

Un'evoluzione dell'Italian Data Transfer System di AMS: dai dati al MonteCarlo

Matteo Boschini

CILEA, Segrate

Abstract

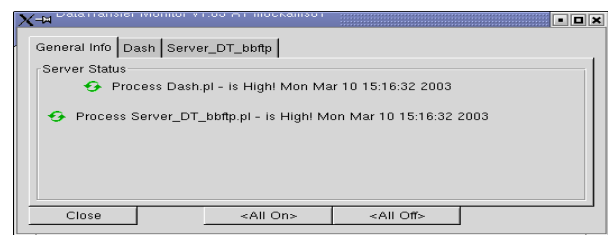
L'Italian Data Transfer System implementato per AMS-02 è stato approvato dalla collaborazione internazionale. Inoltre a partire da esso è stato sviluppato un sistema che permette e semplifica il trasferimento dei dati MonteCarlo, prodotti in 14 laboratori nel mondo, presso la *Central Production Facility* del CERN di Ginevra.

Keywords: Scienze, Database, Networking.

Il sistema di trasferimento dati [1], [2] implementato per l'esperimento AMS nell'ambito della collaborazione CILEA-INFN di Milano, è stato formalmente presentato nel Luglio scorso a tutta la collaborazione internazionale di AMS durante il *Technical Interchange Meeting*, alla presenza dello *spokesman* S. Ting e dei responsabili NASA e CERN. In tale sede, la collaborazione internazionale, composta di più di 200 scienziati provenienti da circa 60 tra Università e centri di ricerca internazionali, visti gli ottimi risultati dello stress test del sistema da noi sviluppato, ha deciso di adottarlo per trasferire i dati dal *Science Operation Center*, situato al CERN, all'*Italian Ground Segment Data Storage*.

Infatti il sistema, basato su architettura *client/server* e implementato con tecnologie *Free Software*, è stato sottoposto ad uno stress-test della durata di 9 mesi, durante i quali sono stati trasferiti 4 Tbyte di dati, dimostrando la sua affidabilità¹ ed efficienza. Durante lo stress-test si sono, ovviamente, sistemate alcune problematiche di implementazione, ma soprattutto si sono ottimizzati i parametri di uso della rete (in particolare il numero di *stream TCP* paralleli e la *TCP_WIN_SIZE*), si sono apportate delle modifiche al protocollo di trasferimento utilizzato (*bbftp*[3]), che ha

dimostrato una inaffidabilità dello 0.03%² e si sono ottimizzate le procedure di *book-keeping*³, fondamentali per il sistema stesso. Completa il sistema un meccanismo di replica dei DB su un terzo host, che nel modello finale sarà installato presso un terzo sito⁴. Attualmente la replica avviene per mezzo di un semplice meccanismo di *export/import*, ma è già in fase di test un meccanismo che usa anch'esso l'architettura *client/server*[4]. Infine il sistema è stato completato con una interfaccia grafica, sviluppata in *Gimp-Perl-Gtk* (Fig. 1), per la gestione dei processi e un sistema di *browsing* web del DB.



² Nella *release* attuale del nostro sistema, le modifiche sono state apportate al sorgente di *bbftp* e comunicate al team di sviluppatori di *bbftp* stesso. Inoltre, è stato implementato un sistema di *consistency check*, che è in grado di ovviare a tale inefficienza.

³ Basato su *MySQL*.

⁴ Presso l'*AMS Data Center* della sezione di Milano degli INFN, installato presso il Dipartimento di Fisica dell'Università degli Studi Milano-Bicocca.

¹ Superiore al 99.8%.

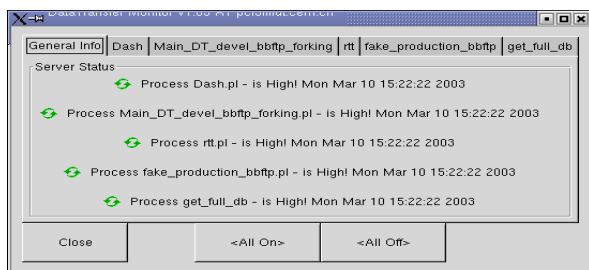


Figura 1 - Esempi della GUI

Parallelamente allo stress-test, durante alcune riunioni dell'AMS *Software Group* tenutesi nell'arco dell'estate del 2003, è emersa l'idea di sviluppare, a partire dal Data Transfer System, un sistema per il trasferimento e la *book-keeping* della produzione MonteCarlo dell'esperimento. La produzione MonteCarlo avviene oggi in modo distribuito in 14 università⁵, ognuna delle quali esegue i job di produzione localmente ed invia alla CPF del CERN i dati prodotti. Ovviamente uno scenario di questo tipo richiede alcune modifiche non trascurabili al sistema. In particolare, le principali differenze sono: a) nel caso del MC si devono gestire 14 sorgenti di dati che trasferiscono ad un'unica destinazione, b) a priori, non si può imporre l'installazione di software aggiuntivo presso i siti remoti. Di conseguenza il software è stato modificato nei seguenti elementi:

