

Euclid au CC-IN2P3

Euclid-France, 07.01.2016

Ken Ganga

Quentin Le Boulc'h



- ▶ Le Centre de Calcul de l'IN2P3
- ▶ La convention CNES / IN2P3
- ▶ Les ressources Euclid au CC-IN2P3

Le Centre de Calcul de l'IN2P3

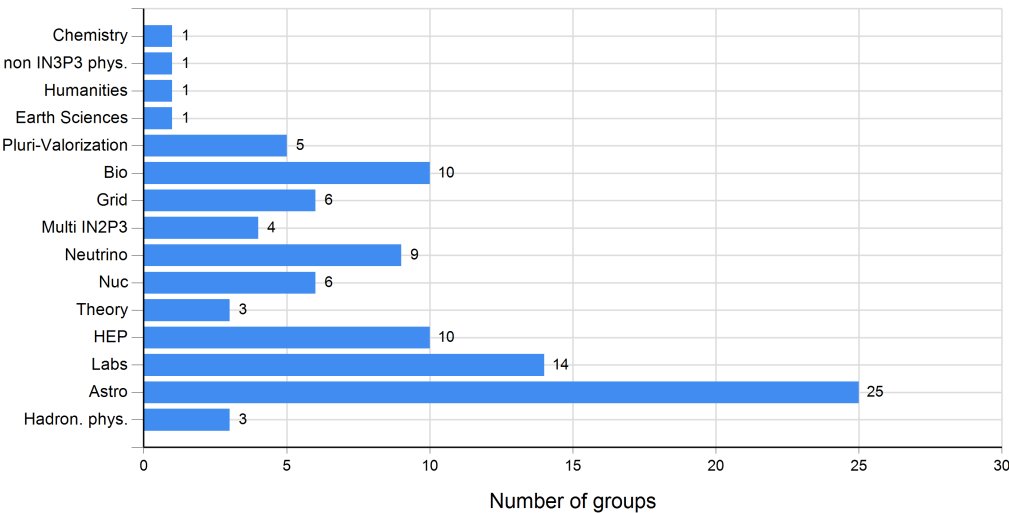


- ▶ Centre HTC dédié aux thématiques de l'IN2P3 : Physique des hautes énergies, Physique nucléaire et Astroparticules
- ▶ Ouverture à d'autres domaines (biologie, sciences humaines, ...)
- ▶ Ressources mutualisées
- ▶ Ressources humaines : 83 personnes (dont 64 ingénieurs et 2 chercheurs)



En partenariat avec le CEA/DSM/IRFU

Number of groups by scientific domain and activity in 2015



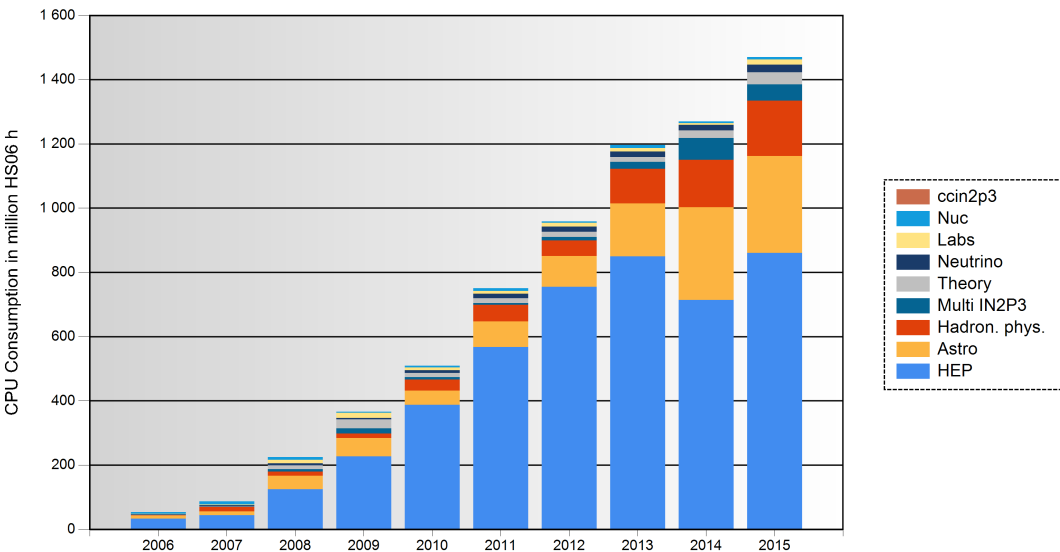
CPU : 25 000 coeurs virtuels
= 253 kHS06

