

Cluster pour études d'erreurs de dynamique faisceau

Romuald Duperrier, Didier Uriot, Nicolas Pichoff

Laboratoire de Modélisation des Accélérateurs et Aimants
CEA/DSM/DAPNIA/SACM

mercredi 17 avril 2002

Introduction

Les besoins en puissance de calcul pour la dynamique de faisceaux intenses vont croissants. Les études d'erreurs nécessaires pour déterminer la sensibilité d'un accélérateur aux imperfections réclament des calculs de plus en plus lourds (mémoire), longs (temps CPU) et nombreux (statistiques). Une façon de répondre à cette demande est évidemment d'utiliser des calculateurs de plus en plus puissants et, par voie de conséquence de plus en plus coûteux. Ces supercalculateurs sont le plus souvent composés de nombreux processeurs (16, 32, 256 et plus parfois). Cette approche fut longtemps la seule utilisée. Elle est d'ailleurs toujours très compétitive pour le calcul parallèle et lorsque un processus unique mais gourmand doit être lancé. Aujourd'hui, une approche plus économique se développe dans de nombreux laboratoires. Elle consiste à relier via un réseau, dédié ou non, des PCs. Cet ensemble constitue un cluster. Selon les logiciels utilisés pour la gestion du cluster, un calcul parallèle ou multi paramètre peut être distribué sur cette ferme de PCs. L'argument économique est évident et peut être poussé à l'extrême en n'utilisant que des machines déjà dans le parc. Il n'est alors pas nécessaire de déboursier le moindre centime supplémentaire pour le matériel. Des entreprises, comme EDF, proposent d'utiliser les PCs de bureautique durant la nuit et les week-ends. Cette approche est pertinente pour le calcul multi paramètre qui se compose de nombreux processus chacun relativement peu gourmands en ressources. L'étude d'erreurs d'un accélérateur rentrant dans cette catégorie¹, le développement d'un cluster utilisant le réseau LAN local a été réalisé au SACM. Ce rapport détaille les principes de la couche logiciel qui gère le cluster : architecture client/serveur, l'interface dans TraceWin pour la gestion du client.

Architecture client/serveur

Le traitement d'un calcul multi paramètre s'effectue grâce à une architecture client/serveur. Un client rassemble les nombreux fichiers d'entrée et exécutables. Il sollicite un ou plusieurs serveurs installés sur différentes machines pour le transfert des fichiers puis pour le lancement des processus. Il interroge les serveurs pour connaître l'état d'avancement des calculs. Le cas

¹ Un linac peut être calculé sur un PC standard mais une bonne statistique réclame plusieurs centaines voire plusieurs milliers de passages.

échéant, il peut demander l'interruption d'un processus. Enfin, il peut rapatrier le ou les fichiers de sorties. Le client doit être lancé selon la nomenclature suivante :

Twclient.exe [*adresse IP*] [*répertoire de travail distant*] [*commande*] [*fichier*]

où:

- *adresse IP* est l'adresse réseau à quatre champs (exemple :132.166.12.255) ou bien le nom (exemple : seapcb255.saclay.cea.fr) de la machine distante.
- *répertoire de travail distant* est le répertoire où seront écrits les fichiers (exemple : si l'argument est *tutu*, sur une machine win32 le répertoire *c:\temp\seapcb255_tutu* sera créé si *C:* est le disque système ; sur une machine Unix, */tmp/seapcb255_tutu* sera créé).
- *commande* décrit le type d'action à exécuter.
- *fichier* sert d'argument à la commande qui le précède.