1. *book-keeping* locale implementato o attraverso DB o file locale (Tie/hash);
2. layout DB presso CPF modificato per tenere traccia dei diversi siti di produzione;
3. DB consistency centralizzata presso la CPF;
4. Integrazione con il MC Production Control System⁶.

In particolare il punto 1) è di importanza fondamentale: non si può infatti imporre l'installazione di un RDBMS ai siti remoti, e si è dovuto quindi fornire un pacchetto auto-consistente. A tal fine, tutte le librerie Perl e C necessarie sono state incluse nel pacchetto stesso, in modo da renderlo il più possibile indipendente da quanto installato sull'host remoto⁷. Il pacchetto risulta in definitiva un *tarball*⁸ che, una volta installato, permette l'uso

⁵ 7 europee, 4 statunitensi e 3 asiatiche.

⁶ Insieme di procedure centralizzate volte a garantire la consistenza del campione MC generato.

⁷ Unica eccezione è la libreria OpenSSL: essa è infatti ormai sempre presente su macchine Unix/Gnu-Linux, e quindi il pacchetto punta alla loro versione dinamica installata sul sistema.

⁸ Del tipo package.tar.gz, conforme allo standard Gnu.

del sistema di trasferimento dati MC e *book-keeping* senza ulteriori requisiti.

L'uso del pacchetto rende completamente trasparente agli utenti dei siti remoti sia la procedura di trasferimento, essendo essa del tutto automatizzata, che quella di *book-keeping*, che essendo centralizzata (v. Fig 2), permette una ulteriore verifica di consistenza interna del MC.

Il sistema è stato sottoposto ad un periodo di stress-test di 3 mesi da parte del gruppo AMS di Milano, che lo ha utilizzato per trasferire e organizzare in catalogo circa 800 file di MC. Attualmente è in fase di beta-test presso la Yale University e Universitaet Karlsruhe, oltre che presso gli INFN Milano. Si prevede la sua adozione da parte di tutti e 14 i siti entro la fine di ottobre.

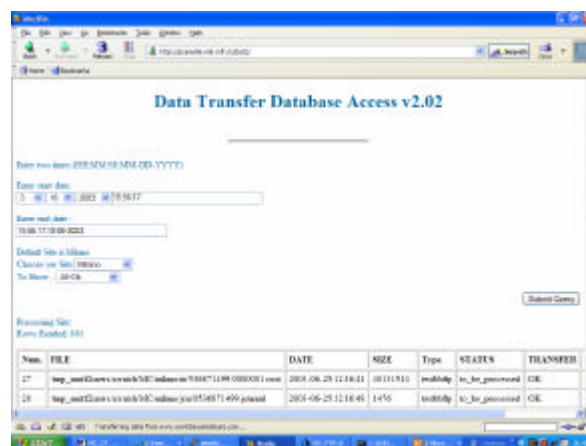


Figura 2 - DB Web browser per il MC DT

Quest'ultima implementazione ed evoluzione del sistema originario di DT dimostra, a nostro avviso, la scalabilità del sistema originario, che si va ad aggiungere alla sua affidabilità ed efficienza.

Acknowledgments

A. Favalli, M. Levchenko, M. Mezzadri, E. Micelotta (INFN – sez. di Milano).

Bibliografia

- [1] M. Boschini, *La Gestione dei dati per l'esperimento AMS*, Bollettino del CILEA, n. 78, Giugno 2001.
- [2] M. Boschini, A. Favalli, *Stato del Progetto AMS*, Bollettino del CILEA, n. 84, Ottobre 2002.
- [3] <http://doc.in2p3.fr/bbftp/>
- [4] M. Boschini et al., *AMS Italian Data Transfer System*, atti 8th ICATPP Conference, Villa Erba, 2003.