cnrs le journal

n° 250
novembre 2010

JUSQU'OÙ IRA l'informatique ?

D'INTERNET À L'ORDINATEUR QUANTIQUE

L'événement
Double Chooz : la traque des neutrinos est lancée

dépasser les frontières

Sommaire

4 | 5 L'essentiel
Le point sur les nominations, les prix, les faits marquants...

6 | 7 L'événement
Lancement du premier détecteur de Double Chooz, qui s'apprête à repérer les neutrinos émis par les réacteurs nucléaires de la centrale ardennaise.

14 | 16 En images
Retour sur la campagne 2010 de fouilles archéologiques à Xanthos, en Asie Mineure.

17 | Décryptage
Yves Dessaux, biologiste, explique les conséquences du fauchage des vignes OGM de l'Inra, le 15 août dernier.

30 | 31 Portrait
Rencontre avec Pierre-Henri Castel, spécialiste de l'histoire et de la philosophie des maladies mentales.

32 | 35 Stratégie
Les innovations, les partenariats et les collaborations internationales.

36 | On en parle
L'actualité de la vie interne du CNRS.

37 | Un jour avec...
Erwan Amice, plongeur sous-marin.

38 | 42 Culture
Livres, expositions, films... La sélection de la rédaction.

43 | Sur le vif
Les coulisses étonnantes d'une photo de science.

Éditorial

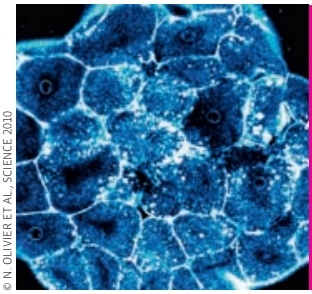
© B. WERNER/ ÉCOLE POLYTECHNIQUE

PAR PHILIPPE BAPTISTE, DIRECTEUR SCIENTIFIQUE DE L'INSTITUT DES SCIENCES INFORMATIQUES ET DE LEURS INTERACTIONS

Les progrès des sciences informatiques ont permis une révolution dont les développements spectaculaires ont bouleversé notre quotidien. Le CNRS doit répondre aujourd'hui à de nouveaux enjeux numériques, notamment dans le domaine de la santé et de l'environnement. En créant un Institut des sciences informatiques et de leurs interactions (INS2I), le CNRS se positionne comme un acteur majeur de l'une des priorités de la Stratégie nationale de recherche et d'innovation.

Avec pour mission première de développer les sciences informatiques, l'INS2I travaille en étroit partenariat avec l'Institut des sciences de l'ingénierie et des systèmes (Insis) sur des sujets comme l'automatique, le signal, l'image, la robotique ou les systèmes sur puce. Plus généralement, l'interdisciplinarité est un enjeu majeur pour l'INS2I, qui diffuse de nouveaux outils et concepts dans toutes les disciplines. Parallèlement, de nouveaux usages scientifiques et sociétaux soulèvent constamment de nouvelles questions fondamentales.

Avec 4 000 permanents travaillant dans les Unités mixtes de recherche (UMR) de l'INS2I pour un peu moins de 400 chercheurs CNRS et autant d'ITA, l'institut est naturellement tourné vers les universités et les grandes écoles. L'Institut national de recherche en informatique et automatique (Inria) est également un partenaire notable de l'INS2I, puisque plus de 60 % des équipes Inria sont communes avec nos unités. Par ailleurs, l'institut souhaite développer ses interactions grâce à de nouveaux laboratoires internationaux, au Japon et au Canada par exemple. Avec l'ensemble du CNRS, en s'appuyant sur une vision nationale et internationale de ses activités, l'INS2I veut donc mener une politique d'excellence au service de la communauté scientifique, tout en encourageant les actions de valorisation et de transfert.



© N. OLIVIER ET AL. SCIENCE 2010

8 | 13 Actualités
Les premières heures de la vie en 3D ; surprise autour de la synthèse de l'ozone ; mieux diagnostiquer la maladie d'Alzheimer ; des révélations sur le climat de l'Ordovicien ; du nouveau sur les mécanismes de notre mémoire ; les secrets du plus célèbre antidiabétique...

18 | 19 Le grand entretien
Lionel Collet, président de la Conférence des présidents d'université, nous livre sa vision du nouveau paysage de la recherche française.

© C. FRESILLON/CNRS PHOTOTHÈQUE

© C. FRESILLON/CNRS PHOTOTHÈQUE ; ESA, HFI ET LFI CONSORTIA

20 | 29 L'enquête
Jusqu'où ira l'informatique ?
21 | L'avènement de la société numérique
25 | Des milliards d'informations à organiser
28 | Ordinateur quantique : l'ultime défi

D'INTERNET À L'ORDINATEUR QUANTIQUE

Jusqu'où ira l'informatique ?

L'avènement de la société numérique **21** | Des milliards d'informations à organiser **25** | Ordinateur quantique : l'ultime défi **28** |

01 © C. FRESILLON/CNRS PHOTOTHÈQUE ; ESA, HFI ET LFI CONSORTIA

De l'ordinateur aux téléphones mobiles dernier cri, le grand public sait combien l'informatique a révolutionné nos modes de communication. Elle a aussi profondément changé le travail des scientifiques dont les recherches réclament d'immenses puissances de calcul. Et ce n'est pas fini. Dans les laboratoires, on s'affaire pour développer un nouvel Internet, inventer des techniques performantes de traitement des données et même concevoir l'ordinateur quantique. À l'occasion du premier anniversaire de la création de l'Institut des sciences informatiques et de leurs interactions, *CNRS Le journal* vous invite à découvrir l'informatique de demain. **UNE ENQUÊTE DE** MATHIEU GROUSSON ET VAHÉ TER MINASSIAN

L'avènement de la **société numérique**

« ***Un mouvement fondamental et inéluctable,*** *comparable à l'arrivée du train à vapeur qui a marqué le début de l'ère industrielle*[1]. » Le constat dressé par Gérard Berry, titulaire pour l'année 2009-2010 de la chaire Informatique et sciences numériques du Collège de France, paraît difficilement contestable : « *Notre civilisation est en train de devenir numérique,* remarque celui-ci. *Des industries classiques comme les télécommunications et la diffusion culturelle sont totalement chamboulées. D'autres grandissent au pas de charge tels l'informatique et les services associés. Internet révolutionne les échanges en abolissant les contraintes de distance, de temps et de volume. Tandis qu'en science la modélisation informatique de tout phénomène est devenue la norme.* » Vingt et un ans après l'invention de la principale application d'Internet, le World Wide Web, énumérer les bouleversements créés par les avancées récentes de l'informatique semble fastidieux et vain, tant il est évident aux yeux de tous que ceux-ci sont déjà incalculables. Et encore peu nombreux au regard de ce que nous réserve l'avenir.

