GENCI

40ème Forum ORAP

18 octobre 2017

Irene un nouveau supercalculateur multi-pétaflopique pour la recherche française et un atout de PRACE2 dans EDI

Philippe Lavocat

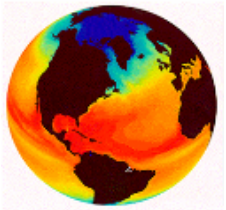
SOMMAIRE

- HPC et Data Analytics : un enjeu majeur
- Evolution du contexte technologique
- Panorama des machines nationales & IRENE
- Evolution du contexte européen
  - EDI & Euro-HPC
- PRACE2

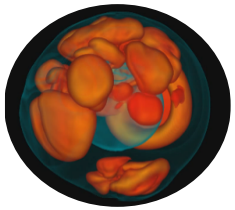
LE HPC : UN ENJEU MAJEUR

Le calcul intensif, un outil stratégique indispensable pour le traitement de données massives

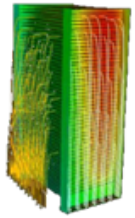
Pour la science



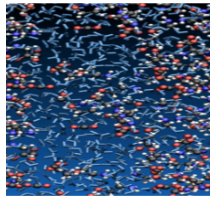
Climat



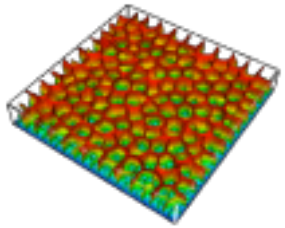
Astrophysique



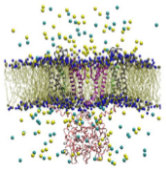
Energie



Chimie



Matériaux



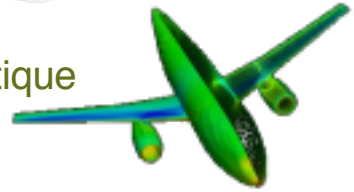
Sciences du vivant

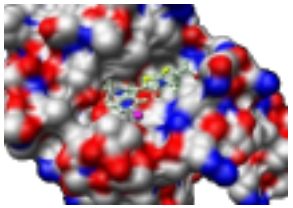
Sciences Humaines et Sociales

Pour l'innovation

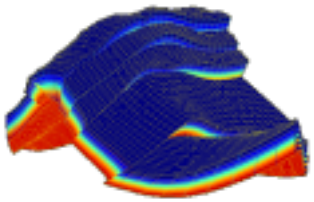
Automobile

Aéronautique





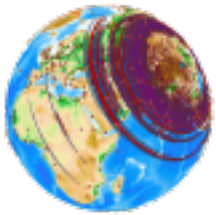
Pharmacologie



Exploration pétrolière

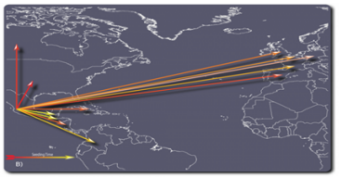
Médecine personnalisée

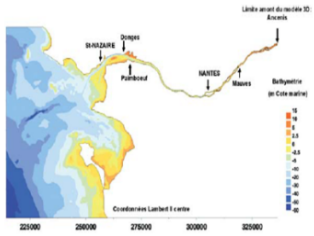
Pour l'aide à la décision



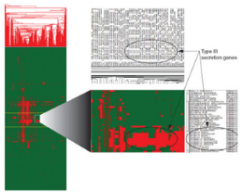
Risques naturels

Risques biologiques et épidémiologiques





Impact des activités industrielles

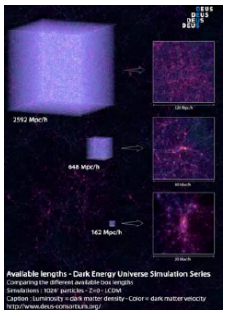


Sécurité

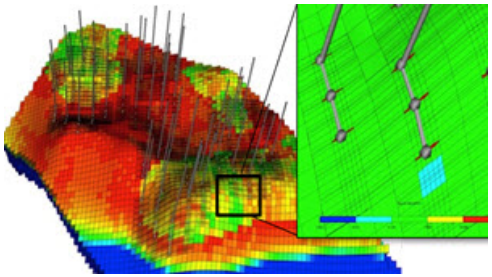
CONTEXTE TECHNOLOGIQUE

Mais aussi convergence HPC / DATA (et IA)

