

L'histoire de l'Internet et du Web

Deux outils qui ont changé le monde

Haute école de gestion de Genève

Filière Information documentaire

2013/12/03

Structure de cette présentation

- Internet – pères, dates, technologie, applications, défis.
- Web – l' inventeur, le lieu, les circonstances, la vie depuis.
- CERN – informatique pour la physique, pas simple.
- Conclusion – Discussion.

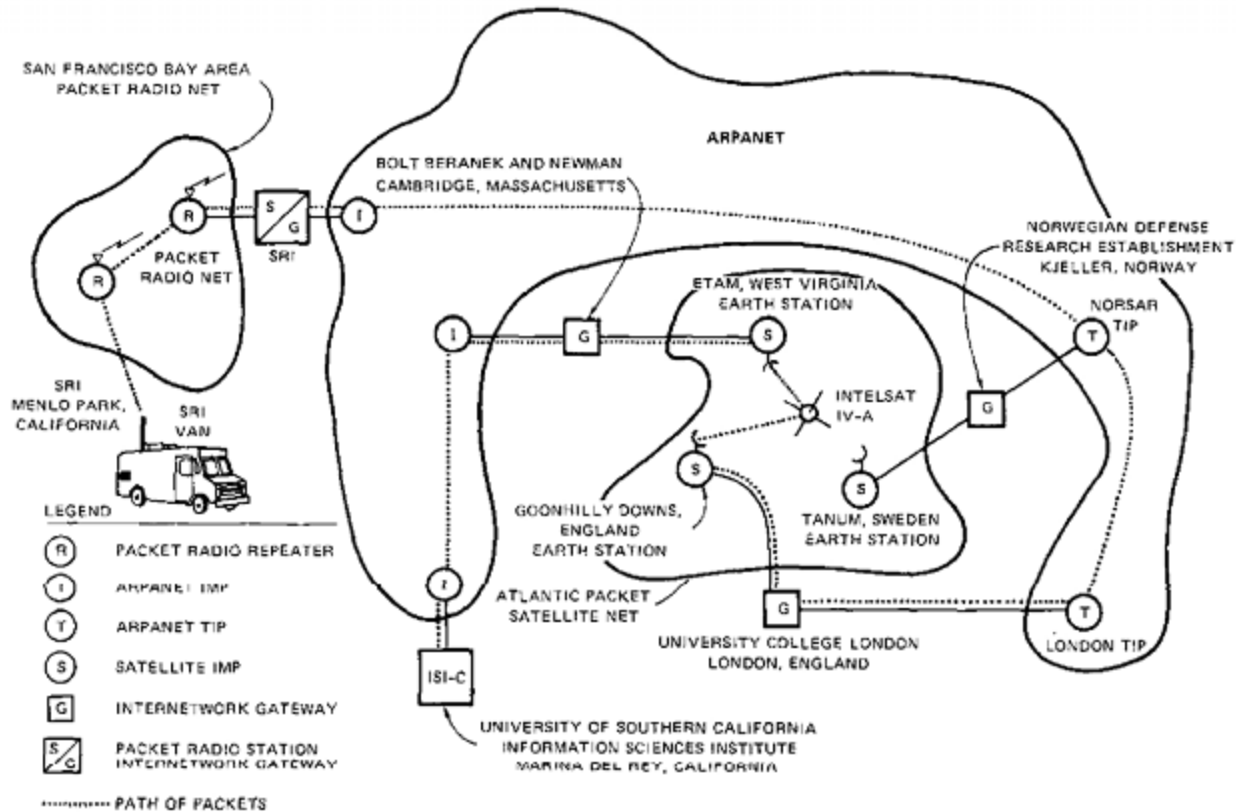
Internet: Qui / Quand

- Vint Cerf est devenu (avec Bob Kahn) l'inventeur de l' Internet en écrivant, au début des années 1970, la série de *protocoles* de réseau TCP/IP (Transmission Control Protocol / Internet Protocol). Il travaille actuellement pour Google, comme vice-président et «évangéliste» principal de Internet.
- Ils se sont basés sur une infrastructure de réseau physique qui reliait les 2 côtes des Etats Units, financé par le Pentagone en 1958, appelé ARPAnet (Advanced Research Projects Agency).
- Les « pères» de ARPAnet étaient L.Kleinrock, P.Baran et L.Roberts grâce à la technologie de transmission de données par l' échange de *paquets*.
- La création de ARPAnet visait à concurrencer les russes pour l' envoi du Sputnik dans l' espace en 1957.



