

SysFera-DS : Un portail d'accès unifié aux ressources des centres de calcul. Mise en application à EDF R&D

Benjamin Depardon (1), Samuel Kortas (2), Boris Daix (2), Renaud Barate (2)

(1) benjamin.depardon@sysfera.com, SysFera, 13 avenue Albert Einstein, 69100 Villeurbanne

(2) samuel.kortas, boris.daix, renaud.barate}@edf.fr, EDF R&D SINETICS, 1 avenue du Général de Gaulle, BP 408, 92141 Clamart Cedex

Overview

There is a complex and widely diverse set of distributed platforms (e.g., workstations, clusters, supercomputers, grids, clouds) available to users. As users often seek ease of use first and foremost, many of these platforms are not used to the maximum of their potential, and users often make do with sub-optimal performance. .

Engineers or researchers should not need to know the characteristics of the machines they have access to, be they static (e.g., performance, capacity) or dynamic (e.g., availability, load). They should be able to launch their application and let the system take care of how and where their application is run to ensure the optimal performance. For the system administrator, managing a set of heterogeneous machines can become a nightmare. Making the most of them, at the lowest possible cost, is close to impossible without using a standard set of tools accessible from a unique interface.

Due to the complexity of the infrastructure and the wide heterogeneity of users, applications and available tools, EDF R&D has to cope with daily problems such: "How do I connect to this remote machine?" "How do I run a job on this specific batch scheduler?" "How do I manage my remote files?" "Why don't I have a single account to access all of these resources"... Clearly, scientists wish to concentrate on their main field of expertise such as CAD modeling, fluid dynamics or structural mechanics simulations without having to deal with the complexity of the infrastructure.

EDF R&D and SysFera¹ have co-developed a solution named SysFera-DS that provides end-users with a unified and simple view of the resources, and which can be addressed from a web portal (the WebBoard), the command line, or through several APIs for advanced users for an easy integration into simulation tools such as Salome. SysFera-DS provides single-sign-on for all of the heterogeneous resources, and functionalities such as resource monitoring, task and file management.

Enjeux scientifiques, besoin en calcul, stockage et visualisation

En 2012, EDF R&D dispose en interne de 5 ressources dédiées au calcul intensif :

- un BlueGene/P de 3768 cœurs, sous LoadLeveler ;
- un BlueGene/Q de 65536 cœurs, sous LSF ;
- un cluster de 1984 cœurs Intel Xeon X7560 Nehalem-EX sous Torque/Maui ;
- un cluster de 384 cœurs Intel Xeon X7560 Nehalem-EX sous LSF ;
- un cluster de 17208 cœurs Intel Xeon X5670 Westmere sous SLURM.

Cette infrastructure évolue régulièrement pour accroître la puissance de traitement (calcul, stockage, échange) d'EDF R&D. Le besoin est ici de développer un logiciel pour :

- Simplifier l'accès à l'infrastructure informatique à haute performance d'EDF aux utilisateurs et les développeurs d'applications scientifiques.
- Uniformiser et de standardiser l'accès à l'infrastructure informatique à haute performance d'EDF.

En particulier, ce logiciel est interfacé avec SALOME (Bergeaud *et al.* 2010). SALOME est co-développé par EDF et le CEA et constitue une plateforme générique open-source (licence L-GPL) qui propose des outils de pré/post-traitement et de couplage de codes de calcul pour la simulation numérique. SALOME est utilisée comme une plateforme d'intégration de codes de calcul numérique afin de créer une nouvelle application à partir des composants de base de SALOME et des codes intégrés dans cette plateforme. SALOME. Basée sur une architecture ouverte et flexible faite de composants réutilisables. SALOME offre en particulier des fonctionnalités :

- De modélisation et de liaison CAO-Maillage,
- De mise en données et préparation des calculs,
- De conception et de supervision de schéma de calcul (couplage, calcul paramétrique, distribution de calcul),
- De gestion de calcul (lancement et suivi) sur machine distante,
- De post-traitement des résultats de ces calculs.