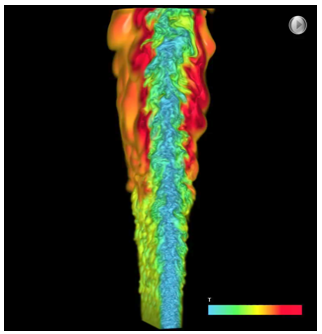
❑Explosion de la donnée computationnelle



Cosmology
DEUS project
150 PB raw data



Reservoir modeling
of gigamodels 350 TB/run



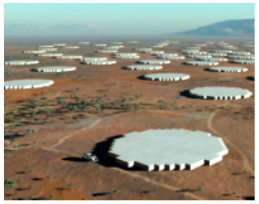
HiFi turbulent
DNS combustion
S3D : 1PB / 30mn

Climate CMIP exercices

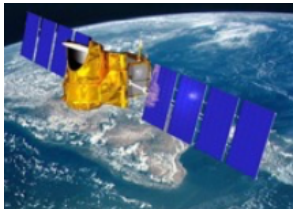
Status CMIP5 data archive:
- 1.8 PB for 590[illegible]0 data sets stored in 4.3 Mio Files in 23 ESGF data nodes
- CMIP5 data is about 50 times CMIP3

Extrapolation to CMIP6:
- CMIP6 has a more complex experiment structure than CMIP5.
- Expectations: more models, finer spatial resolution and larger ensembles
- Factor of 20: 36 PB in 86 Mio Files
- Factor of 50: 90 PB in [illegible]15 Mio Files

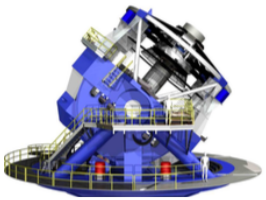
❑Et de la donnée instrumentale



LOFAR/SKA
4 Eo/an bruts

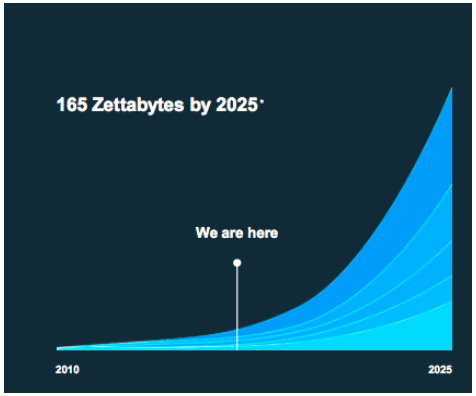


COPERNICUS/SWOT
4 Po/jour bruts



LSST/EUCLID
20 Po/nuit bruts

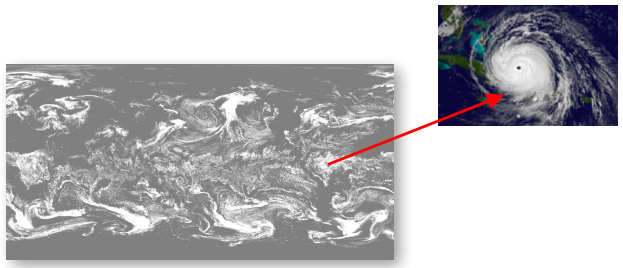
Réseau capteurs
sismiques
100 To/an



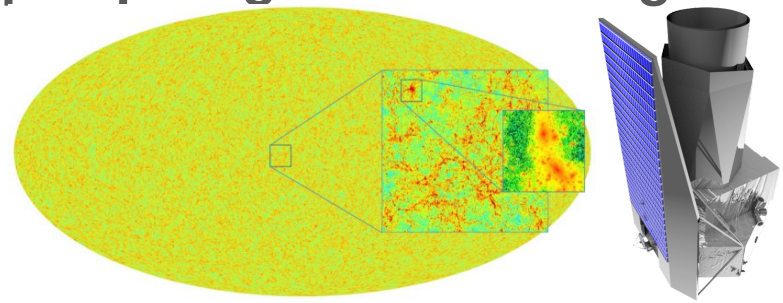
Internet et IoT

CONTEXTE TECHNOLOGIQUE

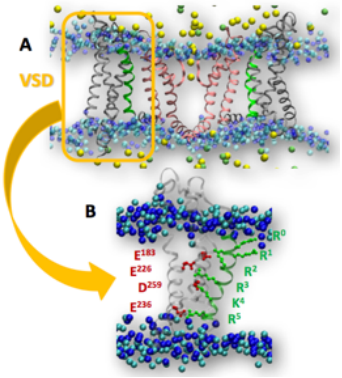
L'Exascale pour quoi : quelques grands challenges