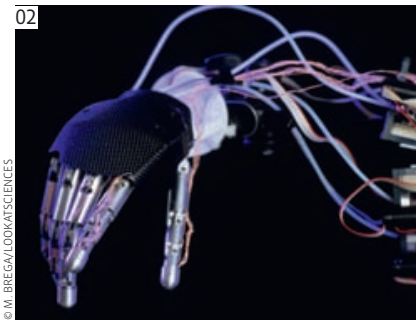
01 Les 32 écrans haute définition de la plateforme Wild permettent d'afficher de très grandes images, ici une partie de la photo la plus détaillée de notre galaxie prise à ce jour.

UN PROCESSUS QUI S'ACCÉLÈRE

Ce monde numérique du futur, justement, de quoi sera-t-il fait ? Bien malin celui qui saurait apporter une réponse définitive à cette question, alors que chaque semaine voit arriver une nouvelle application qui vient bousculer le marché de l'informatique. Les spécialistes du domaine s'accordent toutefois pour prédire un développement spectaculaire d'Internet, lequel devrait à terme relier entre eux non plus seulement les hommes mais aussi les objets situés dans notre environnement, et même dans notre corps.

« *Actuellement, il existe dans le monde quinze à vingt fois plus d'ordinateurs autonomes qu'en interaction avec l'homme,* explique Gérard Berry. *Or ces centaines de milliards de processeurs, disséminés autour de nous – une voiture haut de gamme en compte déjà plus de 80* [qui

02



© M. BREGA/LOOKATSCIENCES

02 À l'avenir, les prothèses électroniques, comme celle du projet Cyberhand, seront directement reliées au système nerveux.

© KAKSONEN/INRIA

03

contrôlent le freinage, la suspension, la combustion ou la jauge de carburant…] –, *sont pour l'instant déconnectés les uns des autres. Demain, avec l'Internet des objets, toutes ces machines communiqueront entre elles sans intervention humaine pour produire collectivement de nouvelles applications. Les infrastructures routières parleront aux véhicules afin de les avertir des limitations de vitesse, de leur signaler les embouteillages et de les protéger des accidents. Les prothèses électroniques seront directement branchées sur le système nerveux, et les circuits dont seront équipés les malades enverront directement des informations sur leur état de santé à l'ordinateur du centre hospitalier. À la limite, ce sera le médecin qui appellera le patient en cas de problème et non le contraire !* »

Dans le même temps, la façon dont nous commanderons aux machines conçues pour recevoir nos instructions changera, elle aussi, radicalement. Écrans tactiles et détecteurs de mouvement pourraient remplacer claviers et souris d'ordinateurs de bureau. Et, avec les progrès du Web sémantique, nous disposerons de moteurs de recherche intelligents, capables de retrouver une information sur la Toile à partir du sens d'une question et non plus sur la base de sa seule syntaxe. Enfin, « *avec le développement des applications de type Twitter ou Facebook, mais aussi avec le succès commercial des smartphones – iPhone ou BlackBerry – le Web a changé de fonction : il n'est plus seulement une bibliothèque où l'utilisateur vient chercher de l'information, mais un espace de communication interactif entre humains auquel certains sont d'ores et déjà reliés en permanence* via *leurs téléphones portables* », observe Serge Abiteboul, membre du Laboratoire de recherche en informatique[2], qui travaille sur la gestion de données et de connaissances sur le Web où l'information est disséminée sur quantité de machines différentes (ordinateurs, téléphones portables, sites Web, Facebook, etc.).

UNE ADAPTATION PERMANENTE

Un tel chamboulement ne saurait se produire sans heurts ni adaptations. « *Malgré sa capacité à intégrer de nouvelles technologies et applications, qui est l'une des clés de son succès, Internet est fragilisé par cette évolution,* confirme ainsi Serge Fdida, professeur au Laboratoire d'informatique de Paris-6[3] et coordinateur de la plateforme européenne OneLab. *Même s'il peut difficilement être cassé, il n'a pas été conçu pour absorber à grande échelle de nouveaux besoins tels que la mobilité, la sécurité et la diversité, dont l'association perturbe son organisation actuelle. Il faut, en effet, se souvenir que le cahier des charges initial de l'Internet était fondé sur l'hypothèse de machines fixes et d'interlocuteurs de confiance, clairement identifiés, ce qui est loin d'être le cas aujourd'hui. De plus, le système s'est petit à petit imposé comme support de nombreux services (distributions de contenus, paiement en ligne…), ce qui a conduit au développement de solutions* ad hoc. *Le problème, c'est que celles-ci sont en général mal intégrées et complexifient le management du réseau et son efficacité.* » Conséquence de ce phénomène : plusieurs pays, dont les États-Unis, le Japon et l'Allemagne, ont lancé voici quatre ou cinq ans d'ambitieux programmes

Une sélection de **photos** dans le cadre de l'exposition itinérante **Un monde numérique** est à découvrir sur le journal feuilletable en ligne > www2.cnrs.fr/journal

03 Les écrans tactiles multipoints utilisés par l'équipe iPARLA (Labri/Inria) permettent de manipuler les objets 3D. 04 La multiplication des terminaux mobiles nécessite d'étudier de nouvelles architectures réseaux. Ici, les systèmes de l'équipe Pops (Lifl/Inria/Ircica).

de recherche dans le but de construire les bases d'un Internet du futur, plus modulable que l'actuel.

Ainsi, le projet européen Fire vise notamment à constituer d'ici à 2015 une plateforme expérimentale sur laquelle des scientifiques, des industriels et des PME pourraient concevoir, déployer et tester en toute sécurité de nouveaux outils et services Internet. One-Lab en constitue la première étape[4]. Opérationnel depuis trois ans, ce prototype fournit un accès à un réseau restreint à 1 000 ordinateurs connectés à travers le monde ainsi qu'à d'autres plateformes de recherche. Il a d'ores et déjà permis de tester de nombreuses applications comme la distribution de contenus (vidéo, eBooks, musique) *via* le réseau mondial ou encore la géolocalisation d'adresses IP, le numéro permettant d'identifier chaque ordinateur qui est connecté à Internet.

L'un des autres problèmes de taille, lié à la mobilité croissante des usagers, réside dans les limites des technologies radio pour les services informatiques mobiles. « *Les réseaux de la téléphonie mobile de la seconde génération, type GSM, ont été conçus pour transmettre de la voix et non des images, de la vidéo ou pour se connecter à la télévision numérique ou à Internet,* rappelle Pierre Duhamel, directeur de recherche au Laboratoire des signaux et systèmes[5]. *Résultat, ils sont souvent à la limite de la saturation dans les grandes villes.* » Plusieurs solutions sont à l'étude, dont le *network coding,* qui consiste à faire transiter les données *via* un réseau formé par les autres mobiles. Ceux-ci joueraient alors, selon les cas, le rôle d'émetteur, de récepteur, de relais ou de routeur. Quoi qu'il en soit, nos chercheurs ont pris le taureau par les cornes. Pour preuve, depuis septembre dernier, Pierre Duhamel coordonne, dans le cadre du Réseau thématique de recherche avancée (RTRA) Digiteo d'Île-de-France, le premier gros projet consacré à ce secteur innovant de la coopération dans les réseaux.

