
L'informatique à l'IN2P3 et le CCIN2P3

Giovanni Lamanna

Prospectives CCIN2P3

23 mai 2013



« Des réflexions et des discussions avec le personnel du CCIN2P3 autour des prochains défis en informatique dans notre Institut, le point de vue des scientifiques et des informaticiens et le rôle international du CCIN2P3 »

alias « les idées du nouveau CMI »?

- Présentation du Charge de Mission pour l'Informatique
- Des défis et des motivations
- L'informatique @/et IN2P3 et le CCIN2P3
- La voix du CCIN2P3
- Discussion

- Directeur de recherche
 - Responsable scientifique de l'équipe HESS & CTA du LAPP.
 - Coordinateur international du projet « CTA Data Management » (FP7-PP).
- Chercheur en physique des astroparticules:
 - Rayonnement cosmique (AMS, ANTARES, HESS, POLAR, CTA) .
 - R&D détecteurs (RPC, Si-Tracker, PMT), projets en électronique, mécatronique et informatique.
- Plus récemment en informatique»:
 - Représentant France dans LHC-CRSC « Computing Resources Scrutiny Group » au CERN.
 - Ateliers internationaux « Aspera & Computing ».
 - Journées de Prospective IN2P3 et IRFU « Astroparticle et computing ».
 - Forum e-Science internationaux: EGI, CE, FP7, ApPEC/ASPERA
 - CTACG, (CORE, HEAPNET), projet international « Data management » de CTA.

[...]

- «... Le chargé de mission pour l'informatique de l'IN2P3... **assure la coordination de la stratégie de la politique de l'institut en termes de développement informatique et sciences associées...** »
- « ...préparer **des revues, avec la direction de l'IN2P3 et du CC-IN2P3, pour évaluer les besoins des nouveaux projets avec les acteurs internes et externes de l'institut...** »
- « ... mettre en place une **cellule de coordination** visant à préparer l'institut pour remettre aux différents appels d'offres ... »
- « ... **favoriser l'organisation en réseau** de la communauté informatique... »
- « ... s'assurer de la communication/visibilité en interne comme à l'externe des activités informatiques de l'institut, et **favoriser les relations avec des partenaires extérieurs...** »

[...]

- Giovanni Lamanna, Pierre Etienne Macchi, Laurent Serin..

[...] + ...:

- L'informatique et/@ l'IN2P3: LCG-FR, France-Grilles, CCIN2P3, (CNRS), SI des Unités...., RI3...CCRI ...

Nombreuses actions, nombreux acteurs (*part-time*)-> travailler ensemble.

- Définir les objectives et les contextes pour maitriser ses actions...

Au niveau international:

Coordination et stratégies:

- IN2P3 -> rôle d'excellence dans des projets en « e-Science »: nouveaux projets (LHC, LSST, CTA, ...)
- H2020 et autres contextes (« *resources-raising* » et « *competences promotion* »).
- CCIN2P3 -> centralité internationale de notre infrastructure
- CCIN2P3 -> correcte conjugaison avec les laboratoires.

Des pistes:

- Synergies -> échange entre les équipes LHC, LSST, CTA et ++.
(*contribution du CCIN2P3!?*)

Des leviers :

- Les futures expériences de « Physique » comme « *use cases* » pour ICT.
- CTA dans toutes les roadmaps EU (e.g. ESFRI)
- LSST synergies extra-EU, mais en cohérence avec le défi EU « big data »
- LHC -> CERN cherche des synergie (?!) et pas que LCG [...]
- Autres partenaires clefs potentiels: ESA, ESO, entreprises privées.
- ApPEC++, CERN++: «H.E. particle and astroparticle Data project » (H2020 Digital Agenda) ?
- Partenariat bilatéral: d'autres institutes EU.

Au niveau international:

Des idées et des actions:

- *Research Data e-Infrastructures: the emergence of data driven science*

Framework for Action in H2020:

- *community support data services*
- *infrastructure for Open Access*
- *storing, managing and preserving research data*
- *Digital Identifier e-infrastructure*
- *RDA, towards global data e-infrastructures*
- *EU Authentication and Authorization e-infrastructures*
- *skills and new professions for research data*

--> Des sujets pertinents et stratégiques pour l'IN2P3 [...]