Climat/météo
Modèles globaux à 1km
Ensembles multiphysiques

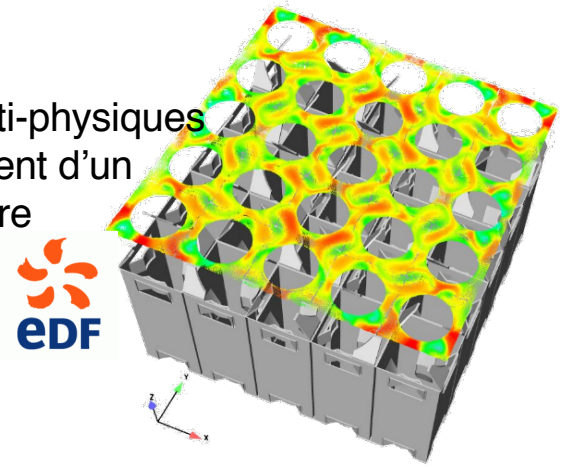


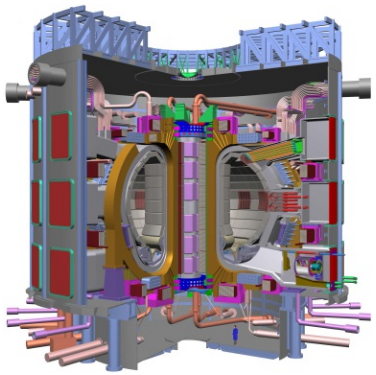
Cosmologie : comprendre origine/évolution Univers
Simulations n-corps sur plus de 10 trillions de particules



Sciences de la vie
Interactions protéines/médicaments
& effets des mutations sur systèmes
> 200k atomes
simulations sur plusieurs µs

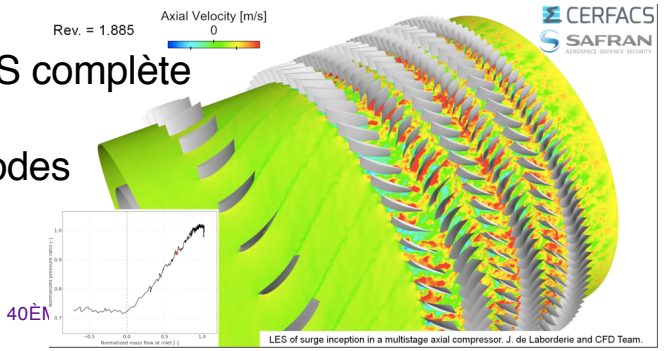
Energie
Simulations multi-physiques
Du fonctionnement d'un
réacteur nucléaire
(digital twin)





Fusion par confinement
Simulation échelle 1 ITER
Dimensionnement, fonctionnement,
Sécurité, performances du futur
réacteur

Combustion
Simulations LES complète
turbine à gaz
Couplage de codes



GENCI

Moyens de calcul nationaux pour la production

❑ 3 centres nationaux (CINES, IDRIS, TGCC)

Occigen depuis 2015

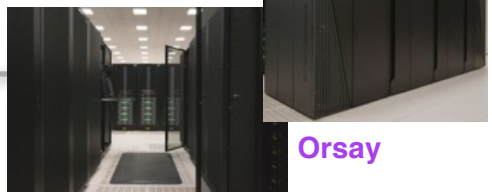
Montpellier

Curie (depuis 2012 : 20 % pour le DARI)

Bruyères-le-Châtel

6,8 Pflop/s

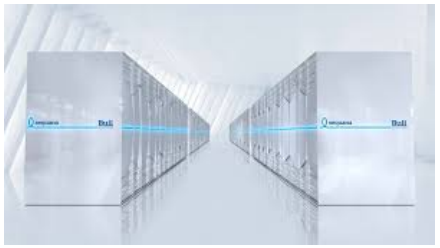
Ada et **Turing** (depuis 2013)



