

Utilisation des infrastructures de la Grille EGI pour les simulations de CTA

Luisa Arrabito(1), Johann Cohen-Tanugi (2), Giovanni Lamanna (3), Nadine Neyroud (4), Matthieu Renaud (5)

(1) arrabito@in2p3.fr, LUPM – CNRS/IN2P3

(2) Johann.COHEN-TANUGI@lupm.in2p3.fr, LUPM – CNRS/IN2P3

(3) lamanna@lapp.in2p3.fr, LAPP – CNRS/IN2P3

(4) nadine.neyroud@lapp.in2p3.fr, LAPP – CNRS/IN2P3

(5) mrenaud@lupm.univ-montp2.fr, LUPM – CNRS/IN2P3

Overview:

The Cherenkov Telescope Array (CTA) project [1] is the next generation of Imaging Atmospheric Cherenkov Telescopes (IACTs), operating in the high- and very high-energy gamma-ray domain (20 GeV-200 TeV). It will consist of two arrays of 50-100 telescopes of different sizes, located in each hemisphere. The CTA consortium gathers more than 1000 scientists and engineers from more than a hundred institutions world-wide. The project is currently in its preparatory phase. The construction is planned to be completed around 2018-2020. CTA will be the first high-energy ground-based astronomical open observatory. As such, it will receive observation proposals from the whole astronomical community and a dedicated Science Data Center will have to provide transparent access to data, analysis tools and user training.

During the current CTA preparatory phase, large computing and storage resources are needed mostly for MC studies. In particular, the selection of the CTA sites (North and South) has a significant impact on the final sensitivity of the instrument, as performance depends on the altitude of the site, its atmospheric conditions, geomagnetic field and night-sky background. The CTA MC working group is studying the impact of these various parameters by means of detailed MC simulations of the detector response to extensive air showers. Large sets of simulated events are generated for different primary particles: protons, electrons and gamma rays in the energy range from 3 GeV to 300 TeV. Moreover, once CTA sites have been selected and the construction phase has started, more detailed simulations will be produced in order to test analysis algorithms and to determine the final performances of the instrument. In order to fulfill present and future requirements for MC massive production and data analysis of CTA, we have proposed the use of the EGI grid infrastructure and of the DIRAC (Distributed Infrastructure with Remote Agent Control) [2] [3] framework. A first deployment of a DIRAC instance for evaluation purpose was followed by the development of a DIRAC extension specific to CTA, called CTADIRAC. In particular, this extension includes the tools for the management of CTA workflows, both for analysis and MC production. This setup has been successfully exploited during the MC campaigns in 2013 enabling, in particular, to achieve stable regimes of 4000-5000 concurrent jobs with a peak of 8000 jobs and producing an overall volume of 600 TB. The analysis of the MC data on the grid has also started in 2012 within the same CTADIRAC framework and first scientific results have been produced. This activity represents a cumulative number of about 1 M jobs run by a tens of users.

1. Enjeux scientifiques, besoin en calcul, stockage et visualisation :

Les objectifs scientifiques de CTA peuvent être regroupés en trois thèmes principaux : 1) la recherche de l'origine des rayons cosmiques et leur rôle dans l'Univers, 2) la compréhension des mécanismes d'accélération des particules au voisinage des trous noirs, et 3) la recherche sur la nature ultime de la matière et sur la physique au-delà du Modèle Standard.

Le thème 1) comprend l'étude des sites d'accélération de particules dans la Galaxie tels que les pulsars et leurs nébuleuses, les vestiges de supernova, les systèmes binaires compacts récemment détectés dans le domaine gamma, mais également les effets cumulés à différentes échelles, depuis les régions de formation stellaire jusqu'aux galaxies dites "starburst". Les signatures radiatives de ces sources en gamma permettent ainsi de contraindre les mécanismes d'accélération de particules et leur interaction avec le milieu environnant.