La sécurité, et en premier lieu celle des hommes, est également une préoccupation majeure des spécialistes. Les myriades de processeurs embarqués qui assurent des fonctions variées dans notre environnement sans intervention humaine offrent déjà des garanties appréciables en matière de réactivité, de disponibilité et d'autonomie. Au point que les ingénieurs n'hésitent plus aujourd'hui à confier à certains d'entre eux, dits critiques, des tâches mettant en jeu la vie humaine : pilotage d'avion, contrôle de centrales nucléaires ou chirurgie assistée par ordinateur. Problème, « *la conception de ces systèmes est extrêmement coûteuse,* signale Joseph Sifakis, directeur de recherche au Laboratoire Verimag[6], à Gières, et titulaire en 2007 du prestigieux prix Turing, l'équivalent du prix Nobel en informatique. *Le développement d'un logiciel critique fait appel à des méthodologies de développement spécifiques, coûte 1 000 fois plus cher que celui d'un code ordinaire et nécessite le passage devant une autorité de certification* ».

© C. LEBEDINSKY/INRIA

SÉCURISER LES SYSTÈMES

Autre complication : si elles ont le mérite d'exister, ces méthodes industrielles de vérification des systèmes embarqués par *model-checking,* dont Joseph Sifakis fut l'un des inventeurs, s'avèrent inopérantes au-delà d'un certain degré de complexité. « *Ce qui interdit l'arrivée de plusieurs technologies nécessitant une disponibilité ou une réactivité importantes,* commente le chercheur. *C'est le cas d'applications touchant à la médecine et à la conduite automobile, mais aussi du Web des objets, où l'on doit franchir une étape supplémentaire en faisant coopérer entre eux des systèmes embarqués dans un environnement Internet non critique, c'est-à-dire peu sécurisé.* »

Face à ces difficultés, certains scientifiques, à l'instar de Joseph Sifakis, se sont résolus à revisiter la théorie afin de rechercher des solutions qui évitent la vérification *a posteriori.* « *Lorsqu'un ingénieur construit un pont, il dispose d'équations mathématiques lui garantissant que son ouvrage d'art ne s'effondrera pas,* note ce dernier. *L'informaticien, lui, n'a rien de tel : il n'a d'autres choix que de fabriquer des systèmes dont il doit tester le bon fonctionnement ensuite. Ce que mes collègues et moi-même tentons de faire, c'est d'essayer d'identifier les bases théoriques qui nous permettront de*

construire au mieux, à partir de composants élémentaires, un système informatique afin d'être en mesure de garantir son bon fonctionnement. »