(Réponse au sondage publique de la CE pour l'IN2P3: [Lamanna Consultation H2020 WEB V1](#))

Au niveau national:

Coordination et stratégies:

- Favoriser les échanges: SIs des Us IN2P3 -> « *département IT virtuel* » ~ 300 informaticiens...
- Communiquer: mutualiser, saisir des choix technologiques, partager pour plus de compétitivité.
- CCIN2P3 cohérence avec la stratégie scientifique et planification (ESC report 2012).

Des pistes:

- Nouvelle organisation info@IN2P3: hybride « top-down » (CMI) & « bottom-up » (RI3)
- Explorer partenariat plus large (e.g. INRIA)
- « Interdisciplinarité » ? En fonction du contexte: *vrais-l., pseudo-l., réduite-l., fonctionnelle-l* ? [...].
(*a discuter aujourd'hui ?!*)

Des idées et des actions:

- Nouveaux défis numériques Ministérielles (actions pour le CCIN2P3 !).
- Des groupes de travail thématiques (en cohérence avec les “actions Ministérielles et CNRS” et EU!)
- Pilotage pour CLOUD et T0 EU/FR.
- Revue et Pilotage des implications interdisciplinaires (extra-IN2P3) du CCIN2P3.
- Revue et pilotage centrale des expériences pour des questions « *e-Science* ».

Groupes de travail

A) Le calcul et les services coopératifs pour l'analyse scientifique

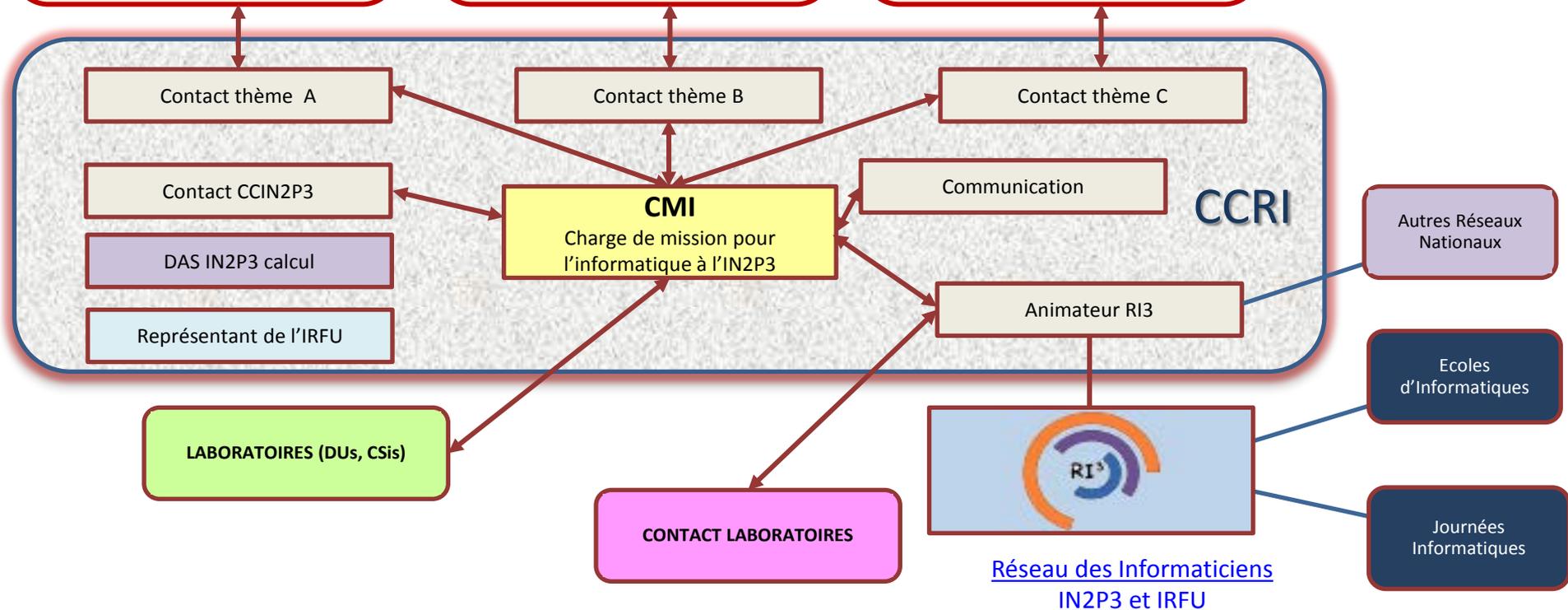
- Programmation parallèle et GPU
- Base de données non-relationnelle
- « Data Access »
- Virtualisation et Cloud
- Préservation des données
- Forge et environnement collaboratif

B) Contrôle commande des instruments, acquisition de données et temps réel.

