

Aristote

Nouvelles technologies de l'information et de la communication

Rapport d'activité 2011



<http://www.association-aristote.fr>
info@association-aristote.fr

Edition du 24 ventôse an CCXX (*vulg.* 14 mars 2012) ©2012 Aristote

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI CEN Saclay
Bât. 474 91191 Gif-sur-Yvette Cedex.
Secrétariat : Aristote, École Polytechnique, 91128 Palaiseau Cedex.
Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 67
Courriel : Marie.Tetard@polytechnique.edu
Site internet <http://www.association-aristote.fr>



Table des matières

1	Aristote en 2011	5
1.1	Éditorial	5
1.2	Vie de l'association	6
1.3	Faits marquants	8
2	Groupes de travail	9
2.1	Les groupes de travail dans Aristote	9
2.2	Groupe PIN	10
2.3	Groupe Gus'G	12
2.4	Groupe Access Grid	12
2.5	Groupe Calcul Hybride	14
2.6	Groupe e-Lab (travail collaboratif)	14
3	Séminaires, conférences et formations	17
3.1	Le cycle 2011	17
3.2	L'interopérabilité dans le <i>Cloud</i>	18
3.3	Journée Calcul Hybride, OpenGPU un an plus tard	22
3.4	Le déluge de données : comment en tirer parti	24
3.5	Travail collaboratif et <i>Cloud</i>	29
3.6	Le Web2.0 : de la théorie à la pratique	34
3.7	Les stages et formations	39
3.8	Formations Cloud et HPC	39
3.9	Pérennisation et communication de l'information numérique	41
4	Perspectives pour 2012	45

Chapitre 1

Aristote en 2011

1.1 Éditorial

En 2011, sur la base des principes fondateurs de l'association, nous avons revisité nos objectifs pour les adapter aux enjeux des NTICs. Grâce à la mobilisation de tous, nous avons réussi à redynamiser l'association. Pour améliorer l'efficacité et assurer une meilleure visibilité à l'association nous avons adopté de nouveaux statuts. De nouveaux membres nous ont rejoint dont l'expérience sera fructueuse pour notre communauté riche de ses savoirs au service de l'innovation. Nous avons réaffirmé l'importance du comité de programme qui est le moteur des activités d'Aristote.

Conformément aux engagements que nous avons pris, l'exercice 2011 est équilibré grâce à la maîtrise des dépenses et à l'augmentation du nombre d'adhérents.

Les perspectives pour 2012 sont bonnes, nous allons conforter la gouvernance, l'anticipation et le modèle économique au service des programmes dans un esprit d'ouverture. Les perspectives, en terme d'innovation, de formation et d'incubation de valeur sur le plateau de Saclay sont une motivation complémentaire. Le dynamisme et l'implication de tous nous permettrons de confirmer nos ambitions d'être un des acteurs de référence au service de l'innovation.



Bernard Rouault
Président d'Aristote

1.2 Vie de l'association

Objectifs et activités

Aristote est une société savante et philotechnique qui regroupe depuis plus de 25 ans organismes de recherche, grandes écoles, entreprises, PME et associations impliqués dans les nouveaux développements et usages des technologies de l'information.

Créée « informellement » en 1984 par l'INRIA, le CEA, EDF et le CNES, Aristote est depuis 1988 une association loi 1901. Le siège de l'association est situé au :

CEA-DSI Centre de Saclay 91191 Gif-sur-Yvette Cedex

Aristote contribue à tisser des liens entre le monde académique et celui de l'industrie et des services à travers ses activités de transfert de technologie ou de veille scientifique et stratégique. L'association propose des groupes de travail qui sont des lieux d'échanges privilégiés autour des NTICs. Elle organise des cycles de séminaires et des **formations**. Le Comité de programme est un lieu de réflexion et de débat : le lien indispensable entre membres, groupes de travail, auditeurs des séminaires et plus généralement la « communauté Aristote ».

Aristote organise chaque année, à l'École Polytechnique, un cycle de séminaires diffusés aussi en direct sur internet. Les thèmes choisis reflètent la diversité et la richesse des membres de l'association. Les programmes privilégient des approches très ouvertes en croisant regards et cultures, recherche fondamentale et retours d'expérience. Ils apportent des éclairages nouveaux et des éléments clefs de réponse aux problématiques traitées. Au programme du cycle 2011 : « Le Web2.0, de la théorie à la pratique », « Travail collaboratif et Cloud », « Séminaire Big Data », « Journée Calcul Hybride », « L'interopérabilité dans le Cloud ». L'expertise des groupes de travail est diffusée à travers des formations, ainsi en 2011 deux sessions « Cloud-HPC » ont eu lieu à l'Institut Henri Poincaré et une session « Pérennisation des informations numériques » a été organisée à l'Institut National du Patrimoine.

Aristote possède des serveurs informatiques hébergés à l'École Polytechnique dans le cadre d'une convention et offre ainsi des services à la communauté, notamment en ce qui concerne les outils collaboratifs (salles de réunions virtuelles, *etc.*).

Aristote fournit à ses adhérents le cadre légal et la structure d'accueil pour la création et pour le fonctionnement des groupes de travail. Au sein d'Aristote, un groupe de travail rassemble une communauté qui se constitue autour d'une problématique ouverte dans le domaine des technologies de l'information et de la communication. Le groupe de travail est un lieu de rencontre et d'échange ; d'entretien, de valorisation et de création de compétences ; ou encore de renouvellement du savoir-faire. Le groupe permet aussi l'émergence de coopérations plus profondes entre certains de ses membres. Il contribue aux activités de l'Association : comité de programme, séminaires, formations. Citons : PIN pérennisation des informations numériques ; Gus'G utilisateurs de grilles et HPC ; Club des utilisateurs du Calcul Hybride ; Utilisateurs d'AccessGrid ; e-Laboratoire, lieux virtuels et travail collaboratif.

Les instances

Le Bureau (depuis l'assemblée générale du 7 juillet 2011)

Président

- M. Bernard Rouault, Directeur Gestion du Patrimoine CEA
bernard.rouault@cea.fr

Vice-Président

- M. François de Castelbajac, Directeur des Systèmes d'Information, École Polytechnique
Francois.de-Castelbajac@Polytechnique.edu

Trésorier

- M. David Bateman, Chef du Pôle SINETICS -EDF-R&D
David.Bateman@edf.fr

Secrétaire

- M. Thierry Roussel, CEA
thierry.p.roussel@cea.fr

Conseiller scientifique et technique

- M. Philippe d'Anfray, CEA
Philippe.d-Anfray@cea.fr

Représentants des membres¹ et administrateurs²

- Denis Attal (Thalès)
- David Bateman (EDF, administrateur)
- Laurent Becquey (CNES, administrateur)
- Dominique Birman (Météo France)
- Luc Boucher (EURIWARE)
- Elba Burity (IFSTTAR, administrateur)
- François de Castelbajac (École Polytechnique, vice-président, administrateur)
- Jean-Michel Cornu (FING, administrateur)
- Pierre Cotty (IFREMER)
- Francis Daumas (Cines)
- Catherine Dhérent (BnF)
- Alain Fabre (Orange Labs)
- Georges Filhol (Institut Telecom)
- Don Foresta (MARCEL)
- Jean-Marc Karçenty (elcca)
- Jean François Lavignon (BULL)
- David Loureiro (SysFera)
- Pierre Malecki (ONERA, administrateur)
- Pascal Michel (IFPEN, administrateur)
- Nicolas Monnier (Cerfacs)
- Hubert Pampouille (INRA)
- Stéphane Requena (GENCI)
- Bernard Rouault (CEA, président, administrateur)
- Joël Sor (CIRAD, administrateur)
- Eric Spiteri (CS-SI, administrateur)
- Jean-Marc Voltini (CNRS)

Secrétariat et communication

- Mme Marie Tétard, Aristote-École Polytechnique 91128 Palaiseau CEDEX
Tél. : +33 (0)1 69 33 99 66 Fax : +33 (0)1 69 33 99 67
Site internet <http://www.association-aristote.fr>
Marie.Tetard@polytechnique.edu

1. Liste mise à jour fin 2011

2. Liste établie lors de l'assemblée générale du 7 juillet 2011

1.3 Faits marquants

Nouveaux adhérents

Thalès, EURIWARE, BULL, elcca, SysFera, INRA et CNRS ont rejoint l'association dans le courant de l'année 2011.

Assemblée générale 2011

L'assemblée générale s'est tenue comme tous les ans au siège du CNES à Paris. À cette occasion, les nouveaux statuts de l'association ont été adoptés.

Activités

Nous avons organisé, cette année, 5 séminaires et 3 sessions de formation. Les réunions du CPG connaissent en outre une affluence particulière, il est prévu d'organiser régulièrement des visio-conférences pour accueillir les participants géographiquement dispersés.

Aristote était présente sur un stand aux Journées du Club des Utilisateurs de l'Informatique du CEA sur le thème du travail collaboratif. JCUIC 2011 CEA Cadarache les 6 et 7 avril 2011.

Aristote était présente aux cotés de GENCI au Forum TERATEC 2011, les 28 et 29 juin 2010 à l'Ecole Polytechnique, centré sur le développement très rapide des usages du calcul intensif.

Un groupe Aristote a été créé sur le réseau social professionnel LinkedIn qui regroupe maintenant (fin 2011) plus de 110 membres.

Partenariats

Aristote était partenaire de l'Open World Forum qui a eu lieu du 22 au 24 septembre 2011 à Paris. À cette occasion, nous avons publié une brochure regroupant des points de vue des adhérents sur l'Open Source.

Aristote était partenaire de la conférence Web3d qui a eu lieu à Paris du 20 au 22 Juin 2011.

Enfin, le partenariat avec le projet OpenGpu a permis l'organisation d'une nouvelle journée commune sur le calcul hybride le 8 juin 2011 (voir la rubrique Séminaires).



Lettre d'information Aristote OWF-2011

Aristote est une société savante et philotechnique qui regroupe depuis plus de 25 ans instituts, grandes écoles, organismes de recherche et entreprises impliqués dans les derniers développements et nouveaux usages des technologies de l'information. Aristote contribue à tisser des liens entre le monde académique et celui de l'industrie et des services à travers ses activités de transfert de technologie ou de veille scientifique et stratégique. L'association organise des séminaires et des formations, elle propose des groupes de travail qui sont des lieux d'échanges privilégiés autour des NTICs et d'animation de la « communauté Aristote ».

OPEN WORLD FORUM 2011

Cette lettre d'information publiée à l'occasion de l'Open World Forum 2011 est l'occasion pour les membres d'Aristote d'exposer leur point de vue ou de présenter un témoignage sur l'Open Source dans leur domaine d'activité

Aristote et l'Open Source

Aristote fait intervenir dans ses séminaires de nombreux acteurs du monde du libre, de tous horizons, chercheurs, chefs d'entreprise, développeurs et responsables impliqués dans leur usage. Des séminaires ont été consacrés à la mise en œuvre des solutions libres (2008 à 2010) dans l'objectif de la réussite des projets et à leurs avantages en termes de coût, de pérennité, de maîtrise et de qualité.

Les actes des séminaires « L'Open Source, réusir son projet » (2008), « L'Open Source, standards et initiatives dans l'enseignement supérieur et la recherche » (2008), « Le libre à coût raisonné » (2010) sont en ligne sur notre site internet.

www.association-aristote.fr

<http://www.association-aristote.fr>

L'ASSOCIATION ARISTOTE porte depuis des années un regard attentif sur l'évolution du paysage des logiciels libres. Cet intérêt découle des nombreux avantages qu'ils offrent :

- les logiciels libres sont construits à partir de normes et standards ouverts qui assurent leur interopérabilité ;
- l'accès au code source permet de comprendre le logiciel et de corriger rapidement les problèmes sans dépendre d'un tiers, assurant ainsi à l'utilisateur une meilleure maîtrise. Cette ouverture offre aux étudiants et aux chercheurs des moyens d'études incomparables et la possibilité de contribuer efficacement à l'évolution des logiciels.

Faut-il rappeler qu'un logiciel est dit « libre » si sa licence accorde la liberté d'adopter le logiciel pour tous les usages, d'étudier son fonctionnement, de diffuser des copies, de modifier et redistribuer la nouvelle version. « Le libre » repose sur des principes très forts, formant un cadre juridique fondé sur le droit d'auteur et le droit au partage.