Orsay

*400 en 10 ans

Irene (S2 2018 50% pour PRACE)



❑ Budget pluriannuel sur 10 ans de 30 M€/an, pour assurer le renouvellement

- 2018 TGCC renouvellement / 2020 extension
- 2019 IDRIS renouvellement / 2021 extension
- 2020 CINES renouvellement / 2022 extension

Une évolution tous les 10 mois
Architectures complémentaires

ACTUALITÉ GENCI (3/8)

Centres de calcul nationaux (2/2)

❑ TGCC

- Configuration d'**Irene**, successeur de Curie:
  - Système **Atos/BULL** SEQUANA
  - **9 Pflop/s** crête pour la 1re phase
  - Mix de technologies Intel Skylake et KNL
  - Près de **400 To** de mémoire distribuée
  - Scratch de **5 Po** utile à plus de **300 Go/s** sous Lustre

IRENE

Configuration équilibrée
puissance calcul / capacité mémoire / capacité I/O

ACTUALITÉ GENCI (4/8)

TGCC (1/4)

- Détails de la configuration Irene
  - Technologie **Intel Skylake** (6,84 PFlops crête)

6 cellules SEQUANA

Cellule SEQUANA X1000
≈ 1,14 Pflops en Intel Skylake

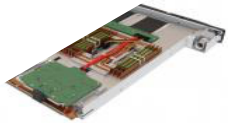
x 92 / cellule

552 lames de calcul

Lame de calcul x1120:
- 3 nœuds Skylake
- Refroidissement direct eau chaude (DLC)

x 3 / lame

1 656 nœuds



Nœud de calcul Skylake composé de:
- 2 processeurs Intel Skylake soit 48 cœurs
- **192 Go de mémoire** DDR4 (soit 4 Go / cœur)

x 2 / nœud

3 312 processeurs
79 488 cœurs

Processeur Intel Skylake 8168
24 cœurs - 2,7 GHz – 205 W
96 Go de mémoire DDR4 (soit 4 Go / cœur)
86,4 Gflops / cœur

ACTUALITÉ GENCI (5/8)

TGCC (2/4)

- Détails de la configuration Irene
  - Technologie **Intel KNL (**2,1 PFlops crête)

3 cellules SEQUANA

Cellule SEQUANA X1000
≈ 0,7 Pflops en Intel Skylake

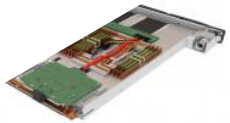
x 74 / cellule

222 lames de calcul

Lame de calcul X1210
- 3 nœuds KNL
- Refroidissement direct eau chaude (DLC)

x 3 / lame

666 nœuds



Nœud de calcul KNL composé de:
- 1 processeur Intel KNL 68 cœurs
- 96 Go de mémoire DDR4 (soit 1,4 Go / cœur)

x 1 / nœud

666 processeurs
45 288 cœurs

Processeur Intel KNL 7250
68 cœurs – 1,4 GHz – 215 W
16 Go MCDRAM (4x débit DDR4)
44,8 Gflops / cœur

ACTUALITÉ GENCI (6/8)

TGCC (3/4)

- Détails de la configuration Irene
  - **20 nœuds de visualisation distante** avec pour chaque nœud :
    - 2 proc. Intel Skylake 8168 de 24-cœurs @ 2,7 GHz
    - 1 **accélérateur GPU** nVIDIA Pascal PCIe
    - 192 Go mémoire DDR4
    - IO: 1 HDD 250 Go + 1 SSD 800 Go/NVMe
  - **5 nœuds à grande mémoire** pour le **pré et post-traitement** avec pour chaque nœud :
    - 4 proc. Intel Skylake 8176 de 28-cœurs @ 2,1 GHz
    - 1 accélérateur GPU nVIDIA Pascal PCIe
    - **3 To de mémoire** partagée DDR4
    - IO: 2 HDD de 1 To + **1 SSD 1600 Go/NVMe**
  - **8 nœuds** de **login** (même processeur que les nœuds de calcul SKL)

ACTUALITÉ GENCI (7/8)

TGCC (4/4)