Stockage sur disque :

Standard = 17 Po

Haute performance = 2 Po

IN2P3 CPU Consumption by activity



Stockage sur bandes magnétiques :

Volume stocké = 32 Po

(Capacité de 340 Po)

Sauvegarde :

Volume stocké = ~ 1,1 Po

(Capacité de 5 Po)

Les challenges à venir côté expériences :

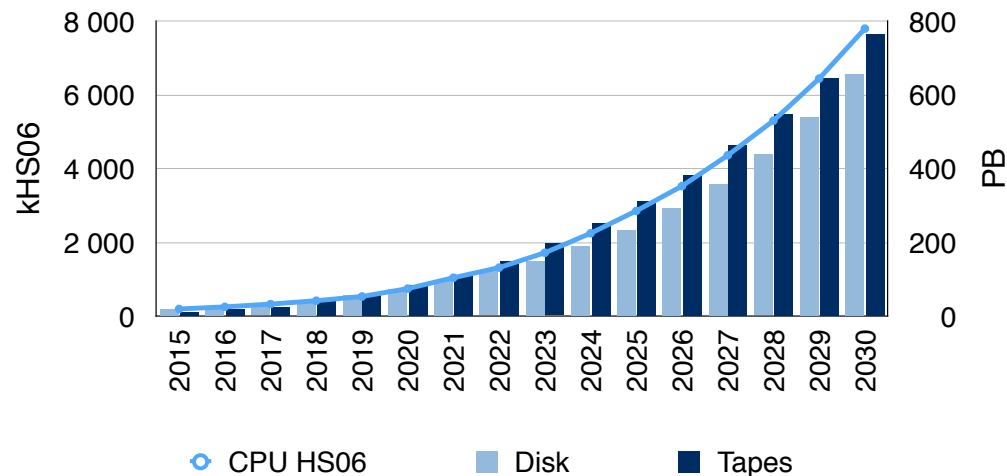
LHC Run 2 et suivants

- Plus d'événements, plus de complexité
- Modèle de calcul à revoir

Astroparticules

- LSST :
 - Le CC-IN2P3 devrait assurer le stockage de la totalité des données et la moitié du traitement
- Euclid :
 - Convention signée : 30% des besoins couvert par le CC-IN2P3
 - **Estimation pour 2030 : 5.25 Po disque, 52.5 Po bande, 67.5 kHS06**
- CTA :
 - 25 % du calcul au CC-IN2P3 ?

Sum of annual capacity needs for these 7 experiments



La convention CNES / IN2P3

- ▶ Périmètre: 30% des besoins informatiques pour la première livraison (DR1) (20% requis par le consortium)
- ▶ Deux approches pour estimer les besoins :
 - ▶ Comparaison avec Planck
 - ▶ Calcul « bottom-up » à partir des estimations des OU
- ▶ Prise en compte de l'impact d'un changement de date de lancement, des prix du matériel, de la fraction de contribution du CC-IN2P3, etc. Avenant en cas:
 - ▶ D'une différence de plus de 20% sur le coût des infrastructures
 - ▶ D'une participation du CC-IN2P3 supérieure à 30 %