Les logiciels libres couvrent aujourd'hui tous les domaines de l'informatique : systèmes d'exploitation, serveurs web, bases de données, langages, navigateurs, suites bureautiques. Depuis quelques années des solutions métiers complètes commencent à voir le jour, en concurrence des progiciels métiers des éditeurs.

Au-delà de l'ouverture du code source, le modèle du logiciel libre propose une logique coopérative et une mutualisation des moyens. A travers ces nouvelles dynamiques de coopérations, certains grands acteurs du numérique proposent un nouvel écosystème des industries du traitement de l'information.

Aristote, grâce à son réseau, exerce une veille active sur le libre. Les actions de l'association contribuent aux échanges des expériences entre les acteurs, pour les choix des logiciels, leur usage, leur déploiement, leur support. La France est très en avance sur ce dernier point, puisque des entreprises et des administrations ont créé des supports contractuels leur permettant d'obtenir des garanties équivalentes à celles des éditeurs tout en contribuant à la vitalité des communautés.

Conception et réalisation : Philippe d'Anfray Édition du 28 tracteur an COXX (juil. 15 sept. 2011) ©2011 Aristote
ARISTOTE Association Loi de 1901 - Siège social : CEA-DSI CEN Saclay Bld. 474, 91191 Gif sur Yvette Cedex.
Secrétariat : Aristote, Ecole Polytechnique, 91128 Palaiseau Cedex.
Courriel : info@association-aristote.fr Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 67

Chapitre 2

Groupes de travail

2.1 Les groupes de travail dans Aristote

Un groupe de travail au sein d'Aristote est une communauté constituée autour d'une problématique ouverte dans le domaine des technologies de l'information et de la communication. Cette communauté est constituée d'experts du domaine et de personnes qui ont à faire face à cette problématique ; elle mêle des représentants du monde de la recherche, de l'industrie et des services. Le groupe de travail est un lieu de rencontre, d'échange, d'entretien et de valorisation des compétences, de création de compétences et de renouvellement du savoir-faire. Les groupes de travail répondent aussi aux besoins exprimés par les organismes membres d'Aristote en terme de veille technologique ou de formation.

Le groupe est une entité qui se veut totalement **neutre** par rapport aux offres commerciales et industrielles du domaine. Les activités du groupe et ses modes d'action sont ouverts et multiples, par exemple :

- interactions au sein du groupe lui-même (réunions thématiques, etc.) pour formaliser les besoins, rendre compte des retours d'expérience, élaborer des synthèses, ... ;
- interactions plus large au sein du réseau Aristote rassemblant l'ensemble des membres des divers groupes ;
- organisation de colloques, séminaires ou ateliers de portée nationale ou internationale ;
- organisation de formations ;
- développement de prototypes, expérimentations.

De part ses activités et ses analyses, le groupe peut contribuer à orienter l'élaboration de méthodologies, de concepts, le développement de produits libres ou commerciaux répondant aux besoins de la problématique posée.

Le groupe est aussi un cadre adéquat pour l'émergence de coopérations plus profondes entre certains de ses membres.

Le groupe contribue aux autres activités de l'Association et notamment au CPG, « Comité de Programme et des Groupes de travail » en charge de la préparation du cycle annuel de séminaires dont les thèmes sont les reflets des travaux menés dans les différents groupes de l'association.

L'association accueille actuellement les groupes PIN (pérennisation des informations numériques) ; Gus'G (utilisateurs de grilles et HPC) ; Club des utilisateurs du Calcul Hybride ; Utilisateurs d'Access-Grid ; e-Laboratoire (lieux virtuels, travail collaboratif, arts et sciences).

2.2 Groupe PIN

Rappel des objectifs du groupe

Le groupe PIN (Pérennisation des Informations Numériques) est un lieu de rencontre et d'échanges entre informaticiens, archivistes et bibliothécaires pour le partage de retour d'expériences, savoir-faire, succès et échecs dans la pérennisation de l'information numérique. Il essaie de contribuer à la définition et au développement de produits répondant aux besoins.

Moyens

Le site Web du groupe PIN (<http://pin.association-aristote.fr>) rassemble les comptes-rendus de réunion et les interventions des réunions plénières depuis dix ans. Il est utilisé par les participants pour obtenir des informations sur la vie du groupe et par les équipes de travail spécifique au sein de PIN, notamment pour l'élaboration et la coordination de la session annuelle de formation. La liste de diffusion du groupe rassemble environ 150 inscrits.

Institutions participant au groupe PIN

L'animation principale est effectuée par O. Rouchon (CINES) avec l'aide d'un comité d'animation regroupant des représentants de six organismes moteurs : BnF, CEA, CINES, CNES, CNRS-INIST, SIAF, et se partageant les tâches de fonctionnement du groupe ; les autres institutions représentées sont :

Organismes à caractère scientifique et technique : CEA, CNES, IGN, CINES, CNRS (INIST), IRD, CIRAD, INSERM, INRIA.

Institutions éducatives, patrimoniales et culturelles nationales et internationales : BnF, SIAF, Unesco, Ecole Nationales des Chartes, Universités.

Ministères : Justice, Education, Ecologie et développement durable, Défense, Culture, Enseignement supérieur et recherche, Cour des Comptes.

Autres secteurs : Sanofi, Institut Pasteur, Hôpitaux de Paris, La Poste, IRCAM, DGAC, Institut National des Télécommunications.

Représentants de Conseils généraux et de communautés territoriales : Conseils Généraux, Archives départementales, Bibliothèques municipales. Industrie : EDF, ATOS.

Divers intervenants étrangers abonnés au forum : NIST, Tessella (UK), Open Planets Foundation, DANS (Pays-Bas), Archives Fédérales (Suisse).

Sujets abordés au cours des trois réunions plénières en 2011

Trois réunions plénières ont été organisées les 18 Janvier, 21 Juin et 2 Décembre, rassemblant en moyenne 45 personnes, au cours desquelles les thèmes suivants ont été abordés :

- le dispositif d'agrément par le SIAF de prestataires en archivage électronique pour la conservation d'archives publiques intermédiaires — conditions d'externalisation ;
- les projets européens pour la conservation de l'audiovisuel — le projet PrestoPRIME ;
- l'archivage des courriels, principes et contexte, retours d'expérience ;
- les métadonnées : les métadonnées et l'interopérabilité, la directive européenne INSPIRE, METS et PREMIS, retour d'expérience ;
- l'implémentation des fonctions de versement, stockage et gestion des données de l'OAIS dans des SAE commerciaux : SDB de Tessella, Worldline archiving d'ATOS, Mnesys de Naoned systèmes et Arcsys d'Infotel.

Formation

Une session de formation de quatre jours a été organisée en collaboration avec l'Institut National du Patrimoine en octobre, à laquelle trente cinq personnes se sont inscrites, dont cinq par l'intermédiaire d'Aristote.

A partir de 2012, une collaboration similaire sera mise en place avec l'Ecole Nationale des Chartes pour une session annuelle de formation de cinq jours à organiser en Novembre ; l'association Aristote y bénéficiera de quinze places de stagiaires dont elle pourra assurer l'inscription.

La collaboration avec l'INP sera maintenue sous une forme à définir — périodicité annuelle ou bisannuelle, à programmer au printemps.

Les travaux du groupe de travail sur le cursus d'enseignement de longue durée, stoppés depuis 2010, n'ont pas été repris.

Travaux et publications

Claude Huc (ex-animateur du groupe PIN) : Préserver son patrimoine numérique, Eyrolles — ISBN 978-2-212-12789-8

PIN - PÉRENNISATION DES INFORMATIONS

Piste: - Dernières actualités

Rechercher

ACTIONS PAGE

- Accueil
- Actualités, Archives
- Fonctionnement
- Ressources
- Réunions plénières
- Contact
- Session de formation
- Espace Membres
- Espace Formation
- Espace Pilotes
- Espaces Groupe de travail
- Dokuwiki, syntaxe
- Bac à sable

ACTIONS WIKI

- INDEX
- DERNIERS CHANGEMENTS

ACTIONS UTILISATEUR

- CONNEXION

PARTAGER CETTE PAGE

- Diigo This!
- Del.icio.us
- Google Bookmark

Le groupe PIN est un lieu d'échange, de réflexion et d'action vers l'émergence de solutions pour ...

- Les informations à pérenniser**
 - Identification des catégories d'informations à pérenniser.
 - Classification des formats de fichiers de données ou de documents électroniques.
 - Processus de collecte des informations à pérenniser.
- Les normes générales du domaine**
- Les questions juridiques**
 - Problèmes juridiques liés à la dématérialisation, l'authenticité et l'intégrité des documents.
- Les systèmes d'archivage**
 - Aspects généraux et architectures.
 - Les fonctions principales : ingestion, stockage, gestion des informations, accès, administration.
 - Sécurité d'accès aux archives.
 - Migrations.
 - Interopérabilité entre les sites distributeurs d'information.
- Les problèmes d'organisation sous-jacents**
 - Les acteurs et leurs rôles.
 - Rôle respectif des professionnels de l'information et des ingénieurs et informaticiens.
 - Coût de l'archivage numérique.

Dernières actualités

Retrouvez toutes les actualités

Formation "Pérennisation et communication de l'information numérique" du 12 au 16 novembre 2012

Comment pérenniser des informations numériques avec des moyens matériels et logiciels dépourvus de toute pérennité ? C'est à ce problème majeur pour la sphère de l'information que cette formation propose d'apporter des solutions concrètes. Cette formation a été conçue par un groupe d'experts, professionnels en charge de la pérennisation d'information numérique dans des contextes variés (archives et bibliothèque nationale, organismes scientifiques, entreprises...). Au sein de l'association Aristote, ces professionnels partagent depuis 10 ans leurs expériences et leur savoir-faire dans le groupe Pérennisation des Informations Numériques.

Par cette formation, le groupe PIN souhaite aujourd'hui proposer à tous les professionnels confrontés à la question de la pérennisation des données numériques : informaticiens, documentalistes, gestionnaires ou administratifs, etc., des solutions fondées sur des expériences concrètes.

Le programme complet et définitif de la session sera très prochainement publié.

L'édition 2012 de la formation est organisée à l'Ecole Nationale des Chartes, 19 rue de la Sorbonne, Paris 5ème - 75005 PARIS

Modalités d'inscription Retourner le [formulaire](#) rempli et signé en tenant compte des dates limites figurant sur la fiche de la session, à l'adresse suivante : Mme Tétard Association Aristote Ecole Polytechnique 91128 PALAISEAU CEDEX

Il est demandé également d'envoyer une préinscription à marie.tetard@polytechnique.edu; un mail contenant vos coordonnées détaillées. Le nombre de places étant limité, les priorités seront basées sur l'ordre de réception de ces mails de pré-inscription.

[Formation "Pérennisation et communication de l'information numérique" du 12 au 16 novembre 2012](#) - [09/11/2012 19:02](#) - [0](#) Commentaires

[Anciens billets >>](#)

2.3 Groupe Gus'G

Présentation et objectifs

Ce groupe a pris en 2004 la succession du groupe CSD « Calcul Scientifique Distribué » avec des objectifs plus « opérationnels ». En effet, depuis quelques années, les grands projets scientifiques et industriels impliquent la **mutualisation et la virtualisation de ressources** dans le cadre d'**applications distribuées** : c'est le concept de grille. Les utilisateurs de ces grilles ont des problématiques souvent liées au calcul haute performance notamment dans l'utilisation des nouvelles architectures. Mais les problématiques ont évolué. Le spectre des problèmes abordés actuellement couvre le *Cloud computing*, les **infrastructures de service** et *via* des activités dédiées le calcul hybride et outils de **travail collaboratif**.

Animateurs :

- Stéphane Requena GENCI ;
- Jean-Michel Batto INRA-Jouy ;
- Philippe d'Anfray CEA.