Le thème 2) concerne l'accélération des particules dans les conditions les plus extrêmes tels que celles observées au voisinage des trous noirs super-massifs. Les objets d'intérêts incluent les blazars, les radio-galaxies et autres classes d'AGN (Active Galactic Nuclei). La capacité de CTA à détecter un grand nombre de ces sources rendra possibles des études de population, ce qui constituera une avancée majeure dans ce domaine. Les études sur le fond diffus extragalactique au travers des spectres mesurés des sources lointaines, les amas de galaxies et sursauts gamma, qui restent encore à détecter dans ce domaine, sont également couvertes par cette thématique.

Enfin, le thème 3) regroupe les recherches liées à la 'Nouvelle Physique', telles que la recherche de la matière noire au travers des possibles signatures radiatives lors de son annihilation et les tests observationnels de la violation de l'invariance de Lorentz qui pourraient défier notre compréhension de

la physique fondamentale. CTA permettra ainsi des avancées majeures dans tous ces domaines. Le consortium CTA est actuellement dans sa Phase de Préparation jusqu'à fin 2013, pendant laquelle plusieurs sites candidats (nord et sud) sont à l'étude. La sélection des sites déterminera la sensibilité de l'instrument, les performances étant directement liées à l'altitude, aux conditions atmosphériques, au champ géomagnétique et au bruit de fond du ciel. Afin d'évaluer les différents sites candidats, le groupe de travail CTA-MC est en charge d'étudier l'impact de ces différents paramètres sur les performances finales de l'instrument, au travers de simulations Monte-Carlo (MC) intensives des gerbes atmosphériques et des différents réseaux de télescopes (optique, caméra, électronique) considérés. De plus, une fois le choix des sites effectué et la phase de construction en cours, de nouvelles simulations MC plus détaillées seront produites afin de tester les outils d'analyse des données et de déterminer précisément les performances finales de l'instrument.

Afin d'estimer les besoins de calcul et de stockage pour les études MC, nous allons considérer les étapes principales de la chaîne de simulation [4] [5], à savoir :

- 1/ génération des gerbes et propagation dans l'atmosphère;
- 2/ simulation de la réponse du réseau des télescopes;
- 3/ reconstruction et analyse.

Ces étapes sont exécutées de façon séquentielle, les produits de chacune d'entre elles servant de données d'entrée pour l'étape suivante. Les deux premières étapes constituent ce qu'on appelle la '*production*' et peuvent être exécutées séparément ou au sein d'une même tâche. La production est une activité gérée centralement par un groupe de 'gestionnaires de production', dont l'objectif principal est de fournir au consortium CTA les produits de l'étape 2/ (fichiers *sim_telarray*). Enfin, l'étape 3/ (reconstruction et analyse) se compose à son tour de plusieurs étapes, que nous allons appeler globalement '*analyse*', qui contrairement à la production sont à la charge des utilisateurs. Les produits de la première étape, très volumineux (fichiers *corsika*), sont sauvegardés sur des périodes de l'ordre de plusieurs mois pour des éventuels '*re-processing*', i.e. la reproduction de l'étape 2/ avec différentes configurations du réseau de télescopes. Les fichiers *sim_telarray*, en revanche sont conservés sur des périodes plus longues de 1-2 ans, afin de permettre leur exploitation de la part des utilisateurs.

La statistique nécessaire pour les études MC est d'environ 10^{10} événements simulés pour différents types de particule primaire : protons, électrons, photons gamma. Pour obtenir une telle statistique, chaque campagne de production est composée d'environ 200 k tâches, chacune produisant 50 k événements. Le temps CPU correspondant à une campagne de production typique est de ~ 8 M HS06 h [6] (environ 1 mois de calcul utilisant 1500 cœurs en parallèle), dont environ 70% est utilisé pour l'étape 1/. Le stockage représente également un défi majeur, car le volume global des données produites par une campagne de production est d'environ 600 TB, dont 30 TB servent à l'analyse des utilisateurs dans l'étape 3/.

