



ACADEMIA ROMÂNĂ
Institutul de Geodinamică "Sabba S.Ștefănescu"
Laboratorul Dinamica Globului Terestru

Str. Jean-Louis Calderon, Nr. 19-21, București-37, România, R-020032,
fax:(4021) 317.2120, tel. (4021) 317.2126; e-mail: inst_geodin@geodin.ro
<http://www.geodin.ro/~prezentare/>

DIRECTOR,

Dr. Crișan DEMETRESCU
Membru corespondent al Academiei Române



UNIUNEA EUROPEANĂ



GUVERNUL ROMÂNIEI



Instrumente Structurale
2007 - 2013

Programul Operațional Sectorial Creșterea Competitivității Economice
Axa prioritară 2: Competitivitate prin Cercetare, Dezvoltare Tehnologică și Inovare
Operațiunea: O.2.1.2 „Proiecte CD de înalt nivel științific la care vor participa specialiști din străinătate”

Proiect: Infrastructură cibernetică pentru studii geodinamice relaționate cu zona seismogenă Vrancea: ID-593, cod SMIS-CSNR 12499

Etapă 3: Construirea unor modele geodinamice tridimensionale de înaltă rezoluție cu ajutorul sistemului HPCC/HPVC/GeoWall

Perioada: 18 decembrie 2011-17 iunie 2012

STUDIU PRIVIND ACTIVITATEA DE PREPARARE SI RULARE MODELE NUMERICE 3D DE CONVECTIE IN MANTA SI DE DEFORMARE LITOSFERICA

Director de proiect,

Dr. Vlad Constantin Manea

A U T O R I:

Dr. Vlad Constantin Manea

Dr. Marina Manea

Drd. Mihai Pomeran

Dr. Lucian Besutiu

Dr. Luminita Zlagnean

**BUCURESTI
2012**

CUPRINS:

1. Introducere	5
2. Dinamica computationala a fluidelor in stiinta pamantului solid - o provocare HPC (High Performance Computing)	6
2.1. Introducere	6
2.2. Configuratie sistem HPCC	7
2.3. Rezultatele testelor de referinta	10
2.4. Discutii si concluzii	14
3. Un cod paralel de urmarire a particulelor pentru simularile de dinamica fluidelor in stiintele pamantului: implementare si testare pe HPCC CyberDyn	14
3.1 Introducere	15
3.2. Folosirea particulelor marculator in contextul stiintelor pamantului	16
3.3. Algoritmul de urmarire a marculatorilor	17
3.4. Teste de referinta ale codului	23
3.5. Concluzii	25

1. INTRODUCERE

Scopul principal al proiectului CYBERDYN este de a construi o infrastructura cibernetica in cadrul Institutului de Geodinamica al Academiei Romane de Stiinte din Bucuresti, Romania, pentru studierea evolutiei geodinamice pe termen lung a zonei seismogene active Vrancea. Aceasta infrastructura cibernetica va fi formata dintr-un HPCC (High Performance Computing Cluster – Grup de Servere pentru Calcule de Inalta Performanta), un HPVC (High Performance Visualization Cluster – Grup de Servere pentru Vizualizare de Inalta Performanta) si un sistem de Vizualizare Stereo in 3D (GeoWall).

Noua infrastructura cibernetica va ajuta la crearea unui corp de cercetatori format din experti cu inalta pregatire obtinuti prin antrenarea tinerilor oameni de stiinta in campul geodinamici computationale, permitand generarea primului centru de excelenta in domeniu din Romania. Activitatea acestui centru de excelenta se va extinde si dupa finalizarea ultimei etape a proiectului prin formarea tinerilor specialisti si prin participarea in proiecte nationale/internationale bazata pe capacitatea si performanta oferite de o asemenea tehnologie.

Paleoreconstructiile reprezinta procesul de reconstructie a pozitilor placilor tectonice unele in raport cu altele (miscari relative) sau in raport cu un alt sistem de referinta, cum ar fi campul magnetic terestru sau pozitia punctelor fierbinti in timp geologic. Acest proces ajuta la determinarea formei si a orientarii supercontinentelor sau a microplacilor si ajuta la conturarea paleoreconstructiilor geografice.