Les secteurs d'activités concernés par le Gus'G couvrent les industries -aéronautique, automobile, nucléaire-, la bio-informatique, la pharmacie, le secteur bancaire, l'ingénierie du logiciel, les « communautés scientifiques » qui déploient des plates-formes type « grille » opérationnelles (physique, climat, astronomie, ...) mais aussi d'autres acteurs (*e-learning*, arts, ...) qui utilisent des infrastructures collaboratives.

Activités

Les activités de ce groupe de travail privilégient :

- la veille technologique ;
- les retours d'expériences ;
- les transferts de technologies.

Il s'agit de regrouper des énergies pour mieux comprendre, utiliser et intégrer dans les pratiques quotidiennes les évolutions technologiques.

Parmi les actions (possibles) de ce groupe de travail :

- réunions d'échange sur des problématiques liées aux **Cloud Computing** ;
- organisation de séminaires, de colloques ou d'écoles thématiques ;
- enquêtes ; diffusion de documents ; rédaction d'articles de synthèse voire de « livres blancs ».

Cette année, le groupe s'est impliqué dans deux séminaires liés aux *Cloud Computing*, beaucoup de travail a été effectué dans les sous-groupes thématiques **Access Grid** et Calcul Hybride décrits plus loin. Le groupe a participé à l'Open World Forum qui a eu lieu du 22 au 24 septembre 2011 à Paris.

Moyens

Le groupe dispose d'une liste de diffusion et d'un site de travail collaboratif.

2.4 Groupe Access Grid

Un groupe créé en 2008

Créé à l'issue du séminaire de juin 2008 sur les outils de travail collaboratif, ce groupe d'utilisateurs français d' **Access Grid** rassemble l'INRA, le CEA, l'IDRIS, l'École Polytechnique, l'IRCAM, le Centre Pompidou, *etc.* Animateurs :

- Jean-Michel Batto INRA-Jouy ;

- Marcolino Pires Rectorat de Paris ;
- Philippe d'Anfray CEA.



Access Grid est un système avancé de travail collaboratif, incluant des éléments de visioconférence et des applications partagées, où les participants se rencontrent dans des lieux de rendez-vous virtuels. Conçus dès la fin des années 90 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 gratuitement, sont maintenus et «supportés» sur toutes les plates-formes (Linux, Windows® et MacOS X®) à l'intérieur de projets bien identifiés.

Les chercheurs et les ingénieurs «connectés» dialoguent et interagissent, ils partagent des instruments scientifiques, conduisent des expériences, manipulent et visualisent des données à travers internet.

Access Grid ne vise pas à supprimer les déplacements mais permet la tenue de réunions régulières assurant un meilleur contact et une meilleure implication des acteurs d'un projet qui tous peuvent faire le «déplacement virtuel». Il suffit d'un PC avec *webcam*, micro et hauts-parleurs pour installer les composants de base du système et l'utiliser.

Access Grid est particulièrement adapté à la collaboration entre groupes de petite ou de grande taille répartis sur un grand nombre de sites : **réunions** informelles, impromptues ou plénières, planifiées ; **revues de projets** ou de programmes ; **formations** et **cours** ; **conférences** et *workshops* ; sans oublier les **séminaires Aristote !**.

Objectifs

Parmi les objectifs du groupe **Access Grid** :

- assurer la promotion du système, travailler à l'interopérabilité avec d'autres logiciels ;
- échanger sur le «savoir faire», les matériels, les logiciels et l'aménagement des salles ;
- monter facilement des sessions «tests» pour valider nos installations ;
- participer au programme «Assurance Qualité» ;
- pour les partenaires «universitaires» participer à certains développements en liaison avec Argonne (projets, stages, ...).

Moyens

Le groupe dispose d'un site de travail collaboratif et a développé de nombreux contacts avec la communauté **Access Grid** (Angleterre, USA, Australie). En 2011 près de 350 sites à travers le monde possèdent des installations spécifiques pour utiliser **Access Grid**. Le groupe dispose d'un serveur et de matériel pour assurer les retransmissions de séminaires.

Réalisations

Suite à l'installation en 2009, avec le soutien de la DSI de l'École Polytechnique, d'un serveur de salles de réunions virtuelles bénéficiant de la connectivité multicast et d'une passerelle unicast-multicast, le groupe offre maintenant un service à la communauté des utilisateurs d'**Access Grid** en France.

La technologie est utilisée avec succès pour retransmettre les séminaires de l'association. Le guide d'utilisation a été remis à jour.

Le groupe a aussi installé un système permettant de recevoir facilement les séminaires Aristote *via* un greffon VLC dans un navigateur.

Un manuel d'installation et d'utilisation est disponible sur le site de l'association.

2.5 Groupe Calcul Hybride

Un groupe créé en 2008

Créé avec GENCI à l'occasion du séminaire d'octobre 2008 sur les outils de travail collaboratif, ce groupe de travail est ouvert à tous, **constructeurs et utilisateurs, chercheurs et ingénieurs**. Animateurs :

- Stéphane Requena GENCI.

Activités

Les accélérateurs de calcul suscitent un fort intérêt en France où de nombreuses expérimentations sont menées à la fois dans le milieu académique et le milieu industriel.

Les membres du groupe proposent donc de fédérer ces initiatives en un club d'utilisateurs afin notamment de mener les actions suivantes :

- la promotion du calcul hybride ;
- la veille technologique autour des architectures logicielles et matérielles ;
- la mise en place d'un benchmark « accélérateurs » ;
- la promotion de solutions innovantes (ex : suite HMPP CAPS Entreprise) ;
- une action de *lobbying* auprès des fournisseurs de solutions visant à définir au plus vite des modèles de programmation qui deviennent des standards (au sens OpenMP du terme) ;
- une mise à disposition de ressources auprès de la communauté (site WEB, voire si possible matériels et logiciels, . . .) ;
- une action auprès des écoles afin de voir comment proposer aux étudiants des cursus autour du calcul hybride.

Le groupe dispose d'un site de travail collaboratif. Le groupe a organisé régulièrement des journées en liaison avec le projet OpenGPU.

2.6 Groupe e-Lab (travail collaboratif)

Un groupe créé fin 2010

Le « travail collaboratif » est un mode de fonctionnement collectif intimement lié aux technologies de l'information et de la communication.

Internet permet la collaboration (de plus en plus) en temps réel : courrier électronique, web interactif, systèmes de gestion de documents et de connaissances, forges logicielles, éléments de visio-conférence, partage d'équipements et de données, réseaux sociaux, . . . Tout cela concourt à la construction des espaces de travail virtuel.

Dans la recherche, les différentes formes de collaboration nécessitent de gérer des échelles de temps très différentes avec des communautés très diverses qui ne se recouvrent pas ou partiellement.

Parmi les points-clefs :

- l'infrastructure de service (c.f. les grilles) qui permet la mutualisation, la fédération et le partage de ressources (intergiciels) et la gestion de la gouvernance (organisations virtuelles) ;
- le portail, au sens portail d'entreprise est un point d'entrée unique à un ensemble de services et de ressources configurable par l'administrateur et personnalisable par l'utilisateur avec bien sûr une authentification unique.

Dans ce contexte, les outils, in fine clients « locaux », profitent de l'infrastructure et sont accessibles à travers le portail.

- la salle virtuelle est un lieu pérenne où les utilisateurs se « retrouvent » pour se voir et dialoguer, exécuter des applications partagées, déposer des “objets”, ... la connexion peut être permanente (télé présence).

Le lieu « concret » où les utilisateurs se « retrouvent » est aménagé de façon spécifique. Ses équipements (écrans, audio, vidéo) reflètent le lieu virtuel et permettent de travailler ensemble sur des applications ou sur des instruments scientifiques avec des interactions « proches du présentiel ».

Activités

Les membres se sont réunis plusieurs fois avsc dans l'idée de produire un livre blanc sur les besoins d'outils et d'infrastructures collaboratifs dans le domaine de la recherche. Les idées du groupe ont été présentées à l'occasion de conférences (Ateliers de l'INRI, Journées JCUIC du CEA, Web3D conférence). Un séminaire Aristote a été organisé sur le thème « Travail collaboratif et Cloud ».



Chapitre 3

Séminaires, conférences et formations

3.1 Le cycle 2011

Les 5 séminaires du cycle 2011 ont eu lieu à l'École Polytechnique :

- « L'interopérabilité dans le Cloud » ;
- « Journée Calcul Hybride, le projet OpenGPU : un an plus tard » ;
- « Séminaire Big Data : le déluge de données, comment en tirer parti » ;
- « Travail collaboratif et Cloud » ;
- « Le Web2.0 : de la théorie à la pratique ».



3.2 L'interopérabilité dans le *Cloud*

Organisation : Aristote, Mercredi 23 mars 2011, École Polytechnique, Palaiseau.

Coordination scientifique :

- Roland Sénéor (Ecole Polytechnique, réseau Saphir) ;
- Philippe d'Anfray (CEA).

Comment s'y retrouver dans les initiatives et les standards émergents du *Cloud*. Quelles sont les garanties (résilience, sécurité) offertes par ces environnements. Quels sont les grands projets en cours, les problématiques des exploitants, et quel est l'avis —le retour d'expérience— des premiers utilisateurs à grande échelle.



Technologies des réseaux et NTICS

L'interopérabilité dans le *Cloud*

Mercredi 23 mars 2011

Coordination scientifique :
– Roland Sénéor (Ecole Polytechnique, réseau Saphir)
– Philippe d'Anfray (CEA)

Amphithéâtre Becquerel, École Polytechnique, Palaiseau

<http://www.association-aristote.fr> info@association-aristote.fr

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI/CEN Saclay Bât. 474, 91191 Gif-sur-Yvette Cedex.
Secrétariat : Aristote, École Polytechnique, 91128 Palaiseau Cedex.
Tél : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 61 Courriel : contact@aristote.polytechnique.edu
Site internet <http://www.association-aristote.fr>

Programme du 23 mars 2011

8h45-9h20	<i>Accueil des participants</i>	
9h20-10h05	Jean-Hugues Lauret (EISTI)	Ouverture du séminaire Etat de l'art de l'offre <i>Cloud</i> en France
10h05-10h40	Chuyen Huynh (Cisco)	Au-delà des standards Le <i>Cloud</i> peut-il être ouvert ou bleu ?
10h40-11h15	Jamil Chawki (Orange Labs)	La normalisation de l'informatique dans le nuage <i>Cloud Computing</i> , état et perspective
11h15-11h35	<i>Pause café</i>	
11h35-12h10	Raphaël Ferreira (eNovance)	Projets, environnements Compatible One : <i>Open source Cloudware Initiative</i>
12h10-12h45	Christian Delbe (ActiveEon)) Denis Caromel (INRIA)	OW2 <i>ProActive Parallel Suite</i> et Interopérabilité
12h45-14h00	<i>Repas (salle Detoef)</i>	
14h-14h20	Stephan Hilby (Intel)	Projets, environnements, retours d'expériences Vision, technologies & initiatives d'Intel pour le <i>Cloud Computing</i>
14h20-14h55	Ahcène Latrèche (Bull)	Le <i>Cloud</i> au service de l'innovation
14h55-15h30	Pierre Beyssac (fondateur d'Eriomem, cofondateur de Gandi)	Le nuage va-t-il s'arrêter à la frontière ?
15h30-15h45	<i>Pause</i>	
15h45-16h20	Jean Pierre Choulet (ESSEC)	Retour d'Expérience du <i>Cloud</i> à l'ESSEC : <i>beyond technology</i>
16h20-16h55	Eric Spitéri (CS-SI)	Interopérabilité dans des systèmes d'information logistique
16h55-17h15	Conclusions, questions-réponses avec les intervenants	

Résumés

Chuyen Huynh (Cisco)

Le Cloud peut-il être ouvert ou bleu ? Le *Cloud* peut-il être ouvert ou bleu ? derrière ce jeu de mots, nous aborderons candidement deux approches qui s'affrontent déjà ouvertement :

1. L'ouverture (architecture, interface, code) qui est l'essence même du *Cloud* mais amène le risque de l'éparpillement.
2. La couleur unique (bleue au hasard, pourrions-nous dire) qui est une transformation du *datacenter* avec l'excuse de l'efficacité et la rapidité de mise en œuvre.