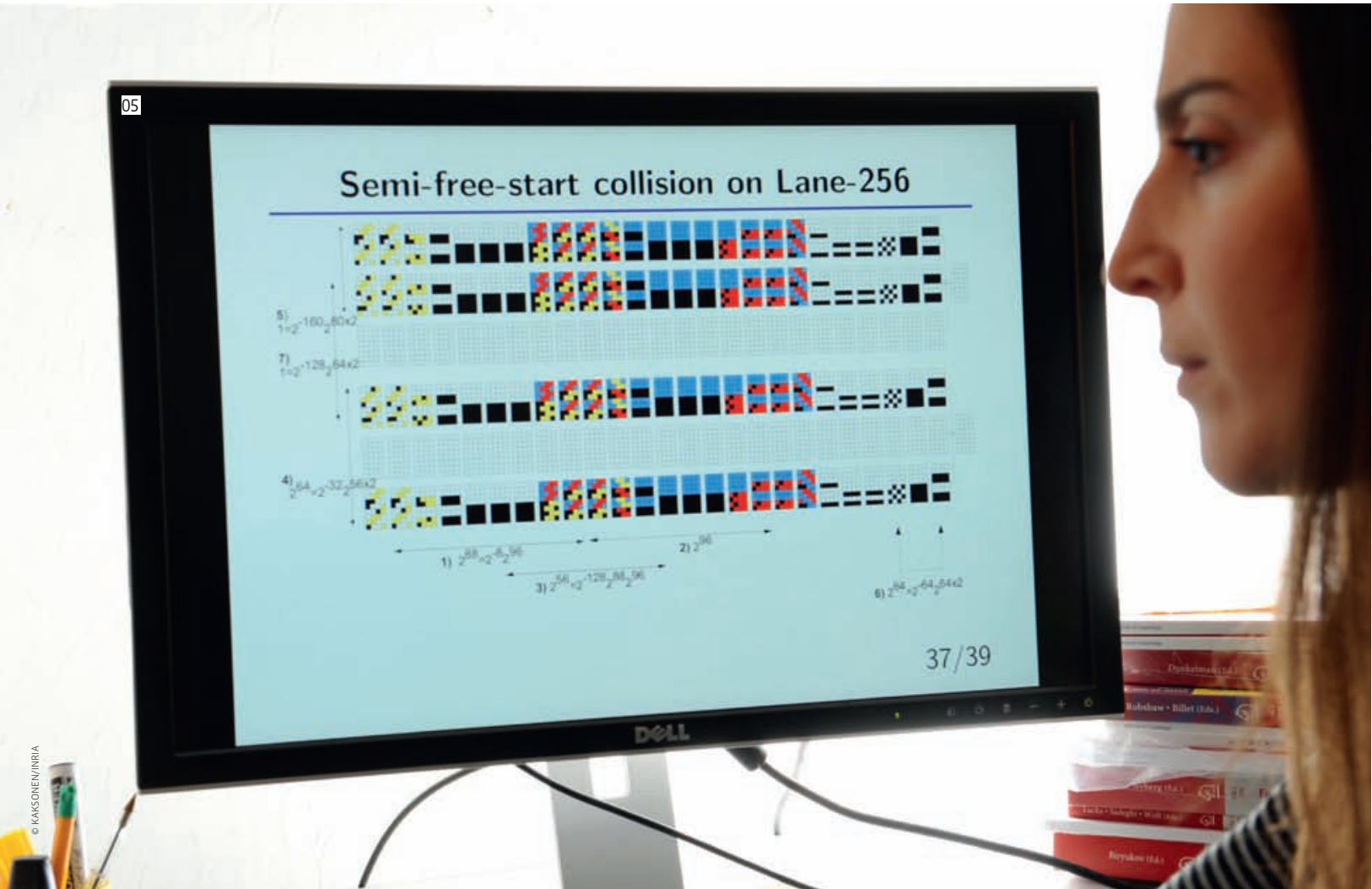
LA DIFFICILE LUTTE CONTRE LE PIRATAGE

Les questions de sécurité informatique concernent aussi la multiplication des objets communicants, de manière un peu plus criante chaque jour. Souvent, les utilisateurs ne se rendent pas compte que leurs ordinateurs sont piratés. Téléphones mobiles, cartes de paiement, consoles de jeu, titres de transport, mais aussi clés électroniques ou télévisions à péage constituent autant de terrains d'étude potentiels pour les cryptographes qui conçoivent les mécanismes de sécurité et les cryptanalystes qui essaient de les prendre à défaut. « *On transmet aujourd'hui de plus en plus d'informations personnelles, mais avec peu ou pas de contrôle. C'est pourquoi l'un des grands problèmes du moment reste la protection de la vie privée et le vol d'identité* », indique Phong Nguyen, directeur de recherche à l'Inria Paris-Rocquencourt et au Laboratoire d'informatique de l'École normale supérieure[7], à Paris. Au sein de l'équipe Crypto de l'ENS, certains s'intéressent à la sécurité prouvée, c'est-à-dire à l'amélioration des garanties de sécurité des programmes cryptographiques. D'autres, au contraire, testent les limites des systèmes de sécurité existants, en étudiant les meilleures formes d'attaque pouvant être mises en œuvre contre tel ou tel procédé cryptographique. « *Et tous les coups sont permis !* », s'exclame le chercheur. Comme essayer de récupérer les données d'une carte à puce en observant sa consommation électrique ou son rayonnement électromagnétique… Un jeu du chat et de la souris qui, selon Phong Nguyen, concernera même à l'avenir des dispositifs futuristes comme l'ordinateur quantique : « *Car, si une telle technologie voit le jour, il faudra nécessairement transformer la cryptographie utilisée actuellement.* »

CRYPTOGRAPHIE Ensemble des techniques de chiffrement qui assurent l'inviolabilité de textes et, en informatique, de données.

1. Tiré de *Pourquoi et comment le monde devient numérique* (Collège de France/Fayard, janvier 2008), de Gérard Berry, membre de l'Académie des sciences et de l'Académie des technologies.
2. Unité CNRS/Université Paris-Sud-XI.
3. Unité CNRS-UPMC.
4. Lire « OneLab2 : l'Internet du futur prend de la vitesse », *Le journal du CNRS*, n° 227, décembre 2008, p. 15.
5. Unité CNRS/Supélec/Université Paris-Sud-XI.
6. Unité CNRS/Université Joseph-Fourier/Grenoble INP.
7. Unité CNRS/ENS Paris/Inria.

CONTACTS :
Serge Abiteboul
› serge.abiteboul@inria.fr
Gérard Berry
› gerard.berry@sophia.inria.fr
Pierre Duhamel
› pierre.duhamel@lss.supelec.fr
Serge Fdida
› serge.fdida@lip6.fr
Phong Nguyen
› phong.nguyen@ens.fr
Joseph Sifakis
› joseph.sifakis@imag.fr



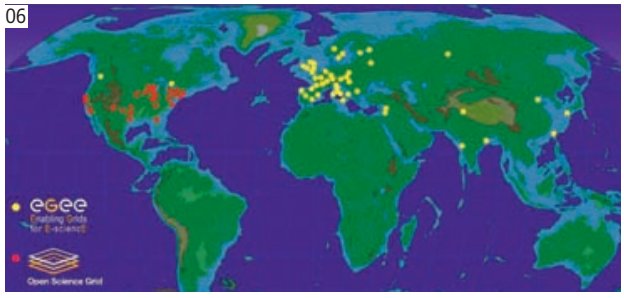
05 En cryptographie, il est courant de condenser, autrement dit hacher, les données. Le condensé ainsi obtenu permet de produire une signature numérique servant à authentifier l'expéditeur d'un message.

© KAKSONEN/INRIA

Des milliards d'informations à organiser

Un touriste à la recherche du voyage au meilleur prix. Un physicien face aux données recueillies par un accélérateur de particules. Une société d'intérim compulsant des CV afin de pourvoir une offre d'emploi. « *Tous ces exemples ont un point commun,* révèle Amedeo Napoli, du Laboratoire lorrain de recherche en informatique et ses applications[1], à Vandœuvre-lès-Nancy. *Ils renvoient à des situations où l'on fait face à un volume colossal de données parmi lesquelles on cherche à extraire une information.* » En principe, la méthode pour y parvenir est simplissime : préparer les données initiales, les confier à un algorithme de fouille et attendre que ce dernier se charge de présenter le résultat sous la forme souhaitée. Mais, dans un univers où le volume des données croît inexorablement, l'extraction de connaissances pertinentes relève de la gageure.

Illustration avec le cas de la recherche d'un séjour, comprenant vol, hôtel et location de voiture, au meilleur prix. Comme le détaille Michel Beaudouin-Lafon, du Laboratoire de recherche en informatique[2], à Orsay, « *mathématiquement, nous savons que la complexité de ce type de problème exclut qu'il puisse être résolu exactement en un temps raisonnable, dès lors que le nombre de données en entrée explose* ». Si bien qu'en pratique les programmateurs doivent ruser afin d'obtenir le résultat le moins mauvais en un temps raisonnable. Et c'est un fait, la fouille de données, à l'heure actuelle en plein essor, agrège des spécialistes de disciplines aussi différentes que l'informatique, bien sûr, mais aussi l'architecture des machines, la linguistique ou les mathématiques. Ces spécialistes empruntant aussi bien à l'intelligence artificielle, aux bases de données, aux techniques d'apprentissage et aux méthodes statistiques.



© CERN

06 Emplacements des sites impliqués dans les deux plus grandes infrastructures de grille aujourd'hui dans le monde : Egee en Europe (en jaune) et OSG aux États-Unis (en rouge).

OPTIMISER LE TRI DES DONNÉES

Une chose est certaine, plus aucun secteur n'échappe à la nécessité de développer des méthodes efficaces pour ne pas crouler sous une montagne de données inexploitables, voire impossibles à stocker. Prenons le projet ANR Midas, dont l'objectif est de réaliser un algorithme capable de résumer un important volume de données produites en temps réel, afin qu'elles puissent être stockées sur une mémoire centrale limitée pour consultation ultérieure. « *C'est typiquement le cas de figure rencontré par France Télécom, EDF ou la SNCF,* précise Pascal Poncelet, du Laboratoire d'informatique, de robotique et de microélectronique de Montpellier[3]. *Par exemple, une rame de TGV enregistre 250 informations par wagon toutes les cinq minutes afin d'anticiper des opérations de maintenance. Or il est impossible de conserver toutes ces informations. Il faut donc sélectionner les événements en fonction de leur intérêt, sachant que celui-ci évolue au cours du temps.* »