SysFera-DS : un intergiciel réparti et scalable d'accès à l'infrastructure

EDF R&D et SysFera co-développent un intergiciel distribué open-source qui simplifie l'accès à des ressources de calcul, et plus particulièrement aux moyens de calcul HPC internes d'EDF. Ce logiciel est constitué de quatre modules :

¹ <http://www.sysfera.com>

- UMS (User Management System) : en charge de l'authentification, gestion des droits d'accès et du single-sign-on en se basant sur ssh, LDAP et une base de données.
- IMS (Information Management System) : s'occupant des métriques sur la plate-forme telles que la charge CPU et mémoire, la surveillance des queues des batch schedulers, la gestion du cycle de vie des processus... Ces métriques sont principalement pour les administrateurs, mais également pour les utilisateurs avancés.
- FMS (File Management System) : gérant les fichiers sur les machines distantes. L'utilisateur peut ainsi créer ou supprimer des fichiers ou dossiers, modifier les propriétés des fichiers (droits d'accès, propriété...), transférer des fichiers de manière synchrone ou asynchrone entre n'importe quelle machine (locale ou distante) en se basant sur les outils scp ou rsync. Cela permet à l'utilisateur de transférer de très grandes quantités de données entre par exemple deux machines distantes sans avoir à se préoccuper de la gestion des processus de copie.
- TMS (Task Management System) : supportant la gestion et l'exécution de jobs sur différents types de batch schedulers (SLURM, LSF, LoadLeveler, Torque, SGE...). Il fournit un identifiant de job unique sur toute la plate-forme, quelque-soit le nombre et le type des batch schedulers, et permet d'écrire des scripts d'une manière générique complètement indépendante du type de batch scheduler sous-jacent. Ainsi, l'utilisateur n'a pas à se préoccuper des variables et commandes spécifiques aux batch schedulers : un seul script peut être utilisé, SysFera-DS s'occupe de gérer la conversion des variables et options en fonction de l'infrastructure sous-jacente. TMS permet également à l'utilisateur de sélectionner sur quelle machine de calcul il souhaite faire tourner son job, ou de laisser SysFera-DS choisir la « meilleure » ressource en fonction de métriques de charge.

Toutes ces fonctionnalités sont disponibles sous forme de commandes UNIX exécutables en ligne de commandes, et via de quatre autres APIs (C++, Java, Python et webservices). Ces dernières permettent une intégration aisée de la solution au sein d'outils de simulation tels que la plate-forme SALOME. Au dessus de cette intergiciel, nous avons également développé un portail web : SysFera-DS WebBoard². Via une interface Web, nécessitant un browser supportant HTML 5, le WebBoard rend accessible à l'utilisateur, quelque soit son OS, de manière graphique toutes les fonctionnalités sous-jacentes. Cela permet aux utilisateurs non-experts en informatique de faire tourner leurs simulations rapidement à l'aide de quelques clics en bénéficiant d'une manière immédiate de toute la puissance de calcul haute performance accessible en interne.

La **Figure 1** présente une vue d'ensemble de la plate-forme déployée à EDF R&D. Elle est actuellement composée de cinq clusters internes, dont quatre sont actuellement gérés avec SysFera-DS et accessibles depuis le WebBoard.

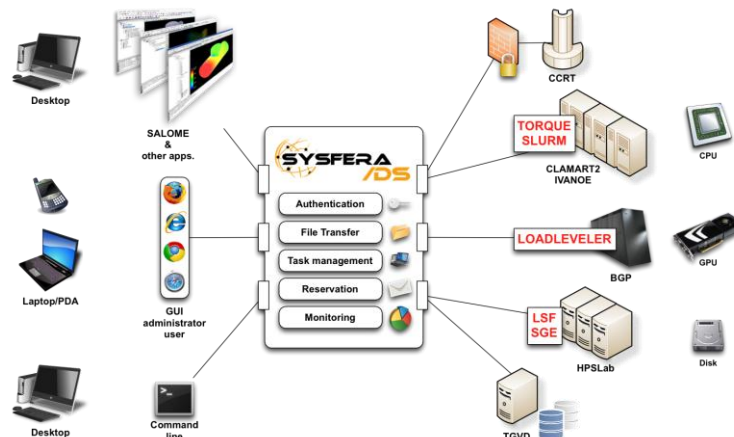


Figure 1. SysFera-DS déployé à EDF R&D.

Contraintes de l'environnement EDF R&D

La **Figure 2** présente la solution à une des contraintes et difficultés rencontrées : la gestion des utilisateurs. L'infrastructure d'EDF R&D ne fournit pas de single-sign-on sur l'ensemble de l'infrastructure : chaque cluster est géré indépendamment, certains utilisent un LDAP pour l'authentification, d'autres non. Les utilisateurs se retrouvent donc avec plusieurs logins différents associés à des mots de passe ou clés ssh à gérer pour se connecter aux différentes ressources. SysFera-DS gère et dissimule cette hétérogénéité, sous la contrainte supplémentaire qu'aucune partie de SysFera-DS ne peut avoir de privilèges root sur l'infrastructure, ce qui est une contrainte forte lorsqu'il faut être capable d'exécuter les commandes à distance avec la bonne identité. Nous avons donc résolu ces problématiques de la manière suivante :