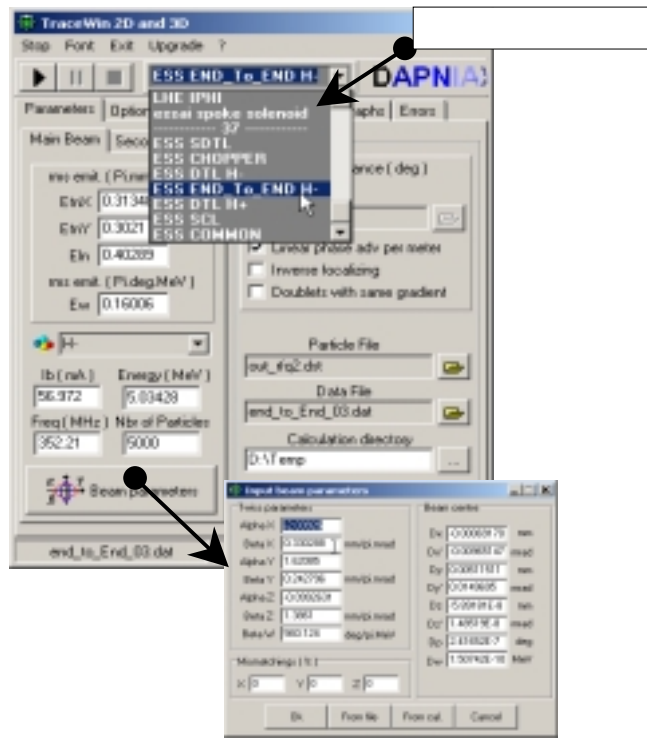
Les différentes commandes sont détaillées dans le tableau ci-dessous.

Commande	Description
<i>upload</i>	Transfert <i>fichier</i> depuis la machine cliente dans <i>répertoire de travail distant</i> sur la machine <i>adresse IP</i>
<i>download</i>	Rapatrie sur le répertoire courant de la machine cliente <i>fichier</i> situé dans <i>répertoire de travail distant</i> sur la machine <i>adresse IP</i>
<i>exec</i>	Exécute <i>fichier</i> sur la machine <i>adresse IP</i> dans <i>répertoire de travail distant</i>
<i>testexec</i>	Teste si <i>fichier</i> tourne sur machine <i>adresse IP</i> dans <i>répertoire de travail distant</i>
<i>clean</i>	Efface <i>fichier</i> dans <i>répertoire de travail distant</i> . Si <i>fichier</i> est absent de la commande, tout <i>répertoire de travail distant</i> sera effacé.
<i>kill</i>	Interrompt <i>fichier</i> sur machine <i>adresse IP</i> si celui-ci tourne dans <i>répertoire de travail distant</i>
<i>getos</i>	Demande à la machine <i>adresse IP</i> son type de système d'exploitation. Le type est écrit dans <i>os.txt</i> sur le répertoire courant du client.

