finaux par une interface Web simple, par exemple, un service de génomique comme [iReceptor](#), [virtTool](#) et [IRIDA](#). Voir aussi le document sur les cas d'utilisation.

c) Surveillance et rapports multinuage, dont le nuage communautaire

Une ou plusieurs applications Web (tableaux de bord qui offrent une vue unifiée de la surveillance et des rapports de l'ensemble des FSI). Affichages possibles pour l'utilisatrice ou l'utilisateur : par chercheuse et chercheur, par projet, par FSI, etc. Une fonctionnalité avancée doit proposer d'autres affichages pour l'Alliance et les administratrices et administrateurs de communauté. Les trois domaines suivants doivent être pris en charge et mis en œuvre en une seule solution ou plus d'une solution :

- i. Tableau de bord financier, rapports, surveillance, alertes, et protection
- ii. Tableau de bord du rendement, des opérations ou de l'infrastructure, rapports, surveillance et alertes
- iii. Tableau de bord de sécurité, rapports, surveillance et alertes

1.2.2 Divers (général)

- Conception et documents de la stratégie d'étiquetage du projet pilote

Hors de portée

1. Transactions financières avec les universités ou les bailleurs de fonds (mais *le suivi* des coûts infonuagiques se trouve bel et bien dans la portée)
2. Protégé B (données ou renseignements sensibles)
3. Marketplace (services tiers exécutés sur le FSI). Par exemple, AWS Marketplace : <https://aws.amazon.com/marketplace>