1er test de 3 réseaux de Internet

Slide de Vint Cerf présenté au CERN le 2009/02/26



November 22, 1977

Le van de l' équipement paquet radio

Slide de Vint Cerf présenté au CERN le 2009/02/26

Packet Radio Van

Google



2013/12/03

Courriel: L' application 'grand public' des 20 premières années

- Au début le courriel était transmis sur des réseaux physiques dissociés, qui appartenaient aux grandes entreprises d' informatique ou des agences de télécommunications.
- Les adresses étaient très incompatibles, e.g.
 - DIMOU AT CERNVM.BITNET (IBM)
 - VXCERN::DIMOU (DEC, devenue Compac, achetée par HP)
 - C=CH/O=CERN/OU=Computing/CN=Maria.Dimou (CCITT X.400)
 - dimou@priam
- Des passerelles avec des logiciel de conversion d' adresses assuraient l' interconnection. On en avait une importante au CERN, à cause de nos multiples plateformes d' ordinateurs.

« Email: Inelegant in its efficiency »

Après l'adoption:

- du protocole SMTP (Simple Mail Transfer Protocol),
- du format d'adresses [user@domain.top-level-domain](#) (RFC822, c.à.d. Request For Comment),
- du DNS (Domain Name Space, inventé par Paul Mockapetris) et
- de l'Internet, comme infrastructure de réseau,

il ne restait qu'un petit pas jusqu'aux adresses faciles à deviner, par ex. [maria.dimou@cern.ch](#) et le transfert de contenu autre que le texte (images, documents formatés, vidéo) grâce au développement de MIME (**Multipurpose Internet Mail Extensions**) par Nathaniel Borenstein et autres.

L'épuisement des adresses IP

- Tout 'appareil' connecté sur Internet s'identifie par son (ses) adresse(s) IP.
- Le format d'adresses, utilisé jusqu'à maintenant, était défini dans la norme RFC791 de 1981 est connu comme IPv4.
- Une adresse IP se compose de 4 octets (32 bits) en chiffres décimales, par ex. 137.138.1.1.
- L'espace d'adressage est réservé le moment où nous achetons un domaine, par ex. cern.ch.
- Certaines de ces adresses sont réservés pour des fonctions standard de l'équipement réseau, par ex. les serveurs de temps.
- Tous les 4 milliards d'adresses IPv4 sont épuisées en 2011.
- La norme IPv6 conçue dans les années 1990, prévoit des adresses à 128 bits des chiffres hexadécimales, i.e. Base 16 (les lettres A à F représentent la partie 11-16).

Le Web: Qui / Quand / Où

- [Tim Berners-Lee](#) (TimBL) a inventé le web en 1989 au CERN, une idée qui a commencé en 1980 avec son projet Enquire , suivi par une proposition écrite en 1989 et première démonstration et distribution du logiciel en 1991.
- Son idée s' est basée sur Hypertext un terme inventé en 1963 par [Ted Nelson](#), lui-même influencé par l' essai de [Vannevar Bush](#) 'As we may think' qui préfère l' association à l' indéxation
- A l'origine, la Toile (le Web) fut conçue et développée pour répondre au besoin de partage d'informations entre scientifiques travaillant dans différentes universités et instituts aux quatre coins du monde.

