



Environnement collaboratif Projet AccessGrid

Philippe d'Anfray CEA
Centre de Saclay 91191 Gif sur Yvette Cedex FRANCE
Philippe.d-Anfray@cea.fr

Version 1.2 du 8 mars 2011

Environnement collaboratif

Plate-forme collaborative

Le concept de travail collaboratif a complètement changé avec les nouveaux instruments disponibles grâce à internet qui permettent aux intervenants de collaborer en temps réel tout en étant dispersés dans des lieux géographiquement distincts. La révolution s'est faite peu à peu (systèmes de messagerie, de gestion et de partage des documents et plus généralement des connaissances, visio-conférence, mise en commun de ressources —au sens large—, ...). Ce qui est vraiment nouveau c'est la dimension de *réseau* portée par les outils les plus récents et qui modifie les modes de collaboration. Tout cela est particulièrement perceptible dans le monde de la R&D où les travaux s'appuient de plus en plus sur des projets impliquant des communautés réunissant des équipes pluridisciplinaires appartenant à des entités distinctes.

Ainsi, la notion moderne de travail collaboratif correspond à un mode de fonctionnement collectif intimement lié aux technologies de l'information et de la communication. L'intégration des différentes technologies permet de construire un *espace de travail virtuel*, un *e-laboratoire* qui facilite les échanges entre les acteurs de la collaboration dans un cadre beaucoup plus large que celui des tâches liées par exemple à un projet bien précis ; la plate-forme collaborative incluant cette dimension sociale qui organise la gestion des groupes et des réseaux d'acteurs et d'entités *via* la notion d'*organisation virtuelle* (voir par exemple [1]).

Grilles et *cloud*

Le concept de grille est aujourd'hui omni-présent dans le paysage des systèmes d'information. Si au départ il s'agissait surtout d'agréger des moyens de calcul ou de stockage, tous les grands projets scientifiques, industriels et commerciaux impliquent maintenant le partage de ressources au sein de communautés d'utilisateurs dispersés dans des organismes et des sites distincts souvent à travers plusieurs pays. Les déploiements opérationnels d'applications distribuées qui ont stimulé le développement des technologies de grille ont entraîné aussi de nouvelles façons de travailler et de coopérer. C'est de là que vient par exemple le concept d'*e-infrastructure*.

En effet, si la collaboration implique la fédération et à la mutualisation de ressources identifiées et mises en commun dans le cadre d'un projet entre organismes, ce qui va nous intéresser ici c'est l'aspect en quelque sorte *dual* des grilles : l'infrastructure de services, qui permet l'authentification des utilisateurs, l'identification des ressources, l'utilisation transparente des infrastructures, *etc.*

Certaines communautés de chercheurs qui *ont fait le choix de la grille* disposent déjà d'infrastructures de services opérationnelles. C'est le cas du projet EGEE [2] devenu EGI [3] avec la pérennisation des *National Grid Infrastructures* (en France, l'Institut des Grilles du CNRS [4]).

En revanche, les problématiques purement *Cloud* qui ciblent d'avantage l'externalisation ou encore l'agrégation mais à travers l'utilisation de moyens qui ne sont pas forcément identifiés ni localisés nous concernent moins. Typiquement un *cloud* —éventuellement privé— peut être considéré comme une ressource (de calcul, de stockage, ...) banalisée qui peut être mutualisée. Nous ne nous intéresserons pas dans aux outils nécessaires pour construire un *cloud* mais plutôt aux interfaces —les services encore une fois— qui permettront d'y accéder ou mieux, d'être encapsulés dans une infrastructure plus large.

Aspect temporel

Nous distinguerons trois types de collaborations possibles selon la durée et le degré d'implication des acteurs.

Les collaborations *sur le long terme* font le plus souvent l'objet d'accords entre instituts et regroupent des équipes dans des domaines scientifiques bien précis (physique fondamentale, climatologie, astrophysique, biologie, *etc.*). Cela se traduit concrètement par le partage de bases de données de résultats de calculs ou d'expérimentations ou encore d'informations issues de capteurs (satellites, ...) et l'exploitation conjointe de grands instruments scientifiques (accélérateurs, télescopes, ...). L'échelle de temps oblige à définir des standards (données, meta-données, ontologies, ...) pour travailler efficacement en réseau.

Ensuite les collaborations liées à des projets de recherche et de développement ont une durée de vie de quelques années, typiquement 2 à 4 ans pour les ANR¹ un peu plus pour les projets européens. Ces projets, le plus souvent de nature pluridisciplinaire, impliquent des communautés qui n'ont pas forcément l'habitude de travailler ensemble. La priorité sera de se doter rapidement d'outils de travail collaboratifs permettant de favoriser le dialogue entre les chercheurs et de faciliter le montage puis la gestion du projet et de ses résultats.

Enfin des collaborations plus ponctuelles (séminaires, *workshops*, *e-learning*, expérimentations, ...) nécessiteront des outils partagés rapidement maîtrisables : visio-conférence, présentation à distance, partage d'outils de visualisation de résultats ou de pilotage d'instrument scientifique,...

L'intérêt de ce *distingo* temporel peut être un peu artificiel est de montrer que l'environnement du chercheur doit lui permettre de s'intégrer rapidement et sans *investissement fort* dans la plate-forme collaborative adoptée pour un projet (réactivité : peut être quelques mois) ou pour un événement plus ponctuel (réactivité : peut être quelques jours). La diversité des acteurs liée à la multiplication des collaborations entre équipes de plusieurs pays rend imprévisible la nature de cette plate-forme. Par ailleurs l'évolution des modes de financement de la recherche —sur contrats— augmente le poids des contraintes liées à des projets à court terme.

Nous observons que les collaborations à long terme capitalisent en quelque sorte la dynamique des projets et bénéficient de la pérennisation de certaines infrastructures. Nous nous orientons bien vers la formation de laboratoires virtuels dans lesquels il importe de favoriser au quotidien les contacts entre des équipes distantes, qu'elles soient dans le bâtiment d'à côté, à

1. Agence Nationale pour la Recherche.

quelques kilomètres, ou bien réparties sur plusieurs continents, avec des outils pouvant inclure par exemple la téléprésence.