- Les utilisateurs peuvent se connecter à SysFera-DS à l'aide d'un simple couple identifiant/mot de passe. Ces identifiants sont vérifiés soit dans une base de données spécifique, soit dans un ou plusieurs LDAPs en même temps. Une fois ces identifiants vérifiés, l'utilisateur a une session SysFera-DS d'ouverte, et il peut alors interagir avec la plate-forme.
- Au sein d'une session, chaque commande devant être exécutée à distance, est réellement exécutée avec le bon compte utilisateur en utilisant des connexions ssh : l'utilisateur donne le droit à SysFera-DS d'accéder à ses comptes locaux sur les machines distantes à l'aide de clés ssh. Ainsi, l'identité est préservée, et aucune surcouche

² Une démonstration en ligne est disponible sur <http://webboard.sysfera.com>

de gestion des utilisateurs et de la sécurité n'est ajoutée, SysFera-DS se repose sur les politiques locales déjà mises en place.

Il est important de souligner que ce cas d'utilisation est courant, particulièrement lorsque vous devez interconnecter plusieurs centres de calcul : chacun peut avoir ses propres politiques de sécurité.

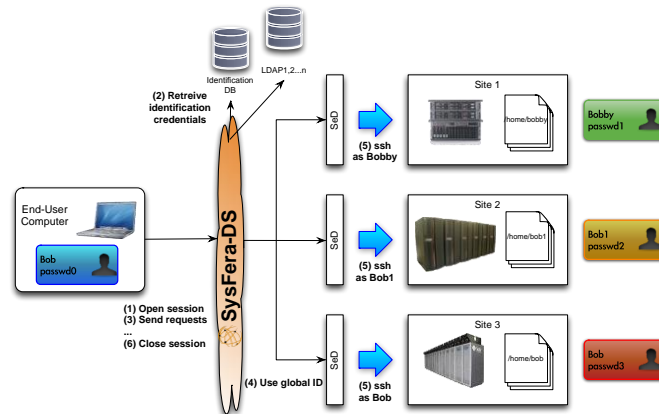


Figure 2. Single Sign On.

Un autre défi courant est l'exécution d'un logiciel distribué sur une infrastructure protégée avec des firewalls. Dans notre cas, seul ssh peut être utilisé entre les différents sites. Ce cas est fréquent à bon nombre de centres de calcul. Nous avons donc introduit des agents spécifiques dans notre plate-forme qui ont le rôle d'encapsuler toutes les communications internes à SysFera-DS dans un flux ssh. Ainsi, même si au sein d'un réseau 1 toutes les connexions restent possibles, mais sont bloquées vers un réseau 2 hormis via ssh, SysFera-DS permet malgré tout des communications globales entre tous les réseaux à l'aide de ces « forwarders ».

Résultats scientifiques

Depuis Juillet 2011, SysFera-DS a été installé avec succès sur quatre calculateurs internes.

Cet Intergiciel nous permet d'y soumettre, suivre ou annuler des travaux en utilisant les mêmes commandes. L'accès à ces ressources en est grandement amélioré (Bergeaud V., October 17-21, 2010) (Kortas S., December 2009) (Kortas S., April 2009) pour l'utilisateur : il n'est plus besoin de se connecter aux frontaux des machines et l'utilisateur peut passer l'ensemble de ses requêtes directement depuis son poste personnel.

L'accès aux fichiers est également simplifié et plus naturel : qu'ils soient locaux ou distants, il suffit de faire précéder leur chemin du nom de la machine qui les héberge pour y accéder de façon transparente sans même avoir à s'authentifier. SYSFERA-DS propose également de nouveaux services comme les transferts asynchrones ou la synchronisation de répertoires.

Pour l'ensemble des utilisateurs, cet accès facilité aux ressources ouvre la voie à leur utilisation plus naturelle et efficace. À noter entre autre, que

- Il n'est plus besoin de se préoccuper de détails comme la syntaxe particulière d'un ordonnanceur, le nom réseau de la machine, son propre login ou mot de passe. Seule l'authentification initiale au middleware est requise.
- le passage d'une machine à une autre est plus immédiat : dans le cas d'un code porté sur les deux supercalculateurs, seul le remplacement du nom de la machine à la soumission d'un job suffit.
- L'utilisateur conserve la même vue des ressources qu'il utilise (fichiers ou jobs) depuis l'ensemble des machines auxquelles il se connecte.
- Le transfert de fichier important en taille est moins problématique car il est asynchrone et entièrement pris en charge par l'intergiciel.

Pour les exploitants, la présence de cette couche d'abstraction rend possible une évolution plus dynamique des ressources pour peu que l'interface avec l'intergiciel SYSFERA-DS soit préservée.

Pour les développeurs d'applications, la standardisation de ces commandes offre l'opportunité d'un développement plus portable, indépendant des spécificités matérielles et logicielles de l'infrastructure.