2. Développements, utilisation des infrastructures :

Compte tenu des besoins de calcul et de stockage importants pour la production MC (cf. section 1), ainsi que de la possibilité de générer la statistique nécessaire en exécutant plusieurs tâches indépendantes sur un ensemble de nœuds, nous avons exploré, dès 2008, au sein d'un groupe de travail dédié, CTA Computing Grid (CTACG), l'adéquation des infrastructures de la grille EGI aux besoins de simulations de CTA [7]. Dans ce cadre, une Organisation Virtuelle CTA (vo.cta.in2p3.fr, VO CTA) [8] créée en 2008 et aujourd'hui est supportée par 20 sites EGI appartenant à 7 pays avec des ressources de l'ordre de quelques milliers de cœurs et plus de 600 TB de stockage dédié.

En 2012 nous avons démarré la migration de la production MC, précédemment basée sur le middleware gLite [9] et GANGA [10], dans le framework DIRAC (Distributed Infrastructure with Remote Agent Control) [11] [12]. La raison principale pour ce choix est l'implémentation du mécanisme du *Pilot Job* dans le Workload Management System de DIRAC, ce qui permet une optimisation de l'utilisation des ressources. En outre, DIRAC permet une gestion centralisée des différentes activités sur la grille, l'intégration de ressources hétérogènes ainsi que la création des services spécifiques à une communauté. Enfin, DIRAC offre une interface utilisateur riche ainsi qu'un catalogue de fichiers et de métadonnées, le DIRAC File Catalog (DFC). Cette migration a été réalisée en deux étapes. La première a consisté dans le déploiement d'une instance DIRAC dédiée à CTA, composée de 4 serveurs faisant tourner tous les services DIRAC et d'un serveur MySQL. La deuxième étape a porté sur le développement d'une extension de DIRAC spécifique à CTA, appelée CTADIRAC. Cette extension a été développée principalement pour fournir aux utilisateurs une interface permettant la configuration des tâches spécifiques à CTA. Elle comprend notamment : les outils de gestion des logiciels sur la grille; les outils de soumission des workflows CTA (production et analyse); l'interface des workflows de production au DFC; un outil permettant la recherche de fichiers en fonction des métadonnées enregistrées dans le DFC.

Du point de vue de l'utilisation des ressources de la grille, nous avons adopté la stratégie suivante pour la production MC : les tâches de simulation sont distribuées sur l'ensemble des 20 sites CTA en fonction de leur capacité, alors que les produits des simulations sont stockés sur 3 sites ayant du stockage dédié à CTA. En revanche, pour des raisons d'efficacité d'accès aux données, le re-processing ainsi que l'analyse des utilisateurs sont exécutés uniquement sur les 3 sites qui hébergent les produits des simulations.

L'instance CTADIRAC a été exploitée avec succès pendant les différentes campagnes de production et de re-processing en 2013 (cf. figure 1), permettant en particulier d'atteindre un régime stable de 4000-5000 tâches concurrentes avec un pic de 8000 tâches concurrentes, en produisant environ 600 TB. En 2012, l'activité d'analyse de la part des utilisateurs a également démarré avec le portage des premiers logiciels d'analyse dans CTADIRAC. Cette activité, en constante augmentation (cf. figure 2), représente plus de 1 M de tâches exécutées par une dizaine d'utilisateurs. Les différents cas d'utilisation seront décrits dans la section 4.

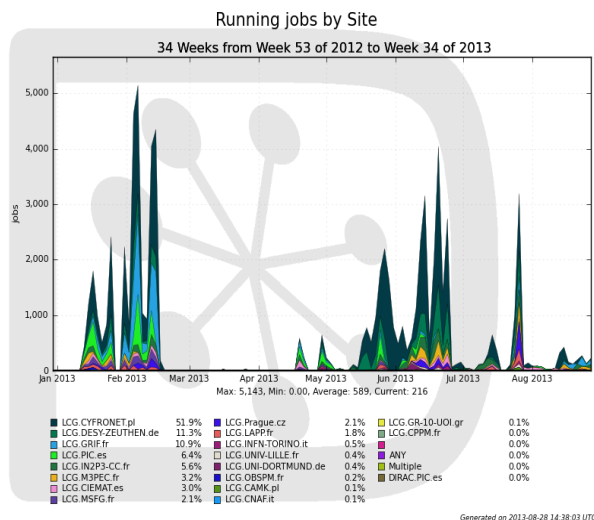


Figure 1. Tâches exécutées depuis Janvier 2013 sur la grille EGI en utilisant l'instance CTADIRAC. Les pics correspondent aux différentes campagnes de production et de re-processing. L'activité d'analyse est également incluse. Les codes de couleurs représentent les différents sites EGI sur lesquels les tâches ont été exécutées.