Ressources et infrastructure

Dans l'optique collaboratif, le terme *ressource* doit être pris au sens le plus large : calculateurs, espaces de stockage, réseaux informatiques, informations en tous genres, bases de données, réseaux de capteurs, instruments scientifiques, ..., voire aussi les ressources humaines !

La problématique majeure, qui ne peut être contournée, est l'*hétérogénéité*. Aussi les travaux sur les *intergiciels* se sont-ils rapidement focalisés sur les *couches d'interopérabilité* puis sur les problèmes de *gouvernance* d'où le concept clef d'*organisation virtuelle* (authentification des utilisateurs, identification des ressources, droits et devoirs des différents acteurs, vie et mort des projets, ...). La priorité actuellement n'est plus d'agréger des ressources connues et disponibles —ce que l'on sait très bien faire— mais bien d'offrir une *infrastructure de services* qui permette la fédération et la mutualisation des ressources non identifiées *a priori* pour les projets futurs en créant une nouvelle organisation virtuelle transversale aux organismes d'appartenance des équipes impliquées.

L'*infrastructure de services* peut bien sûr intégrer des outils collaboratifs ou mieux, des passerelles entre différents outils de même nature (la couche d'interopérabilité, typiquement entre systèmes de visio-conférence), ou encore l'accès à des outils de dépouillement, *etc.* L'interopérabilité passe par l'adoption des standards ouverts. Dans cette optique, l'environnement de travail du chercheur découle de l'infrastructure de service tant pour les outils que pour l'aspect « réseau » ou social.

La boîte à outils du chercheur

Un autre point de vue est de regarder comment travaillent les chercheurs dans les laboratoires : l'intersection au quotidien de leur métier avec les technologies de l'information. L'environnement de travail du chercheur, au sens des outils informatiques comprend :

- des outils *proposés* par son institution ;
- des outils *culturels* utilisés traditionnellement dans un domaine d'application ;
- des outils *imposés* dans le cadre d'un projet ou par des collaborations extérieures ;
- sans oublier... le *pot commun* utilisable par tous : *l'outil libre* qui dépanne et que l'on finit par adopter (et retour au deuxième item...).

Le poste de travail contient donc une boîte à outils très complète (messagerie, annuaires, listes de diffusion, suites de bureautique, agendas partagés, VPN et clients en tous genres, sites internet, Wiki, GED, CMS, portail *web*, gestionnaires de versions, forges logicielles, ...), mais ne dispose pas encore des instruments qui lui permettraient —sauf expérimentations spécifiques— de sortir des modes de fonctionnement et d'« organisation classiques » des projets. Si chaque brique du laboratoire virtuel existe bien, elle n'est pas intégrée dans une infrastructure plus large. Elle n'est pas toujours utilisable dans le cadre d'un projet donné pour des raisons (par exemple) de sécurité. Ainsi, pour schématiser, l'outil « le plus crucial » (par exemple une forge logicielle) va imposer son mode de travail dans le cadre de la collaboration. Or il faut arriver à se projeter

dans un environnement où tous les outils auront en quelque sorte fusionné et seront accessibles à travers un portail (au sens *portail d'entreprise*) sans hiérarchisation *a priori* de leurs usages.

Le chercheur évolue dans sa structure mais aussi, nous l'avons vu, dans des organisations virtuelles multiformes. Il a besoin d'un accès « uniforme », avec bien sûr une authentification unique type SSO (*Single Sign On*), aux outils et aux ressources locales comme aux outils et aux ressources mutualisées dans le cadre de plusieurs projets. Conclusion de cette section, la somme des outils ne constitue pas un environnement, cela d'autant plus qu'il y a beaucoup de redondance à cause du recouvrement des fonctionnalités des divers outils utilisés.

L'environnement réseau : une nouvelle problématique pour l'utilisateur

Petit point (pas de) détail, l'environnement réseau est bien devenu pour le chercheur une problématique nouvelle. Jusqu'à un passé récent, la question de la connectivité réseau ne se posait pas. Il s'agissait d'un service standard (comme le téléphone, ...) qui répondait parfaitement aux besoins de l'outil informatique et des applications. Or l'essor du collaboratif conduit à formuler des besoins « nouveaux » et à entamer un difficile dialogue avec les équipes réseau et celles chargées de la sécurité. Difficile car les besoins précis sont le plus souvent mal cernés (en terme de débit, d'ouverture, *etc.*), imposés par des outils « extérieurs » et aussi parce que le dialogue avec ces équipes est peu habituel. Faisons par exemple le parallèle avec les personnes en charge du dimensionnement des centres de calcul qui eux ont l'habitude de recueillir des *benchmarks* représentatifs des besoins des applications en terme de puissance ou de stockage. Il n'existe pas vraiment de *benchmark réseau* qui validerait les besoins des applications (qualifiées de *network intensive*) ou encore ceux des outils collaboratifs en terme de connectivité.

Des concepts clefs (1)

Nous en avons déjà rencontré trois dans cette courte analyse :

Les infrastructures de service sans infrastructure, la fédération des ressources et la mutualisation des efforts nécessiterait des développements hors de portée de la plupart des projets.

Les organisations virtuelles qui permettent la gouvernance et le contrôle des groupes et des réseaux, des personnes et des entités.

Le portail qui autorise un accès uniforme —authentification unique— et transparents aux outils informatiques mutualisés.

Cela permet d'envisager un environnement de travail qui permet l'authentification (unique) comme membre d'une ou plusieurs organisations virtuelles ; rend visibles les ressources allouées aux travaux de ces organisations ; rend possible *via* un système de gardes-barrières l'accès à des ressources localisées ou non (calculateurs, bases de données, *cloud*, ...); permet la soumission de travaux, la recherche d'informations, les réunions virtuelles, *etc.*

Nous allons maintenant changer un peu de sujet et étudier de plus près les interactions entre les personnes et l'aspect temps réel.