In acest studiu vom prezenta activitatea de preparare si rulare a modelelor numerice 3D de convecție in manta si de deformare litosferica, activitate care s-a concentrat pe doua directii principale:

- 1) Realizarea unor modele de convecție de inalta rezolutie si implementarea lor pe infrastructura de calcul CyberDyn. In acest scop s-a urmarit si studierea comportamentului supercomputerului CyberDyn atunci cand este folosit pentru rulare de modele numerice de convecție in manta. Astfel, s-au rulat o serie de modele numerice de convecție la nivel global, cat si la nivel regional, in care s-a folosit un numar crescator de procesoare de la 12 la 1024. Pentru eficientizarea procesului de cercetare cu ajutorul supercomputerului CyberDyn, s-a masurat randamentul real al infrastructurii CyberDyn, avand ca scop principal gasirea numarului optim de

procesoare in functie de rezolutia modelului de convecție in manta folosit. In acelasi timp, s-au comparat rezultatele obtinute cu sisteme de calcul similare, cum ar fi de exemplu, supercomputerul Horus din cadrul Laboratorului de Geodinamica Computationala, Centrul de Geostiinte, UNAM, Mexico.

2) A doua directie principala de cercetare, a fost realizarea unui cod paralelizat de particule (tracers) care sa fie integrat in codul de simulare numerica de convecție in manta (CitcomS). Acest cod de particule este folosit intensiv in simularile numerice de convecție, datorita faptului ca ofera posibilitatea ca de a monitoriza in timp geologic miscarile termo-chimice din interiorul Pamantului.

2. DINAMICA COMPUTATIONALA A FLUIDELOR IN STIINTA PAMANTULUI SOLID – O PROVOCARE HPC (HIGH PERFORMANCE COMPUTING)

2.1. INTRODUCERE

În ultimii ani, modelarea si calculul numeric au ajuns să joace un rol central în științele moderne ale pământului, și unul din motive se datorează dependenței lor de rețelele de discretizare de rezoluție mare și de pașii mici de discretizare necesari pentru integrare, acestea fiind utilizate pentru rezolvarea sistemelor de ecuații numerice ce exprimă matematic procesele fizice. În prezent, științele pământului au început să se îndrepte către implementarea facilităților de cercetare care dispun de capacități de calcul foarte performante. Când aceste capacități de calcul sunt optimizate pentru nevoile specifice comunității oamenilor de știință din domeniul științelor pământului, și cuplate cu programele open source deja disponibile comunității, o astfel de infrastructură de calcul performantă, cu siguranță oferă o unealta cheie care permite progrese majore în acest domeniu captivant de cercetare. Metodele numerice au progresat până la punctul în care simulările numerice au devenit o parte centrală a științelor moderne ale pământului, în special pentru geodinamică. Astfel de sisteme de calcul sunt construite în mod special pentru a raspunde nevoilor de simulare ale comunității științelor pământului. Alcătuirea acestora include numărul mare de nuclee de calcul, capacitate de stocare rapidă și fiabilă, cantitate considerabilă de memorie, totul configurat într-un sistem conceput

pentru o funcționare de lungă durată. Una dintre capabilitățile cheie ale HPC este performanța, iar proiectarea unei soluții HPC adaptate la un domeniu specific de cercetare precum științele pământului, care au nevoie de un raport optim preț / performanță, este de multe ori o provocare. Performanța sistemului HPC depinde foarte mult de interacțiunea software-hardware, și prin urmare, este importantă cunoașterea cât mai bună a performanțelor software-ului paralel specific pe diferite arhitecturi HPC pentru că aceste cunoștințe pot cântări semnificativ în alegerea configurației finale.

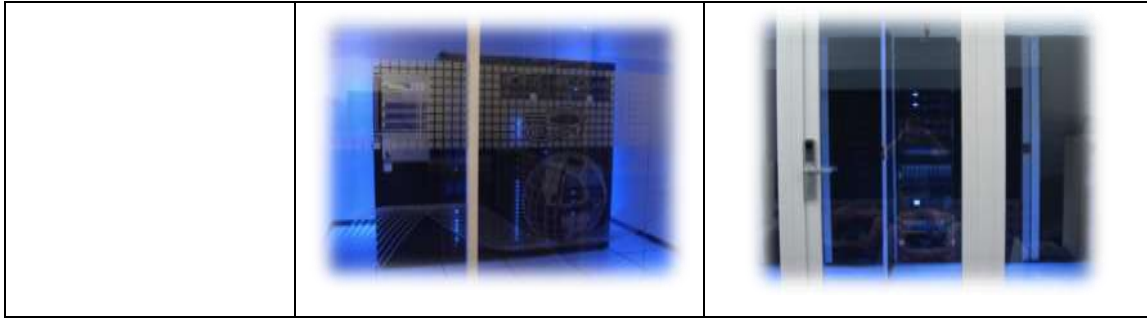
2.2. CONFIGURAȚII DE SISTEME HPCC

În această lucrare se prezintă și evaluează rezultatele testelor de performanță de la două tipuri diferite de sisteme HPC: una low-end HPCC (Horus), cu 300 de nuclee (Tabelul 1), cu 1.6 teraflopi performanță teoretică de vârf, și un high-end HPCC (CyberDyn), cu 1344 de nuclee și 11.7 teraflopi performanță teoretică de vârf (Tabelul 1). Horus utilizează sistemul de operare CentOS 5 și ca mediu de administrare software-ul open source Rocks Cluster Distribution-5.1 (<http://www.rocksclusters.org>). CyberDyn folosește Bright Cluster Management (<http://www.brightcomputing.com>) ca mediu de administrare și Scientific Linux (www.scientificlinux.org) ca sistem de operare.