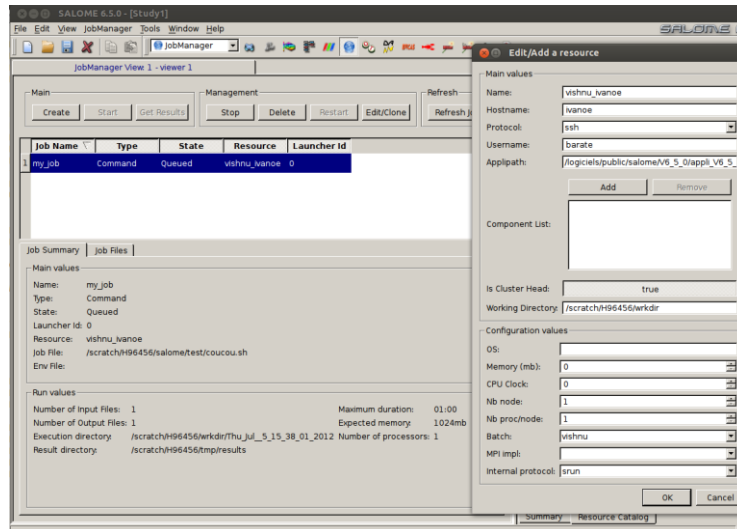


Figure 3. Utilisation de SysFera-DS depuis SALOME.

En particulier, le client SYSFERA-DS est désormais livré avec la plate-forme SALOME (voir **Figure 3**). Via l'interface graphique du gestionnaire de calcul (module JOBMANAGER), de cette plate-forme, la soumission de travaux sur les moyens de calculs adressés par l'intergiciel est transparente et apparaît comme un choix supplémentaire aux gestionnaires de batch classiques déjà disponibles dans SALOME.

L'intérêt de cet interfaçage est à court terme pour SALOME de pouvoir immédiatement accéder par ce biais à toute nouvelle machine intégrée à SYSFERA-DS sans nécessiter la moindre réinstallation ou mise à jour nécessaire.

À moyen terme, cette stratégie de répartition des rôles présente l'avantage pour SALOME de bénéficier de fonctions avancées ultérieurement disponibles dans SYSFERA-DS (réservation de ressources, de répartition de charge entre différents clusters, publication d'application...).

Conclusion et travaux futurs

De nombreuses nouvelles fonctionnalités sont actuellement à l'étude pour les inclure dans les prochaines versions de SysFera-DS, par exemple :

- Visualisation distante : télécharger les fichiers de résultats sur le poste de travail pour les post-traiter peut être très long, et donc inefficace. Il est souvent plus rentable de les traiter directement à distance, et de fournir une solution de visualisation distante pour accéder aux logiciels de post-traitement. Nous prévoyons d'ajouter des fonctionnalités de visualisation distante dans SysFera-DS, plus particulièrement dans le WebBoard.
- Desktop computing : les clusters sont efficaces pour les grosse simulations, mais pour des calculs plus petits et indépendants, les ressources de calcul que représentent les postes de travail peuvent être utilisées quand leurs propriétaires ne les utilisent pas. Des solutions existent pour gérer ces ressources, telle que BOINC (Anderson, 2004). Nous prévoyons d'interfacer TMS avec BOINC de telle façon qu'il soit vu comme une sorte de batch scheduler particulier.

Comme montré sur la **Figure 1** deux autres centres de calcul pourraient être rendus disponibles pour EDF R&D à travers SysFera-DS : le CCRT ("Centre de Calcul Recherche et Technologie") et le TGVD ("Très Grand Volume de Données"). Via SYSFERA-DS, ces ressources externes seraient alors accessibles.

Bibliographie

- Anderson, D. (2004). BOINC: A system for public-resource computing and storage. *Grid Computing, 2004. Proceedings. Fifth IEEE/ACM International Workshop on* (pp. 4-10). IEEE.
- Bergeaud, V., & Lefebvre, V. (October 17-21, 2010). SALOME: a software integration platform for multi-physics, pre-processing and visualization. *Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010 (SNA + MC2010)*. Hitotsubashi Memorial Hall, Tokyo, Japan..
- Kortas, S. (April 2009). Fonctionnalités attendues d'un portail scientifique d'accès aux ressources de Calcul Haute Performance. *EDF Internal Report H-I2D-2009-01134-FR*.
- Kortas, S., Mouton, C., Ribes, A., & Dutka-Malen, I. (December 2009). Accès standardisé aux ressources HPC : rôles, cas d'utilisations, et premiers éléments d'architecture fonctionnelle. *EDF Internal Report H-I2D-2009-01173-FR*.