Autres gros consommateurs de techniques de fouille, les scientifiques eux-mêmes. Archétype du genre, le LHC, le collisionneur de particules géant du Cern, à Genève. Lorsqu'elle fonctionnera à plein régime, cette machine projettera des protons les uns contre les autres 40 millions de fois par seconde. Mais les physiciens estiment que seule une centaine de ces événements présenteront un intérêt et devront être enregistrés. Or ces

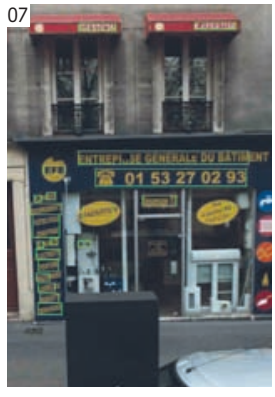
COMMENT FAIRE PARLER LES IMAGES

Désormais, nous possédons tous des milliers de photos. Les plus grosses banques d'images en recèlent des millions. Pour s'y retrouver, des outils existent. Tels ceux permettant à certains logiciels d'identifier un visage. Mais, comme le fait remarquer Matthieu Cord, du Laboratoire d'informatique de Paris-6, *« le taux de réussite est seulement compris entre 50 et 60 % ».* Typiquement, un algorithme spécialisé s'y retrouve très bien avec des informations dites de bas niveau : couleur, contraste, vecteurs de déplacement des pixels dans le cas d'une vidéo, etc. Plus délicate est leur transformation en informations de haut niveau qui rendent possible l'identification à coup sûr d'un objet ou d'un événement particulier. Ce qui n'empêche pas des applications de plus en plus performantes. Par exemple celle développée par l'équipe de Jenny Benois-Pineau, du Laboratoire bordelais de recherche en informatique[1], à Talence, en collaboration avec l'Inserm, dans le cadre du projet ANR Blanc Immed. Comme elle le précise, *« il s'agit de filmer des actions de patients atteints de la maladie d'Alzheimer chez eux et d'identifier des comportements associés à la maladie et qui sont utiles aux soignants pour suivre l'évolution des malades. »* De son côté, Matthieu Cord collabore au projet ANR iTowns, une carte numérique de Paris construite à partir de photographies, tel le service de Google Street View, à la précision du centimètre ! « *Nous développons des outils pour détecter automatiquement les personnes et les voitures afin de flouter les données personnelles,* détaille celui-ci. *Mais aussi une multitude d'objets plus ou moins enfouis dans ces images – les enseignes, les panneaux de signalisation, la végétation, les façades, etc. – pour faciliter des navigations avancées. »*

1. Unité CNRS/Université Bordeaux-I/IPB Enseirb-Matmeca Bordeaux/Université Victor-Segalen.

CONTACTS :
Jenny Benois-Pineau
› jenny.benois-pineau@labri.fr
Matthieu Cord
› matthieu.cord@lip6.fr

07



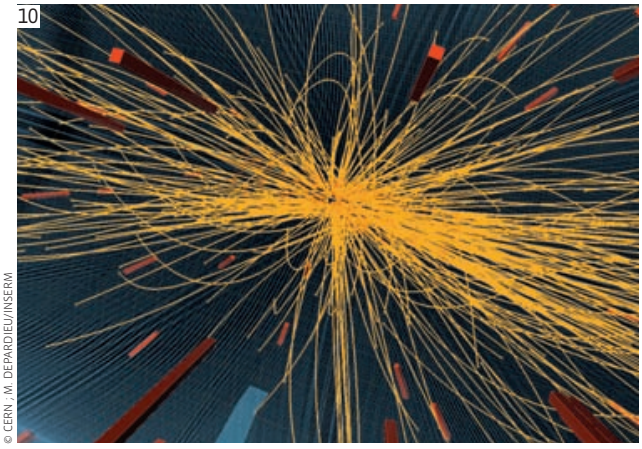
08

09

© IGN - STEREOPOLIS

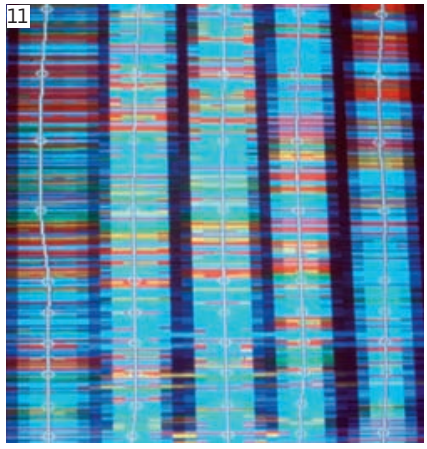
07 08 09 iTowns extrait automatiquement des informations présentes dans l'image.

10



© CERN - M. DEPARDIEU/INSERM

11



12

© C. LEBEDINSKY/INRIA

10 11 Certaines expériences, comme les collisions de particules ou le décryptage du génome, produisent d'importants volumes de données qu'il faut pouvoir trier et analyser.
12 L'étude des données scientifiques nécessite parfois de très gros moyens de calcul ainsi que la mise en réseau de machines, ici le projet Grid 5000.

DES RÉSEAUX POUR CALCULER

Les grilles informatiques sont des infrastructures virtuelles constituées d'un ensemble d'ordinateurs ou de grappes de PC géographiquement éloignés mais fonctionnant en réseau. Apparues voici quelques années sous l'impulsion de la physique des particules, elles permettent aux chercheurs et aux industriels d'accéder à moindre coût à d'importants moyens de calcul dans des domaines aussi variés que l'ingénierie, l'étude des maladies neurodégénératives ou la biochimie.
En France, l'Institut des grilles du CNRS, dirigé par Vincent Breton, fédère depuis trois ans l'activité dans ce domaine. Aux côtés de la Grid 5000, un outil spécifiquement dédié à la recherche dans le secteur des grilles, il met à la disposition des scientifiques et des industriels une grille de production rassemblant une vingtaine de milliers de processeurs disséminés dans une vingtaine de centres du CNRS, du CEA et d'universités.
Le 24 septembre dernier, ce dispositif déjà conséquent a franchi une étape supplémentaire avec la création par plusieurs organismes de recherche et universités[1] du GIS (Groupement d'intérêt scientifique) France Grilles, dont le but est de coordonner le déploiement d'une infrastructure de grille d'envergure nationale, puis de l'intégrer dans une grille européenne. Avec un objectif chiffré, annonce Vincent Breton, qui a été nommé à sa tête : *« Doubler les ressources et le nombre d'utilisateurs d'ici à 2015. »*

1. CEA, Conférence des présidents d'université (CPU), CNRS, Inra, Inria, Inserm, Renater et ministère de la Recherche.

CONTACT :
Vincent Breton
› vincent.breton@idgrilles.fr

derniers devront être sélectionnés en temps réel par des algorithmes spécialisés. « *Ce sont typiquement des algorithmes d'apprentissage, où l'ordinateur, au fur et à mesure qu'il est confronté à de nouvelles données à conserver ou à rejeter, accomplit sa tâche de mieux en mieux* », explique Michel Beaudouin-Lafon, dont l'unité collabore avec le Laboratoire de l'accélérateur linéaire[4] d'Orsay, sur la fouille de données d'accélérateurs.