Interactions entre groupes

Les réseaux sociaux

Point non négligeable à ne pas laisser de côté, l'impact des réseaux sociaux et de toute la communication non institutionnelle entre les personnes. Certains aspects des réseaux sociaux ressortissent aux laboratoires virtuels. Nous connaissons les *Facebook d'entreprise* qui ont une fonction essentiellement sociale (justement) en rapprochant les membres des équipes mais dont le périmètre forcément limité compromet le succès (...). Nous rencontrons maintenant de plus en plus de chercheurs sur les réseaux professionnels type LinkedIn où ils sont très impliqués dans des groupes de discussions scientifiques ou techniques. Nous retiendrons aussi les aspects messagerie instantanée, *blogs* (avec la dimension *brainstorming*), et aussi le *flux continu de nouvelles* qui est diffusé aux membre connectés du réseau.

Dans ce domaine, il existe des outils (dont les archétypes sont Facebook, *LinkedIn*, *tweeter*, les *WebTV*, *etc.*) tout le problème est de définir un environnement dans lequel ils peuvent s'intégrer et interagir : une infrastructure ? Notons, par exemple, que la commission européenne demande explicitement aux projets scientifique des actions de communication et de dissémination qui dépassent les communautés impliquées et touchent même *le grand public*. Cela est déjà courant aux États-Unis (voir la NASA par exemple) et passe par ce type d'outils.

La visio-conférence

Le terme *visio-conférence* est un peu réducteur dans ce contexte car il doit couvrir outre la communication point à point, l'organisation de réunions « distribuées » à grande échelle, de séminaires, de cours voire même l'utilisation au quotidien de lieux virtuels et la télé-présence.

« Point à point » ou « groupe à groupe »

L'utilisation courante de la visio-conférence n'est finalement qu'un prolongement de celle du téléphone. Des salles de réunion équipées d'un dispositif *ad hoc* permettent d'inviter des personnes installées dans des sites similaires. Même s'il s'agit de salles de réunion le concept n'est guère éloigné de Skype (tout au moins dans sa perception par les utilisateurs) et reste basé sur une communication de personne à personne, en point à point. Ajoutons qu'il faut généralement se déplacer, trouver un « expert » qui sache faire fonctionner le système et que l'écran, le plus souvent placé dans un coin, oblige à tourner la tête pour se rappeler qu'il y a quelqu'un en ligne.

Une vision plus *collaborative* (justement) des choses consiste à s'intéresser plutôt aux interactions de *groupe à groupe* et à configurer l'installation, les équipements et... les lieux pour obtenir une qualité de contact plus proche du présentiel. L'étape suivante consiste à situer ces réunions dans des salles virtuelles pérennes dans lesquelles un interlocuteur « isolé » pourra, lui aussi, s'intégrer facilement et où les utilisateurs pourront partager des objets et des outils.

Logiciels et passerelles

Hors installations spécifiques (systèmes type TANDBERG, ...) nous trouvons beaucoup de logiciels dans ce domaine. Citons Webex [5] ou la suite *Multimedia Conference* [6] de chez Orange et coté *open source*, EVO [7] (successeur du système VRVS [8]) ou encore Access Grid [9]. Ces deux derniers sont très utilisés dans la communauté scientifique. En un mot, Access Grid est plus orienté interactions de groupe à groupe (*room based system*) et permet l'intégration d'applications scientifiques ; plus de 300 salles spécialement pensées et équipées pour Access Grid sont recensées dans le monde ainsi qu'un réseau de serveurs et de passerelles. En revanche, EVO qui part de l'interaction personne à personne (*desktop based system*) est plus dynamique sur certains aspects (création de salles virtuelles à la volée, ...) ; EVO est très largement répandu dans certaines communautés (physique fondamentale, ...).

Un aspect important coté réseau, est l'utilisation par Access Grid du protocole *multicast* qui permet d'économiser les ressources locales sur les serveurs et la bande passante sur le réseau.

Enfin et c'est le plus important, la plupart de ces logiciels « se causent » sans problème car ils utilisent les mêmes composants de base et les mêmes protocoles, il est facile d'imaginer des passerelles entre les différents systèmes (voir par exemple [10]).

Lieux virtuels et réels

Les voyages trop fréquents dans une collaboration *au quotidien* sont toujours coûteux (au moins en temps), fatigants mais surtout ils ne concernent *in fine* qu'une fraction des personnes concernées par la collaboration. Les visio-conférences —gardons ce terme pour le moment— ne visent pas à supprimer les déplacements mais à permettre la tenue de réunions régulières assurant un meilleur contact et une meilleure implication des acteurs d'un projet qui tous peuvent faire ce *déplacement virtuel*. Il est clair que cela facilite aussi l'implication des personnes qui travaillent sur les sites *en marge des projets* (équipes support, ...).

Dans ce cas une salle de réunion *équipée d'un écran dans un coin* ne suffit plus. La qualité des échanges devient essentielle et il faut travailler sur la disposition spatiale des équipements et l'aménagement d'un lieu spécifique qui est le point d'entrée de la salle de réunion virtuelle. En cas de collaboration sur le long terme, un tel lieu que l'on peut situer alors entre la salle de réunion et la salle documentaire (voire « de détente » ou « café ») peut rester « en ligne » en permanence et permettre des rencontres non planifiées (*modulo* quand même les décalages horaires !).

Tout cela ne décrit plus vraiment un système de visio-conférence mais plutôt une sorte d'infrastructure à la fois logicielle et matérielle de travail collaboratif. Coté logiciel, nous trouverons bien sûr des greffons de type son et vidéo mais nous pouvons imaginer d'autres fonctionnali-

tés généralistes (partage d'écran, de présentation, tableau blanc, *etc.*) ou plus orientées métier (visualisation de résultats de calcul, pilotage d'instruments scientifiques, *etc.*). Tout cela doit s'inscrire dans un contexte plus large : une infrastructure de service.