Nous présenterons ainsi OpenStack, d'Open Cloud Forum, Open Cloud Consortium, *etc.*

Jamil Chawki (Orange Labs)

La normalisation de l'informatique dans le nuage *Cloud Computing*, état et perspective L'informatique en nuage *cloud computing* est une nouvelle façon de fournir en mode partagé des ressources informatiques depuis le réseau Internet. C'est une évolution des modes de livraison de Service d'Application *ASP* et la fourniture de logiciel comme un Service *Software as a Service SaaS*. La présentation portera sur :

- la définition de l'informatique en nuage proposée par les organismes NIST et l'ITU-T ;
- l'identification des besoins pour la normalisation du *Cloud* ;
- l'état des activités d'une quinzaine d'organismes de normalisation (ISO, NIST, ITU-T, DMTF, CSA, ...) et leur positionnement par rapport à l'écosystème *Cloud*.

Nous aborderons enfin les perspectives des nouvelles normes de l'informatique en nuage.

Raphaël Ferreira (eNovance)

Compatible One : *Open source Cloudware Initiative* Compatible One is one of the major R&D projects about Cloud Computing in Europe. *Compatible One* objective is to provide an open "cloudware" allowing to create, deploy and manage private, public or hybrid cloud platforms. It addresses the three layers of services (IaaS, PaaS, SaaS) with a global and unified solution. The key points of the project are :

- OPEN : Source code, interfaces, and documentations ;
- COMPATIBLE : with most platforms to provide maximum freedom to users and developers ;
- SCALABLE : offering maximum abstraction of resources and facilitating the creation of applications and services.

Christian Delbe (ActiveEon) & Denis Caromel (INRIA)

OW2 ProActive Parallel Suite et Interopérabilité ProActive Parallel Suite is an innovative Open Source solution (OW2) for private and hybrid Cloud computing. ProActive features Java parallel programming seamlessly integrated with scheduling and resource management. ProActive simplifies the programming and execution of parallel applications and workflows, together with the optimization of infrastructure resources, allowing to create a single portal for diverse hardware and applications.

With respect to interoperability, ProActive runs on Linux, Windows and Mac and features management of resources such as desktop, servers, clusters, enterprise Grids and Clouds. Dynamically upon demand, it can burst onto public Clouds (e.g. Amazon, EC2). It allows to deploy applications with various protocols (ssh, rsh, sshGSI), and to manage most VMs (VMware, KVM, Xen, Xen Server, QMU, Microsoft Hyper-V). It also features interoperability with most native schedulers (PBS, LSF, SGE, Prun,

EGEE gLite, IBM LL), and supports CPUs, GPUs, and remote visualization. Finally, all types of applications are supported : native applications (including MPI and OpenMP), Java, Scripts (Ruby, Python, JavaScript), Matlab and Scilab. The presentation will feature a live demo on a 1 200 cores production Cloud. <http://ProActive.inria.fr>

Stephan Hilby (Intel)

Vision, technologies & initiatives d'Intel pour le *Cloud Computing* Intel vous propose de partager sa vision pour le *Cloud Computing*, soulignant les enjeux inhérents tant aux serveurs qu'aux multiples périphériques qui s'y connecteront. Intel vous exposera également quelques initiatives vous permettant de mieux appréhender ce nouveau paradigme et d'anticiper les futurs impacts sur les architectures de vos *datacenters*.

Pierre Beyssac (fondateur d'Eriomem & cofondateur de Gandi)

Le nuage va-t-il s'arrêter à la frontière ?

- qu'est ce que le *Cloud* ;
- vrais et faux problèmes ;
- infrastructure, plateforme, standards ouverts ;
- contrats : conditions générales de vente, SLA ;
- le point de vue du fournisseur ;
- particularités françaises ;
- évolutions prévisibles.

Jean Pierre Choulet (ESSEC)

Retour d'Expérience du *Cloud* à l'ESSEC : *beyond technology* Au delà des projets et des recherches technologiques pour construire un écosystème *Cloud* européen, l'innovation est d'ores et déjà disponible pour le grand public et pour les entreprises. À travers son axe stratégique *Let Outside Innovation Be Inside* l'ESSEC a fait le choix de s'appuyer sur l'écosystème Google Apps for Education pour déployer son offre de services en ligne pour supporter les processus pédagogique et de gestion. La présentation vise à faire le point sur cette démarche, ses enjeux et ses retour sur investissement.

Eric Spitéri (CS-SI)

Interopérabilité dans des systèmes d'information logistique Les systèmes d'information logistique (SIL) sont des outils qui amènent des gains de productivité et de disponibilité pour les matériels suivis. De tels système sont mis en œuvre dans les organisations étatiques. Ces systèmes ont besoin d'évolutivité, de résilience, d'ouverture pour prendre les évolutions des organisations, des métiers et des matériels. Après une présentation des moyens de définition (fonctionnel et architecture) d'un SIL, les aspects relatifs à la problématique de l'interopérabilité sont abordés, d'abord d'un point de vue générique, et ensuite au travers de deux projets. Ses systèmes d'information étatiques intègrent, aussi, des aspects qui tiennent du *Cloud*.

3.3 Journée Calcul Hybride, OpenGPU un an plus tard

Organisation : Aristote, Mercredi 8 juin 2011, École Polytechnique, Palaiseau.

Coordination scientifique :

- Aristote et le projet OpenGPU.

Un an après son lancement au mois de mars 2010, les partenaires du projet OpenGPU organisent un colloque national pour partager avec la communauté scientifique les premiers résultats quantitatifs et opérationnels obtenus par les différentes équipes de recherche.

Cette journée, organisée en partenariat avec l'association Aristote, sera suivie d'une journée consacrée aux nouvelles tendances du *Big Data* pour lesquelles les architectures hybrides GPU/CPU joueront un rôle essentiel.

Les partenaires de l'événement





Technologies des réseaux et NTICS

Journée Calcul Hybride

« le projet  : un an plus tard »

Mercredi 8 juin 2011

Coordination scientifique :
 – Association Aristote et le projet OpenGPU
 Amphithéâtre Gay-Lussac, École Polytechnique, Palaiseau
<http://www.association-aristote.fr> info@association-aristote.fr

Membre de l'association en CC BY-NC-SA (only 27 Jan 2011) ©2011 Aristote

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI/SEN Sclay BB. 474, 91191 Gif-sur-Yvette Cedex.
 Secrétariat : Aristote, Ecole Polytechnique, 91128 Palaiseau Cedex.
 Tél : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 65 Courriel : secretariat@polytechnique.edu
 Site internet <http://www.association-aristote.fr>

Programme du 8 juin 2011

8h30-9h15	<i>Accueil des participants, café</i>	
9h15-9h40	Ouverture du séminaire Jean-Noël de Galzain (Wallix) Eric Mahé (Minds Planet)	OpenGPU : un projet au cœur des évolutions technologiques
9h40-10h45	Les outils d'aide à la parallélisation Mathias Bourgoïn (LIP6) Michel Barreteau (Thales Research & Technology) Tarik Saidani (AS+)	Le langage CAML et la programmation des GPU De l'OpenCL performant par couplage d'outils de parallélisation Retours d'utilisation du langage OpenCL
10h45-11h15	<i>Pause café</i>	
11h15-12h10	Philippe Coucaud (Ateji) Guillaume BARAT (CAPS) Basile Starynkevitch (CEA)	Java sur GPU avec Ateji PX Migration <i>manycore</i> : méthodologie et outils OpenGPU et GCC-MELT
12h10-13h45	<i>Repas (salle Detoef)</i>	
13h45-16h00	Applications opérationnelles des architectures hybrides Vivien Clauzon (Numtech) Jean-Pierre Panziera (Bull) Rémi Barrère (Thales Research & Technology) Pierre Leca (CEA - DAM) Antoine Petitet et Stefanos Vlachoutsis (ESI-Group) Fouad Boumezbeur (INRA)	Retours d'expérience d'une PME sur le GPGPU pour des applications de dispersion atmosphérique Vision de Bull sur les architectures hybrides. Évaluation d'OpenCL et intégration des enseignements pour une programmation performante à haut niveau Le prototype de <i>cluster</i> OpenGPU au CEA-DAM Contexte et retour d'expérience Premiers retours d'expérience sur l'utilisation de GPU pour des applications de mécanique des structures Applications du GPU à l'investigation métagénomique : illustration avec le projet européen MetaHIT
16h00-16h20	<i>Pause</i>	
16h20-17h25	Parallélisation : la place de l'open source Ronan Keryell (HPC Project) Sylvestre Ledru (Scilab) Albert Cohen (INRIA)	L'outil de parallélisation automatique Par4All Retour sur l'intégration de fonctionnalités GPGPU dans Scilab Parallélisation automatique de noyaux de calcul pour GPGPU
17h25-18h15	Conférences invitées Denis Caromel (INRIA/ActiveEON) Yann Le Du (ENSCP)	Orchestration de <i>Workflows</i> CPU et GPU avec OW2 ProActive Parallel Suite HPU4Science, calcul haute performance sur CPU/GPU avec du matériel grand public : conception et choix logiciels
18h15-19h30	<i>Cocktail (Salon de Marbre)</i>	

3.4 Le déluge de données : comment en tirer parti

Organisation : Aristote, Jeudi 9 juin 2011, École Polytechnique, Palaiseau.

Coordination scientifique :

- David Menga (EDF R&D) ;
- Jean-Michel Batto (INRA) ;
- Pierre Léonard (INRA).

De nos jours, les données sont partout et constituent la matière première de notre monde numérique. Elles redéfinissent la façon dont on crée de la connaissance scientifique et offrent aux entreprises de nouveaux leviers de croissance et plus de performances opérationnelles aux gouvernements.

Décoder le génome signifie analyser 3 milliards de paires de base. Walmart traite chaque heure 1 million de transactions clients avec une volumétrie totale de 2,5 Petaoctets. Nous vivons « une révolution industrielle » des données et nos outils actuels de stockage et de traitement des données sont inadaptés pour traiter de tels volumes en un temps acceptable.

Ce séminaire a pour objet de donner les clefs de compréhension de cet univers, de mettre en perspective les défis scientifiques et opérationnels et d'offrir des éléments de réponse à travers des travaux scientifiques de pointe et des solutions marché innovantes.

Le matin sera consacré à la définition de la problématique et donnera des réponses. L'après midi, nous aborderons les expériences des acteurs confrontés à ce défi, comme les électriciens, les biologistes et les astrophysiciens.



Technologies des réseaux et NTICS

**Le déluge de données
comment en tirer parti**

Jeudi 9 juin 2011

Coordination scientifique :
– *David Menga (EDF R&D)*
– *Jean-Michel Batto (INRA)*
– *Pierre Léonard (INRA)*

Amphithéâtre Becquerel, École Polytechnique, Palaiseau

<http://www.association-aristote.fr> info@association-aristote.fr

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI CEN Saclay Bât. 474, 91191 Gif-sur-Yvette Cedex.
Secrétariat : Aristote, École Polytechnique, 91128 Palaiseau Cedex.
Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 67 Courriel : Sec.CE - Sec.ar@dpolytechnique.edu
Site internet <http://www.association-aristote.fr>

Programme du 9 juin 2011

9h00-9h30	<i>Accueil café</i>	
9h30-9h45	Pierre Léonard (INRA) Jean Michel Batto (INRA) David Menga (EDF R&D)	Présentation du séminaire
9h45-10h30	Xavier Dalloz (Consultant)	<i>Big data</i> , le fuel de l'économie du XXI ^e siècle
10h30-11h00	<i>Pause café</i>	
11h00-11h45	Vincent Heuschling (Affini-Tech)	Cartographie des solutions <i>big data</i> du marché
11h45-12h30	Alexandru Costan (INRIA)	Analyse des systèmes de stockage à grande échelle pour les applications de traitement intensif des données
12h30-14h00	<i>Repas (salle Detoef)</i>	
14h00-14h35	Joerg Bienert (ParStream)	An innovation solution to manage heterogeneous big data
14h35-15h10	Ronan Keryell (HPC-Project)	Environnement de programmation pour traitements massifs sur architectures modernes
15h10-15h20	Denis Caromel (ActiveEon-INRIA)	Solutions ProActive pour <i>Workflows</i> , <i>Map/Reduce</i> , <i>Matlab</i> / <i>Scilab</i> , CPU/GPU
15h20-15h45	Nicolas Pons (INRA)	La métagénomique, un défi supplémentaire pour la loi de Moore
15h45-16h00	<i>Pause</i>	
16h00-16h35	Patrick Fuhmann (DESY-Hamburg)	dCache : scaling out affordable storage.
16h35-17h10	Marie-Luce Picard (EDF R&D et ENST-Bilab)	Données massives pour les <i>smart-grids</i>
17h10-17h40	David Konopnicki (IBM-Haïfa Research Labs)	Massive Scale Analytics for a Smarter Planet
17h45-18h00	Conclusions, questions-réponses avec les intervenants	

Résumés des présentations

Xavier DALLOZ (Consultant)

Big data, le fuel de l'économie du XXI^e siècle Régulièrement une innovation majeure change tout avec à chaque fois de nouvelles technologies, de nouveaux métiers et de nouveaux enjeux. Après l'ère des *mainframes*, l'ère des *mini computers*, l'ère des PC, l'ère du *software* pour améliorer la productivité personnelle, l'ère de l'Internet, voici celle du « *Big Data* ».