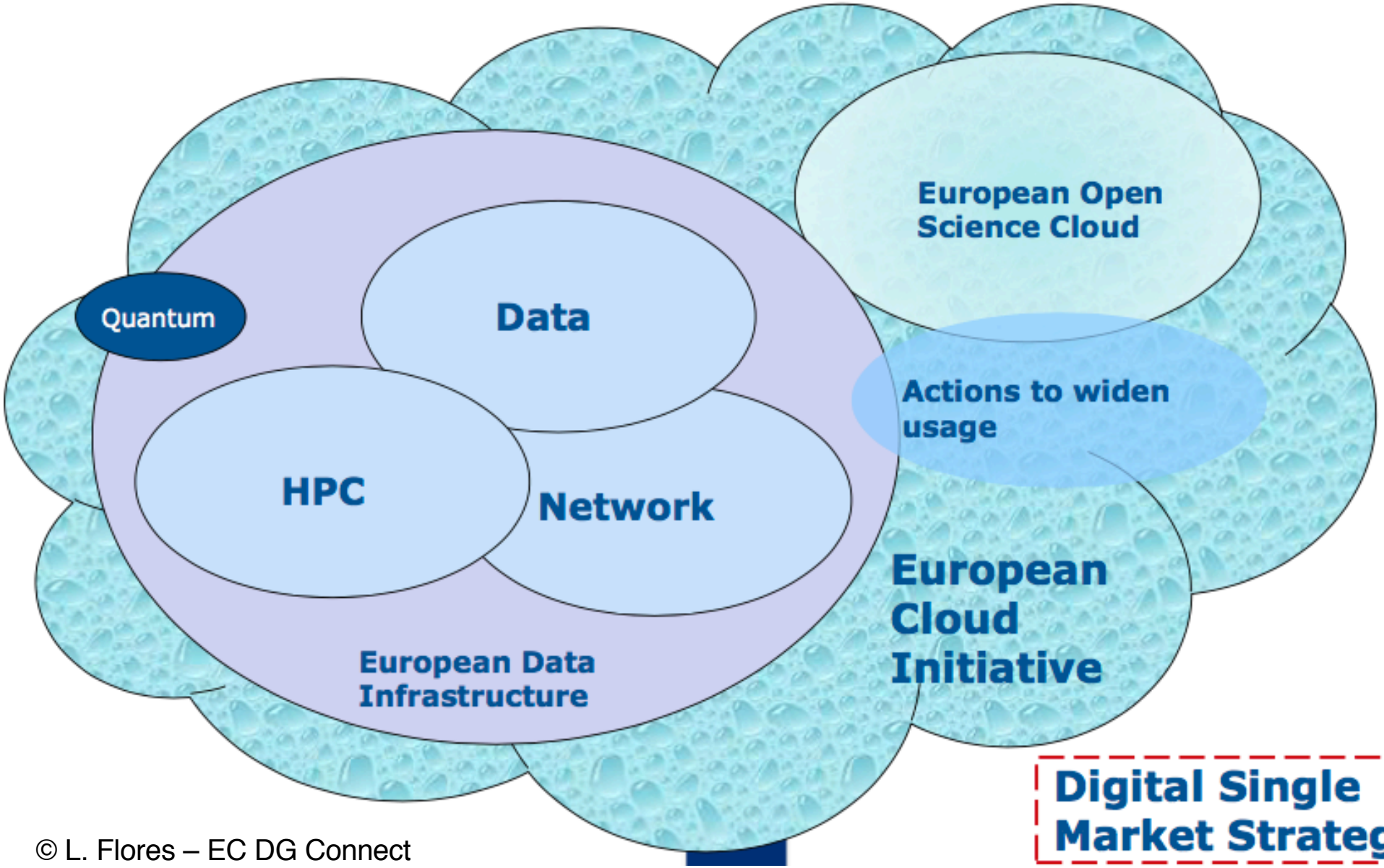
- Détails de la configuration Irene
  - **Réseau d'interconnexion**
    - Partition Skylake: Infiniband **EDR**
    - Partition KNL: nouveau réseau Bull **BXI**
  - **Stockage scratch**: cluster Seagate
    - **5 Po** utile
    - Bande passante de **300 Go/s** (+ 200 Go/s pour les autres espaces)
    - Environnement **Lustre Intel v3** basée sur version communautaire Lustre 2.7
    - 42 serveurs OSS, 1,88 Md inodes
    - Accélérateurs NytroXD pour les petites I/O (cache SSD)
  - Environnement **logiciel**
    - Distribution **Bull SCS5** basée sur OS RHEL7 (noyau Linux 3.10), SLURM
    - Gestionnaire de travaux **SLURM**
    - Compilateurs INTEL (Intel Parallel Studio XE) et GNU
    - Bibliothèque OpenMPI (norme MPI3) et support d'OpenMP 4.0
  - Intégration **physique**
    - consommation électrique totale : **1.1 MW** ~~(production)~~
    - 9 cellules + 10 racks = 60m^2, un peu plus de 52 tonnes