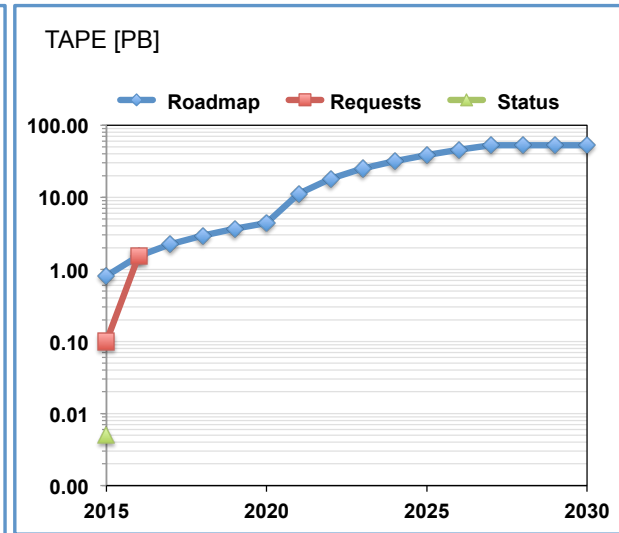
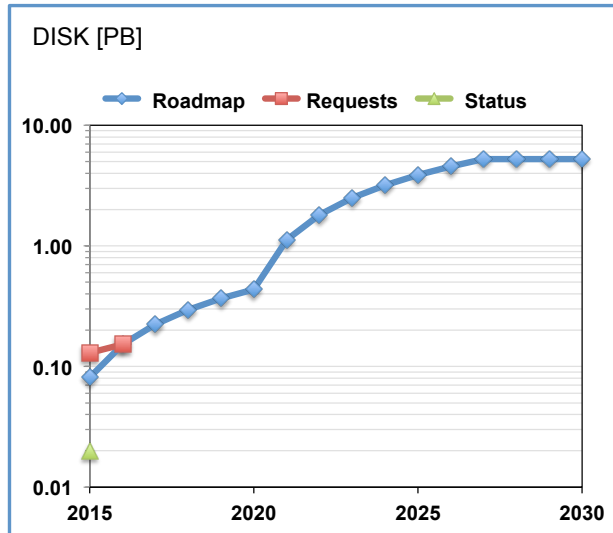
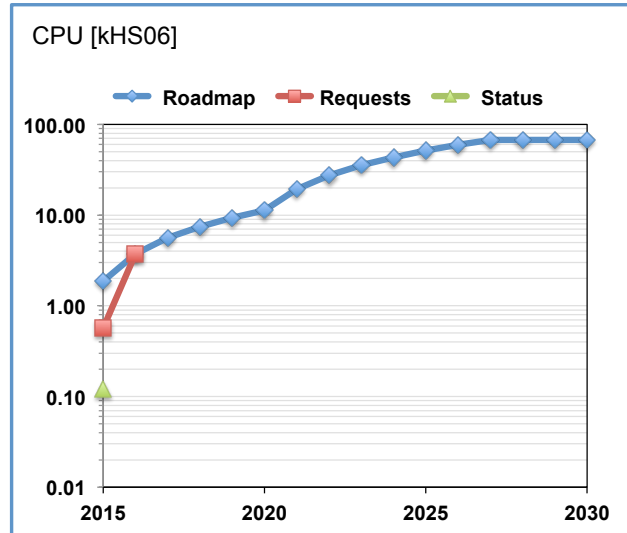
- ▶ Contribution du CC-IN2P3 : 2,660 M€
 - ▶ Ressources de calcul, stockage et réseau
 - ▶ Administration, maintenance et exploitation de l'infrastructure
 - ▶ Electricité, climatisation, etc.
 - ▶ Gestion des ressources humaines
- ▶ Contribution du CNES : 1,495 M€ : financement de 3 CDD ingénieurs (+ missions)
- ▶ Pilotage
 - ▶ Réunion de démarrage : 16 septembre 2015 puis 29 janvier 2016
 - ▶ Réunion annuelle du comité directeur : 29 janvier 2016
 - ▶ Points-clés annuels (jalons)
 - ▶ Téléconférence SGS-France