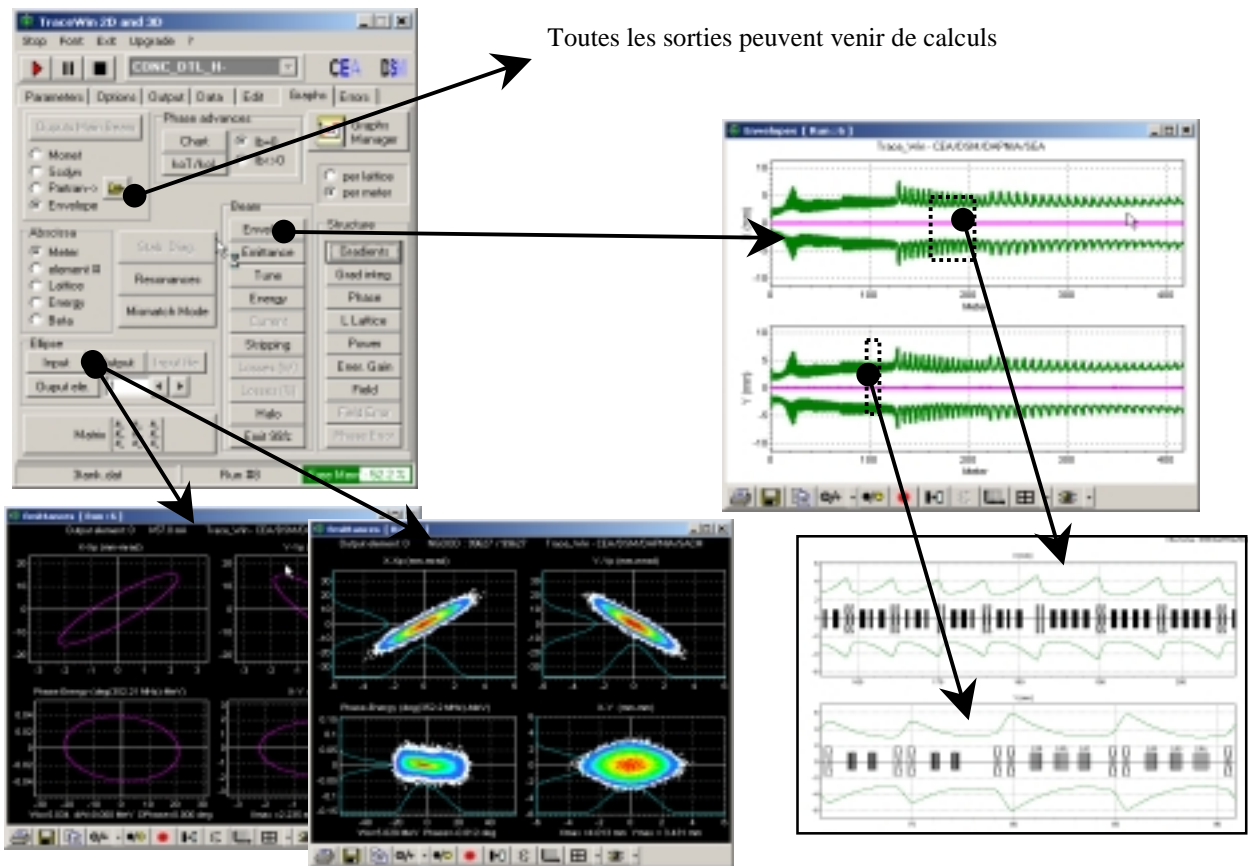
Gestion du client par TraceWin

TraceWin [1]

TraceWin est un programme de calcul d'enveloppes d'un faisceau groupé ou continu à travers différents éléments de transport. Il peut simuler un accélérateur linéaire de la source à la cible. C'est un code linéaire incluant les effets 3D de force de charges d'espace. Une de ces principales caractéristiques est qu'il capable de mener à bien l'étude complète d'un linac incluant le calcul des adaptations entre les différentes sections, l'intégration des erreurs et les calculs des correcteurs nécessaires. De plus, Il peut exécuter PARTRAN [2] et TOUTATIS [3] (codes multiparticules) et exploiter directement les résultats. Ces fonctionnalités lui permettent donc de mener des études d'erreurs statistiques, c'est à dire, simuler N linacs incluant des distributions d'erreurs, traiter les résultats aussi bien issus d'un calcul d'enveloppe que d'une simulation multiparticulaire.



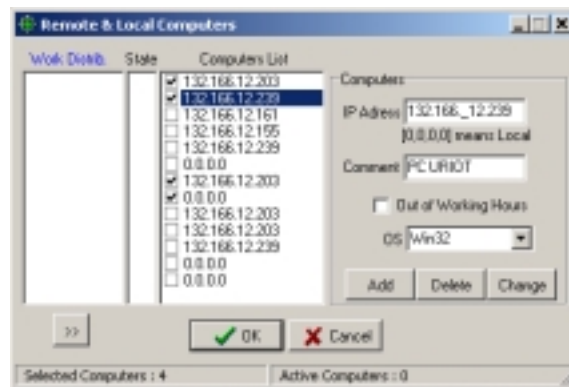
Interface pour la saisie de projets et de paramètres faisceaux de TraceWin