débit global >2x CURIE

-40% / CURIE pour 4.5x en performance

HPC strategy in a wider context

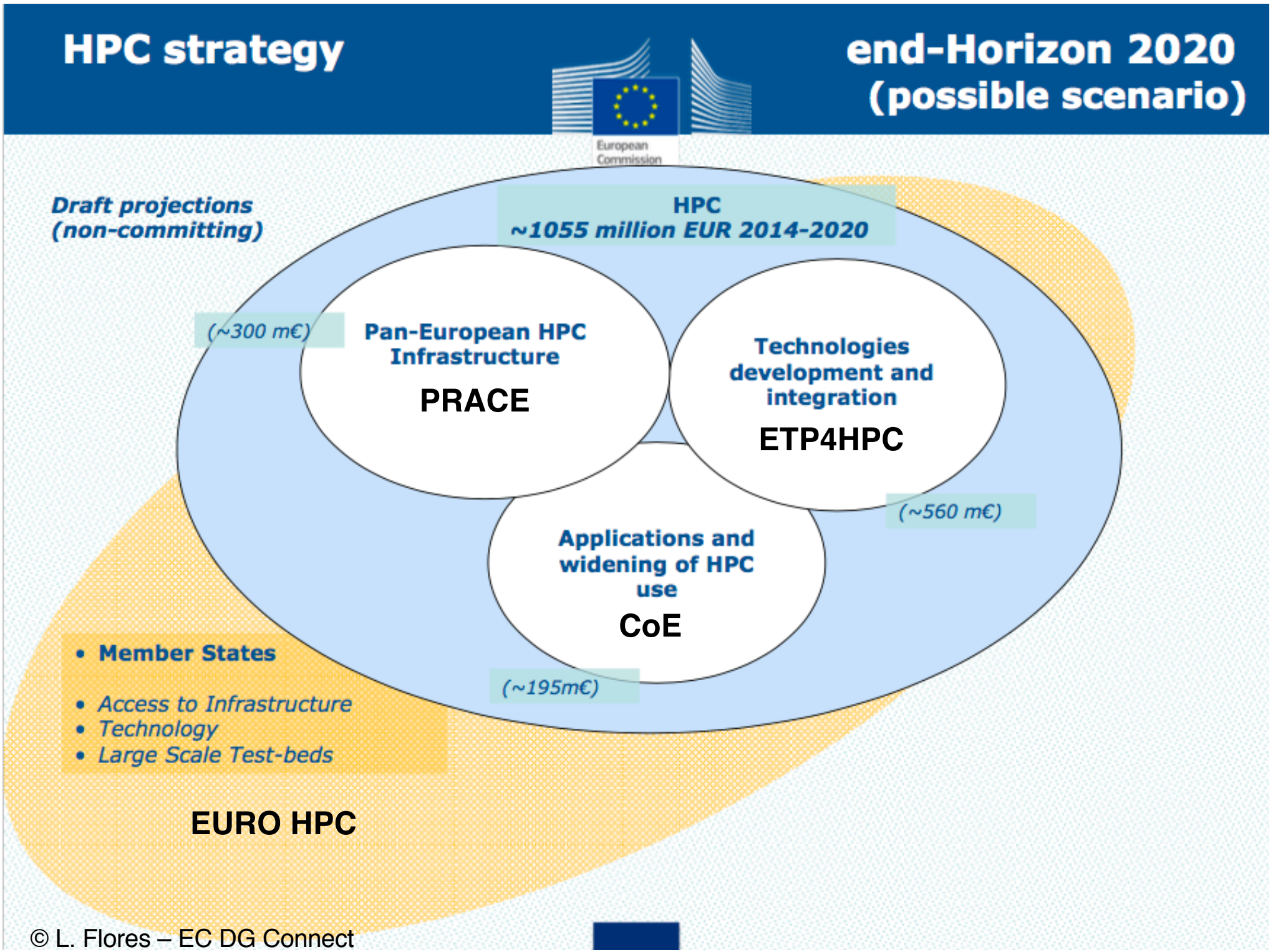
Digital Single Market European Cloud Initiative