La convention CNES / IN2P3 : roadmap

PM2

Units : CPU [kHS06] / Storage [PB]

Years	Total Euclid			CC-IN2P3			CC-IN2P3 increment			Requests			Status [% requests]					
	CPU	Disk	Tape	CPU	Disk	Tape	CPU	Disk	Tape	CPU	Disk	Tape	CPU	Disk	Tape	CPU	Disk	Tape
2015	6.24	0.271	2.706	1.87	0.08	0.81	1.87	0.08	0.81	0.570	0.13	0.10	0.122	21%	0.020	15%	0.005	5%
2016	12.49	0.508	5.081	3.75	0.15	1.52	1.88	0.07	0.71	3.747	0.152	1.524						
2017	18.73	0.746	7.457	5.62	0.22	2.24	1.87	0.07	0.71									
2018	24.98	0.983	9.832	7.49	0.29	2.95	1.88	0.07	0.71									
2019	31.22	1.221	12.208	9.37	0.37	3.66	1.87	0.07	1.52									
2020	37.50	1.458	14.583	11.25	0.44	4.37	3.76	0.15	1.43									
2021	64.29	3.750	37.500	19.29	1.13	11.25	9.91	0.76	7.59									
2022	91.07	6.042	60.417	27.32	1.81	18.13	9.91	0.76	7.59									
2023	117.86	8.333	83.333	35.36	2.50	25.00	9.91	0.76	8.40									
2024	144.64	10.625	106.250	43.39	3.19	31.88	9.91	0.76	8.30									
2025	171.43	12.917	129.167	51.43	3.88	38.75	11.79	0.84	14.46									
2026	198.21	15.208	152.083	59.46	4.56	45.62	17.95	1.45	14.46									
2027	225.00	17.500	175.000	67.50	5.25	52.50	17.94	1.45	15.27									
2028	225.00	17.500	175.000	67.50	5.25	52.50	9.91	0.76	8.30									
2029	225.00	17.500	175.000	67.50	5.25	52.50	9.91	0.76	14.46									
2030	225.00	17.500	175.000	67.50	5.25	52.50	11.79	0.84	14.46									



- ▶ Roadmap revue chaque année :
 - ▶ Evolution du modèle de calcul : Data Challenges
 - ▶ Evolution du modèle de coûts : nouveaux achats de matériel, évolution du prix de l'électricité
 - ▶ Besoins, requêtes et consommations effectives

- ▶ Requêtes annuelles :
 - ▶ Estimations des besoins à fournir chaque année par les équipes PF (première estimation à mi-année puis confirmation en novembre)
 - ▶ Cette année : les requêtes ont été ajustées sur la roadmap, collecte des besoins en cours pour consolidation
 - ▶ Un document décrivant le processus de demande de ressources par les équipes PF au CC-IN2P3 est en préparation

- ▶ Les estimations datent de 2013 et ne prennent pas en compte les besoins HPC de LE3.
- ▶ La convention ne concerne que la DR1: Data Releases 2 et 3 à négocier plus tard.
- ▶ La convention ne concerne que le SDC-France / SGS : le travail des SWGs n'est pas inclus. Un autre accord peut être négocié.
- ▶ Ressources fournies en best-effort aux SWGs (procédure en cours de rédaction).

Les ressources Euclid au CC-IN2P3

- ▶ Ressources Euclid 2015 : 130 To stockage disque, 100 To stockage bandes, 570 HS06
- ▶ 570 HS06 ~ 50 coeurs en moyenne = ~ 5 000 000 heures HS06 sur un an
- ▶ Requêtes 2016 : 3747 HS06, 152 To disque, 1.5 Po bandes
- ▶ Sous-projets batch :
 - ▶ Accounting (suivi de l'utilisation des ressources)
 - ▶ Priorités selon activité

