

Internal Report INAF/IASF-BO 551/2009

**PORTING DELLA PIPELINE DI PLANCK SULLA
PIATTAFORMA POWER6**

DE ROSA A, FRANCESCHI E., FATIGATI G.

INAF / IASF – Bologna

Dicembre 2009

PORTING DELLA PIPELINE DI PLANCK SULLA PIATTAFORMA POWER6

DE ROSA A., FRANCESCHI E., FATIGATI G.

Abstract

1. INTRODUZIONE

Lo scopo di tale attività è quello di installare e verificare la funzionalità della pipeline di analisi dati di Planck sulla piattaforma Power6 IBM di ultima generazione disponibile presso CINECA. La pipeline era già stata in parte testata su piattaforme simili come i Power5+ disponibili presso il centro di calcolo del NERSC. L'importanza fondamentale di tale attività è quella di permettere ai ricercatori afferenti al progetto PLANCK di poter usufruire per l'analisi dei dati della missione delle risorse di calcolo messe a disposizione da CINECA, il maggiore centro di calcolo italiano (il quale risulta essere tra l'altro un consorzio interuniversitario), fornendo quindi un'alternativa ai centri di calcolo americani nell'affiancare il Data Processing Center di Trieste.

2. AMBIENTE DI SVILUPPO DISPONIBILE SU POWER6

Il sistema Power6 di IBM è costituito da nodi a 32 processori con 128GB di RAM a nodo, gestiti dal sistema operativo AIX 6.1 di IBM. I compilatori disponibili sul cluster sono i compilatori IBM xlf e xcc nella versione 11.1, sono inoltre disponibili le librerie di calcolo scientifico ESSL di IBM nelle versioni seriale, parallela open, parallela MPI. L'ambiente di sviluppo è corredato poi delle più diffuse librerie di calcolo open source utilizzate in ambito scientifico, come scalapack, blacs, lapack, gsl. È stato quindi necessario aggiungere le sole librerie cfitsio e healpix ampiamente utilizzate in ambito astrofisico per la gestione dei dati, in particolare delle mappe.

Occorre inoltre ricordare che il file system utilizzato sul cluster è GPFS, e che i nodi del cluster risultano interconnessi con una rete a bassa latenza Infiniband 10GB/s.

3. COMPILAZIONE

La compilazione della pipeline è stata effettuata preferendo i compilatori xlf_r, per i sorgenti in Fortran 77, xlf90_r, per i sorgenti in Fortran 90, xcc_r per i sorgenti in C standard, xCC_r per i sorgenti in C++. La scelta è stata fatta in modo da avere una maggiore sicurezza sulla generazione di codice thread safe, i compilatori IBM contrassegnati da _r hanno difatti la caratteristica di tentare una compilazione thread safe.

Utilizzando come primo tentativo di compilazione quello già utilizzato per la piattaforma power5+, si è poi proceduto ad una successiva variazione delle opzioni di compilazione in modo da permettere al compilatore di eseguire delle ottimizzazioni sul codice che meglio si prestano all'architettura del Power6.

Abbiamo tentato di compilare il codice infine utilizzando anche le ottimizzazioni interprocedurali presenti nell'ultima release del compilatore IBM, in particolare le opzioni `-O5 -qipa=level=2`, le quali prevedono una compilazione in più fasi in modo da permettere un'analisi interprocedurale che permetta la generazione di un eseguibile che tenga conto delle dipendenze fra le funzioni presenti nel codice permettendo quindi di aumentare in modo considerevole l'inlining delle funzioni. Tale modalità di compilazione non è andata tuttavia a buon fine a causa di una serie di errori in fase di linking dei codici in C++. Allo stato attuale quindi la pipeline risulta compilata con un livello di ottimizzazione `-O3 -qarch=pwr6 -qtune=pwr6 -qsmp=omp`.

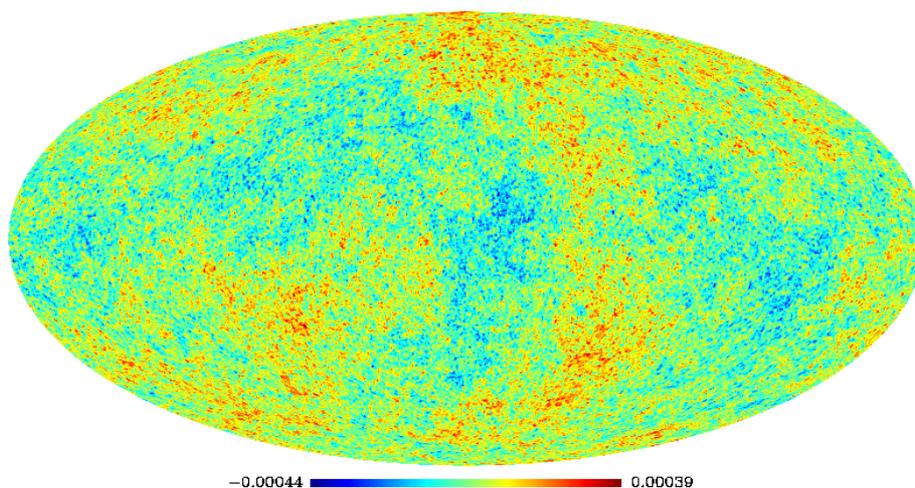
Le performance dei codici testati, in particolare `totalcovolver`, `multimod`, `madam` presentano performance superiori a quelle ottenute su cluster basati su piattaforme AMD o INTEL. I tempi di calcolo ottenuti con 960 core su piattaforme AMD sono leggermente superiori a quelli ottenuti con 256 core IBM Power6.

