Mathrice, une communauté, une organisation, un réseau, des projets pour les mathématiques

$Rev: 151 $

Laurent Azema
Institut Camille Jordan – Lyon / CNRS

Jacquelin Charbonnel
LAREMA – Angers / CNRS

David Delavennat
Centre de Mathématiques Laurent Schwartz – Polytechnique Paris / CNRS

Laurent Facq
Institut de Mathématiques de Bordeaux / CNRS

Damien Ferney
Laboratoire de Mathématiques – Clermont-Ferrand / Université Blaise Pascal

Sandrine Layrisse
Institut de Mathématiques de Bordeaux / Université de Bordeaux

Albert Shih
Observatoire de Paris / CNRS

Romain Théron
MAPMO / Université d'Orléans

Résumé

L'exposé a pour but de vous faire découvrir une organisation humaine soudée et particulièrement dynamique à travers l'architecture et les choix techniques qu'elle a mis en œuvre.

Mathrice, groupement des informaticiens des laboratoires de mathématiques français, a su créer une dynamique forte au service de la communauté mathématique française (mais pas seulement). C'est aujourd'hui un incontournable pour le mathématicien concernant l'accès aux ressources documentaires, au catalogue de services informatiques et aux informations institutionnelles de la recherche.

Les services en ligne sont proposés aux enseignants-chercheurs au travers de la Plateforme en Ligne pour les Mathématiques (PLM), reposant sur 70 serveurs (la plupart virtualisés) répartis sur 4 campus universitaires. La PLM compte plus de 2000 utilisateurs en France et à l'étranger. Elle est gérée à distance par une équipe d'administrateurs système disséminée dans les laboratoires.

Or le déploiement de nouveaux services devient aujourd'hui complexe, et l'équipe doit être de plus en plus experte dans tous les domaines d'intégration. Face à ce problème, une démarche d'urbanisation de l'infrastructure est entamée : des mécanismes basés sur des API `REST` sont déployés en périphérie sur les serveurs, permettant de récupérer et d'agir sur l'état de chaque service. À l'autre bout, via son navigateur, l'enseignant-chercheur consulte et agit sur les services auxquels il a souscrit, depuis son tableau de bord personnel. Entre les deux, un « hub de services », passage obligé des requêtes, prend en charge les aspects authentification et autorisations. On obtient ainsi une architecture de type MVC 3-tiers : l'interface utilisateur, l'abstraction des services et habilitations, et un backend homogène des services.

Mots-clefs

mutualisation, travail collaboratif, fédération d'identités, framework `Sencha`, `Sinatra`, web-services, virtualisation, catalogue de services, architecture n-tiers, paradigme MVC

1 Présentation de Mathrice

Mathrice[1] rassemble les acteurs informatiques des laboratoires de mathématiques français, principalement des administrateurs système et réseau (ASR), mais aussi quelques développeurs et enseignants-chercheurs[2] qui contribuent au fonctionnement des moyens informatiques de leur laboratoire. Sa finalité est double :

- animer et faire évoluer une communauté d'informaticiens, c'est la facette *réseau métier ;*
- proposer des services pour la communauté mathématique, c'est la facette *groupement de services* (GDS).

C'est une initiative prise en 1999 par les responsables de la discipline mathématiques du CNRS (institutionnalisée aujourd'hui par l'INSMI[3], l'Institut National des Sciences Mathématiques et de leurs Interactions). Elle s'explique par le fait que la communauté mathématique est disséminée et mobile[4] :

- disséminée car on trouve des mathématiciens partout : un laboratoire dans quasiment chaque université, des mathématiciens affectés dans des laboratoires de physique théorique, d'informatique, de biologie, etc. ;
- mobile par obligation, parce qu'il n'y a pas de recrutement local (donc la carrière du chercheur progresse en changeant de laboratoire), ce qui le conduit à collaborer avec des personnes n'appartenant pas à son laboratoire.

Dans cette communauté constituée de longue date, où le travail scientifique est réalisé à l'échelle de l'individu et souvent en interaction avec d'autres individus physiquement éloignés, les outils de communication et de travail collaboratif sont essentiels[5]. Mais pour les déployer et les maintenir, il faut des informaticiens. Or à l'époque (fin des années 90), les laboratoires de mathématiques comptent peu d'ITA/ITRF, et quand il y en a, ils sont isolés. C'est ainsi que naît l'idée d'une structure ─ un GDS pour *groupement de services,* sur le principe d'un GDR, *groupement de recherche* ─ qui faciliterait le partage de connaissances et mettrait en place une offre nationale de services à l'attention de la communauté mathématique[6]. Mathrice est donc un appui à la recherche et non pas la DSI de l'INSMI.

Son organisation administrative est légère et non hiérarchique, composée seulement d'un directeur nommé par la direction de l'INSMI. Les informaticiens (ITA CNRS, ITRF universitaires et personnels d'école d'ingénieur) restent géographiquement et administrativement affectés dans leur laboratoire au plus près des chercheurs et participent en même temps à des tâches d'intérêt national. Le niveau de participation, sur la base du volontariat, est très variable d'un individu à l'autre. Chacun contribue à son rythme, selon ses disponibilités, ses compétences, avec pour ligne directrice : « ce que je fais dans mon labo, je peux le faire (sans trop de surcharge) pour la communauté ». Le choix pour un informaticien de participer (ou non) induit, par construction, une équipe motivée, évoluant dans des environnements similaires et confrontée à des problématiques proches, au service d'une même thématique scientifique. S'y impliquer est une façon de sortir du périmètre du laboratoire, d'accroître son domaine de compétence et son expertise pour une meilleure visibilité. Tous les 4 ans, les unités de recherche décident d'exprimer ou non leur soutien au renouvellement du GDS sous forme d'une adhésion officielle (adhésion du laboratoire et adhésions individuelles des personnels qui contribuent).