UNE DÉMARCHE EMPIRIQUE

Mais les physiciens des particules ne sont pas les seuls à manipuler d'importantes quantités de données. Ainsi, l'équipe de Pascal Poncelet, en partenariat avec une équipe de l'Inserm, a développé un algorithme capable de caractériser les gènes impliqués dans différentes catégories de tumeurs du sein à partir de données de patients (informations génétiques, âge, poids, taille de la tumeur, traitement, devenir du malade…). « *Il offre aux cliniciens des informations sur les évolutions possibles d'une tumeur* », ajoute le chercheur. De même, l'équipe d'Amedeo Napoli, dans un projet en collaboration avec des astronomes, a mis au point des logiciels de fouille afin d'explorer des données sur des étoiles, dans le but de relever des caractéristiques ou des associations qui auraient pu échapper à un opérateur humain.

La fouille de données accomplit-elle pour autant des miracles ? Pas exactement. Car la discipline, qui a émergé à la fin des années 1980, est encore dans sa prime jeunesse. Conséquence, les chantiers sont légions. Pour Michel Beaudouin-Lafon, « *la plupart des démarches sont aujourd'hui empiriques. On ajuste des paramètres à la main et, lorsque cela fonctionne, on ne sait pas très bien pourquoi. Or, dans beaucoup de cas, il n'existe pas de critère quantitatif pour juger de la qualité d'informations extraites d'une base de données. Cela est laissé à l'appréciation des spécialistes du domaine* ». Et Amedeo Napoli de renchérir : « *Il y a encore beaucoup de travail à faire pour appréhender les très gros volumes. Actuellement, on peut gérer quelques milliers*

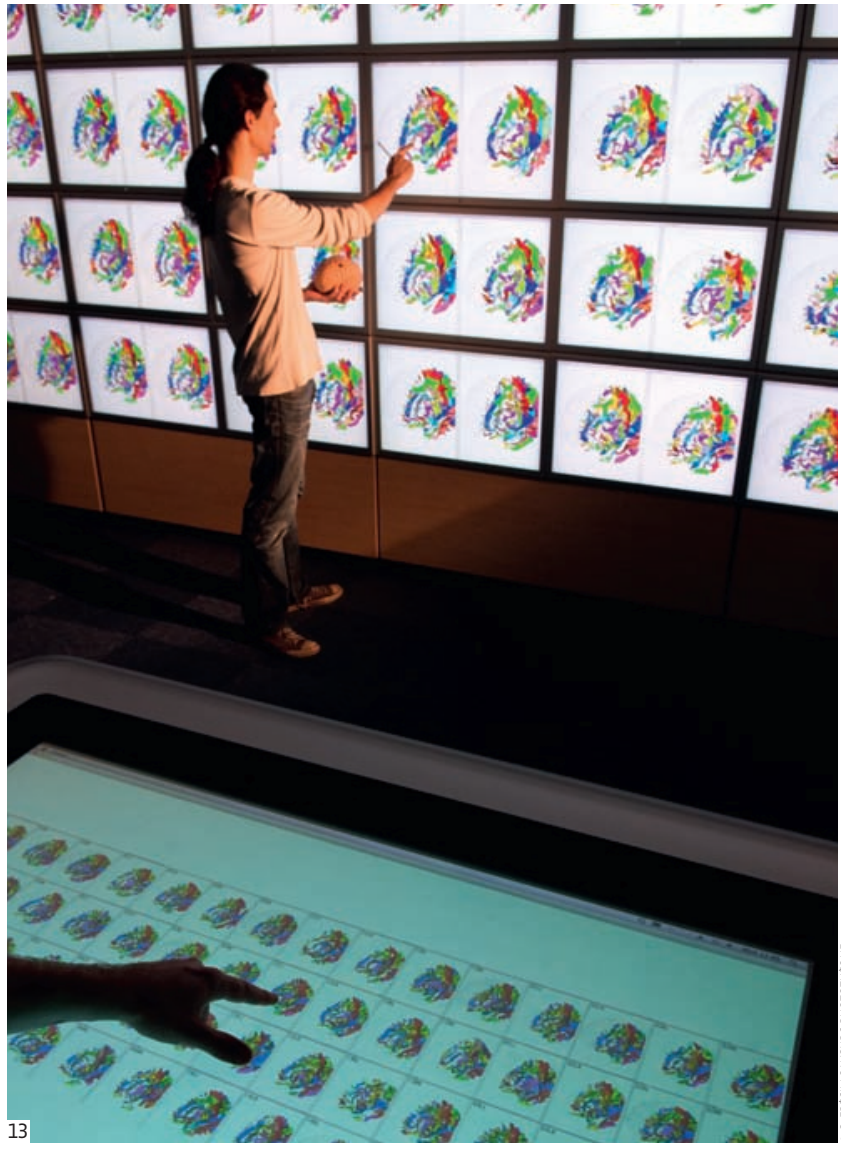
d'objets possédant quelques centaines d'attributs. Mais au-delà, on est confronté aux limites physiques des machines. »

Pour pallier cette difficulté, deux approches complémentaires sont possibles. Tout d'abord, là où une seule machine ne suffit pas, on peut faire travailler en parallèle plusieurs ordinateurs. C'est le principe de la grille *(lire l'encadré ci-contre),* poussé à l'extrême au LHC, qui dispose de 50 000 PC dispatchés dans différents centres de recherche à travers le monde, afin d'analyser l'équivalent des 3 millions de DVD de données dont les scientifiques disposeront au terme de l'expérience. Autre option, le supercalculateur, tel celui dont dispose depuis 2008 l'Institut du développement et des ressources en informatique scientifique (Idris) du CNRS, à Orsay[5]. Un monstre informatique capable de réaliser 207 milliers de milliards de calculs par seconde sur des nombres à virgule. *« Dans certains cas, typiquement la simulation d'armes nucléaires ou celle de la météo, il est difficile de morceler les données. Le superordinateur reste donc la solution »,* complète Michel Beaudouin-Lafon.

LA GESTION DU FACTEUR HUMAIN

Cependant, développer des ordinateurs ne suffit pas. De fait, à l'autre bout de la chaîne d'un processus de fouille se trouve un utilisateur humain. Se pose donc la question de la meilleure façon de lui présenter le résultat d'une recherche. Il suffit pour comprendre la problématique de penser à Google : le programme peut faire remonter plusieurs milliers d'adresses pour une requête, mais ne peut en afficher qu'une dizaine à l'écran. Comme le regrette Michel Beaudoin-Lafon, *« c'est dommage de bénéficier d'algorithmes sophistiqués pour faire remonter de l'information et de ne pas être capable de la présenter de façon correcte ».*

Pour ce faire, le Laboratoire de recherche en informatique a mis au point une plateforme d'un nouveau genre, baptisée Wild.