« Standard interface instrumentation »

C) Les services de soutien informatique aux divers métiers de la recherche..

- Sécurité
- Active Directory
- Qualité des infrastructures
- Travail et collaboration a distance



A) Le calcul et les services coopératifs pour l'analyse scientifique

Programmation parallèle et GPU

Base de données non-relationnelle

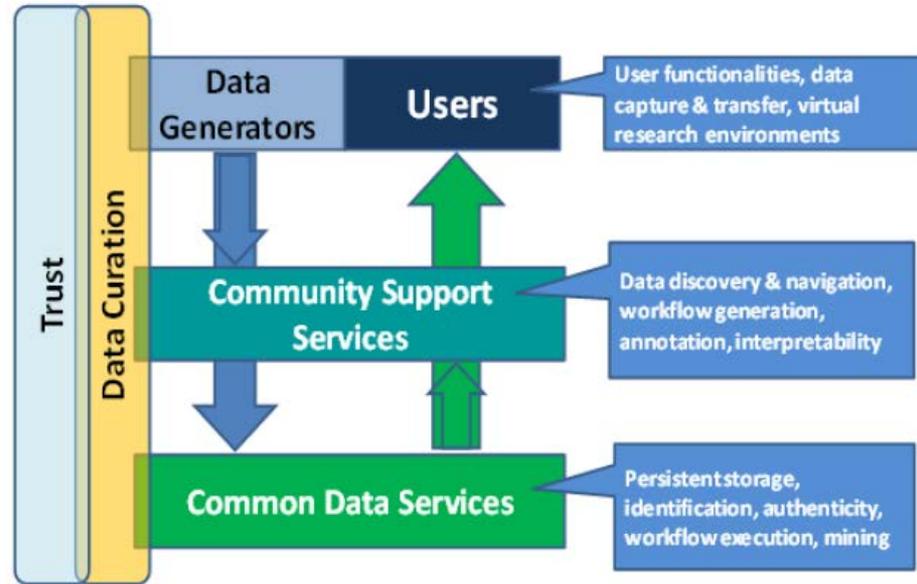
« Data Access »

Virtualisation et Cloud

Préservation des données

Forge et environnement collaboratif

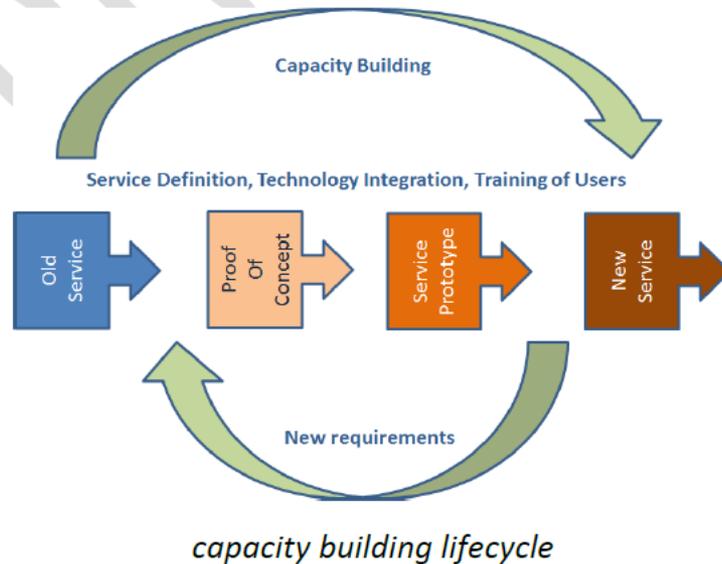
Le scientifique: SAS, Datacentre, DCI, Archive,...