Le *Big Data* va permettre notamment de repenser les modèles économiques en misant sur de nouvelles créations de valeur avec notamment l'intelligence collaborative et la *shazamisation* de notre environnement de telle sorte qu'il y ait une meilleure efficacité de « notre » capital : santé, énergie, éducation, équipements, stocks... Avec le *Big Data*, la chasse aux gaspillages va enfin devenir une réalité. Nous n'avons encore rien vu... Tout va s'accélérer.

Que d'opportunités ! Que de leviers de croissance pour nos économies !

Vincent HEUSCHLING (Affini-Tech)

Cartographie des solutions big data du marché Face au déluge de donnée, que nous vivons aujourd'hui, quelles sont les réponses des grands acteurs du marché ? Quels défis présente cette explosion du volume de données pour les infrastructures. Décryptage des offres des grands constructeurs que sont EMC, IBM, Oracle, ..., et des architectures innovantes du monde open-source.

Alexandru Costan (INRIA)

Analyse des systèmes de stockage à grande échelle pour les applications de traitement intensif des données Avec l'augmentation rapide des volumes de données dans de nombreux domaines d'application de la science de l'ingénierie et des services d'information, les défis posés par les traitements intensifs des données présentent une importance croissante. Avec l'émergence des infrastructures récentes (plates-formes de type *Cloud*, architectures massivement parallèles pétaflopiques) réaliser une gestion des données capable de passer à l'échelle dévient un enjeu crucial car les performances globales des applications dépendent des propriétés du service de gestion des données.

Nous définissons un ensemble de principes pour la conception de systèmes de stockage distribués, optimisés pour pouvoir passer à large échelle, tout en permettant autant de manipulations concurrentes des données que possible. Combinés, ces principes peuvent aider les développeurs de systèmes de stockage distribués à répondre aux exigences strictes de gestion de données à grande échelle.

Nous analysons ensuite plusieurs systèmes de stockage représentatifs afin d'évaluer la façon dont ils se conforment à ces principes de conception. Nous nous concentrons sur les systèmes de fichiers spécialisés qui ont été introduits pour cibler spécifiquement les besoins des applications de traitement intensif des données : HDFS, la couche de stockage standard utilisé par Hadoop MapReduce, GPFS proposé par IBM, ainsi que les systèmes de fichiers distribués massivement parallèles, comme Lustre ou PVFS, généralement utilisé pour les *clusters* de calcul à grande échelle. Avec l'émergence du calcul de type *Cloud*, des solutions de stockage spécialement conçues pour s'adapter à ce contexte ont été développées : nous présenterons textttAmazon S3.

Nous détaillons en particulier les avantages potentiellement importants du versionnage pour améliorer les performances d'accès hautement concurrents aux données des applications. Dans ce contexte, nous proposons une interface d'accès basé sur la gestion de versions des données, matérialisée au sein de la plate-forme BlobSeer développée par l'équipe KerData de l'INRIA à Rennes. Cette approche qui permet d'exploiter d'une manière efficace le parallélisme inhérent des flux de données : nous en illustrons utilisation avec une application d'analyse conjointe de données génétiques et de neuro-imagerie.

Joerg Bienert (ParStream)

An innovation solution to manage heterogeneous big data Analyzing Big Data is getting more and more important for companies in all industries, *e.g.* web analytics, fraud detection, smart metering, telco, *etc.* Current established database products are not able and not designed to perform large scale data analytics. New approaches like Map/Reduce lack important features like short response times. ParStream is a novel innovative database technology focusing on processing large data sets (billions of records) in milliseconds and with low latency. Parstream is a columnar in memory database running on multiprocessor architectures and, as first product, on GPU based HPC-Servers.

Ronan Keryell (HPC-Project)

Environnement de programmation pour traitements massifs sur architectures modernes L'ubiquité de l'informatique déclenche une avalanche de données à traiter de manière rapide et économique. Malheureusement, pour des raisons physiques, la loi de Moore ne fournit plus de processeurs plus rapides (dissipation thermique, vitesse de la lumière, ...) mais fournit néanmoins toujours plus de transistors. Le seul moyen d'utiliser ces transistors est d'utiliser du parallélisme massif, mais cela remet en question les architectures et les modes de programmation. Les architectures actuelles (GPU, MP-SoC, Tilera, FPGA, ...) seront présentées avec leurs avantages et leurs contraintes ainsi que leurs pendant logiciels permettant de les utiliser au mieux. L'environnement de compilation source à source Par4All de HPC Project est un moyen de s'abstraire de certains détails de programmation.

Denis Caromel (ActiveEon-INRIA)

Solutions ProActive pour Workflows, Map/Reduce, Matlab/Scilab, CPU/GPU ProActive Parallel Suite (<http://ProActive.inria.fr>), un projet *Open Source* d'OW2, offre une solution flexible pour regrouper des ressources de calcul et offrir aux entreprises un accès simple et unifié à ces ressources par le biais de Portails et d'API. ProActive optimise l'exécution des applications les plus exigeantes, les *workflows* d'entreprises, les simulations numériques et financières, l'analyse des données (avec un *Map/Reduce* qui supporte les APIs Hadoop sans nécessiter un *cluster* dédié). Un mécanisme de sélection de ressources permet de combiner les exécutions sur CPU et GPUs, sur des tâches simples ou au sein même de *workflows* dynamiques. Des analyses de données parallèles sur *Cloud* peuvent être initiées directement sans quitter les environnements Matlab et Scilab. Des *benchmarks* dans les *biotechs* et des démonstrations interactives de *Map/Reduce* seront présentées sur une plate-forme en production.

Nicolas Pons (INRA)

La métagénomique, un défi supplémentaire pour la loi de Moore Depuis le séquençage du premier génome en 1995, la production de données de séquençage d'ADN a révolutionné les possibilités de compréhension du vivant par la biologie moléculaire. Avec l'arrivée des technologies de séquençage à très haut-débit, on assiste aujourd'hui à une explosion des volumes de données avec un doublement des bases de données de séquence tous les 6 mois et une augmentation du débit d'acquisition d'un facteur 1000. Ce déluge de données ouvre de nouvelles perspectives scientifiques notamment dans le domaine de la métagénomique qui vise à caractériser l'ensemble des génomes bactériens d'un écosystème complexe : il est désormais possible de quantifier les génomes, gènes et fonctions de ces écosystèmes. Le traitement de ces *big data* constitue un défi majeur tant en matière d'optimisation des calculs qu'en matière de stockage et de leur mise à disposition aux biologistes. Nous illustrerons ces défis à travers l'exemple des projets MetaHIT et MicroObes qui proposent d'étudier le génome de l'ensemble des bactéries constituant la flore intestinale humaine afin de caractériser ses fonctions et ses implications sur la santé.

Patrick Furhmann (DESY-Hamburg)

dCache : scaling out affordable storage. The presentation will briefly walk through the various facets of the dCache storage technology and its supporting collaborations. Functional objectives will be discussed, as well as some bits and pieces of the technical implementation. The presenter will touch upon the results of an ongoing detailed evaluation of supported file access protocols at the DESY Grid-Lab facility including a discussion on application level behaviour and pitfalls. Finally the most prominent dCache deployments will be presented and the involvement of dCache in other projects.

Marie-Luce Picard (EDF R&D et ENST-Bilab)

Données massives pour les *smart-grids* De nombreux projets *smart-grids* voient le jour à travers le monde, motivés par des aspects réglementaires, des contraintes économiques ou la prise en compte de besoins environnementaux ou sociaux. Ces projets reposent sur le déploiement de compteurs communicants et la mise en place d'une infrastructure de communication adéquate. Mais il ne s'agit là que de la première étape de la mutation technique et économique du secteur énergétique. Cette vision long terme de la problématique des réseaux intelligents sous-tend une capacité à gérer et traiter de larges volumes de données, provenant en particulier des compteurs intelligents ou encore de différents capteurs sur le réseau. Dans cette perspective, un certain nombre de travaux ont été menés à EDF R&D et seront présentés dans cet exposé, en particulier : le stockage de grandes quantités de séries temporelles, le traitement temps-réel de courbes de charge imparfaites, les perspectives d'évolution des approches de prévision de consommation en présence de données individuelles massives.

David KONOPNICKI (IBM Haifa-Research)

Massive-scale Analytics for a Smarter Planet Everyday, we create 2.5 quintillion bytes of data—so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere : from sensors used to gather climate information, posts to social media sites, digital pictures and videos posted online, transaction records of online purchases, and from cell phone GPS signals to name a few. This data is Big Data. Big Data is more than a challenge ; it is an opportunity to find insight in new and emerging types of data and to answer questions that, in the past, were beyond reach. Until now, there was no practical way to harvest this opportunity. Today, IBM's platform for Big Data opens the door to a world of possibilities, giving organizations a solution that is designed specifically with the needs of the enterprise in mind and provides the infrastructure of a Smarter Planet : intelligence is being infused into the systems and processes that make the world work—into things no one would recognize as computers : cars, appliances, roadways, power grids, clothes, even natural systems such as agriculture and waterways.

3.5 Travail collaboratif et *Cloud*

Organisation : Aristote, mardi 11 octobre 2011, École Polytechnique, Palaiseau.

Coordination scientifique :

- Drissa Houatra (Orange Labs) ;
- Philippe d'Anfray (CEA).

En séparant totalement les problématiques *hardware* et *software* l'informatique en nuage *cloud computing* a finalement rapproché les utilisateurs, qui, au sein de communautés, peuvent collaborer, en temps réel avec leurs outils métier en partageant l'information pertinente. La complémentarité des deux termes *travail collaboratif* et *cloud computing* est évidente, et le champ immense pour un seul séminaire.

Nous avons choisi d'aborder le matin, grâce aux interventions de l'INRIA, des problématiques pointues liées à l'infrastructure *cloud* et aux modèles de développement puis de nous projeter dans des applications et des réflexions avancées autour de la collaboration proprement dite, environnement virtuels 3D à Telecom SudParis, conception de services innovants en environnement collaboratif à l'Université de Technologie de Troyes, et enfin navigation dans les données avec SLIDERS-Lab.

L'après-midi permettra de mettre tout cela en perspective avec « ce qui se fait », l'état de l'art dans le domaine, XWiki logiciel de collaboration pour les entreprises ; une solution de portail scientifique collaboratif chez EDF avec la société SysFera et enfin un nouveau portail Intranet Collaboratif pour la R&D à l'IFPEN. Nous concluons par une réflexion sur les usages du temps réel et de la collaboration.