Le lieu virtuel n'existe qu'à travers ses contreparties concrètes. Le lieu réel de la collaboration n'est plus un bureau ni une salle de réunion, c'est un lieu de travail et d'échange où les équipements (écrans, audio, vidéo, *etc.*) qui peuvent être connectés en permanence (télé-présence) doivent refléter le lieu virtuel (la pluralité des lieux réels connectés), permettre un contact et des interactions proches du présentiel, et permettre de travailler sur des applications, des instruments, des *objets informatiques* (fichiers, ...) partagés.

Des concepts clefs (2)

Nous en avons rencontré deux nouveaux :

Les lieux virtuels des salles pérennes où l'on peut certes se voir et dialoguer mais aussi exécuter des applications partagées, accéder à des ressources, déposer des *objets*.

Les lieux réels spécialement pensés et équipés pour les divers aspects du travail collaboratif.

Nous allons continuer maintenant en prenant comme exemple la plate-forme Access Grid qui apporte des éléments de réponse aux problématiques du travail collaboratif. En effet, Access Grid implémente le concept de salle virtuelle pérenne ; réalise une infrastructure de service en proposant la gestion des salles et des accès (authentification) ainsi que le transfert transparent des données en utilisant au mieux les ressources matérielles disponibles (protocole *multicast*). La communauté d'utilisateurs a tissé un réseau mondial de serveurs et de passerelles (*unicast-multicast*) et propose des développements et des services (par exemple un programme d'assurance qualité, des reflecteurs pour tester la réception à distance de ses équipements, ...).

AccessGrid

Access Grid est une boîte à outils qui comprend un ensemble de ressources logicielles qui s'appuient sur une infrastructure de services pour faciliter l'organisation de sessions de travail collaboratif. Access Grid permet l'organisation de *réunions distribuées* à grande échelle, de séminaires et de cours utilisant des salles de réunion virtuelles. Conçus dès les années 2000 au sein de l'Argonne National Laboratory aux États-Unis, les différents composants du système Access Grid sont développés par un réseau de laboratoires spécialisés. Les modules disponibles (audio, video, *plugins* applicatifs, *etc.*) forment un ensemble cohérent et hautement sécurisé basé sur les technologies les plus avancées et performantes. Ces outils, diffusés librement, sont *supportés* et maintenus sur toutes les plates-formes (Linux, MacOS X® et Windows®) à l'intérieur de projets bien identifiés.

L'idée maîtresse dans la conception d'Access Grid est le concept d'interaction de groupe à groupe dans des lieux virtuels pérennes plutôt que d'individu à individu et ponctuellement (cf. la visio-conférence). Les outils d'Access Grid sont conçus pour utiliser et interconnecter les équipements multimedia : audio et video, et plus généralement d'interaction et de présentation en s'appuyant sur des services d'infrastructure (authentification, transferts de données, ...) du type de ceux que l'on trouve dans les intergiciels de grille. Le système offre aussi des interfaces qui facilitent l'établissement de passerelles vers les applications distribuées et les ressources qu'elles permettent d'accéder (typiquement, les environnements de visualisation). Access Grid est distribué sous forme *open source* et bénéficie des apports d'une très nombreuse communauté d'utilisateurs qui développe des services et des applications (par exemple l'enregistrement des sessions, les systèmes de réservation, ...) et offrent aussi des lieux d'échange virtuels et des réseaux d'expertises. Si le logiciel Access Grid propose une solution efficace pour la visio-conférence sur IP, il va beaucoup plus loin en répondant aux besoins de collaborations plus étroites des communautés regroupées dans des laboratoires virtuels ou impliquées dans les projets d'*e-infrastructure*. Access Grid est particulièrement adapté à la collaboration entre groupes de petite ou de grande taille répartis sur un grand nombre de sites :

- travail collaboratif, laboratoires virtuels, *e-infrastructures* ;
- réunions informelles impromptues ou plénières planifiées ;
- préparation et revues de projets ou de programmes ;
- formations et cours ;
- conférences, séminaires et *workshops*.

Concrètement, il suffit d'un PC avec *webcam*, micro et hauts-parleurs pour installer les composants de base du système, l'utiliser. Pour faciliter l'acceptabilité de ces technologies, les groupes de travail formés autour du projet Access Grid s'intéressent aussi aux aspects techniques et pratiques des installations (choix des matériels, réglages du son, ...) ainsi qu'aux

principes d'aménagement des salles pour obtenir un contact *proche du présentiel*. Dans cet ordre d'idée, l'université Central Queensland [11] en Australie a initié un programme de certification des nœuds Access Grid *Global Quality Assurance Program* [12]. Rappelons aussi que l'utilisation du protocole multicast permet de tirer le meilleur parti de l'infrastructure matérielle (serveurs et bien sûr le réseau).



La communauté Access Grid est très vivante et de nombreuses équipes contribuent activement au développement du logiciel ou des paquetages associés. Sur le serveur d'Argonne « la *default venue* par défaut », il est possible de partager ses éventuels déboires avec tous les primo-utilisateurs du monde entier ! La *default venue* du serveur de l'association Aristote [13] ambitionne, à terme, jouer ce rôle tout aussi efficacement pour la communauté française (pas de problème de langue et de décalage horaire). Début 2011, 325 sites à travers le monde possèdent des installations spécifiques pour utiliser Access Grid.



Map of 324 worldwide Access Grid nodes (2011)

Il existe aussi une liste de diffusion AG-TECH avec ses précieuses archives et le site internet (qui est un *wiki*) comprend, bien sûr, un forum. Des « communautés régionales » ont aussi vu le jour (Pacifique, Corée, Canada, *etc.*). Pour l'Europe, l'université de Manchester [14] est très active dans le projet Access Grid.