framework suggested by the High-Level Group on Scientific Data

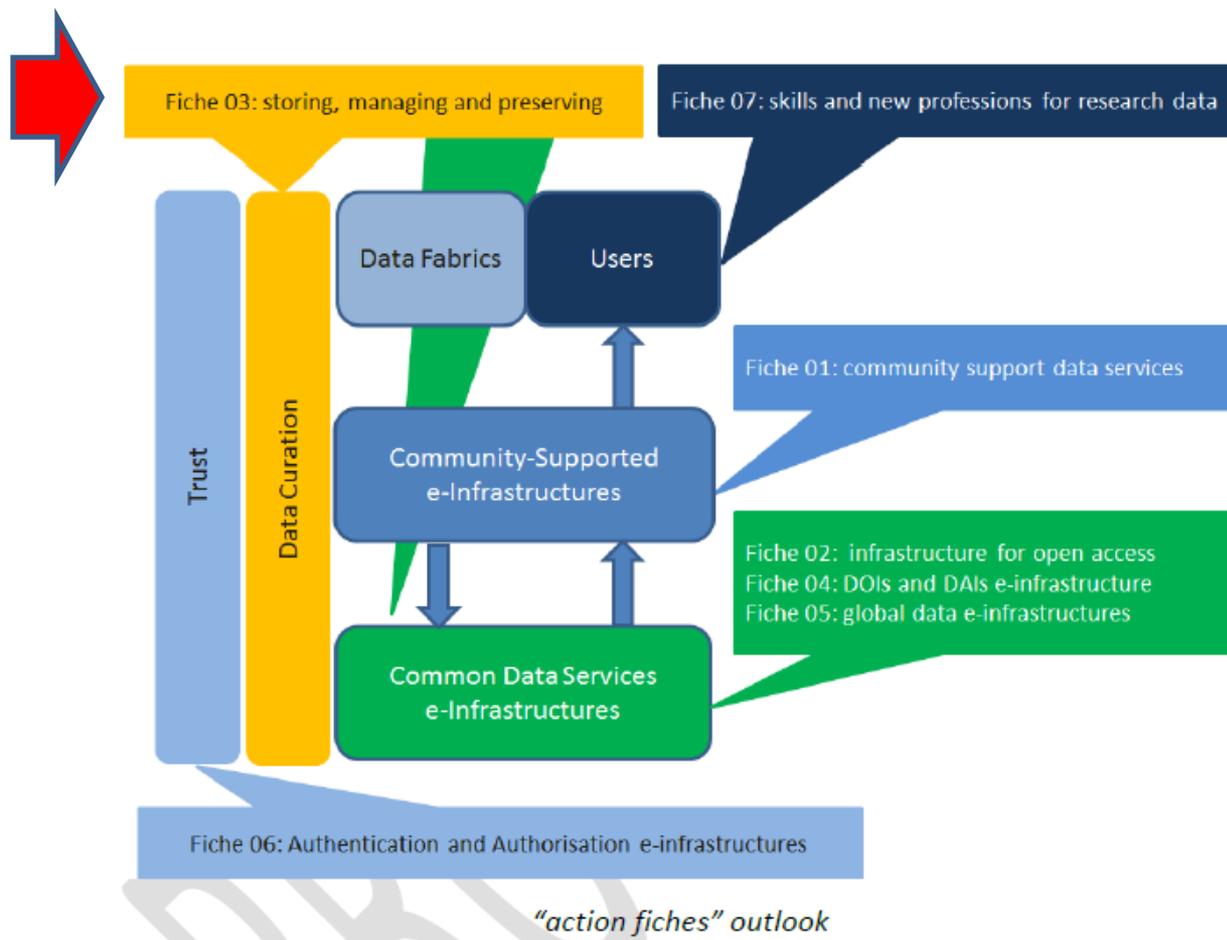
Le personnel du CCIN2P3: ... ?

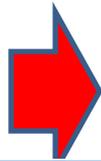
All these initiatives have a common aspect: they are the **biggest research data factories** of the present and the future. Some are led by large research infrastructures undertakings (ESA, EBI, ESO, CERN, EMBL,...) and others by collaborative undertakings of e-infrastructure service providers (university and national libraries, data-centers, super-computing centers, etc).



an immense global data factory: the forthcoming Square Kilometre Array (SKA) radio telescope will generate more data per day than the entire internet when it comes online in 2020; it will detect radio waves using dishes with a collecting area of 1 million square metres, distributed over a distance of 3000 km; SKA will drive innovation in query and knowledge creation on large databases, energy efficient computing (as very large computers like the SKA massive energy demands), and improved communications and networking technologies.