Technologies des réseaux et NTICs

Travail collaboratif et *Cloud*
Mardi 11 octobre 2011

Coordination scientifique :
– *Drissa Houatra (Orange Labs)*
– *Philippe d'Anfray (CEA)*

Amphithéâtre Gay-Lussac, École Polytechnique, Palaiseau

<http://www.association-aristote.fr> info@association-aristote.fr

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI CSN Sudp, Bld. 71, 91191 Gif-sur-Yvette Cedex.
Secrétariat : Aristote, Ecole Polytechnique, 91128 Palaiseau Cedex.
Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 87 Courriel : tar.ar@polytechnique.edu
Site internet : <http://www.association-aristote.fr>

Programme du 11 octobre 2011

9h00-9h20	<i>Accueil des participants</i>	
9h20-9h30	Drissa Houatra (Orange-labs), Philippe d'Anfray (CEA)	Ouverture du séminaire
9h30-10h15	Christian Perez (INRIA)	De l'évolution <i>cloud</i> des environnements d'exécution dans le contexte HPC
10h15-11h00	Roberto Cascella (INRIA)	Enabling a collaborative environment via the <i>Contrail</i> project
11h00-11h15	<i>Pause café</i>	
11h15-12h00	Patrick Horain (Telecom Sud Paris)	Perception et restitution des actions des utilisateurs dans un environnement 3D virtuel
12h00-12h40	Florie Bugeaud (Univ. de Technologie de Troyes)	Modélisation collaborative de systèmes de services. Cas de l'innovation de services dans les télécoms
12h40-13h00	Jean-Marie Dallet (SLIDERS-Lab. EESI)	Architectures de mémoire
13h00-14h15	<i>Repas (salle « aquarium »)</i>	
14h15-15h00	Ludovic Dubost (Xwiki))	XWiki, améliorer le partage d'informations avec un wiki applicatif
15h00-15h45	Benjamin Depardon (Sysfera) et Samuel Kortas (EDF)	SysFera-DS : vers une solution de portail scientifique collaboratif chez EDF
15h45-16h10	<i>Pause</i>	
16h10-16h55	Hervé Durand et Nathalie Triquet (IFPEN)	PRISME, le nouvel Intranet Web 2.0 d'IFP Energies nouvelles
16h55-17h15	Alexandre Eisenchteter (af83)	Temps réel et collaboration : des usages à inventer
17h15-17h30	Conclusions, questions-réponses avec les intervenants	

Résumés des présentations

Roberto Cascella (INRIA Rennes)

Enabling a collaborative environment via the Contrail project Cloud computing has changed the way the resources are shared and used. We envision that companies will rely on highly dynamic distributed IT infrastructures, where the infrastructure can be provided through a collaborative effort between existing resource providers. In this scenario, federation models enable a given organization to be both a Cloud provider during periods when its IT infrastructure is not used at its maximal capacity and a Cloud customer in periods of peak activity. In the framework of the Contrail project funded by the European Commission (<http://www.contrail-project.eu>), a consortium of eleven academic and industrial partners develops an integrated approach to virtualization, offering Infrastructure as a Service services (IaaS), services for federating IaaS clouds and Contrail Platform as a Service services (ConPaaS) on top of federated clouds. Contrail aims at developing this comprehensive cloud computing software stack in open source.

This talk presents an overview of Contrail project. Then it discusses the research challenges tackled in the project : the implementation of a federation of clouds providing Quality of Service (QoS) and Protection (QoP) under the control of SLAs, security, an efficient virtual platform management, and cloud storage.

Patrick Horain (Telecom SudParis)

Perception et restitution des actions des utilisateurs dans un environnement 3D virtuel Dans une collaboration à distance médiatisée par l'ordinateur, la communication peut-être renforcée par un contact visuel entre utilisateurs. La vidéoconférence ne permet pas restituer les actions des participants sur les objets partagés. Nous proposons de renforcer la perception mutuelle des participants et de leurs actions au moyen d'un environnement virtuel en réseau. Des représentations virtuelles des utilisateurs distants (avatars) et des objets qu'ils partagent y sont rassemblées de sorte que les actions des utilisateurs sur les objets sont utilisées pour animer les avatars qui miment ces actions. Les gestes de communication sont eux acquis par vision artificielle et restitués virtuellement en temps réel. Nous décrivons un système d'acquisition par vision artificielle des gestes et expressions du visage en 3D à partir d'une unique webcam. Celui-ci intègre un modèle statistique des gestes de communication qui fournit des contraintes a priori pour guider le suivi 3D des gestes, prendre en compte les ambiguïtés propres à la vision monoculaire et augmenter les données acquises par des détails de gestes appris. Nous avons développé GpuCV, une bibliothèque logicielle à code ouvert pour accélérer l'analyse d'image sur processeur graphique (GPU). Ce système à bas débit et bas coût permet d'envisager des applications grand public de téléprésence virtuelle en 3D.

Florie Bugeaud (Univ. Techno. Troyes)

Modélisation collaborative de systèmes de services. Cas de l'innovation de services dans les télécoms Nos travaux s'intéressent à la notion de "service" et aux verrous théoriques et opérationnels qui l'entourent. Le cas de la recherche d'opportunités en amont d'un processus de conception de services chez un opérateur de télécommunication est un exemple d'activité nécessitant une compréhension et une collaboration particulièrement fortes autour de ce concept. Mais le travail des acteurs de l'innovation est freiné par le caractère pluridisciplinaire de la communauté qu'ils forment et la complexité du concept de "service" qu'ils peinent à définir. Après avoir respecifié le "service" en tant que phénomène dynamique et émergent, nous avons défini un cadre théorique et conceptuel pour sa formalisation. Nous avons également élaboré une démarche de co-modélisation et simulation pour aider ces concepteurs à se projeter

ensemble dans une représentation commune des situations de services vécues par les clients et à identifier des solutions innovantes. Enfin, nous avons créé “iSamsara”, la maquette d’un environnement web collaboratif de modélisation et simulation de systèmes de services. La présentation aura pour objectifs de reprendre les points clés de ces travaux et d’en montrer l’apport pour la collaboration des acteurs de l’innovation.

Jean-Marie Dallet (SLIDERS-Lab. EESI)

Architectures de mémoire L’augmentation considérable des données en circulation à travers les réseaux allié à l’augmentation des capacités de stockage pose la question de l’accès à l’information et donc de sa visualisation. Il faut inventer les nouvelles grammaires gestuelles, les nouvelles architectures de mémoire et les nouvelles formes de navigation dans les bases de données. Elles doivent être plus sensibles et intuitives que verbales comme le propose aujourd’hui la plupart des moteurs de recherche.

Le SLIDERS_lab développe des stratégies de navigation sémantique dans des espaces interactifs en 3D. Les architectures inventées sont des structures « localisantes » qui permettent de circuler dans des « inventaires » d’images, de textes et de sons décrits dans des dictionnaires sous la forme de mots clés. Ces structures localisantes sont des formes symboliques d’appréhension du réel au même titre que pouvait l’être la perspective classique. Elles sont temporelles et nécessitent donc la création d’espaces pour ce développer. Créer sera donc l’invention d’espaces, la constitution de « lieux » qui rassemblent et dans lesquels on peut puiser. La construction de ces lieux se fait par l’élaboration de schémas à partir desquels il est possible de réaliser un nombre infini de groupements et de concordances au sein du matériau.

Cette spatialisation des relations nous entraîne dans la construction d’architecture 3D qui, sur le modèle de la partition de puisque de Guido d’Arezzo, nous permet de « spatialiser le temps ». Nos modèles sont parfois anciens comme la colonne Trajane et la tour de Babel, par exemple, mais ils peuvent être aussi des formes plus abstraites comme des tores, des diabolos, des cyclides. Ils présentent tous un point commun, celui de mettre au centre de la recherche le corps du spectateur. Ce qui nourrit notre imaginaire, c’est que l’on puisse concevoir une alternative à la consultation linéaire des médias (cf. Villem Flusser). C’est cette idée directrice qui nous a entraînée à imaginer un cinéma non linéaire, qui reposerait sur le concept de dé-linéarisation.

Une autre idée importante est la volonté de lier l’image et le texte sans séparation puisque dans beaucoup de nos réalisations ces deux entités sont indissolublement jointes (cf. Edward Tufte).

Ludovic Dubost (XWiki)

XWiki, améliorer le partage d’informations avec un Wiki Applicatif” On parle aujourd’hui beaucoup d’Enterprise 2.0 et de Réseaux Sociaux d’Entreprises, qui cherchent à placer l’utilisateur au coeur du logiciel collaboratif. Un Wiki Applicatif (Application Wiki) comme XWiki propose un approche complémentaire et cherchant à rassembler les utilisateurs autour des informations avec lesquelles ils travaillent afin de non seulement améliorer la façon dont l’information est diffusée mais surtout afin d’améliorer la participation autour de ces informations et leur organisation.

Benjamin Depardon (Sysfera), Jean-Marc Laymajoux et Samuel Kortas (EDF)

SysFera-DS : vers une solution de portail scientifique collaboratif chez EDF Dans cette présentation nous détaillerons notre solution nommée SysFera-DS permettant la gestion et l’utilisation d’infrastructures HPC distribuées et hétérogènes. Proposée en tant que portail scientifique collaboratif, elle fournit un accès unifié, simple, robuste et sécurisé aux ressources. Elle permet une collaboration fortes entre utilisateurs, applications et ressources de calcul. De multiples interfaces en facilite l’adoption (Ligne

de commande, Web services et différents langages) par différents types utilisateurs et l'intégration dans l'environnement logiciel client, notamment au sein de la plate-forme logicielle SALOME.

La présentation réalisée conjointement par EDF et SysFera présentera le déploiement de SysFera-DS dans l'environnement EDF R&D à travers le projet de collaboration Vishnu.

Hervé Durand et Nathalie Triquet (IFPEN)

PRISME, le nouvel Intranet Web 2.0 d' IFP Energies nouvelles IFP Energies nouvelles a souhaité améliorer l'accès et le partage de l'information par ses collaborateurs, faciliter la publication de contenu et favoriser la collaboration et l'émergence de communautés tant en interne qu'avec ses partenaires extérieurs.

Le projet PRISME (plate-forme pour la recherche, l'innovation, les services, la mutualisation et les échanges), lancé à l'initiative de la direction générale apporte un cadre moderne, accessible à l'ensemble du personnel, tout en maintenant un bon niveau de sécurité et de cohérence de l'ensemble. Il offre de nouvelles possibilités issues du Web 2.0 pour le travail collaboratif, ainsi que les fonctionnalités nécessaires au travail inter-directions.

Paradoxalement, l'appropriation de ces nouveaux outils par les acteurs de la R&D et leur intégration dans les activités du quotidien n'est pas une évidence et nécessite un effort et une organisation appropriée pour favoriser l'adoption des nouvelles technologies de l'information dans l'entreprise. Au-delà du projet fonctionnel et technique, c'est aussi ces préoccupations que nous essaierons de partager avec vous dans cette présentation.

Alexandre Eisenchteter (af83)

Temps réel et collaboration : des usages à inventer L'arrivée de technologies temps-réel pour le web permet d'envisager la création de nouveaux usages de collaboration. C'est ce que nous appelons l'Instant Collaboration. Cependant la conception de telles expériences est un challenge car l'instantanéité doit être appliquée à bon escient afin de servir les équipes et éviter de reproduire les usages que nous avons fait de l'email. Les propos seront illustrés avec la présentation de Stormz, une plateforme d'idéation en temps-réel et de sa technologie U.C.Engine.

3.6 Le Web2.0 : de la théorie à la pratique

Organisation : Aristote, mercredi 12 octobre 2011, École Polytechnique, Palaiseau.

Coordination scientifique :

- Jean Michel Batto (INRA) ;
- Yohanan Winogradsky (INRA).

En très peu de temps, le web est passé d'un contenu mis à disposition par des experts informaticiens à un web alimenté par l'ensemble de ses utilisateurs. Dans le même élan, le web 2.0 participe à l'idée que l'on puisse faire d'internet un centre de profit. Voilà le tournant que marque le web 2.0 : la création de contenu sans connaissances techniques puis sa valorisation. Quels rapports entretiennent mise à disposition de contenu et création de valeur ? Dans quelle mesure assiste-t-on à une nouvelle structuration du web ? Le caractère critique de la production de données, en tant que substrat du web 2.0, est indiscutable, quantitativement et qualitativement. Il nous semble que c'est à travers une lecture interrogeant les concepts de confiance, de travail et de risque que nous pouvons décrire la portée de cette production de données massive et généralisée. Nous souhaitons faire de cette journée une rencontre entre des acteurs d'horizons variés - qu'ils soient du monde politique, culturel, technologique, des affaires, industriel ou entrepreneur - ayant en commun une contribution significative au web 2.0.