Un conseil d'orientation, composé de la direction de l'INSMI, de représentants des sociétés savantes et de personnalités de la communauté, se réunit régulièrement pour assurer l'adéquation entre les actions du réseau et les attentes de la communauté. En interne, le positionnement du réseau et les actions stratégiques sont décidées par un noyau dur (une quinzaine de membres actifs) appelé le *groupe « admin »*.

Bien que l'initiative soit venue du CNRS, le GDS ne fait pas de distinction entre les statuts de ses membres (CNRS, universitaire ou école d'ingénieur). Les financements sont exclusivement CNRS, mais la contribution implicite des universités, via leur personnel (1/3 des *mathriciens* sont universitaires) et leurs infrastructures réseau, est capitale.

---

1. Mathrice, pour **MATH R**éseau d'**I**nformation, de **C**ommunication et d'**E**changes.

2. Dans la suite, le terme « chercheurs » regroupera sans distinction les directeurs de recherche, chargés de recherche et enseignants-chercheurs (professeurs des universités et maîtres de conférences).

3. L'INSMI est le plus petit institut du CNRS avec ses 400 chercheurs (directeurs et chargés de recherche) et ses 200 ITA en personnel propre (ce qui n'empêche pas les mathématiques françaises d'être très bien placées dans la compétition internationale, premières ex-æquo en nombre de médailles Fields).

4. Nous estimons la communauté mathématique française à environ 8000 personnes en 2013. Cela comprend les chercheurs et enseignants-chercheurs académiques, ainsi que le personnel technique et administratif des unités de recherche et équipes d'accueil universitaires.

5. Ces outils sont de préférence des logiciels libres en environnement `UNIX`. En particulier, la bureautique s'appuie essentiellement sur `TeX` et `LaTeX`.

6. Les deux premiers GDS à avoir été créés au CNRS sont Mathrice et le RNBM, le Réseau National des Bibliothèques de Mathématiques.

Depuis le début, une attention particulière est accordée à la qualité des relations humaines au sein du groupe, afin que chaque membre se sente à l'aise pour s'exprimer, apporter ses idées, et puisse contribuer à la hauteur de ce qu'il veut et peut fournir. Les initiatives personnelles sont encouragées et accompagnées. Ainsi, les idées naissent souvent sur le terrain, aux interfaces ingénieurs/chercheurs, et sont partagées et débattues au sein du réseau pour y être concrétisées le cas échéant.

Les premiers services se sont mis en place au début des années 2000, au travers d'une maquette qui deviendra la PLM, la *Plateforme en Ligne pour les Mathématiques*. Pour en savoir plus sur l'historique et les activités de Mathrice à l'origine, on pourra se reporter à [JRES2005].

Aujourd'hui, Mathrice compte plus de 200 membres et est devenu un soutien incontournable à la communauté mathématique. La PLM offre des services à 2300 chercheurs, répartis dans 70 structures de recherche. Le niveau de production de Mathrice, 13 ans après, reste constant et dépasse parfois le périmètre des mathématiques, que ce soit par l'implication de ses membres[7] ou par la cible de ses projets (par exemple les projets `FaDDef`[8] [JRES2009] et `PLACO`[9] [JRES2009b][JOSY2009]). De nouveaux projets sont régulièrement initiés, le dernier en date étant une plate-forme d'édition en ligne de revues, utilisée pour la conception de 5 journaux scientifiques internationaux.

L'annuaire national de la communauté mathématique est un exemple de réalisation illustrant l'intérêt de la présence de Mathrice au plus proche des structures de recherche. C'est ainsi que, soutenu par l'INSMI, Mathrice a pu construire ce référentiel inter-organismes unique, qui recense aujourd'hui plus de 8000 personnes (membres d'UMR, d'équipes d'accueil universitaires, de sociétés savantes, mais aussi les mathématiciens disséminés dans des UMR de thématique différente, et même dans d'autres organismes, tels INRIA, le CEA, le CNAM). Cet annuaire est maintenu à jour par une équipe (3 personnes) en contact avec des correspondants sur le terrain (un réseau de 82 correspondants annuaire, un dans chaque structure). Elle développe des procédures automatiques pour collecter les informations produites par ces correspondants (sous forme de fichiers `LDIF`) et détecter les anomalies (problèmes de disponibilité, de péremption et d'incohérence des informations récoltées). Ces informations sont ensuite normalisées, mises en forme et publiées quotidiennement.

Cette irrigation du terrain et son savoir-faire reconnu explique pourquoi aujourd'hui le GDS est de plus en plus sollicité pour participer à des projets communautaires. Citons-en quatre qui illustrent bien cet aspect :

- l'« agenda national des maths », destiné à récolter et publier le calendrier et le programme des séminaires de la communauté ;
- la refonte de l'« annuaire des masters de mathématiques en France » ;
- le projet PTICREM[10] (Projet de Tableau Interactif pour les Collaborations de Recherche En Mathématiques) ;
- le projet « portail des maths » de l'INSMI[11], dont l'objectif est de proposer aux chercheurs un guichet d'accès unique et simplifié :
    - à la documentation scientifique (en accès libre ou contrôlé par des licences ;
    - aux services en ligne (pour le travail nomade et collaboratif) ;
    - aux informations institutionnelles et professionnelles.