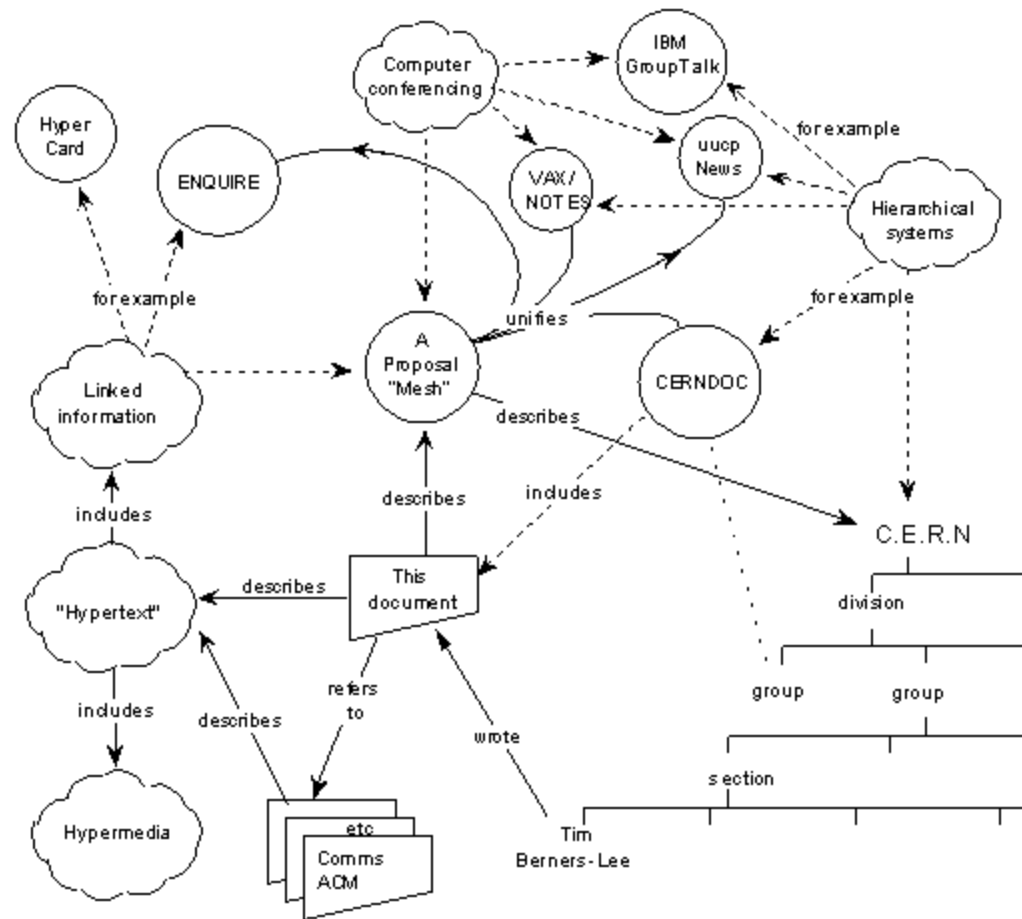


Un vrai « one man show »

- TimBL a dû soumettre sa proposition (document) 2 fois au département informatique du CERN pour obtenir l'accord de développer le Web.
- Hypertext, dans ce document, est décrit comme 'human-readable information linked together in an unconstrained way'.
- Il avait déjà décidé sur le nom World Wide Web en 1990.
- Il a écrit le langage HTML (HyperText Markup Language) fortement inspiré par SGML (Standard Generalized Markup Language), un standard européen de description de texte à balises.
- Il a conçu le protocole http (HyperText Transfer Protocol) et écrit le logiciel httpd.

•2013/12/03

Le dessin sur la proposition originale de TimBL



L'ordinateur où le logiciel du web a été développé



Influences et évolution dans ces 20 ans

- Enquire, le programme développé pour le système de contrôle de l'accélérateur PS au CERN, était inspiré d'un livre de l'époque victorienne «Enquire Within Upon Everything » qui montre l'intérêt d'utiliser des *associations* plutôt que des *hiérarchies* pour gérer l'information.
- A l'époque (1990), au CERN et dans toute la communauté de physique des particules, il y avait plusieurs marques d'ordinateurs et systèmes d'exploitation. L'accès à l'information à *distance*, l'*hétérogénéité* et la *non-centralisation* font partie de la définition du web.
- En 1994 TimBL a accepté l'offre de [Michael Dertouzos](#) (MIT Computer Science Lab) d'y travailler pour fonder le Web Consortium (W3C), qu'il dirige depuis. Sujets de développement:
 - [Web sémantique \(1994\)](#) ([une présentation de 2009 ici!](#))
 - **Web foundation (2008)**
 - **Web science (2008)**