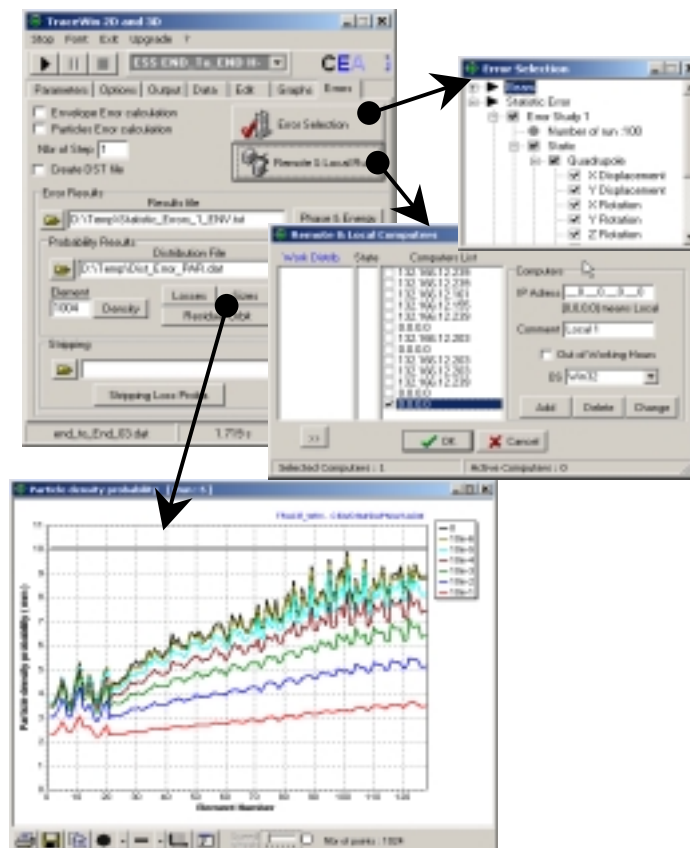
Exemples de sortie graphique de TraceWin

Interface et gestion du client dans TraceWin

Dans le cadre d'un fonctionnement de type cluster, le rôle de TraceWin est de gérer la distribution des calculs sur les différentes machines en fonction de leurs disponibilités et de leurs puissances, d'interfacer le module client et de permettre à l'utilisateur d'administrer simplement le parc de machines. TraceWin est en quelque sorte le maître d'orchestre du calcul de linac distant. Chaque calcul est indépendant et le TraceWin client se charge de transférer ses « fils », de les gérer, de rapatrier les résultats et enfin de les exploiter.



Interface permettant d'implémenter les machines et de contrôler le bon fonctionnement des processus distants



Interface pour l'exploitation des résultats (ici courbes en densités du faisceau le long du linac).

Portabilité

Afin de s'affranchir des systèmes d'exploitations, le serveur est écrit en langage multiplateforme TCL. Il a fallu également porter les codes « TraceWin remote », PARTRAN et TOUTATIS en C ou C++ ANSI. Ceci a permis de compiler différentes versions pour chaque système d'exploitation (Windows, Linux, ...). Le TraceWin client ne fonctionne pour le moment que sous Windows, mais il permet de manière transparente la distribution des calculs sur un parc de machines hétérogène (PC, station ; Windows, Linux, Solaris,...).

Sécurité

Le danger dans l'utilisation de tels logiciels est la possibilité pour un pirate informatique d'intercepter les messages entre le client et le serveur circulant en clair sur le réseau. Cette interception peut permettre, une fois le principe du fonctionnement du cluster compris, l'envoi de fichiers parasites ou pire de lancer des exécutables endommageant le système hôte. Afin de minimiser le trou de sécurité que représente un tel outil, chaque commande est précédée d'une clé cryptée. L'authentification de cette clé par le ou les serveurs permet l'exécution de la commande. Pour des raisons évidentes, le détail de ce cryptage ne sera pas détaillé dans ce rapport.