  En s'identifiant une seule fois via la fédération d'identités RENATER, le chercheur pourra consulter la documentation à laquelle les multiples contrats signés par ses différentes tutelles lui donnent droit, tout en accédant à ses outils de travail en ligne et à ses données.

2 La PLM

D'un point de vue fonctionnel, on retrouve sur la PLM d'aujourd'hui la plupart des services créés à l'origine comme :

7. Par exemple, 25 % des intervenants à la JOSY/Plume « Les outils libres de base utiles à tout ASR », ESPCI, 22 novembre 2010 étaient membres de Mathrice. De même pour 50 % des intervenants à l'ANF « Installation et configuration des logiciels libres de base, utiles à tout ASR d'un laboratoire », Fréjus 5-9 décembre 2011.

8. `FaDDef` est un système de déploiement rapide et simplifié de systèmes GNU/Linux sans disque, `http://projets.mathrice.org/faddef`

9. `PLACO` est un générateur de plateformes collaboratives (`subversion`, `sympa`, `apache`, `openldap`), `http://placodev.mathrice.fr`

10. `http://pticrem.math.cnrs.fr`

11. Mathrice en collaboration avec Mathdoc et le RNBM, `http://www.mathdoc.fr` et `http://www.rnbm.org`

- l'accès aux revues mathématiques ;
- la messagerie `@math.cnrs.fr` ;
- le service VPN, qui permet aux chercheurs d'accéder aux ressources de leur laboratoire, particulièrement depuis des réseaux locaux filtrés en sortie ;
- la messagerie instantanée (`jabber`), utilisée notamment en interne pour l'administration de la PLM et le développement de projets.

Quelques-uns (par exemple les sessions interactives `Windows`) ont disparu, faute d'utilisateurs. Les années 2008-2011 ont été celles de l'explosion de la demande d'espaces collaboratifs, se traduisant par une montée en charge importante :

- de l'hébergement web (avec plus d'une centaine de sites institutionnels), qui s'ouvre aux projets ANR, aux GDR, aux pages professionnelles individuelles, aux collaborations et colloques, et aux sites « grand public »[12] ;
- des listes de diffusion, à destination de la communauté au sens large, mathématique mais aussi enseignement/recherche[13] (un serveur `SYMPA` multi-domaines avec plusieurs centaines de listes) ;
- des espaces de stockage versionnés (350 dépôts `subversion` et `git`) ;
- une plateforme de web-conférences, interconnectée aux agendas personnels sur la PLM.

Depuis 2012, une plateforme de planification et d'organisation d'événements (basée sur `Indico`[14]) et un service `PLMbox` de partage de dossiers et de documents (basé sur le logiciel `seafile`) sont venus compléter l'offre. Un projet d'édition collaborative de documents `LaTeX` est en cours de validation et une forge intégrant un outil d'intégration continue en cours de développement.

De nouveaux services apparaissent donc régulièrement sur la PLM. Ainsi, l'architecture « bicéphale » originelle de la PLM (localisée dans les universités de Bordeaux et Paris VI/Paris VII) s'étend dans les années 2007-2009. Après la migration des infrastructures de Paris à Lille en 2006, de nouveaux services sont déployés à Angers en 2007 sur des serveurs propres, puis à Lyon en 2010. En 2011, les sauvegardes et l'archivage sont centralisés et migrés à Grenoble. Répartir les services de la PLM sur quatre sites géographiques présente l'avantage de minimiser l'impact des coupures (opérations de maintenance, pannes réseau ou électriques) et évite de mettre une pression écrasante sur l'informaticien sur place. Mais cela oblige à rationaliser, en essayant de converger vers une architecture identique sur chaque site (Figure 1: Architecture nationale), pour appliquer des procédures d'administration identiques.

Les sites ont eu tendance à se spécialiser au fil du temps, en fonction des possibilités offertes par l'environnement. Par exemple, le service *hébergement de machines virtuelles pour les utilisateurs* est localisé à Lille, car c'est là qu'historiquement il y avait le plus gros réservoir d'adresses IP publiques disponibles. Lorsque des données utilisateurs sont partagées entre plusieurs services, ceux-ci sont logiquement implantés sur le même site géographique. On évite ainsi des synchronisations ou des partages de données entre les sites.

L'architecture physique de chaque site est la suivante :

- un sas (`ssh`) pour accéder au réseau d'administration des machines (pour les opérations d'arrêt, de redémarrage, d'installation ou de configurations matérielles ou logicielles des systèmes) ;
- un serveur de stockage et son secours sous `DragonFly BSD` et son système de fichiers `HammerFS` ;
- une infrastructure virtualisée reposant sur au moins 2 hyperviseurs en équilibrage de charge (sous `KVM`) ;
- un filtrage réseau à plusieurs niveaux (le filtrage PLM par défaut, auquel s'ajoute un filtrage spécifique propre à chaque site, auquel peut s'ajouter un filtrage particulier par service).