Deux types d'évènements sont programmés à intervalles réguliers :

- l'*Access Grid Townhall Meeting* (conseillé !!) rencontre mensuelle et virtuelle —*via* bien sûr Access Grid— pour échanger avec les membres de la communauté ;
- l'*Access Grid Retreat* colloque annuel —en présentiel— de trois jours. Les programmes et présentations sont accessibles en ligne sur le site d'Access Grid.

D'autres sessions de travail plus techniques sont parfois annoncées pour échanger, par exemple, sur les problèmes liés à l'utilisation du multicast, *etc.*...

Enfin un planificateur partagé de réunions et une liste des nœuds Access Grid (salles spécialement aménagées, voir ci-dessus) sont disponibles sur le site.

Brève description

Venues

Access Grid considère en tout premier lieu des salles virtuelles ou *venues* dans lesquelles les participants se regroupent pour une session de travail. Ces salles sont pérennes, si l'on y dépose des documents (fichiers, *etc.*) ceux-ci seront toujours accessibles lors de la session (réunion) suivante. Un serveur de salles virtuelles est accessible *via* une adresse de type `https` à travers le *venue client* que chacun peut installer sur son poste de travail. Un serveur peut proposer plusieurs salles virtuelles (*venues*, dont par exemple une « de test »), et des portes de sortie (*exits*) vers d'autres serveurs. Ainsi certains serveurs jouent-ils le rôle de points d'entrée privilégiés et, en quelque sorte, d'annuaires. Le logiciel comprend un certain nombre de modules destinés aux administrateurs (*venue server*, *venue management*, *etc.*) mais l'utilisateur pourra lui les ignorer et se contenter du *venue client*.

Profils

Lorsque l'on se connecte à un serveur de salles, plusieurs profils sont possibles : *user* ou *node*. En profil *user*, typiquement l'utilisateur dispose d'un système audio (au moins sur son PC, ...) et d'une caméra (*webcam*, ...). Mais si l'on connecte une salle de conférence, le profil *node* permettra de regrouper plusieurs ressources audio et vidéo sous un même identifiant. Une des machines connectées joue alors le rôle de pilote. Sur les machines *auxiliaires* un module *service manager*, exécuté en tâche de fond, peut lancer à la demande du pilote les services (audio, vidéo, ...) qui vont s'y exécuter en local. La machine pilote contrôle l'ensemble des périphériques et sert ainsi de régie.

Client, services et applications

Une fois *présent* dans une salle virtuelle, l'utilisateur peut accéder, selon la configuration de son client, à une sélection de services et d'applications. Parmi les services proposés dans le paquetage de base, nous trouverons au moins tout ce qu'il faut pour faire de la visio-conférence. Citons rapidement :

- service audio, en standard avec le logiciel *rat*[15], issu d'un projet fort ancien mais encore bien vivant repris dans le cadre du projet SUMOVER[16] devenu en 2008 AVATS[17] ;
- service *consommateur* ou *producteur* de vidéo, en standard avec le logiciel *vic* [18] issu des mêmes projets que *rat*(voir [16][17] déjà cités). Une version spécifique pour la haute définition[19] a été développée au laboratoire VISLAB[20] à l'université du Queensland en Australie ;
- *VPCscreen*[21] qui permet d'envoyer comme un flux video (avec *vic*) tout ou partie de ce qui affiché sur un écran ;
- sans oublier un service de *chat*, ou bavardage qui permet de régler rapidement bien des problèmes.

Les logiciels *rat* ("*Robust Audio Tool*") et *vic* ("*Videoconferencing Tool*") sont tout à fait standards, portables et correspondent à l'état de l'art dans le domaine.

L'utilisateur peut aussi partager des documents et activer ou se connecter à des «applications partagées» entre participants. Trois sont incluses dans le paquetage de base :

- présentation partagée, *via* le logiciel PowerPoint ou la suite OpenOffice ;
- document pdf partagé, *via* le lecteur Xpdf ;
- navigateur partagé, *via* le navigateur Firefox.

Enfin AGVCR[22], *enregistreur de sessions* qui permet de rejouer la réunion virtuelle est maintenant inclus dans la distribution standard d'Access Grid. D'autres outils permettent l'annotation ou encore la production de DVD à partir des enregistrements.

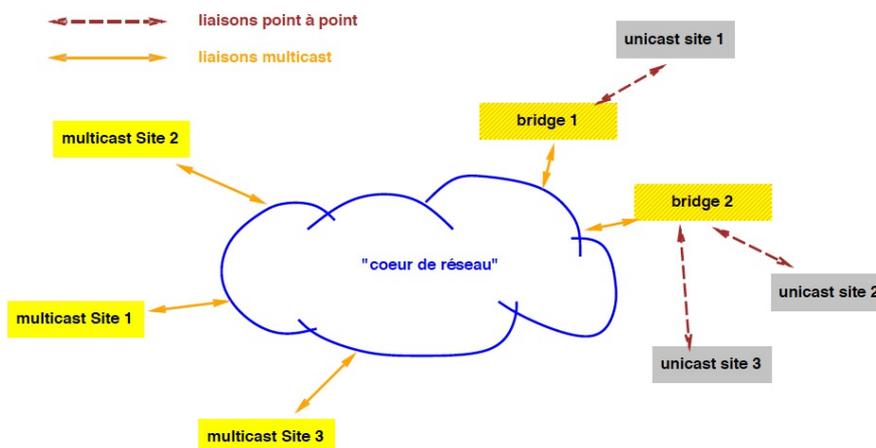
Proposer des salles virtuelles “Venue Server”

Un serveur de salles virtuelles doit être installé sur une machine visible sur internet avec la connectivité multicast. Cette machine, pour être intégrée dans l'architecture de grille d'Access Grid, doit posséder un certificat serveur qui peut être demandé auprès de l'autorité de certification *ad hoc* à l'Argonne National Laboratory. L'outil *venue server* sert ensuite à démarrer le serveur ; les ports et les plages d'adresses utilisés pour communiquer avec les clients ou pour les communications entre clients peuvent être configurés.