© L. Flores – EC DG Connect

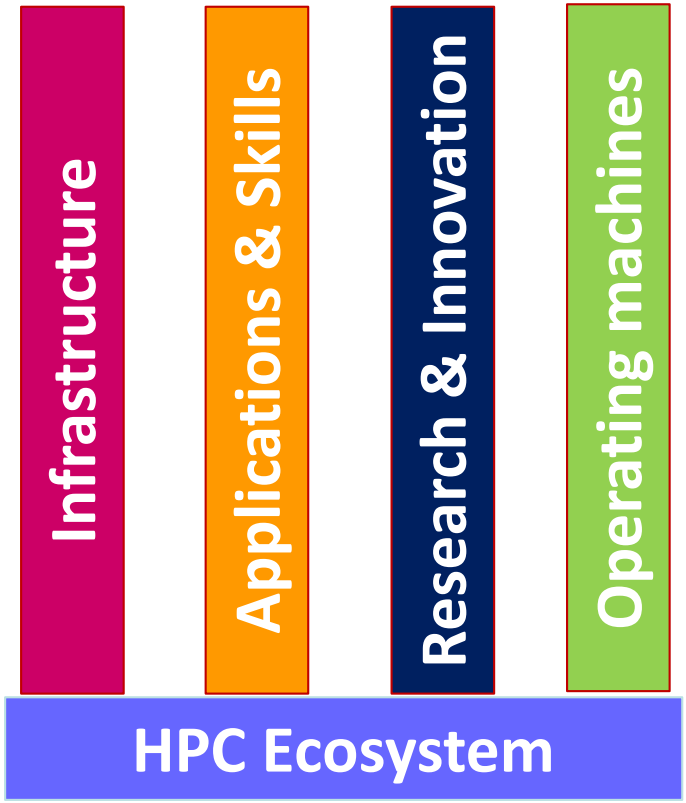
HPC strategy

end-Horizon 2020 (possible scenario)



© L. Flores – EC DG Connect

New instrument: Four Pillars Approach



- **Infrastructure**
  - Acquisition of infrastructure (linked to Research and Innovation) and access to research users
- **Applications & Skills**
  - Excellence in HPC applications; Supporting SMEs; Training and Outreach
- **Research and Innovation**
  - Technologies and systems developed in Europe
- **Operating the Machines**
  - Installation, deployment and operation via hosting entities

EVOLUTION DU CONTEXTE EUROPEEN

France : acteur du numérique européen

Genci
Représentation en HPC de la France au niveau européen

PRACE

EDI

Euro-HPC

STRATÉGIE EUROPÉENNE

Evolutions de l'infrastructure PRACE

❑PRACE 2 (2017-2020)

MareNostrum : IBM
BSC
Barcelona, Spain
10 Pflop/s

JUQUEEN : IBM BlueGene/Q
GAUSS/FZJ
Jülich, Germany
6 Pflop/s

Marconi : Lenovo
CINECA
Bologna, Italy
13 Pflop/s

68 Pflop/s
en 2017

SuperMUC : IBM
GAUSS/LRZ
Garching, Germany
6,8 Pflop/s

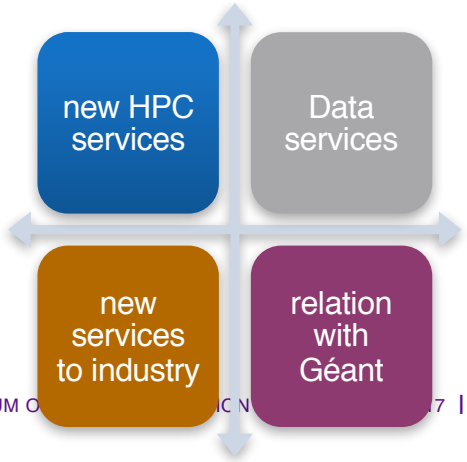
CURIE : Bull Bullx
GENCI/CEA
Bruyères-le-Châtel, France
1,7 Pflop/s →9→20