Technologies des réseaux et NTICs

Le Web2.0 : de la théorie à la pratique
Mercredi 12 octobre 2011

Coordination scientifique :
- Jean Michel Batto (INRA)
- Yohanan Winogradsky (INRA)

Amphithéâtre Cauchy, École Polytechnique, Palaiseau

<http://www.association-aristote.fr> info@association-aristote.fr

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI CEN Saclay Bât. 474, 91191 Gif-sur-Yvette Cedex.
Séminaire : Aristote, École Polytechnique, 91128 Palaiseau Cedex.
Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 65 Courriel : info@aristote.polytechnique.edu
Site internet : <http://www.association-aristote.fr>

Programme du 12 octobre 2011

9h00-9h20	<i>Accueil café</i>	
9h20-9h30	Jean-Michel Batto (INRA)	Ouverture du séminaire
9h30-10h00	Ludovic Hirliman (Mozilla)	Hirliman, témoin de l'évolution du web 0.0 au web 3.0
10h00-10h30	Patrick Murzeau (DGFIP)	Certificats et identité déclarative
10h30-11h	<i>Pause café</i>	
11h00-11h30	Vincent Puig (IRI)	Web 2.0 : Contribuer n'est pas collaborer
11h30-12h00	Matthieu Soulé (BNP Paribas)	Le web, démocratisation des outils de distribution et bientôt de production
12h00-12h45	David François et Pierre Noizat (société Paymium)	Bitcoin, logiciel libre et devise complémentaire
12h45-14h00	<i>Repas (salle « aquarium »)</i>	
14h00-14h30	Finhur Palmi Magnusson (Parlement Islandais)	Présentation annulée : How to write a constitution in 2011
14h30-15h00	Michael Chekroun (Carenity)	Carenity, le premier réseau social pour les patients et leurs proches
15h00-15h30	Laurence Foucher (Danone)	Le 2.0 dans l'entreprise
15h30-16h00	<i>Pause</i>	
16h00-16h30	Max Bingham (24 Media Labs)	Communicating Science 2.0 : Getting the message out in a wired-up world
16h30-17h00	Stephane Simon (Mendeley)	De la gestion de la bibliographie au réseau social académique
17h00-17h30	Carl Zimmer	"Arsenic Life" Lessons for journalists and scientists
17h30-17h45	Conclusions, questions-réponses avec les intervenants	

Résumés des présentations

Ludovic Hirliman (Mozilla)

Ludovic Hirlimann, témoin de l'évolution du web 0.0 vers le web 2.0 Historique de l'évolution des technologies du web, comparaison avec l'adoption de l'utilisation d'Internet dans notre société. Comparaison entre les technologies du web et les technologies déjà présentes sur Internet lorsque le web est arrivé. La période couverte commencera à la fin des années 80 et jusqu'à nos jours.

Patrick Murzeau (DGFIP)

Certificats et identité déclarative En matière de sécurité informatique, il existe de nombreux paradoxes. Le premier n'est pas le moindre, puisqu'il existe une tendance irréprouvable à forcer les solutions avant même d'avoir évalué le risque. Le second relève des postulats, le principal étant que la signature électronique serait équivalente à la signature manuscrite. Pourtant, à l'usage, l'acte de signer de manière manuscrite n'a rien à voir avec l'acte de signer électroniquement. Les conditions dans lesquelles se déroulent ces signatures sont très éloignées l'une de l'autre. Une signature manuscrite, lorsqu'elle comporte une valeur probante, est recueillie dans un cadre de confiance, et le signataire, s'il recule, ne le fait en raison même de l'acte de signer. Tel n'est en général pas le cas pour une signature électronique, où le cadre offert peut ne pas être de confiance ; quant à la signature elle-même ses modalités techniques sont hors la vue et la compréhension du signataire, sauf à être spécialiste en cryptographie et en informatique, et d'être capable d'analyser les algorithmes mis en oeuvre ainsi que les modules activés pour la réaliser.

La présentation a pour objet, après une expérience d'une dizaine d'années du conférencier dans le cadre de la sécurisation de téléprocédures pour le grand public, de faire le point sur les concepts et de décrypter ce qui va et ne va pas avec les certificats et l'identité. Des premières mises en oeuvre, au début des années 2000, jusqu'à l'actualité récente, et particulièrement la révocation des certificats de l'autorité de certification néerlandaise nommée Diginotar. Reposer les concepts ce sera d'abord faire la distinction entre risque évalué et réalité du risque, puis à distinguer identité, vérification de l'identité et preuve, et nous verrons que sans cadre de confiance, les conséquences qui conduisent à faire peser sur l'internaute l'évaluation des dispositifs de sécurité qui se présentent à lui est peu raisonnable.

Ce parcours sera l'occasion de faire le point sur les projets en cours, qu'il s'agisse d'Idénium, de la carte nationale d'identité électronique, de la fédération d'identité et de la préservation du contenu des documents stockés dans des coffres-fort numériques.

Vincent Puig (IRI)

Web 2.0 : Contribuer n'est pas collaborer L'Iri s'attache depuis sa création en 2006 au sein du Centre Pompidou à étudier les enjeux culturels et cognitifs des dispositifs numériques collaboratifs favorisant de nouvelles formes d'attention, de nouveaux champs de collaboration et une nouvelle implication du corps. Malheureusement le Web 2.0 est encore majoritairement centré sur la contribution (modèle YouTube) ou sur la conversation (modèle Facebook), parfois dans une grande et désespérante solitude, mais très peu sur la collaboration au sens du "construire ensemble" (modèle Wikipedia). Des exemples d'outils emblématiques ou prototypés par l'Iri seront présentés dans le séminaire.

David François et Pierre Noizat (Sté Paymium)

Bitcoin, logiciel libre et devise complémentaire Introduction à Bitcoin Technologie du réseau de « pair à pair » (P2P) Traitement des transactions Nouvelle devise ou méta-devise ? D'où viennent les bitcoins et leur valeur d'échange ? Pourquoi et comment utiliser les bitcoins ? Qu'est ce qui a changé avec Bitcoin ?

Michael Chekroun (Carenity)

Carenity, le premier réseau social pour les patients et leurs proches Carenity est le premier réseau social en France dédié aux patients et à leurs proches. Carenity est né de la volonté de permettre aux millions de personnes concernées par des pathologies chroniques de partager témoignages, conseils pratiques et soutien moral via une plateforme communautaire sécurisée et cautionnée par des médecins et des associations de patients. Au-delà des possibilités de mise en relation avec d'autres patients, l'internaute peut suivre l'évolution de ses indicateurs médicaux (symptômes, traitements, analyses biologiques) grâce à des applications développées en partenariat avec des professeurs de médecine. Gratuit pour l'utilisateur, Carenity génère ses revenus via la commercialisation d'études statistiques anonymisées auprès d'opérateurs de santé qui, grâce à ces informations, pourront mieux appréhender les besoins et usages des patients.

Dr. Max Bingham (24 Media Labs)

Communicating Science 2.0 : Getting the message out in a wired up world Arguably the biggest innovations in academic publication in the past 350 years have been the removal of "Dear Sir" as an introduction (around 1750) and the move to electronic publication in the past ten years. This second point has significantly changed the landscape of science. For example, in biomedical sciences the corpus of publications has doubled since 1995 to approximately 20 million papers. This year could see one million papers published alone which is about two publications per minute. This makes navigating the scientific literature a challenge with some estimates suggesting that a full 30% of researchers' time is taken up with information management. Going beyond scholarly publication and getting the message out to a broader audience from this ocean of information is proving hard too. This is particularly troubling (for some) as more and more funding bodies now insist on demonstrable outreach and communications activities as a condition of funding. With main stream media business models collapsing (the development of the internet has had a part to play in this) and journalists having to deal with hundreds of press releases in a day, many have been left wondering what to do. This is where Web 2.0 technologies will make a difference. With examples of some success stories, I will demonstrate how web 2.0 approaches to communications may well be the only viable channel for most scientists to get their message out from behind lab doors.

Stéphane Simon (Mendeley)

Mendeley : de la gestion de la bibliographie au réseau social académique La gestion des références bibliographiques est un réel souci pour les étudiants, les chercheurs et les enseignants. Plusieurs logiciels ont été développés ces dernières années, mais l'absence d'interfaces entre ordinateurs et utilisateurs limitent considérablement leur utilisation. Mendeley est un logiciel en web 2.0 qui autorise le stockage des références en ligne, une synchronisation entre ordinateurs, mais également smartphone, iPhone, iPad, tablettes, etc. Le partage des références et documents au sein de groupes d'utilisateurs et les applications de discussion entre membres sont de réelles avancées dans le domaine de la communication entre chercheurs, mais également entre l'enseignant et ses étudiants. Mendeley est probablement à ce jour le logiciel le plus abouti dans son domaine. Au cours de cette présentation, nous proposons de décrire les principales solutions offertes par le logiciel.

Carl Zimmer

Arsenic Life : Lessons For Scientists and Journalists Intervention à distance depuis les USA.

In December 2010, a team of NASA-funded scientists published a paper in which they claimed bacteria could substitute arsenic for phosphorus. They triggered a worldwide media frenzy. But through new forms of social media, like blogs, Facebook, and Twitter, scientists aired doubts and even began to attempt to replicate the findings in an open fashion, posting updates on a blog. I will discuss what this episode signifies for the future of science journalism.

3.7 Les stages et formations

L'association a organisé en 2011 trois sessions de formation :

- deux « Formations *Cloud* et HPC » à l'Institut Henri Poincaré, Paris ;
- « Pérennisation et communication de l'information numérique » à l'Institut National du Patrimoine, Paris.

3.8 Formations Cloud et HPC

Organisation : Aristote et ActiveEon, Institut Henri Poincaré, Paris.

Deux sessions : 18-19 mai 2011 et 28-29 septembre 2011.

Coordination scientifique :

- Aristote, ActiveEon SAS en collaboration avec EIT ICT Labs et l'INRIA.

Cette formation a été conçue par les meilleurs experts français dans le domaine du CLOUD Computing. Elle se veut à la fois pratique, en proposant des solutions applicables et en traitant sur machine des exemples concrets, pédagogique en alternant des concepts généraux et des exercices d'application.

La formation vise principalement les développeurs et responsables d'applications HPC et CLOUD, les responsables de Système d'Information en charge de la consolidation de leur infrastructure, et les personnes souhaitant acquérir une compréhension précise des solutions disponibles aujourd'hui en Open Source pour le CLOUD Computing.

Il est demandé aux participants de venir avec leur PC portable (Linux, Windows, ou Mac) pour bénéficier de l'installation des outils, accéder à une plateforme de Cloud, et participer à la mise en œuvre d'applications et/ou à l'administration d'un Cloud.

Programme de la formation

9h00-10h00	<i>Première journée, accueil café</i>
10h00-10h15 10h15-11h00	Présentation des 2 journées Introduction générale au CLOUD Computing
11h00-11h15	<i>Pause et café</i>
11h15-12h00	Présentation de la solution Open Source ProActive Parallel Suite – Uses cases : Engineering, IT, BioTech, Finance, etc. – Technical Computing : Interface et utilisation directe d'un CLOUD depuis Matlab/Scilab – Management de Scheduler natifs : LSF, PBS, SGE, LoadLeveler, – Virtualisation : VMware Server/ESX, XenServer, VirtualBox, KVM/Qemu/Xen, Microsoft Hyper-V
12h00-14h00	<i>Déjeuner libre</i>
14h00-14h45	Présentation d'une plateforme en production (1 200 cœurs) – Démonstrations de Use Cases
14h45-16h00	Session pratique : installation guidée des clients sur les PCs des participants
16h00-16h30	<i>Pause et café</i>
16h30-18h00	Utilisation de base des clients et d'un Portail – Tests avec applications fournies
18h00	<i>Fin de la première journée</i>
9h00-9h30	<i>Deuxième journée, accueil café</i>
9h30-11h00	Workflows avancés et utilisation d'un Editeur Graphique de Workflows – Test, Boucle et Variation Paramétrique, Soumission et Exécution sur le Cloud
11h00-11h15	<i>Pause et café</i>
11h15-12h00	Administration d'un Cloud : Présentation générale – Ressources Serveurs/Clusters/Desktops, Machines Physiques et Virtuelles – Politiques Statiques et Dynamiques on-demand
12h00-14h00	<i>Déjeuner libre</i>
14h00-16h00	À la carte : exercices pratiques avec le support des intervenants – « Construisez et lancez sur le Cloud votre propre application » ; – ou « Créez et Administrez votre propre Cloud ».
16h00-16h30	<i>Pause et café</i>
16h30-17h30	Suite de la formule À la carte
17h30	<i>Fin de la formation</i>

3.9 Pérennisation et communication de l'information numérique

Organisation : Groupe PIN d'Aristote, du 18 au 21 octobre 2011, Institut national du patrimoine, Paris.