Detailed Accounting *Units : CPU [HS06]*

	Request	Total		VIS		SPE		LE3		SIM		EXT		MER		SGS		SCI		OTHER	
2015	570	122	21%	24	20%	0	0%	0	0%	67	55%	0	0%	0	0%	0	0%	25	20%	6	5%
2016	3747																				
2017																					
2018																					
2019																					
2020																					
2021																					
2022																					
2023																					
2024																					
2025																					
2026																					
2027																					
2028																					
2029																					
2030																					

- ▶ Différentes queues adaptées aux différents profils de jobs. Exemples :
 - ▶ La queue « huge » permet d'utiliser jusqu'à 16G de mémoire par job
 - ▶ La queue « longlasting » autorise des jobs de 7 jours
- ▶ Détails sur http://cctools.in2p3.fr/mrtguser/info_sge_queue.php

List of Grid Engine queues defined at Computing Center:

If you want to use a queue which has a restricted access, please, contact user support at CC-IN2P3.

Queue name	Host list	Max CPU time (hh:mm:ss)	Max elapse time (hh:mm:ss)	Max memory	Max scratch	Restricted access	Slots used	Available slots	Used/available
admin	@multicores @sequentials @multiseqs	00:15:00	20:00:00	500M	5G	admins	0	733	0.0 %
admin_test	@sequentials @multicores @multiseqs	00:05:00	01:00:00	500M	5G	admins	0	733	0.0 %
demon	@sequentials	24:00:00	INFINITY	2G	2G	demonqueue	10	341	2.9 %
huge	@multiseqs @sequentials	48:00:00	72:00:00	16G	110G	hugequeue	10582	12133	87.2 %
interactive	@interacts	24:00:00	24:00:00	4G	20G	NONE	0	48	0.0 %
long	@multiseqs @sequentials	30:00:00	47:30:00	4G	30G	NONE	2606	8507	30.6 %
longlasting	@sequentials	168:00:00	192:00:00	6G	30G	longlastingqueue	1094	1364	80.2 %
mc_interactive	@interacts	24:00:00	24:00:00	3G	20G	mcqueue	72	72	100.0 %
mc_long	@multicores @multiseqs	30:00:00	47:30:00	5G	30G	mcqueue	2626	11240	23.4 %
mc_longlasting	@multicores @multiseqs	202:00:00	226:00:00	3G	30G	mc_longlastingqueue	612	6272	9.8 %
mc_long_grid	@multicores @multiseqs	30:00:00	47:30:00	7G	30G	mcgridqueue	4200	11240	37.4 %
mc_medium	@multicores @multiseqs	05:00:00	12:00:00	3G	30G	mcqueue	0	6272	0.0 %
medium	@multiseqs @sequentials	05:00:00	12:00:00	3G	30G	NONE	220	1505	14.6 %
pa_long	@parallels	46:00:00	47:30:00	3G	30G	paqueue	184	928	19.8 %
pa_longlasting	@parallels	168:00:00	192:00:00	3G	INFINITY	palonglastingqueue	0	928	0.0 %
pa_medium	@parallels	05:00:00	12:00:00	3G	30G	paqueue	0	928	0.0 %

- ▶ Fermes interactive, séquentielle, multicore et parallèle.
- ▶ Détails sur http://cctools.in2p3.fr/mrtguser/info_sge_parc.php

- Host group name: **multicores** with 150 workers, 3 616 processors. Power: 3 753 494 UI(h)/day, **31 279** HS06-hours.
 - Platform name: **sl6** with 150 workers, 3 616 processors. Power: 3 753 494 UI(h)/day, **31 279** HS06-hours.

Model type	Processors per worker	Power (UI) per processor	Power (HS06-hours) per processor	Physical memory per worker (GiB)	Total number of workers	Total number of processors	Overall UI power	Overall HS06-hours power
Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz	40	56,8	11,35	126	1	40	2 270	454
Intel(R) Xeon(R) CPU X5650 @ 2.67GHz	24	43,1	8,62	71	149	3 576	154 126	30 825
All sl6					150	3 616	156 396	31 279

- Host group name: **multiseqs** with 242 workers, 7 640 processors. Power: 9 224 304 UI(h)/day, **76 869** HS06-hours.
 - Platform name: **cl7** with 1 workers, 40 processors. Power: 54 480 UI(h)/day, **454** HS06-hours.