13 L'application Substance Grise utilisée sur la plateforme Wild sert à comparer simultanément les reconstructions 3D des cerveaux de 64 patients.

© C. FRÉSILLON/CNRS PHOTOTHÈQUE

Concrètement, un mur tapissé de 32 écrans d'ordinateurs représentant 130 millions de pixels et qui permet d'appréhender en un coup d'œil d'importantes quantités d'information. *« Nous travaillons avec huit laboratoires du plateau de Saclay sur ce projet »,* indique Michel Beaudouin-Lafon. En neurosciences, Wild permet d'afficher 64 IRM de cerveaux, *« ce qui présente un avantage indéniable lorsqu'il s'agit d'identifier une pathologie alors même que l'on observe une variabilité importante parmi les cerveaux sains »,* poursuit l'informaticien. De même, en astrophysique, certains observatoires fournissent désormais des images dont la taille excède largement celle d'un écran. Pour visualiser ces images en entier à leur résolution maximale, des outils tel que Wild font la différence. *« Je suis convaincu que ce type d'approche est amené à se développer, dans la recherche, mais aussi dans le monde industriel,* conclut Michel Beaudouin-Lafon. *Tout simplement parce que les données ne cessent d'augmenter, et les questions que l'on veut leur poser sont de plus en plus complexes et mal définies. »* Bref, il s'agit ni plus ni moins que d'éviter à la société de l'information de crouler sous son propre poids !

1. Unité CNRS/Université Henri-Poincaré/Université Nancy-II/Inria.
2. Unité CNRS/Université Paris-Sud-XI.
3. Unité CNRS/Université Montpellier-II.
4. Unité CNRS/Université Paris-Sud-XI.
5. Lire « Le CNRS s'offre un supercalculateur », *Le journal du CNRS,* n° 218, mars 2008, p. 34-35.

800 000 petaoctets,

c'est l'estimation du volume mondial de données numériques en 2009. Les experts s'attendent à une croissance de 45 % par an d'ici à 2020.

CONTACTS :
Michel Beaudouin-Lafon
› michel.beaudouin-lafon@lri.fr
Amedeo Napoli
› amedeo.napoli@loria.fr
Pascal Poncelet
› pascal.poncelet@lirmm.fr

© IBM

14

Ordinateur quantique : l'ultime défi

C'est un rêve d'informaticien… Un ordinateur si rapide que casser un code, prévoir la météo à long terme ou battre à plate couture n'importe quel grand maître des échecs ne lui prendrait pas plus d'une seconde. Disons le tout net, ce fantasme est loin d'être une réalité. Ce qui n'empêche pas mathématiciens et physiciens de commencer à esquisser les contours de ce que sera peut-être un jour cette extraordinaire machine. Son nom ? L'ordinateur quantique. Son concept ? Tirer partie des étonnantes lois quantiques qui autorisent une particule, un atome ou une molécule, à occuper deux états en même temps. À la manière du chat imaginé en 1935 par Erwin Schrödinger, l'un des pères de la mécanique quantique, à la fois mort et vivant. Ainsi, alors que, dans un ordinateur ordinaire, les informations sont stockées sous la forme de bits prenant les valeurs 0 ou 1, des bits quantiques (ou qubits) pourraient simultanément prendre les valeurs 0 et 1. L'intérêt : la possibilité de stocker, en principe, sur la même mémoire des informations représentant un grand nombre de solutions potentielles d'un problème. Et, en appliquant des algorithmes adaptés, traiter toutes ces solutions de concert. De quoi renvoyer les plus puissants calculateurs d'aujourd'hui à la préhistoire de l'informatique.

FACTORISATION Décomposition en facteurs premiers des grands nombres.

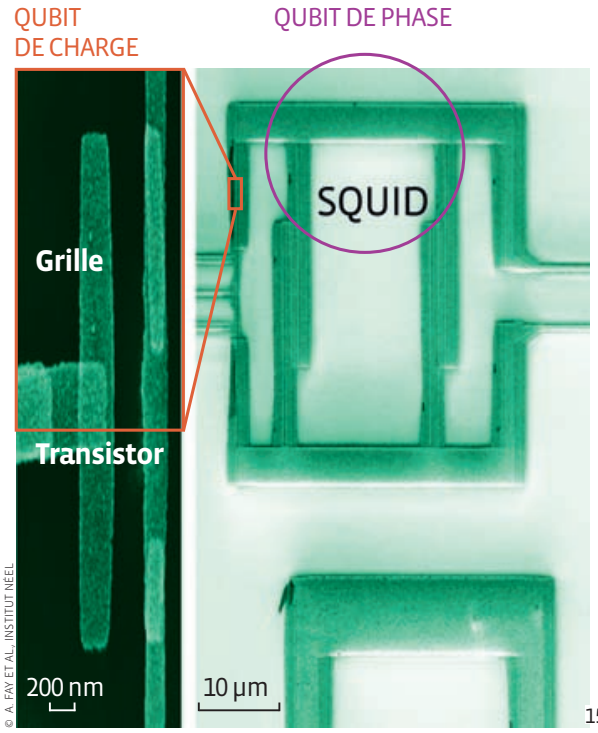
QUBIT Bit quantique qui a la particularité d'avoir un état dit de superposition où les valeurs 0 et 1 sont prises en même temps, en plus des valeurs standard 0 et 1 du bit classique.

UNE IDÉE QUI A FAIT SON CHEMIN

Pour autant, un tel ordinateur sortira-t-il jamais des laboratoires ? Et si c'était un jour le cas, serait-il vraiment capable de tous les prodiges ? Rien n'est moins sûr. Après tout, au début des années 1980, l'ordinateur quantique n'était qu'une idée lancée en l'air par le prix Nobel de physique Richard Feynman. Comme le raconte Julia Kempe, du Laboratoire de recherche en informatique (LRI), à Orsay, élue Femme en or de la recherche 2010, « *Feynman a fait remarquer qu'avec un ordinateur quantique on pourrait calculer bien plus rapidement les propriétés d'une assemblée de particules quantiques, des électrons par exemple, qu'avec un ordinateur classique. On pourrait en effet encoder chaque électron sur un qubit, alors qu'il faut une grande quantité de bits classiques pour encoder les nombreux états dans lesquels il peut se trouver en même temps. Mais ce n'était qu'une idée.* »

À dire vrai, une très bonne idée. Car, en 1994, Peter Shor, alors aux Laboratoires AT & T, aux États-Unis, montre formellement qu'un ordinateur quantique pourrait factoriser un nombre, c'est-à-dire le décomposer en un produit de nombres premiers en un temps record. De quoi faire de l'ordinateur quantique la bête noire de tous les cryptographes, puisque, du fait de sa gourmandise en temps de calcul, la factorisation est actuellement la clé de tous les codes secrets, de celui de nos cartes bleues à ceux permettant d'échanger des secrets d'État. De même, en 1997, Lov Grover, des laboratoires Bell, démontre qu'un ordinateur utilisant des qubits pourrait considérablement augmenter l'efficacité des algorithmes utilisés pour la recherche d'informations dans une base de données.