Une fois le serveur lancé, l'outil *venue management* permet de créer les différentes salles et les portes de sorties vers d'autres serveurs. Pour chaque salle, il est possible de définir qui en est l'administrateur (une personne qui peut donc être éventuellement extérieure à l'organisme qui héberge le serveur) et les conditions d'admission dans la salle (libre, ou *filtrage* par exemple avec des certificats, ...) ainsi que les modalités d'un éventuel cryptage des flux audio et video.

Passerelle “Bridge”

Pour les utilisateurs qui ne disposent pas du multicast, un service de passerelle *bridge* peut être lancé sur une machine elle aussi visible sur internet qui doit s'enregistrer auprès d'un serveur global. La communauté Access Grid propose un grand nombre de passerelles « ouvertes ». L'intérêt est quand même de disposer d'une passerelle « géographiquement proche » pour optimiser les transferts. Le choix multicast ou unicast +passerelle (à choisir dans la liste des passerelles accessibles) se configure, pour chaque utilisateur, dans le client.



Exemple de configuration Access Grid

Déployer AccessGrid au CEA

Prérequis

L'utilisateur doit pouvoir accéder à un serveur de salles de réunion virtuelles.

- soit directement en `multicast` ;
- soit en `unicast` *via* une passerelle.

L'accessibilité des machines qui supportent le serveur ou la passerelle déterminent la taille de la communauté potentiellement impliquée. Notons que divers ports doivent aussi être ouverts pour établir des communications vers le serveur ou encore pour la fonctionnalité de *chat*, ... tout cela sera récapitulé par la suite.

En intranet

Au CEA, le système peut être déployé en `intranet` sans aucun problème. Le `multicast` est *a priori* disponible, mais une passerelle peut être aussi proposée. Il suffit de configurer les modules serveur d'`Access Grid` pour que toutes les ressources accédées soient trouvées en local (notamment la localisation des passerelles et le *chat*). Un simple PC suffit pour mettre immédiatement à disposition les fonctionnalités d'`Access Grid` comme service `intranet` (audio, video, partage de documents, *etc.*).

En extranet

Pour s'ouvrir sur l'extérieur (et bénéficier éventuellement du réseau existant de serveurs et de passerelles), il faut bien sûr disposer d'un accès réseau configuré spécialement pour être capable de contacter un serveur extérieur et d'émettre et recevoir en `multicast` ou à travers une passerelle. Une configuration type a déjà été réalisée dans une pièce abritant une machine de l'association Aristote (Bt. 474 p. 39). De nombreuses pistes existent pour sécuriser le système, par exemple autoriser l'accès vers l'extérieur uniquement à travers une *passerelle CEA*, ...

Projet

L'intérêt d'un tel déploiement est de montrer les points forts du système en travaillant sur la dualité lieux réels-lieux virtuels. Pour cela il faut identifier des salles de réunion (de taille raisonnable) et y installer les matériels nécessaires avec la connectivité réseau *ad hoc*. L'aménagement des lieux nécessite :

- des caméras (dont au moins une haute définition) ;
- un ou des microphones (avec de préférence un dispositif de suppression d'écho) ;
- un PC raisonnablement puissant pour l'acquisition des divers flux ;
- des vidéo-projecteurs (2 ou 3) : pour les personnes, les lieux connectés, les documents (transparents, *etc.*) ; dans un premier temps un grand mur blanc est le bienvenu ;
- une disposition *ad hoc* du mobilier.

Si la salle est déjà équipée pour la visio-conférence *classique* il est normalement possible d'utiliser les dispositifs d'acquisition (micro, caméra) déjà présents dans la salle. Les vues ci-dessous montrent des aménagements typiques de lieux (de petite taille) pensés pour utiliser Access Grid.



Exemples d'utilisation d'Access Grid

Mise en œuvre

Il est possible de procéder en deux temps. Tout installer en intranet, ce qui permet d'avancer sur les aspects sécurité avant de s'ouvrir vers l'extranet lorsque tout les problèmes ont été abordés et résolus.

Besoins : serveur et passerelle

Un PC avec les logiciels Access Grid installés et configurés pour le serveur et la passerelle (qui peuvent *tourner* sur la même machine). Une connexion réseau spécifique (cf. bulle Aristote ou autre solution) mais dans un premier temps tout peut être installé et testé en intranet.

Besoins : lieux et matériels

Il faut identifier des lieux (salles de réunion) et rassembler les matériels nécessaires. Si la salle est déjà équipée pour la visio conférence il suffira d'un PC, de quelques *webcams* et de vidéo-projecteurs. Sinon il faut prévoir en plus une caméra (raisonnablement) HD et un microphone avec suppression d'écho. Tout cela connecté à un PC, installé dans la salle, qui doit pouvoir contacter le serveur et éventuellement la passerelle ; ce qui est évident si tout fonctionne en intranet.

Sécurité

Nous regroupons dans cette section, les informations qui permettent de configurer l'accès réseau pour que les utilisateurs puissent utiliser le client `Access Grid` sur leur poste de travail. Cela permet de configurer notamment les *firewalls*. Notons que la restriction de l'utilisation à certains serveurs de salles ou encore à certaines passerelles (par exemple un hypothétique *bridge* CEA) permet d'affiner ces réglages et de ne pas ouvrir des brèches dans les *firewalls*. Toutes ces données sont susceptibles d'être modifiées (et plutôt simplifiées) au fil des versions ; les informations à jour sont disponibles sur le site internet du projet `Access Grid`. Le serveur et la passerelle peuvent être installés sur la même machine.

Configuration pour le serveur

Le serveur sert juste à mettre les clients en relation, en particulier, les divers flux audio et video ne transitent pas par lui. Un certain nombre de ports sont dédiés aux communications avec les clients et les autres composants du système.