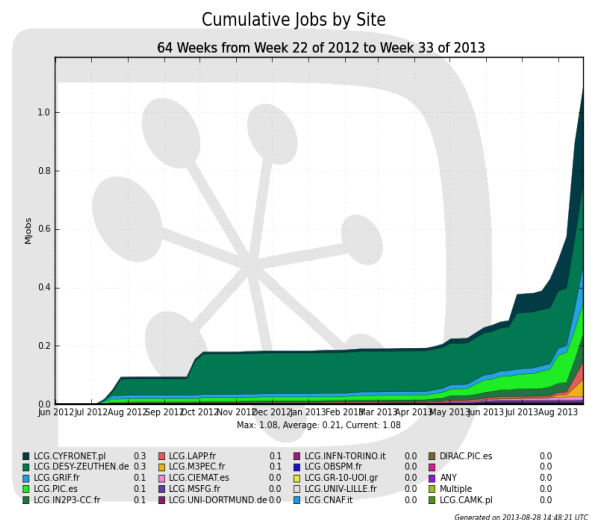


Figure 2. Nombre cumulé des tâches exécutées par les utilisateurs pour l'analyse des données MC depuis Juillet 2012. Les codes de couleurs représentent les différents sites EGI sur lesquels les tâches ont été exécutées. Une augmentation importante dans le nombre de tâches est visible à partir de Mai 2013, ce qui correspond au début des analyses des productions 2013.

3. Outils, le cas échéant complémentarité des ressources, difficultés rencontrées :

Malgré les bonnes performances obtenues dans l'utilisation des ressources de la grille avec notre système de production CTADIRAC, nous avons rencontré certaines difficultés. Une première difficulté est liée aux ressources et concerne une capacité de stockage limitée, comparée aux besoins de la production MC. En effet, le stockage dédié à CTA est de l'ordre de 600 TB, ce qui correspond à l'espace nécessaire pour stocker les produits d'une seule campagne de production (cf. section 1). Pour faire face à cette limitation, une partie de la statistique a été produite de façon indépendante sur des clusters locaux. Une autre conséquence de cette limitation est l'impossibilité de stocker plusieurs répliques des produits de simulation, afin d'augmenter leur disponibilité vis-à-vis des utilisateurs.

Une deuxième difficulté concerne l'accès en lecture et en écriture de grands volumes de données. En effet, la saturation des services de stockage (Storage Element, SE) est la cause principale des erreurs les plus fréquemment observées. Nous avons néanmoins allégé les effets de saturation des SE en régulant l'accès aux fichiers selon la capacité d'absorption de chaque SE, au travers d'un simple réglage des paramètres de soumission dans DIRAC.

Une dernière difficulté concerne l'utilisation optimale des ressources des 20 sites CTA. Bien que DIRAC, au travers du mécanisme des 'pilots jobs', minimise le taux d'échec dû à l'instabilité des ressources, le résultat net est une réduction des ressources disponibles. L'intervention du gestionnaire de production est alors demandée. La charge de travail de ce dernier pour le suivi de la production et de par l'activité croissante des utilisateurs n'est pas négligeable. Plus précisément, il doit s'assurer

que les ressources sont utilisées de la façon attendue, en investiguant les causes des erreurs éventuelles et en travaillant à leur résolution en interaction avec les administrateurs des sites.