Depuis... le défi informatique du LHC

Input par F. Hemmer, CERN IT Leader

- L'accélérateur est en opération depuis 2009.11.21 et fonctionnera pendant 20 ans
- Les expériences produisent environ **25 millions de Gigaoctets** (Petabytes) de données chaque année (environ 3 millions de DVDs = 850 ans de films!)
- L'analyse des données du LHC requiert une puissance informatique équivalente à **±100'000 processeurs actuels les plus puissants**
- Cela nécessite la coopération de plusieurs centres de calcul étant donné que le CERN ne peut fournir **que ± 20% de cette capacité**

La solution: la Grille

Slide de F. Hemmer, CERN IT Leader

- Utiliser la Grille pour réunir les ressources informatiques des instituts en physique des particules du monde entier

Le **World Wide Web** fournit un accès simplifié à de l'information stockée en des millions d'endroits

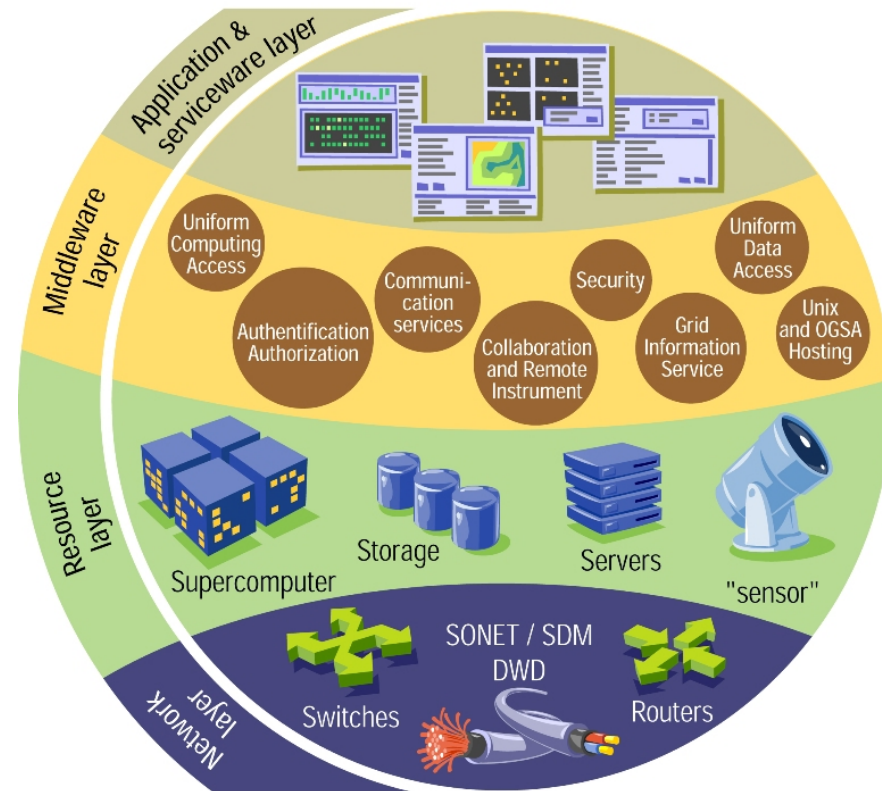
La **Grille** est une infrastructure fournissant un accès simplifié à de la puissance de calcul et de stockage distribué tout autour du globe