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
8000	TCP	IN/OUT	Virtual Venue Server Port
8002	TCP	IN/OUT	Virtual Venue Server Event Port
8004	TCP	IN/OUT	Virtual Venue Server Text Port
8006	TCP	IN/OUT	Virtual Venue Server Data port
50000-50020	TCP	IN/OUT	Virtual Venue Server Data transfers (fichiers, <i>etc.</i>)

Le serveur `jabber`[24] pour le *chat* utilisé par défaut est celui d'Argonne, mais il est facile d'installer un serveur en local :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
5223	TCP	IN/OUT	Jabber (par défaut <code>jabber.mcs.anl.gov</code>)

Le serveur doit aussi pouvoir contacter les registres de passerelles (par défaut Argonne, ou Manchester en Europe mais il est possible d'installer un registre local) :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
8030	TCP	IN/OUT	Bridge registry (par défaut <code>vv3.mcs.anl.gov</code>)

Enfin le serveur doit disposer de la connectivité `multicast` pour établir les connexions entre les clients. La plage d'adresses `multicast` va de `224.0.0.0` à `239.255.255.255` (c'est à dire `224.0.0./4`) mais les adresses effectivement utilisées peuvent être configurées.

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
5353	TCP/UDP	IN/OUT	nécessaire pour le DNS multicast
224.0.0.0/4	UDP	IN/OUT	trafic multicast

L'idéal, pour vérifier la connectivité *multicast* est de pouvoir lancer un *beacon* sur sa machine[23]. Cela nécessite :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
10002	UDP	IN/OUT	nécessaire pour beacon multicast
10003	UDP	IN/OUT	nécessaire pour beacon multicast
10004	UDP	IN/OUT	nécessaire pour beacon multicast

Configuration pour la passerelle

La passerelle communique avec les clients *via* un port spécifique et utilise une autre plage de ports pour les flux de données UDP (audio, vidéo, *etc.*). Ces valeurs sont configurable.

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
2000	TCP	OUT	communication avec les clients (valeur par défaut)
50000-52000	UDP	IN/OUT	trafic avec les clients (valeurs par défaut)

Configuration pour le client

Le client (*venue client*) doit bien sûr pouvoir contacter et communiquer avec le serveur (*venue server*) qui utilise les ports suivants (cf. plus haut) :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
8000	TCP	OUT	Virtual Venue Server Port
8002	TCP	OUT	Virtual Venue Server Event Port
8004	TCP	OUT	Virtual Venue Server Text Port
8006	TCP	OUT	Virtual Venue Server Data port
50000-50020	TCP	IN/OUT	Virtual Venue Server Data transfers (fichiers, <i>etc.</i>)
5223	TCP	IN/OUT	Jabber (par défaut jabber.mcs.anl.gov)

Pour que les divers services fonctionnent (mais tous ne sont pas indispensables), le client doit pouvoir utiliser aussi :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
21	TCP/UDP	IN/OUT	nécessaire pour FTPS
443	TCP/UDP	IN/OUT	nécessaire pour HTTPS, TLS, SSL
5909	TCP/UDP	OUT	nécessaire pour VenueVNC (et IN pour partager son écran)

et aussi, mais le plus souvent cela concerne des machines qui sont toutes sur le réseau local :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
11000	TCP	IN/OUT	nécessaire pour le NodeService Manager

Il est souhaitable que les *firewall* laissent passer certains trafics :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
22	TCP/UDP	IN/OUT	nécessaire pour SSH
80	TCP	OUT	nécessaire pour HTTP
	ICMP	IN/OUT	nécessaire pour Ping

Enfin il faut aussi que le *firewall local* accepte tout le trafic avec *localhost* pour que certains modules puissent fonctionner (typiquement rat) :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
localhost	TCP/UDP	IN/OUT	nécessaire pour rat (et autres applications)

Mode multicast Le client Access Grid fonctionne mieux en mode *multicast* qui est la solution idéale pour distribuer des contenus de un à plusieurs utilisateurs. L'utilisation du *multicast* n'introduit pas *a priori* de faille de sécurité dans la mesure où l'utilisateur doit d'abord s'abonner à un groupe pour recevoir des données. En revanche, potentiellement, tout le monde peut recevoir les données *multicast* émises par un utilisateur d'Access Grid, pour cela l'accès à la salle de réunion virtuelle peut être soumise à la présentation d'un certificat X.509, ou encore les flux peuvent être cryptés. D'où les configurations supplémentaires :

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
5353	TCP/UDP	IN/OUT	nécessaire pour le DNS multicast
224.0.0.0/4	UDP	IN/OUT	trafic multicast (défaut)

Si le trafic *multicast* est ouvert « en entrée seulement », il est possible de suivre une conférence ou un séminaire retransmis *via* Access Grid mais sans pouvoir intervenir. C'est le cas sur l'intranet du CEA actuellement.

Mode unicast Pour les sites qui ne bénéficient pas de la connectivité *multicast*, Access Grid prévoit l'installation de machines passerelles (*bridges*). Les flux (audio, vidéo, *etc.*) liés à la session Access Grid sont alors échangés sous forme de flux UDP entre le client et la passerelle. Chaque passerelle est configurée pour écouter sur un port précis et fonctionner avec une plage de ports donnée. La liste des passerelles disponibles détectées par le système (à partir d'une source configurable) avec leurs caractéristiques précises est disponible dans le menu préférences (puis bridging) du client.