4. Résultats scientifiques :

Plusieurs groupes au sein du consortium CTA sont actuellement à l'œuvre pour extraire de la campagne de simulation massive précédemment présentée des informations importantes dans la phase actuelle cruciale de choix définitif du site où sera installé CTA (dans sa version 'hémisphère sud'). En effet, des différences dans les performances du réseau de télescopes Cherenkov sont attendues en fonction de l'altitude du site, de l'intensité et direction locales du champ magnétique, différences que seule une large campagne de simulation Monte-Carlo permet d'étudier, et dont le déploiement sur la grille à l'aide de DIRAC a en grande partie assuré le succès. Les performances étudiées incluent le seuil en énergie, les limites de sensibilité en flux, la résolution angulaire atteignable, ainsi que la résolution spectrale à différentes énergies. Les sites étudiés, tous dans l'hémisphère sud (San Antonio de los Cobres et El Leoncito en Argentine, et Aar en Namibie), semblent tous satisfaire au cahier des charges défini par le consortium CTA. Toutefois, des différences significatives sont apparues, qui seront résumées à l'achèvement de ces travaux, par une 'figure de mérite' globale qui facilitera la prise de décision subséquente. Par ailleurs, cette campagne de simulations a permis de tester différentes configurations pour le réseau et différentes options de conception des différents télescopes. Par exemple, l'analyse des performances du modèle ASTRI (Astrofisica con Specchi a Tecnologia Replicante Italiana) des télescopes de petite taille a récemment été présentée [11]. ASTRI est un projet financé par l'INAF (Istituto Nazionale di Astrofisica), dont l'objectif principal est la construction d'un prototype de télescope de petite taille pour le sous réseau de CTA dédié au domaine des plus hautes énergies.

5. Perspectives :

Bien que notre système de production actuel basé sur DIRAC nous ait permis d'obtenir de très bonnes performances (cf. section 2), vue son utilisation croissante et les difficultés de gestion qui en découlent (cf. section 3), nous envisageons des améliorations dans l'automatisation des différentes tâches de production, en développant des outils de haut niveau au sein de CTADIRAC. Ces outils comprendront la gestion automatique des tâches de production (création et suppression des tâches, réplique des fichiers, etc.) en fonction des scénarios prédéfinis.

Du point de vue de l'activité d'analyse, suite aux retours d'expérience des utilisateurs CTA, nous prévoyons également d'enrichir l'interface utilisateurs avec des outils pour la gestion massive de tâches et des fichiers. En effet, nous rappelons que chaque campagne d'analyse consiste à exécuter des centaines de milliers de tâches pour le traitement d'environ 30 TB (cf. section 1).

Ensuite, afin de centraliser la production MC dans un environnement unique, nos perspectives consistent également à intégrer au sein de CTADIRAC, des ressources de calcul, à savoir des clusters locaux, actuellement utilisées de façon indépendante pour la production MC.

Enfin, le modèle d'utilisation des ressources de la VO CTA tire bénéfice à la fois de ressources dédiées, et de ressources obtenues avec une approche opportuniste. Or, compte tenu de l'activité croissante depuis 1 an au sein de la VO CTA, nous prévoyons d'évoluer du point de vue organisationnel en formalisant ultérieurement les accords avec les sites sur l'attribution des ressources et en définissant des règles d'utilisation des ressources à l'intérieur de la VO (priorités entre les groupes, quotas sur l'utilisation du stockage, etc.).

Références :

- [1] Actis M et al. (CTA Consortium) 2011 *Experimental Astronomy* 32 193
- [2] Tsaregorodtsev A et al. 2008 *Journal of Physics Conference Series* 119 062048
- [3] Arrabito L et al. 2012 *Journal of Physics: Conference Series* 396 032007
- [4] www-ik.fzk.de/corsika/
- [5] Bernlöhr 2008, *AIP Conf. Proc.*, Vol. 1085, 874-877
- [6] <http://hepik.caspur.it/benchmarks/doku.php>
- [7] Komin N et al. <http://hal.archives-ouvertes.fr/hal-00653017>
- [8] Barnacka A et al. 2012 *Lecture Notes in Computer Science* 7136/2012 301-313
- [9] <http://glite.cern.ch/>
- [10] Moscicki J T et al. 2009 *CoPhC* 180 2303
- [11] Bigongiari C., et al., In *Proceedings of the 33rd International Cosmic Ray Conference (ICRC2013)*, Rio de Janeiro (Brazil). arXiv:1307.5006