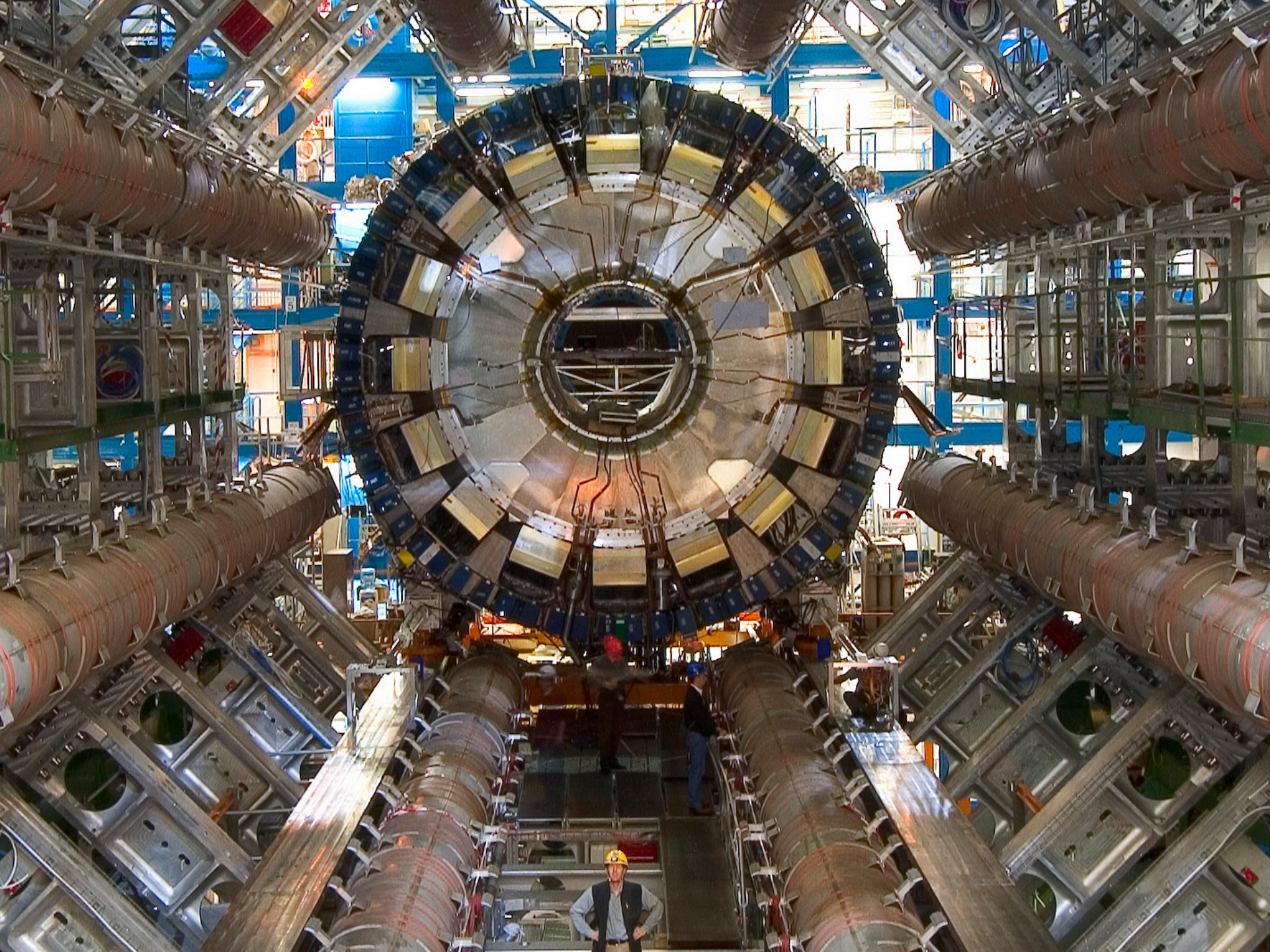


Comment fonctionne la Grille?

Slide de F. Hemmer, CERN IT Leader

- Pour l'utilisateur final, elle donne à de nombreux centres de calcul l'apparence d'un seul système
- L'intergiciel, un logiciel de pointe, trouve automatiquement les données et la puissance de calcul nécessaires aux scientifiques
- L'intergiciel distribue la charge entre les différentes ressources. Il gère également la sécurité, la comptabilisation, la surveillance et bien plus...