Les hyperviseurs sont banalisés et ne servent qu'à faire tourner des machines virtuelles[15] (VM), dont les disques sont situés sur le serveur de stockage, qui les exporte par `NFS`. Les données utilisateurs, également localisées sur le serveur

12. Par exemple `http://audimath.math.cnrs.fr` sur la diffusion des mathématiques, `http://experiences.math.cnrs.fr` sur l'expérimentation numérique interactive, et `http://images.math.cnrs.fr` la recherche en mathématique en mots et en images.

13. dont notamment un domaine de listes @`listes.resinfo.org`.

14. `http://indico-software.org`

15. On distingue deux catégories de machines virtuelles : celles qui font *tourner* les services de la PLM (la grande majorité d'entre elles) et celles qui sont livrées clé en main aux utilisateurs.

de stockage, sont montées sur les VM par `NFS` au travers d'une autre pile IP. Pour une description détaillée de l'architecture de virtualisation de la PLM, on pourra se référer à [JOSY2011] et [ARAMIS2012].

Toujours dans cette démarche de normalisation, les configurations ont été factorisées et centralisées (par `puppet` + `kickstart` + `subversion`) selon des procédures affinées au cours du temps. Les sauvegardes ont été désenchevêtrées, centralisées puis délocalisées avec `BackupPC`. La PLM est surveillée *de l'intérieur* (`hobbit` est utilisé pour surveiller les espaces disques, le filtrage réseau, l'utilisation de la mémoire et les mises à jour système) et *de l'extérieur* (`nagios` est utilisé pour surveiller la connectivité des machines et l'écoute sur les ports des services). Le support aux utilisateurs s'est professionnalisé autour d'un guichet unique de signalement d'incidents et de demandes de support technique basé sur un système de gestion de tickets (`RT`).