Model type	Processors per worker	Power (UI) per processor	Power (HS06-hours) per processor	Physical memory per worker (GiB)	Total number of workers	Total number of processors	Overall UI power	Overall HS06-hours power
Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz	40	56,8	11,35	126	1	40	2 270	454
All cl7					1	40	2 270	454

- Platform name: **sl6** with 241 workers, 7 600 processors. Power: 9 169 824 UI(h)/day, **76 415** HS06-hours.

Model type	Processors per worker	Power (UI) per processor	Power (HS06-hours) per processor	Physical memory per worker (GiB)	Total number of workers	Total number of processors	Overall UI power	Overall HS06-hours power
Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60GHz	32	48,4	9,67	94	71	2 272	109 851	21 970
Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz	40	56,8	11,35	126	78	3 120	177 060	35 412
Intel(R) Xeon(R) CPU X5650 @ 2.67GHz	24	43,1	8,62	71	92	2 208	95 165	19 033
All sl6					241	7 600	382 076	76 415

- Host group name: **parallels** with 57 workers, 912 processors. Power: 937 901 UI(h)/day, **7 816** HS06-hours.
 - Platform name: **sl6** with 57 workers, 912 processors. Power: 937 901 UI(h)/day, **7 816** HS06-hours.

Model type	Processors per worker	Power (UI) per processor	Power (HS06-hours) per processor	Physical memory per worker (GiB)	Total number of workers	Total number of processors	Overall UI power	Overall HS06-hours power
Intel(R) Xeon(R) CPU E5540 @ 2.53GHz	16	42,9	8,57	47	57	912	39 079	7 816
All sl6					57	912	39 079	7 816

- Host group name: **sequentials** with 351 workers, 12 048 processors. Power: 14 653 095 UI(h)/day, **122 109** HS06-hours.
 - Platform name: **sl6** with 351 workers, 12 048 processors. Power: 14 653 095 UI(h)/day, **122 109** HS06-hours.

Model type	Processors per worker	Power (UI) per processor	Power (HS06-hours) per processor	Physical memory per worker (GiB)	Total number of workers	Total number of processors	Overall UI power	Overall HS06-hours power
Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60GHz	32	48,4	9,67	94	268	8 576	414 650	82 930
Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz	40	56,8	11,35	126	64	2 560	145 280	29 056
Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz	48	55,5	11,10	142	19	912	50 616	10 123
All sl6					351	12 048	610 546	122 109

- ▶ Remarques :
 - ▶ Les machines multicore ont jusqu'à 130 GB de mémoire. Il est possible de fournir plus de capacité, mais les besoins doivent être connus et testés.
 - ▶ Renouvellement de la ferme parallèle (1024 coeurs, infiniband). Utile pour les tests LE3 ?
 - ▶ Bandes magnétiques encore peu utilisées pour Euclid.
 - ▶ Ressources Cloud disponibles pour tests et développements.
- ▶ Il y a des ressources disponibles, ne pas hésiter à les utiliser !
- ▶ Ne pas hésiter à nous contacter si besoin.

- ▶ Voir le *SDC-FR (CC-IN2P3) guide for the Euclid Consortium users* :

http://euclid.roe.ac.uk/attachments/download/5151/Euclid_CCIN2P3_guide.pdf

Comment ouvrir un compte ? Comment soumettre un job ? Etc.

- ▶ Contacts :
 - ▶ <https://cc-usersupport.in2p3.fr>
 - ▶ Quentin Le Boulc'h / Rachid Lemrani
 - ▶ Maurice Poncet
 - ▶ Ken Ganga / Eric Aubourg