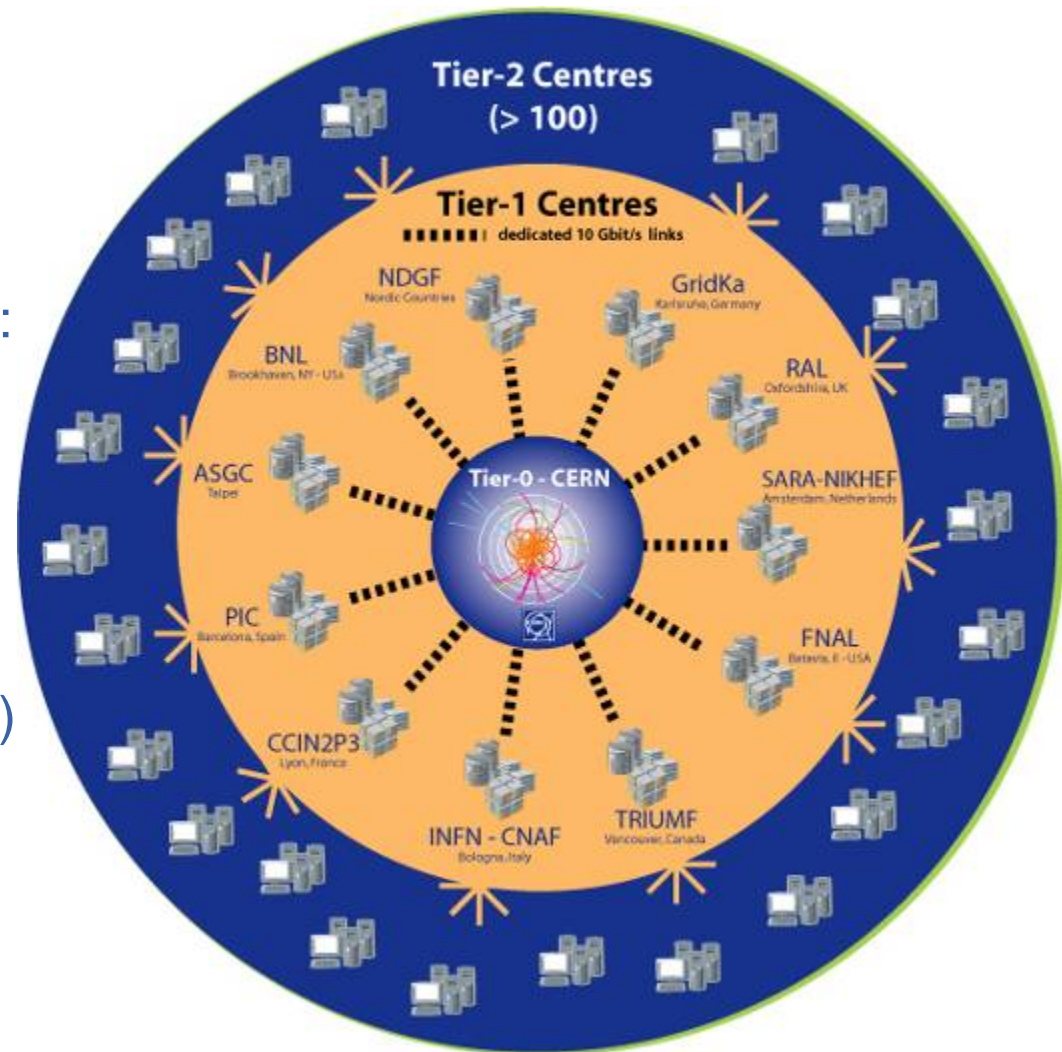




Le projet LHC Computing Grid (LCG)

Slide de F. Hemmer, CERN IT Leader


- Plus de 160 centres de calcul
- 13 grands centres pour la gestion initiale: le CERN (Tiers-0) et 12 Tiers-1 (récente inclusion de la Corée du Sud).
- 38 réseaux de plus petits centres (Tiers-2)
- \pm 35 pays impliqués
- 2 millions jobs/jour.



Conclusion – Qu'est-ce qu' on a compris

- Il y a des processus irréversibles dans la nature et dans la société == les étudiants des années 1980 n' avaient que la bibliothèque.
- Une bonne invention peut donner les moyens à des mauvaises applications; c' est la responsabilité personnelle qui doit jouer == toute la connaissance peut s' afficher sur le web mais toute la perversité aussi.
- Il y a des paradoxes partout, vivre avec rend plus tolérant == l' Internet était indirectement un produit de la guerre froide, il a pourtant servi aussi de voie pour l' ouverture et le liberté.
- Une solution technique peut être géniale et éphémère en même temps == la question d' adresses IP.
- Méfions-nous des outils qui semblent nous faciliter la vie mais qui nous enferment dans une 'société privée' == préférons le code source libre ([open source](#)).

Plus d'information:

TimBL	w3.org/People/Berners-Lee
Web consortium	w3.org
 W3C [®] Semantic Web	w3.org/standards/semanticweb
Webscience	webscience.org
Web Foundation	webfoundation.org
A propos de la Grille	gridcafe.org/index_FR.html
LHC Computing Grid	cern.ch/lcg

Merci de votre attention!

cern.ch/dimou