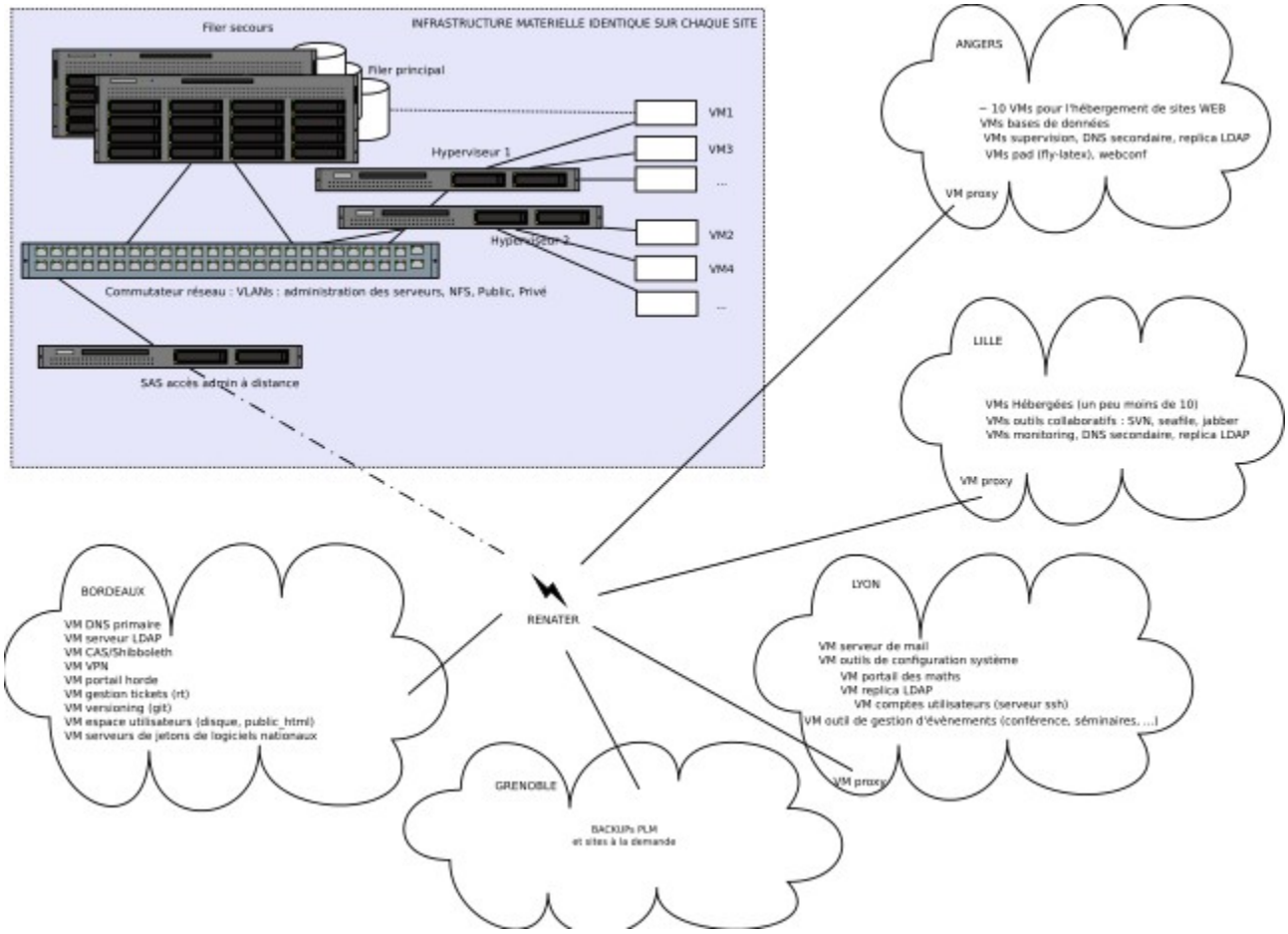

Figure 1: Architecture nationale

La PLM est gérée par une équipe de 12 informaticiens en poste dans les laboratoires de mathématiques[16] d'Angers, Bordeaux, Clermont-Ferrand, Lille, Lyon, Orléans, Poitiers, Tours, Paris et sa région, qui coopèrent en utilisant les outils collaboratifs de la PLM (ce qui lui donne l'occasion de valider les services qu'elle propose). Il n'y a pas de cloisonnement géographique : toute l'équipe intervient sur tout système quelle que soit sa localisation. La gestion des comptes utilisateurs (création, gestion des mots de passe, suppression, migration) est déléguée à des correspondants locaux. Un réseau de 95 correspondants PLM irrigue la communauté, chaque unité de recherche ayant au moins dans ses murs un correspondant, relais d'informations et support de premier niveau.

3 Vers une nouvelle architecture de services évolutive

Au fil des années, le nombre et la diversité des outils proposés par Mathrice à la communauté mathématique n'a cessé de croître, rendant nécessaire une administration plus automatisée et des mécanismes de délégation plus puissants. Dans

16. Pour être tout à fait objectif, il faut également citer les « anciens » de l'équipe, qui ne sont plus dans un laboratoire de mathématiques, mais qui contribuent toujours très activement à administrer la PLM.

le même temps, un besoin d'intégration s'est fait sentir pour améliorer la lisibilité du catalogue des services et permettre le déploiement d'outils synchronisés (forge logicielle).

Deux aspects de cette intégration sont problématiques :

- l'hétérogénéité et la complexité des technologies utilisées ;
- le besoin de compétences en développement parfois pointues alors que l'essentiel des ressources humaines informatiques des laboratoires de mathématiques se trouve être des administrateurs systèmes et réseaux.

Pour résoudre ce problème nous avons travaillé sur plusieurs axes :

- la conception d'une plate forme exposant des technologies d'intégration les plus simples possibles au regard de la complexité du problème à résoudre ;
- la formation des ASR des laboratoires de mathématiques au développement sur ces technologies [17] [18].

L'architecture en cours de développement (Figure 2: Nouvelle architecture PLM 2.0) s'appuie sur un guichet unique de type web gérant l'activation de services et le paramétrage des applications en lien avec le cœur du système d'information Mathrice (notamment la gestion des groupes, des rôles et des droits d'accès).

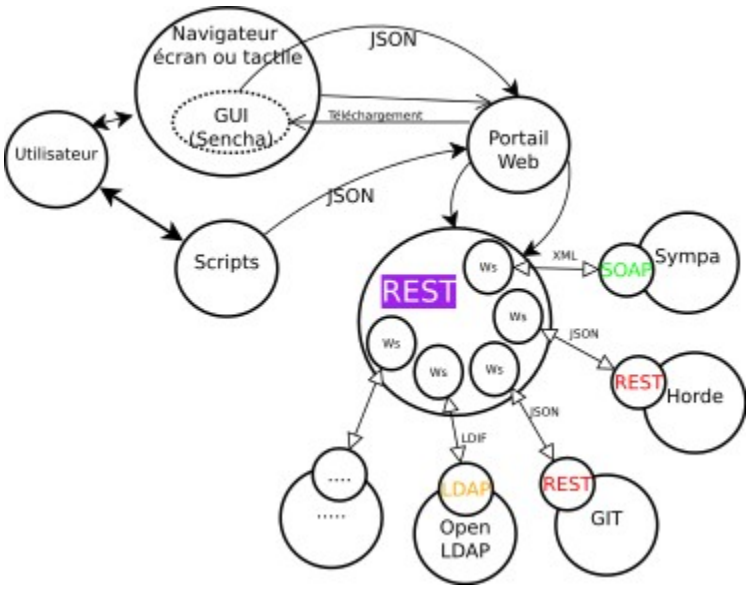

Figure 2: Nouvelle architecture PLM 2.0

Cette architecture permet de regrouper les opérations de contrôle et de paramétrage des applications, indépendamment du logiciel choisi. Les applicatifs peuvent ainsi changer mais l'interfaçage avec les utilisateurs (*guichet unique*) et les scripts de traitement automatique restent stables.

3.1 Guichet unique de services

Du point de vue de l'utilisateur, le nombre, la nature et la localisation des services mis à sa disposition importent peu : il n'a affaire qu'à l'URL du guichet. Le frontal web est un mandataire inverse. Il partitionne l'espace de nommage et aiguille les requêtes vers les web-services du guichet. À chaque service de la plateforme correspond un web-service du guichet. Cela normalise l'interface avec les applications clientes de la PLM.

Lorsque l'utilisateur se connecte à la plateforme, le module d'authentification du guichet lui détermine une identité. Cette information est rendue disponible à l'ensemble des requêtes de sa session via un cookie de session. Ainsi, le développeur d'un web-service n'a pas à prendre en charge l'authentification de l'utilisateur et se contente de gérer seulement les autorisations basées sur cette identité.

Les web-services du guichet sont indépendants. La défaillance de l'un d'eux ne peut pas porter atteinte à l'ensemble de

17. Action nationale de formation « Développement logiciel pour l'administration système et réseau », Angers mai 2012.

18. Présentations aux Journées Mathrice, Caen mars 2013 et Rennes octobre 2013.

la plateforme. Cette indépendance n'empêche pas chaque web-service d'accéder aux données communes relatives à la session de l'utilisateur.