Limites et exemple

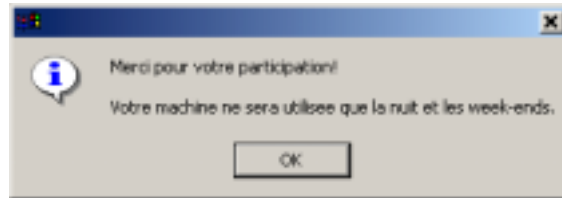
Par son principe de fonctionnement, le système mis en place autorise aussi bien l'utilisation d'une machine du réseau LAN local que d'une machine située au JAPON. Cependant, la gestion des processus via le réseau, scrutation, exportation puis importation des données rallonge la durée d'un calcul. Ces logiciels sont donc pertinents uniquement pour des calculs plus longs que le temps de transfert. Ceci est toujours le cas pour les calculs multiparticulaires mais pas toujours pour les calculs d'enveloppes. L'utilisateur peut alors utiliser Tracewin en mode local en utilisant l'adresse IP 0.0.0.0.

A titre d'exemple, le linac du projet IFMIF se compose d'un RFQ suivit de 7 tanks DTL. Le temps de calcul en enveloppe incluant le calcul des correcteurs est de l'ordre 10 s, alors qu'une simulation multiparticulaire (100.000 macroparticules) demande environ 1 h. Une grosse partie de ce temps est due à la simulation du RFQ (calcul des champs extérieurs à l'aide de grilles 3D). La même étude concernant le projet ESS demande 3 à 4 fois plus de temps, l'intérêt du cluster prend alors tout son sens. Pour obtenir une statistique correcte, en particulier si on cherche des pertes faisceau de l'ordre de 10^{-7} , au moins 1000 simulations sont nécessaires. Ceci représente 42 jours de calcul sur une machine. Par contre, si 30 % du parc des machines du SACM non utilisés la nuit, sont reliées au cluster, à peine une nuit suffit.

Appel à contribution

L'intérêt de ce type de distribution de calcul ne réside pas dans la puissance des machines mais dans leur nombre, toute personne possédant un PC sous Windows et acceptant que celui-ci soit mis à contribution **en dehors des heures de service** est la bienvenue. Elle contribuera à la noble cause des linacs hautes intensités et nous lui en serons infiniment reconnaissants ! A noter qu'un minimum de 128 Mo de RAM est tout de même nécessaire.

La procédure est très simple: Se rendre sur « DAPDIV\\Dapnt7\\Duperrrie_Pub\\TWEErrors» puis exécuter le programme « INSTALL.EXE », l'installation est automatique et ne prend que quelques secondes et se termine en cas de réussite par la fenêtre ci-dessous. Il est inutile de nous avertir une fois l'installation effectuée, un fichier est automatiquement mis à jour. Le serveur est relancé automatiquement par la suite, il est donc inutile de relancer l'installation à chaque session. En cas de problème, n'hésitez pas à nous contacter (<mailto:rduperrier@cea.fr;durriot@cea.fr>).



Lien direct : \\Dapnt7\\Duperrrie_pub\\TWEErrors\\INSTALL.exe

Conclusion

Cet outil a été développé dans le but de partager les études d'erreurs d'un linac. Il pourrait être utilisé pour un autre calcul multi paramètre. Il est disponible pour toute personne ou application désirant partager des processus indépendants.

[1] D. Uriot & N. Pichoff, « TraceWin Documentation », rapport interne DSM/DAPNIA/SEA 2000/45.

[2] N. Pichoff et al., « Simulations results with an alternate 3D space charge routine, PICNIC », MO4042, Linac 1998, Chicago

[3] R. Duperrier, « TOUTATIS, a RFQ code », Physical Review Acc. & Beams, décembre 2000.