HAZELHEN : Cray XC40
GAUSS/HLRS
Stuttgart, Germany
7,4 Pflop/s

Piz Daint: Cray XC50 CSCS Lugano, Suisse
25 Pflop/s

❑Réflexions en cours autour de PRACE+ (dans EDI)

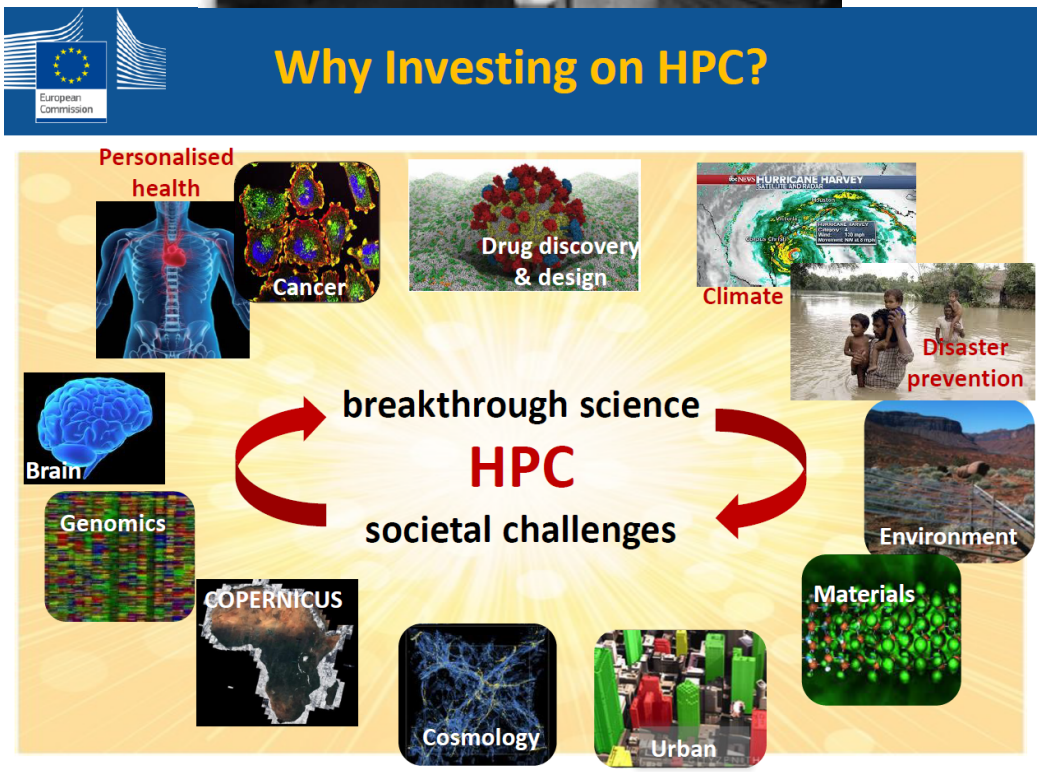
- PRACE comme un des piliers d'EDI
- Services étendus vers grands instruments/programmes, **urgent computing**, industrie, cloud, HPDA/IA, …



ACTUALITÉ GENCI (8/8)

PRACE

- PRACE 2 (2017 – 2020)
  - Organisation PRACE 2 officiellement approuvée le 3 mars 2017
  - Taux d'ouverture d'un tier0 PRACE 2: 40 % d'une machine 5 000 nœuds
    - ➔ Taux ouverture Curie **50 %** et Irene **54 %** (2 300 nœuds)

- Reprise de PRACE 2 au TGCC
  - Call 14 (avr. 17 – mars 18) - 100 % sur Curie
    - **12** projets FR (**2e** en projets et en heures)
    - 4 projets sur Curie
  - Call 15 (oct. 17 – sept. 18) – 100 % sur Curie
    - **5** projets FR (**3e** en projets, **8e** en heures)
    - 7 projets sur Curie
  - Call 16 (avr. 18 – mars 19) –
    - Ouverture de l'appel à projets : du **26 septembre** au **21 novembre**
    - Démarrage sur Curie puis transfert sur Irene à partir de juillet 2018

Merci pour votre attention !