Les web-services peuvent être développés dans n'importe quel langage/framework. Notre choix s'est porté sur `Ruby`/`Sinatra`. Une validation du code est faite à l'aide de `RuboCop` pour la forme et de `Rspec` pour le fond. Le premier garantit la lisibilité du code par l'ensemble des développeurs. Le second favorise le travail à plusieurs développeurs en intégrant des jeux de tests de non-régression. Il facilite le développement d'interfaces et la maintenance du code.

3.2 Interface des services en *backend*

Le guichet interagit avec l'applicatif en *backend* via un format de données et un protocole de communication. Par exemple, le couple `LDIF`/`LDAP` peut remplir ce rôle pour interroger un annuaire. Un web-service est la solution lorsque l'application n'apporte pas son propre protocole.

Ce web-service est placé au plus près de l'applicatif. Certains applicatifs fournissent déjà un web-service : `Sympa` avec `SOAP`, `OpenMeeting`, `RT` avec son interface `REST`. Dans les autres cas, nous développons le web-service avec les interfaces nécessaires au contrôle et au paramétrage par le guichet.

Le web-service fourni par l'applicatif peut imposer le format d'échange et le protocole de communication. Par exemple `Sympa` utilise le couple `XML`/`SOAP`. Bien que `SOAP` soit plus robuste, nous choisissons le couple `JSON`/`REST` pour sa simplicité quand cela est possible.

Tout comme les web-services du guichet, le développement est fait en `Ruby`/`Sinatra`.

3.3 Interface utilisateur

La plateforme a été conçue pour être indépendante de l'interface d'accès en utilisant les mêmes technologies que ci-dessus (format d'échange `JSON` et protocole de communication `REST`). Il est à la fois possible d'interagir via une interface utilisateur riche, une interface utilisateur tactile, ou par script. Ce dernier point permet notamment d'automatiser facilement les tests de régression des API et de superviser le bon fonctionnement des services.

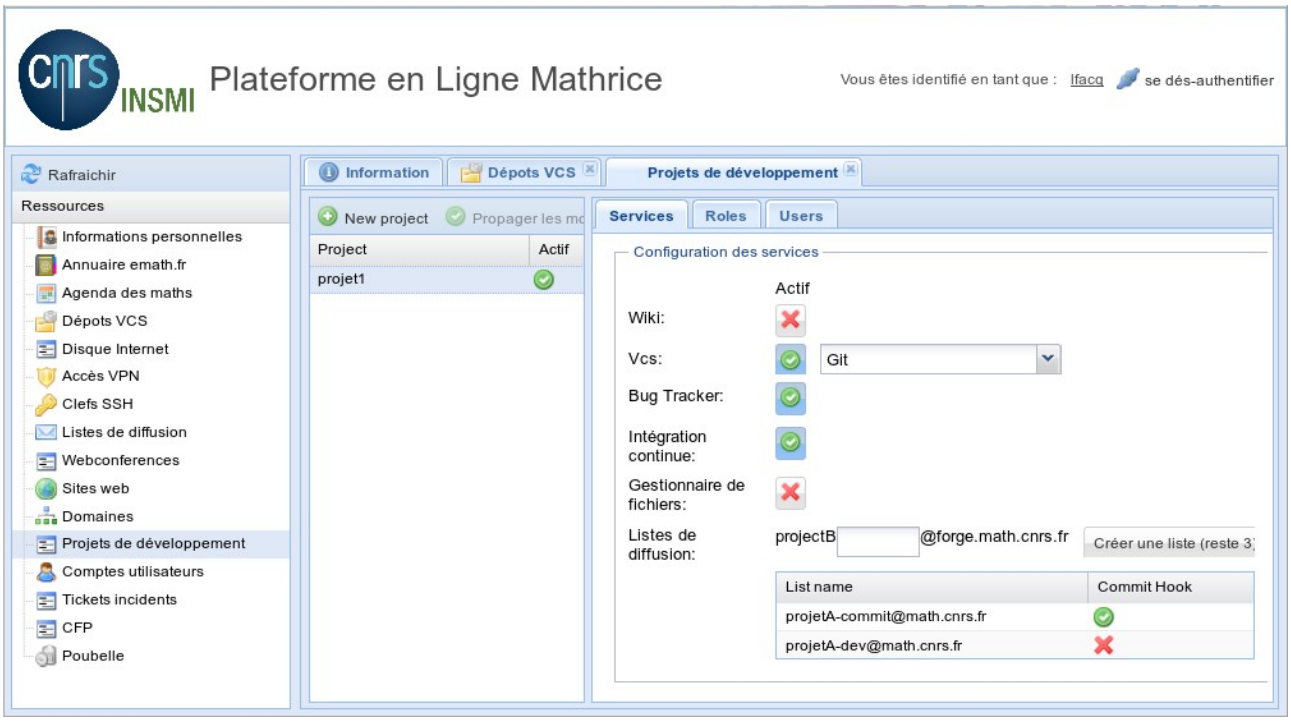

Figure 3: Portail web PLM 2.0

Le framework `javascript` le plus intégré actuellement et que nous avons utilisé est le framework `Sencha`[19] (voir Figure 3: Portail web PLM 2.0)

Il dispose d'une large bibliothèque de composants graphiques (arborescence de fichiers, onglets, tableaux de type `excel`, etc...) le rendant concurrentiel avec les interfaces utilisateurs natives.