Port ou Hôte	Protocole	Entrant-Sortant	Commentaire
2000	TCP	OUT	communication avec la passerelle (défaut)
50000-52000	UDP	IN/OUT	trafic avec la passerelle (défaut)

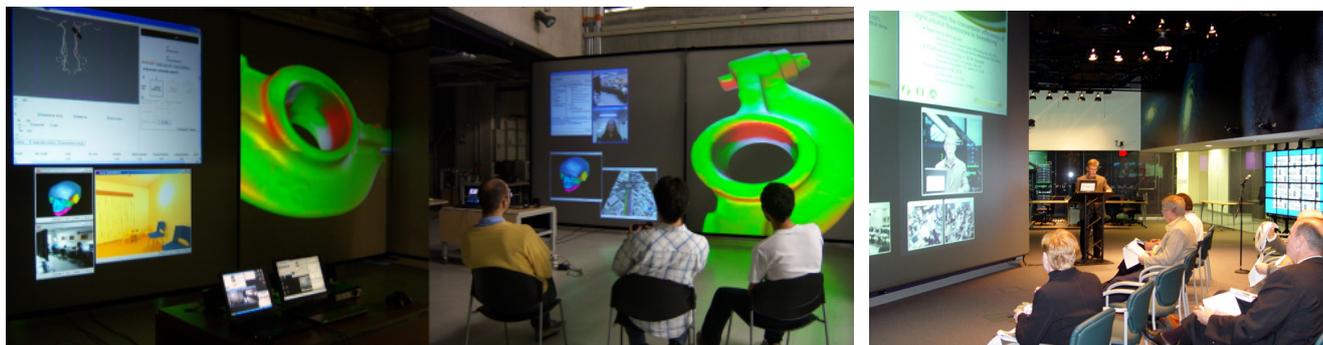
Conclusions

Un déploiement expérimental d'Access Grid au CEA permettrait d'appréhender les problématiques liées aux futurs environnements de travail collaboratif : usages (aménagement spécifiques, télé présence, ...), approche infrastructure de service-couche d'interopérabilité, problèmes de déploiement et de sécurité. Il faut pour cela identifier une *communauté* qui serait prête à investir un minimum pour installer et tester le système puis configurer au moins deux « lieux Access Grid » au CEA (dont un USCI). Pour un calendrier s'étalant sur (environ) 6 mois, les différentes étapes seraient les suivantes :

1. Installation en intranet d'un serveur, d'une passerelle et tests sur des postes de travail (micro ou casque et *webcam*) ; cela permet d'effectuer rapidement des démonstrations.
2. En parallèle, étude sécurité pour envisager des lieux ouverts vers l'extérieur.
3. En parallèle, identifier des utilisateurs potentiels au CEA.
4. Identifier des lieux à aménager (un USCI, un *utilisateurs*).
5. Aménagement des lieux, tests en interne « en vrai grandeur ».
6. Ouverture du système, utilisation avec des acteurs externes au CEA.

Le projet nécessite un budget matériel mais pas de développement logiciel : juste la configuration correcte des différents modules. Dans un premier temps l'aménagement des lieux peut rester *minimal*. Parmi les suites possibles d'un tel projet :

1. Intégration dans un portail, qui permet un accès simple sans installer un client.
2. Interopérabilité avec les autres systèmes déployés ou utilisés au CEA (Visio « standard », logiciels comme EVO, *etc.*)
3. Utilisation dans un cadre plus large (séminaires, cours, ...).
4. Le système constituant une infrastructure de service, il est facile de développer des greffons applicatifs dédiés —il en existe déjà— à des utilisations spécifiques (partage d'outils de visualisation de données scientifiques, *etc.*).



Exemples d'utilisation avancées d'Access Grid

Références

- [1] BEVO 2008, Building Effective Virtual Organizations
<http://www.ci.uchicago.edu/events/VirtOrg2008/>
- [2] EGEE, Enabling Grids for E-scienceE
<http://www.eu-egee.org>
- [3] EGI, European Grid Initiative
<http://www.egi.eu/>
- [4] Institut des grilles du CNRS
<http://www.idgrilles.fr>
- [5] WebEx, Vidéoconférence, net meeting et web conférence en ligne
<http://www.webex.fr/>
- [6] Orange Multimedia Conference
<http://www.orange-business.com/fr/entreprise/travail-collaboratif/reunion/multimedia-conference/>
- [7] EVO, Enabling Virtual Organizations
<http://evo.caltech.edu> et <http://evo.vrvs.org>
- [8] VRVS, Virtual Rooms Videoconferencing System
<http://www.vrvs.org>
- [9] Access Grid is a collection of resources and technologies that enables large format audio and video based collaboration between groups of people in different locations.
<http://www.accessgrid.org>
- [10] The UK's Education and Research Network, Access Grid Support Centre
<http://www.ja.net/services/video/agsc/AGSCHome/>
- [11] The University of Queensland (AU) <http://www.uq.edu.au> et
<http://www.hpc.cqu.edu.au/accessgrid.html>
- [12] Jason Bell, The AccessGrid Global Quality Assurance Program, 2007 Access Grid Retreat, Chicago, 14-16 Mai 2007.
- [13] Groupe de travail « Access Grid Travail collaboratif » de l'association Aristote <http://e-lab.association-aristote.fr>
- [14] Research Computing Services, University of Manchester
<http://www.rcs.manchester.ac.uk/services/accessgrid>
- [15] RAT, Robust Audio Tools
<http://http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/>

- [16] SUMOVER, Support for MbOne Videoconferencing for the Research community
<http://http://www.cs.ucl.ac.uk/research/sumover/>
- [17] AVATS, AccessGrid Video and Audio Tools Support
<http://www.cs.ucl.ac.uk/research/avats/>
- [18] VIC, Videoconferencing Tool
<http://www-mice.cs.ucl.ac.uk/multimedia/software/vic/>
- [19] Access Grid Advanced Video Project
<http://www.vislabs.uq.edu.au/research/accessgrid/software/advideo>
- [20] ViSLaB, A Computational Science & Visualisation group, The University of Queensland
<http://www.vislabs.uq.edu.au/research/accessgrid>
- [21] VPCScreen, Video Producer Service for the AccessGrid <http://www.accessgrid.org/project/VPCScreen>
- [22] AGVCR : AccessGrid Video (Cassette) Recorder, Indiana University
<http://iri.informatics.indiana.edu/~dcpiper/agvcr>
- [23] Multicast Beacon <http://dbeacon.innerghost.net>
- [24] Jabber.org is the original IM service based on XMPP, the open standard for instant messaging <http://www.jabber.org>