Occorre tenere presente che tale aumento di performance non è legato alle sole maggiori performance del processore ma anche alle elevatissime performance del file system che permette di scrivere ben 144GB di dati in meno di 5 minuti con 256 processi (32 per nodo – il che implica che 32 processi in scrittura devono condividere l'interfaccia verso il file system), ovvero con una velocità media stimata di circa 480MB/s complessiva, ovvero quasi 2MB/s per processo.

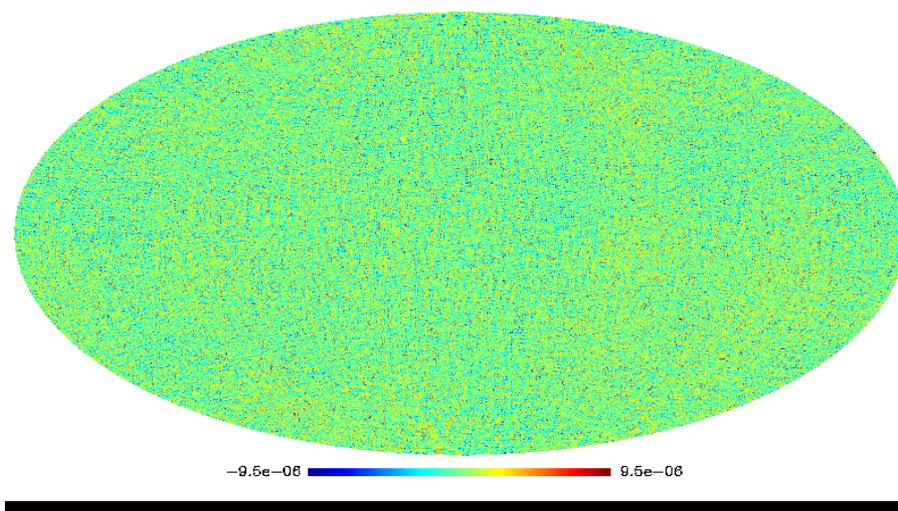
4. TEST

I test effettuati sono stati basati sulla simulazione di una mappa di segnale e di una mappa di rumore generata utilizzando i dati simulati per la missione.

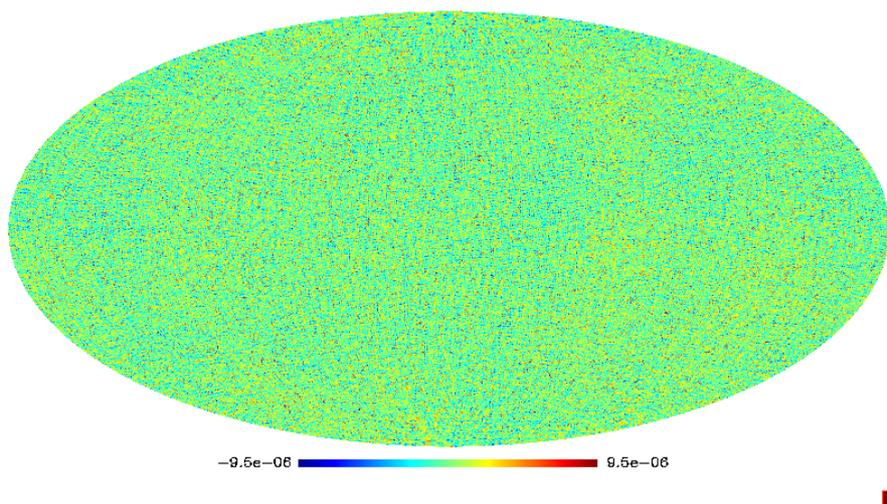
`lfi_mc1_madam_emb.all.1024map.fits: 0s`



lfi_mc1_madam_emb.all.1024map.fits: 0s



lfi_mc1_madam_emb.all.1024map.fits: 0s



Le mappe qui rappresentate sono relative ad una particolare realizzazione di segnale in T,Q,U, l'unita' di misura di tali mappe e' il K.

5. UTILIZZO DI UN DATABASE PRESSO CINECA

La pipeline di Planck prevede la possibilita' di interfacciarsi direttamente con un sistema di database per la gestione dei dati di missione. Tale interfaccia, denominata DMC, e' sviluppata dal il gruppo di MPA-Garching in collaborazione con il DPC di Trieste per essere utilizzata sulle machine del DPC per facilitare l'interfaccia fra pipeline e il sistema di database che gestisce i dati della missione. Allo stato attuale l'accesso al database e' limitato alle machine del DPC e i dati devono essere esportati in formato fits sui cluster esterni al DPC. La decisione di investigare la possibilita' di utilizzare un sistema di database e di interfacce analogo a quello impiegato a Trieste e' legato quindi semplicemente alla necessita' di doversi comunque preparare a poter coordinare e gestire elevati quantita' di dati con piu' utenti, sarebbe quindi inutile non sfruttare quanto gia' fatto presso

il nostro DPC. Tale soluzione permetterebbe inoltre un piu' facile scambio di informazioni fra i database, aumentando quindi la fruibilita' delle informazioni fra gli scienziati afferenti al progetto. Tale attivita' e' ancora in fase di svolgimento.

6. CONCLUSIONI

L'attivita' svolta per il porting della pipeline di analisi dati di Planck sulla piattaforma Power6 ha permesso e permette di poter sfruttare le risorse di calcolo messe a disposizione da CINECA sia tramite la comunita' virtuale DEISA-Planck (circa 500000 ore cpu) sia direttamente ad IASF-Bologna.