Coordination scientifique :

– Olivier Rouchon (CINES).

Comment pérenniser des informations numériques avec des moyens matériels et logiciels dépourvus de toute pérennité ? C'est à ce problème majeur pour la société de l'information que cette formation propose d'apporter des solutions concrètes. Cette formation a été conçue par un groupe d'experts, professionnels en charge de la pérennisation d'information numérique dans des contextes variés (archives et bibliothèque nationale, organismes scientifiques, entreprises...). Au sein de l'association Aristote, ces professionnels partagent depuis 10 ans leurs expériences et leur savoir-faire dans le groupe Pérennisation des Informations Numériques.

Par cette formation, le groupe PIN souhaite aujourd'hui proposer à tous les professionnels confrontés à la question de la pérennisation des données numériques : informaticiens, documentalistes, gestionnaires ou administratifs, etc., des solutions fondées sur des expériences concrètes.

Programme

	18 octobre 2011
9h30 - 10h00	<p>Accueil des participants</p> <ul style="list-style-type: none"> – tour de table ; – présentation de l'association Aristote et des activités du groupe « Pérennisation des Informations Numériques » ; – présentation des journées.
10h15 - 1300	<p>Vue d'ensemble du contexte normatif Le Modèle de Référence OAIS et les autres normes généralistes du domaine</p> <p>Panorama du paysage normatif. Classification des domaines, rôle et objets des différentes normes et présentation plus détaillée des normes généralistes dédiées à l'archivage des informations sous forme numérique :</p> <ul style="list-style-type: none"> – le Modèle de Référence OAIS (norme ISO 14721 : 2003) est aujourd'hui internationalement reconnu et utilisé défini le vocabulaire, les concepts et les modèles indispensables à la compréhension de la problématique de l'archivage long terme ; – les normes dérivées du Modèle OAIS et qui abordent l'interface entre le Producteur et l'Archive ainsi que la certification de l'Archive.
13h00 - 14h30	<i>Déjeuner libre</i>
14h30 - 17h30	<p>La stratégie globale et l'organisation de l'archivage numérique, les facteurs critiques d'un projet d'archivage numérique</p> <ul style="list-style-type: none"> – comment résoudre le problème en termes organisationnels ? – comment mettre en place une archive numérique sur la base de services indépendants ayant chacun des responsabilités, des fonctions, des ressources et des compétences bien définies ? – quelles sont les possibilités de mutualisation et de coopération qui sont envisageables ? – comment gérer les risques liés au projet ? – comment maîtriser les coûts, les modèles d'évaluation.
17h30	<i>Fin de la première journée</i>

19 octobre 2011	
9h30 - 13h00	<p>Cycle de vie de l'information et archivage /records management</p> <ul style="list-style-type: none"> – qualifier et classer l'information numérique : caractéristiques internes et externe ; – cycle de vie des documents engageants (records) : problématique de nommage, de versionnage, de copie, de validation et de validité ; – contraintes réglementaires et risques : archiver pour prouver et pour tracer ; – la gestion de l'archivage au sens du records management : principes et processus (capture, conservation, destruction, mise à disposition).
13h00 - 14h30	<i>Déjeuner libre</i>
14h30 - 17h30	<p>Les questions juridiques et leurs conséquences Le nouveau cadre juridique de la preuve.</p> <ul style="list-style-type: none"> – le nouveau cadre juridique du droit de la preuve ; – l'e-administration : administration de la preuve et archivage – les actes authentiques ; – ADELE et l'archivage ; – quelles sont les possibilités de mutualisation et de coopération qui sont envisageables ? – un exemple : le contrôle de légalité ; <p>La protection des données personnelles et l'archivage.</p>
17h30	<i>Fin de la deuxième journée</i>

20 octobre 2011	
9h30 - 13h00	<p>Les Métadonnées de pérennisation et l'accès aux objets</p> <ul style="list-style-type: none"> – introduction sur les Métadonnées en général : définition, aspect théorique, utilité des Métadonnées ; – le modèle d'information de l'OAIS : place des Métadonnées dans le modèle ; – quelles Métadonnées pour la pérennisation ? – les droits de propriété intellectuelle ; – encoder les Métadonnées ; – l'accès à l'information grâce aux métadonnées.
13h00 - 14h30	<i>Déjeuner libre</i>
14h30 - 17h30	<p>Les plateformes logicielles existantes sur la pérennisation Présentation de l'offre disponible sur le marché, qui est à présent plus mûre, et peut être comparée avec le Modèle de Référence OAIS.</p> <ul style="list-style-type: none"> – les SAE orientés planification de la préservation : PAC, SPAR, pil@e ; – pour tous, préciser chaque fois les fonctionnalités, s'il s'agit de logiciel ou de stockage, leur périmètre. ; – outils libres de référencement/indexation : DSPACE, LOCKSS, Fedora Commons.
17h30	<i>Fin de la troisième journée</i>

	21 octobre 2011
9h30 - 12h00	<p>Problématique de la gestion des supports de l'information numérique</p> <p>Pourquoi le support peut-il se dégrader sans perte immédiate d'information ; notion de codage, analogie avec un texte écrit :</p> <ul style="list-style-type: none"> – effet de seuil dans la perte d'information et codes de correction d'erreurs ; – qualités attendues d'un support de stockage : robustesse, durée de vie de la technologie, outils de contrôle d'état du support disponibles ; – éléments de réflexion stratégique : risques budgétaires, risques environnementaux... Les supports aujourd'hui : disque dur, bandes magnétiques, disque optique enregistrable ; – le stockage de très gros volume, technologies à venir.
12h00 - 13h30	<i>Déjeuner libre</i>
14h30 - 17h30	<p>La représentation de l'information sous forme numérique, les formats de données</p> <ul style="list-style-type: none"> – Le codage de l'information numérique aux différents niveaux dans les systèmes d'information (périphériques de stockage, système de fichier, encodage des données primaires, formats applicatifs ; – exemples : le format PDF d'Adobe, XML du W3C, PNG du W3C ; – l'information de représentation selon le modèle OAIS : contenu, intérêt, sources ; – les initiatives en cours pour mutualiser l'information sur les formats (UDFR, PRONOM).
16h30-17h00	Session question réponse avec les intervenants de la semaine
17h00-17h30	Évaluation et bilan des journées
17h30	<i>Fin de la formation</i>

Chapitre 4

Perspectives pour 2012

Forte de ce bilan très positif, l'association Aristote confortera en 2012 son rôle de catalyseur entre le monde de la recherche académique et celui l'industrie et des services, tout en continuant à s'ouvrir vers les petites et moyennes entreprises.

Nos centres d'intérêts évoluent, certains restent au sein de groupes bien constitués (PIN) mais de nouveaux partenariats, de nouveaux groupes de travail se profilent autour des technologies émergentes (GPU, ...). L'association travaille aussi à la pérennisation de services à l'usage de la communauté par exemple autour d'**Access Grid**. L'association mettra tout en œuvre pour la réussite des projets engagés pour l'année à venir.

L'association et ses groupes de travail sont ouverts à toutes les propositions de collaboration pouvant contribuer à la réalisation de ses objectifs.

<http://www.association-aristote.fr>

Il est prévu, fin juin 2012, un événement sur le thème « SI immobilier, développement durable » organisé par Bernard Rouault (CEA). Nous devrions aussi renouer avec le concept d'assemblée générale publique incluant, en 2012, un débat sur le thème « Politique et Réseaux sociaux ».

Groupes de travail

Les participants confirment leur intérêt pour de nouvelles thématiques initiées l'an passé :

Des systèmes d'information communicants en toute sécurité comment concilier le travail collaboratif, la mobilité,... et la sécurité ?

e-Laboratoire type de collaborations, projets, aspects temporels, quelles sont les ressources partagées ou à partager, nomadisme, environnement de travail et outils.

Séminaires 2012

Le programme prévisionnel du cycle de séminaire 2012 comporte huit sessions :

1. Traçabilité des individus et protection de la vie privée, avec Djordje Djokic (Bearstech) et Pierre Leonard (INRA) ; mars 2012.
2. Applications en mode SaaS pour les entreprises : sécurité et tiers de confiance avec Jean-Claude Sabattier du CEA ; avril 2012.
3. Sécurité et nomadisme : nouveaux usages des technologies avec F. Morris (CNRS), T. Roussel (CEA) et R. Seneor (Ecole Polytech.) ; mai 2012.

4. Green IT, Green Cloud avec Drissa Houatra (Orange Labs) ; juin 2012.
5. Quels services pour le bâtiment intelligent ? Avec D. Menga (EDF) et B. Rouault (CEA) ; septembre 2012.
6. Atelier de la simulation numérique : outils autour des codes et visualisation, avec l'ONERA (Thien Hep LE) ; octobre 2012.
7. Portails d'application avec aspects de gestion des identités et de la preuve, avec P. d'Anfray (CEA) et Patrick Murzeau (DGFIP) ; novembre 2012.
8. Evolution du numérique et de l'édition avec Roland Senéor (Ecole Polytechnique) et V. Heuschling (Affini-Tech) ; décembre 2012.

Partenariats 2012

Notre partenariat avec OpenGPU devrait déboucher sur une nouvelle journée dédiée au calcul hybride pendant l'édition 2012 du forum Ter@Tech. Des contacts avec l'ONERA devraient conduire à une organisation commune de certaines éditions des « ONERA Scientific Days ». Dans le cadre des activités avec l'*Open World Forum*, il est prévu une journée dédiée à la place de l'Open Source dans le monde éducation-recherche en partenariat avec le projet PLUME du CNRS. Enfin, coté « Arts et Sciences » (réseau MARCEL), Aristote est sollicitée pour assurer le pilotage scientifique et technique dans le cadre de projets financés par le ministère de la culture.

The screenshot shows the website for Aristote, an association focused on Information and Communication Technologies. The header features the title "[[Nouvelles Technologies de l'Information et de la Communication]]" and the tagline "Penser & Innover ensemble". A navigation menu includes "Admin", "Mettre à jour le profil", "Index", and "Déconnexion".

The main content area is titled "Nouvelles Technologies de l'Information et de la Communication" and describes the association as "Association de Réseaux Interconnectés en Systèmes Totalement Ouverts et Très Elaborés". It states that the association is francophone, non-profit, and independent, aiming to group users of new technologies (NTIC) for research and development. Its mission includes sharing experiences, mutualizing actions, and promoting the use of new technologies in a human and economic context.

A prominent banner for 2012 reads "L'association ARISTOTE vous présente ses meilleurs vœux pour l'année 2012!!" with the slogan "Anticiper... Partager... Innover...". The banner features a network diagram and the year "2012".

On the left side, there is a search bar and a sidebar menu with links to "Accueil", "Qui sommes nous", "Evénements récents", "Séminaires", "Suivi à distance", "Groupes de travail", "Comité de programme", "Conseil d'administration", "Espace 'bureau'", "Communauté", "Archives", "Dokuwiki, syntaxe", and "Bac à sable". It also displays visitor statistics: "Visiteurs: 44312" and "Connectés: 1".

<http://www.association-aristote.fr>

info@association-aristote.fr

ARISTOTE Association Loi de 1901. Siège social : CEA-DSI CEN Saclay Bât. 474, 91191 Gif-sur-Yvette Cedex.
Secrétariat : Aristote, École Polytechnique, 91128 Palaiseau Cedex.
Tél. : +33(0)1 69 33 99 66 Fax : +33(0)1 69 33 99 67 Courriel : Marie.Tetard@polytechnique.edu
Site internet <http://www.association-aristote.fr>