Idem pour CTA, LSST!

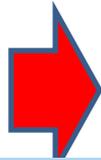




Fiche 03: storing, managing and preserving

e-Infrastructure fiche 03 storing, managing and preserving research data

Scenario: Diversity is likely to remain a dominant feature of research data – diversity of formats, types, vocabularies, and computational requirements – but also of the people and communities that generate and use the data. Europe needs to develop an integrated and service-driven approach to e-infrastructures for the data of a wide range of research communities. Europe should therefore step-up the available capacity to cope with extremely large, heterogeneous and complex datasets incorporating advanced computing and software. Furthermore, costs of storing and preserving data can be significant if one needs to keep it well managed for long periods to be used by others. Different institutions archive their research data in different ways - making access difficult from outside the institution - whereas storage and computing media evolve and become quickly obsolete. How will we preserve and maintain future access to priceless research data? Data management plans are intended to help researchers, funders and data repositories to get maximum value from research data at minimum cost.



Fiche 03: storing, managing and preserving

e-Infrastructure fiche 03 storing, managing and preserving research data

Scenario: Diversity is likely to remain a dominant feature of research data – diversity of formats, types, vocabularies, and computational requirements – but also of the people and communities that generate and use the data. Europe needs to develop an integrated and service-driven

Deploying *service-driven data e-infrastructure* to provide a tiered architecture for data management and long term preservation accommodating the need for replication and the reusability of information in different research and education contexts. To achieve cost efficiency in preservation infrastructures - tier zero will be developed by a limited number of data centres operating at the European level, associating the function of traditional storage preservation of information with the advanced curation of information for re-use, complemented by national, institutional or thematic networks of data centres.

data repositories to get maximum value from research data at minimum cost.

The image shows the cover of a brochure for a government seminar on digital technology. At the top, it features the French national emblem and the text 'LIBERTÉ • ÉGALITÉ • FRATERNITÉ RÉPUBLIQUE FRANÇAISE' and 'PREMIER MINISTRE'. The main title is 'Séminaire gouvernemental sur le numérique'. Below this, there are three numbered points: 01 'Faire du numérique une chance pour la jeunesse', 02 'Renforcer la compétitivité de nos entreprises grâce au numérique', and 03 'Promouvoir nos valeurs dans la société et l'économie numériques'. The website 'www.gouvernement.fr' is listed vertically. At the bottom left, it says '28 16 mai 2013' and provides contact information for the 'Service de presse de Matignon'. A QR code and the slogan 'CONSTRUISONS ENSEMBLE LE NOUVEAU MODÈLE FRANÇAIS' are at the bottom.

Plusieurs technologies méritent d'être particulièrement mises en avant :

- les objets connectés et intelligents et les logiciels embarqués, qui sont un domaine d'excellence pour les entreprises françaises et occuperont une place croissante dans l'économie numérique,

- les technologies de calcul intensif et leur application à la simulation numérique dans les entreprises, essentielles pour la conception de systèmes complexes,
- l'informatique en nuage maîtrisée,
- les technologies d'exploitation massive des données toujours plus nombreuses collectées par les services numériques ("big data"),

- la sécurité des systèmes d'information, qui touche de manière transverse toute l'économie numérique.

Communication:

... des petites actions pour commencer:

- WEB IN2P3

OLD (<http://www.in2p3.fr/>)

NEW (<http://www.in2p3.fr/actions/informatique/presentation.htm>)

- Espace collaboratif et de documentation:

« Redmine » -> Projet Informatique-IN2P3

<https://forge.in2p3.fr/projects/informatique-in2p3>

A suivre....

« Le personnel du CCIN2P3: le point de vue des informaticiens et le rôle du CCIN2P3 »

- Votre professionnalité...
- Votre aperçu des défis...
- Vos implications dans le contexte internationale ...
- Vos implications dans le contexte national ...
- Votre aperçu des chercheurs ... et des idées du CMI ...



PHYSICIENS