Sauf que si, dans les années 1990, mathématiciens et physiciens commencent à démontrer l'intérêt de disposer d'un ordinateur quantique, la "bête" elle-même n'est encore qu'une chimère. De fait, aujourd'hui comme hier, personne ne sait concrètement de quoi seront composés les fameux qubits : des atomes ou des ions, des molécules, des électrons, des



14 L'ordinateur quantique, comme celui des chercheurs du Massachusetts Institute of Technology, à base de molécules organiques, reste pour le moment très expérimental.
15 Certains circuits supraconducteurs permettent d'analyser et de tester les nouvelles propriétés de la nanoélectronique quantique.

circuits supraconducteurs ? Sur un support solide, liquide ou gazeux ? Mystère. De nombreuses équipes à travers le monde expérimentent actuellement toutes sortes de supports matériels susceptibles d'être utilisés comme composants de base d'un futur processeur quantique. Par exemple, explique Bernard Barbara, de l'Institut Néel, à Grenoble, *« nous étudions actuellement des qubits dont les deux états 0 et 1 correspondent aux états de spin* [sorte de rotation de la particule sur elle-même] *de molécules ou d'ions de certains métaux dans des matrices solides ».*

PRINCIPAL OBSTACLE : LA DÉCOHÉRENCE

Mais, loin d'être en mesure de proposer un ordinateur clé en main, les physiciens tentent pour le moment de comprendre et, dans la mesure du possible, de contrôler l'écueil principal sur le chemin du calculateur quantique : la décohérence. Comme le détaille le spécialiste, *« tout système dans une superposition quantique de différents états est extrêmement fragile. Ainsi, sous l'effet de ses interactions avec l'environnement, il peut perdre en une fraction de seconde les propriétés nécessaires à tout calcul quantique. Et cela est d'autant plus vrai que ce système contient plus de qubits ».*

DÉCOHÉRENCE Temps pendant lequel les propriétés d'un système quantique ne sont pas corrompues par l'environnement extérieur.

À ce jour, la plus belle prouesse calculatoire réalisée avec des qubits est l'œuvre d'Isaac Chuang, de l'Institut de technologie du Massachusetts. En 2001, en utilisant le spin du noyau de sept atomes d'une molécule, ce chercheur est parvenu à factoriser 15, soit à montrer que ce nombre se décompose en 3 fois 5. *« Or, pour être performant,* indique Bernard Barbara, *un ordinateur quantique devra comporter quelques milliers de qubits. Et offrir la possibilité de les coupler afin de réaliser des calculs logiques. »*

De l'avis général, deux systèmes offrent aujourd'hui les perspectives les plus intéressantes. D'une part, les qubits supraconducteurs, soit de microscopiques circuits électroniques dans lesquels un courant électrique peut en même temps circuler dans un sens ou dans l'autre : *« Ils offrent l'avantage d'une grande facilité de fabrication. Il est donc aisé de les dupliquer et de disposer de puces comprenant de nombreux qubits supraconducteurs »,* explique le physicien. Mais surtout, d'autre part, *« les ions gazeux piégés par de puissants faisceaux lasers, avec lesquels on obtient des temps de cohérence de plusieurs minutes malgré des systèmes encore relativement restreints ». « L'ordinateur quantique n'est pas pour demain,* confie Bernard Barbara. *Mais je pense que d'ici à quelques dizaines d'années il pourrait devenir une réalité. »* Miklos Santha, lui aussi du LRI, est plus nuancé : *« Qui sait si nous ne finirons pas par découvrir que la nature interdit la possibilité même d'un ordinateur quantique… »*

LES RECHERCHES CONTINUENT

Et, quand bien même, celui-ci ne serait pas exactement l'ordinateur ultime. Car seules certaines catégories de problèmes pourraient voir leur résolution accélérée par un ordinateur quantique. *« Certes, le gain est considérable dans le cas de la factorisation. Mais il l'est déjà moins dans le cas de la recherche de données non triées,* reconnaît Miklos Santha, *de même que pour déterminer l'itinéraire le plus court sur une carte, ou bien pour le jeu d'échec ou le Go. Et quasi nul pour d'autres types de données. Il y a quelques grands miracles, mais ils sont rares. »* De quoi rendre vaine toute recherche sur l'ordinateur quantique ? Loin de là. En effet, comme le précise Bernard Barbara, *« que nous construisions ou pas un ordinateur quantique, nos recherches permettent d'apprendre à maîtriser les lois quantiques et de mieux en comprendre les fondements ».*

Quant à Julia Kempe, elle insiste sur l'intérêt de développer des algorithmes quantiques : *« Ils constituent des outils mathématiques très performants pour aborder des questions fondamentales liées à la complexité. Mais aussi pour étudier ce qu'un ordinateur classique peut faire ou ne pas faire. Enfin, les algorithmes quantiques de factorisation sont à la base du développement de la cryptographie quantique qui est déjà utilisée pour l'échange de données secrètes. »* Ainsi, personne ne sait si l'ordinateur quantique sortira un jour des laboratoires. Peu importe, même inatteignable, il demeure une source d'inspiration sans fin. Bref, un véritable rêve de scientifique.

Pour en savoir +

À LIRE |

L'Informatique en France
De la Seconde Guerre mondiale au Plan Calcul
Pierre-Éric Mounier-Kuhn, Pups, coll. « Roland Mousnier », 2010

Pourquoi et comment le monde devient numérique
Gérard Berry, Collège de France/ Fayard, 2008

À VOIR |

Jacques Stern ou la science du secret
(2006, 15 min), réalisé par François Tisseyre, produit par CNRS Images

Marc-Olivier Killijian roboticien
(2010, 5 min), réalisé par Didier Boclet, produit par CNRS Images

Émergence d'un nouveau monde
(2006, 53 min), réalisé par Jean-Pierre Mirouze, produit par Flight Movie et CNRS Images

CONTACT | Véronique Goret, CNRS Images-Vidéothèque
Tél. : 01 45 07 59 69
› videotheque.vente@cnrs-bellevue.fr
› http://videotheque.cnrs.fr

+ WEB
Des **photos** et des **films** sont à découvrir sur le journal feuilletable en ligne
› www2.cnrs.fr/journal

CONTACTS :
Bernard Barbara
› bernard.barbara@grenoble.cnrs.fr
Julia Kempe
› julia.kempe@lri.fr
Miklos Santha
› miklos.santha@lri.fr