Dans sa version 4, il utilise le paradigme de développement *Modèle-Vue-Contrôleur* qui s'est avéré être un réel avantage pour le développement en `javascript`.

3.4 Evolution de l'authentification

Depuis sa création, la PLM utilisait naturellement son propre référentiel de comptes utilisateurs[20]. Ces dernières années, l'émergence puis l'adoption par la plupart des établissements des systèmes de fédération d'identités ont fourni des outils d'authentification mutualisés beaucoup plus pratiques pour les usagers :

- *Single Sign On* : l'utilisateur se connecte une seule fois pour accéder à un grand nombre de services ;
- diminution du nombre d'identifiants/mots de passe à retenir : dans l'idéal, un seul identifiant/mot de passe par établissement.

Afin de bénéficier de ces avantages, la PLM est en cours de migration pour refonder ses mécanismes d'authentification/identification sur la fédération d'identités nationale opérée par RENATER, avec une problématique particulière propre aux communautés transversales : comment identifier les membres de la communauté mathématique parmi l'ensemble des utilisateurs de la fédération ?

Une première idée a consisté à exploiter les attributs (notamment l'affectation à un laboratoire de mathématique) fournis par la fédération d'identités. Mais à ce jour, les informations délivrées par les établissements au niveau de leur fournisseurs d'identités ne le permettent pas (attributs non diffusés ou pas assez précis).

Le mécanisme finalement retenu s'appuie sur l'annuaire national de la communauté mathématique qui recense tous ses membres avec, entre autres informations, leur adresse e-mail. On recherche alors l'adresse e-mail de la personne authentifiée par la fédération dans cet annuaire pour établir son appartenance à la communauté.

Dans la mesure où certains membres de notre communauté ne sont pas dans la fédération nationale, nous avons mis en place un fournisseur d'identités « local » (qui n'est pas dans la fédération nationale mais qui est reconnu par nos fournisseurs de services) qui s'appuie sur l'*annuaire historique* de la PLM.

3.5 La convergence d'identités

Bien qu'être membre de la communauté suffise pour accéder à certains services de la plateforme (comme les web-conférences ou les ressources documentaires), d'autres services (comme par exemple les sessions `UNIX` interactives) nécessitent un véritable compte (un `login`, un groupe principal, etc.). Ces comptes sont conservés dans l'*annuaire historique* mais doivent être mis en relation avec les identités provenant de la fédération, sachant qu'un même utilisateur peut disposer de plusieurs identités au sein de différents établissements (par exemple EPST+Université ou 2 Universités).

Via un outil web développé à cet effet, les membres de la communauté peuvent eux-mêmes relier leurs différentes identités à leur unique compte PLM.

Du côté de nos applications (fournisseurs de services), cela implique la mise en place d'un mécanisme pour réaliser cette convergence d'identités, juste après la phase d'authentification par la fédération, de façon à ce que chaque application dispose des attributs les plus pertinents correspondant à l'usager connecté (y compris les attributs stockés dans l'*annuaire historique*).

Sachant que ce mécanisme de convergence nécessitera des ajustements réguliers, plutôt que de l'implémenter dans chaque service (ce qui rendrait les mises à jours complexes), nous avons opté pour un mécanisme centralisé et redondé.

En parallèle de cette problématique, la multiplicité des services offerts sur la plateforme impliquait un grand nombre de déclarations de fournisseurs de services au sein de la fédération.

19. `ExtJS/SenchaTouch`

20. Un annuaire LDAP des comptes de la PLM, appelé dans la suite « annuaire historique ».

Pour résoudre ces deux points nous avons finalement adopté un schéma n'utilisant qu'un seul fournisseur de services `Shibboleth` pour l'ensemble de la plateforme de services, associé à une seconde authentification par cookie via une implémentation dérivée de `AuthMemCookie`, baptisée `Mod_Auth_PLM`, qui utilise une base `MemCache`.

Les services de la plateforme sont donc protégés en premier lieu par `Mod_Auth_PLM` (Figure 4: Authentification et convergence d'identités).

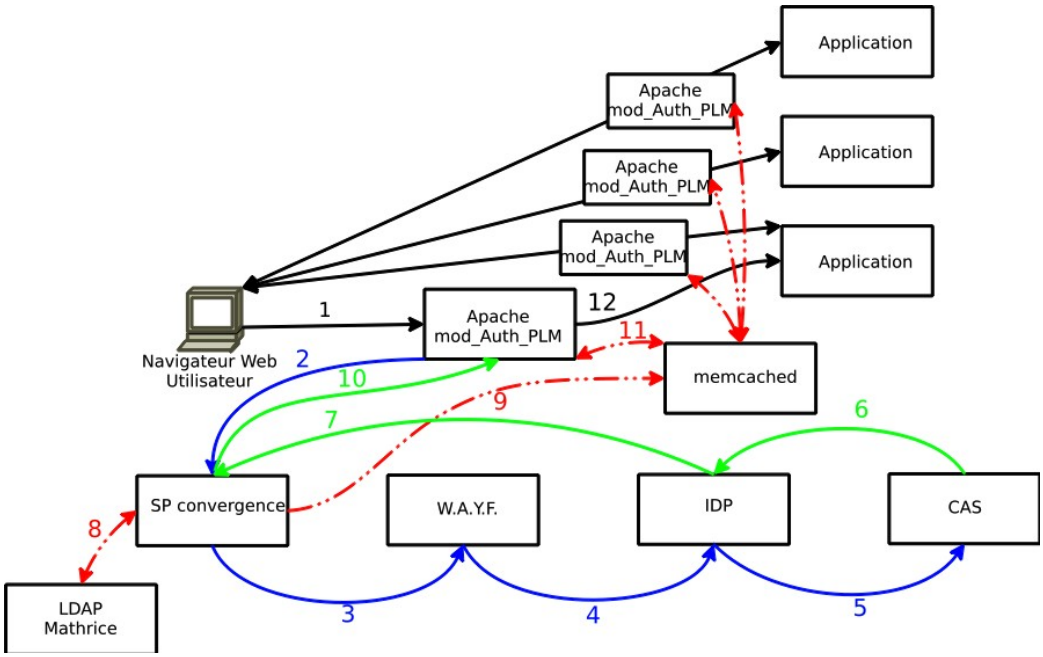

Figure 4: Authentification et convergence d'identités

Quand un client se connecte sans cookie (1), il est redirigé (2) vers le service de convergence d'identités, lui-même protégé par la fédération d'identités nationale (3). Une fois l'authentification réalisée sur l'établissement d'origine (4,5,6,7), le service de convergence récupère (8) dans l'annuaire PLM les informations concernant l'utilisateur et canonise son e-mail. Toutes ces informations sont stockées (9) dans une base `MemCache`. Le client est alors renvoyé (10) vers le service demandé initialement, avec un hachage positionné dans l'URL. Au niveau du service, le hachage est transformé en cookie. Ce cookie est ensuite utilisé pour récupérer (11) les attributs stockés dans la base `MemCache`. Tous ces attributs sont fournis généralement à l'application (12) sous forme de variables d'environnement.

3.6 Perspectives

Ce mécanisme d'authentification avec convergence d'identités est également utilisable pour le contrôle d'accès à toute application hors PLM. Il est notamment en cours de déploiement sur le futur portail documentaire des mathématiques de Mathdoc.

La fédération d'identités nationale, telle qu'elle est implémentée, fonctionne très bien pour les applications de type web. Il serait extrêmement intéressant de la généraliser à d'autres services (par exemple `ssh`).

C'est l'objectif du projet européen Moonshot[21] : étendre l'utilisation de la fédération d'identités à d'autres protocoles. Et pour tester le concept, il est intéressant de disposer d'une communauté géographiquement répartie et informatiquement organisée : c'est le but de la collaboration RENATER/Mathrice mise en place en 2013, destinée à expérimenter une maquette sur la communauté mathématique française.

4 Conclusion

Le modèle d'organisation de Mathrice est celui d'une équipe de volontaires fortement dispersée, consacrant des

21. `https://community.ja.net/groups/moonshot`

fragments d'ETP[22] pour le développement opérationnel d'une infrastructure mutualisée au service d'une communauté scientifique mono-disciplinaire. Ses acteurs sont soudés par une adhésion identitaire forte, car évoluant dans des environnements similaires. Leur implication particulièrement forte et leur aptitude à obtenir l'adhésion du groupe autour de leurs idées et de leurs projets est la force motrice du GDS. Pour autant, rien n'aurait été possible sans le soutien affiché et indéfectible de la direction de l'INSMI au long de ces treize années.

Un effort important est mené actuellement pour uniformiser l'accès aux outils, simplifier leur configuration et faciliter leur usage, tout en augmentant l'autonomie des chercheurs. L'accueil positif des utilisateurs est un encouragement à poursuivre le développement dans ce sens.

Ouvrir la PLM à d'autres communautés, en donnant la possibilité aux mathématiciens d'inviter leurs collaborateurs étrangers (ou d'autres disciplines) à venir travailler sur les outils collaboratifs de la PLM, est un objectif qu'il nous semble important de suivre pour faciliter les travaux de recherche interdisciplinaire aux frontières des mathématiques.

Bibliographie

[ARAMIS2012] Architecture de serveurs virtualisés pour la communauté mathématique ─ Jacquelin Charbonnel ─ Lyon, ARAMIS 2012

[JOSY2011] `KVM` retours d'expériences ─ Jacquelin Charbonnel ─ Strasbourg, JOSY 2011

[JRES2009] Déploiement simplifié de stations sans disque avec `FaDDeF` ─ Philippe Depouilly, Zouhir Hafidi ─ Nantes, JRES 2009

[JOSY2009] Le projet `PLACO` ─ Jacquelin Charbonnel ─ Strasbourg, JOSY 2009

[JRES2009b] `PLACO`, un générateur de plateformes collaboratives ─ Jacquelin Charbonnel, Philippe Depouilly, Francois Ducrot, Damien Ferney, Jacques Foury, Mickael Marchand, Benoit Métrot, Albert Shih, Olivier Thibault ─ Nantes, JRES 2009

[JRES2005] Mathrice, un réseau métier pour les Mathématiques ─ Philippe Depouilly, Gérard Grancher, Joël Marchand ─ Marseille, JRES 2005

[JRES2005b] Les clients légers 4 ans après ─ David Bonnafous, David Delavennat, Philippe Depouilly, Zouhir Hafidi, Gérard Henry, Joël Marchand, Bernard Perrot, Albert Shih ─ Marseille, JRES 2005.

[JRES2001] Clients légers ─ Denis Auroux, Jean-Luc Bellon, Philippe Depouilly, Joël Marchand, Albert Shih ─ Lyon, JRES 2001.

22. Équivalent Temps Plein