Tabelul 1

	HPCC <i>Horus</i>	HPCC <i>CyberDyn</i>
Nod principal		
Model server:	1 x Dell PE2970	2 x Dell R715 (configurat failover)
Tip procesor:	AMD Opteron 2 x Dual core 2.8 GHz	AMD Opteron 2 x Twelve core 2.1 GHz
Memorie:	8 GB RAM	68 GB RAM
Plăci:	PERC 5i Integrated Raid Controller	PERC H200 Integrated Raid Controller Broadcom NetXtreme Dual port SFP + Direct Attach 10

		GbE NIC
Noduri de calcul		
Model server:	38 x DELL Sc1435/PE1950	28 x DELL R815
Tip procesor:	AMD 2 x Quad core 2.6 GHz Intel 2 x Quad core 2.6 GHz	AMD 4 x Twelve core 2.1 GHz
Memorie:	16-32 GB/node	96-128 GB/node
Infrastructură de rețea:	1000T Ethernet (1 Gbit/sec)	InfiniBand QDR (40 GBit/sec)
Rețea		
Trafic MPI	HP Procurve 48G – 2900 1000T-48 port fără management, switch performant, 10 GbE și posibilitate de stacking	Qlogic 12300-BS01 36 port InfiniBand Quad Data Rate
Management/IPMI	Folosim același Switch pentru traficul MPI și administrarea clusterului.	2 x Dell PowerConnect 6248 – 48 Port cu management Layer 3, 10 GbE și posibilitate de stacking
Stocare		
	2 x dispozitive Direct Attached Storage , conectate la nodul principal. Fiecare având un disk RAID 5 cu o capacitate de 15 TB.	1 x Panasas 8 Series cu 40 TB capacitate de stocare.
Sistem de răcire (temperatura 21°±2°C, umiditate 50%±5%).		
	Liebert 8 ton	Liebert 2 x 7 ton
Vedere în interior		



În afară de numărul nucleelelor de calcul, principala diferență dintre cele două sisteme HPC este infrastructura de rețea de mare viteză. Sistemul HPC Horus utilizează un nod central de rețea Gigabit Ethernet atât pentru traficul de administrare, scrierea / citirea de date (NFS sau alte protocoale) traficul MPI cât și pentru de traficul necesar distribuirii sarcinilor. Al doilea sistem HPC CyberDyn, care este de dimensiuni mai mari, folosește două tipuri de infrastructură de rețea. Pe de o parte, rețeaua Gigabit Ethernet este folosită pentru distribuirea sarcinilor, întreținerea nodurilor, autentificarea în sistem, în timp ce a doua rețea internă QDR InfiniBand este dedicată exclusiv traficului aferent calculului paralel MPI. Pentru studiul aprofundat al deplasărilor mantalei pământului sunt necesare infrastructuri importante de HPC și instrumente software specifice. Software-ul utilizat pentru testele de referință din această lucrare este pachetul open source CitcomS, care este utilizat pe scară largă în comunitatea de geodinamică (www.geodynamics.org). CitcomS este un cod paralel de calcul al dinamicii fluidelor care folosește metoda elementelor finite, și este proiectat pentru a rezolva problemele termice de convecție relevante pentru mantaua Pământului. Scris în limbajul de programare C, codul rulează pe o diversitate de arhitecturi de computere capabile să ruleze aplicații paralele, inclusiv platforme cu memorie partajată și platforme cu memorie distribuită și se bazează pe metoda de descompunere a domeniului. După cum s-a menționat mai sus, software-ul FEM (software de rezolvare a ecuațiilor diferențiale cu metoda elementelor finite) folosit pentru acest test comparativ de referință este CitcomS. Acest cod numeric paralel are nevoie de o bibliotecă de funcții, care implementează standardul MPI (Message Passing Interface), pe cele două sisteme HPCC CitcomS este compilat cu ajutorul versiunii OpenMPI 1.4.2. Pe ambele sisteme HPCC vom folosi Sun Grid Engine (SGE), ca software de distribuire a sarcinilor pe nucleele de calcul. Pentru aceste teste de referință specifice, considerăm o serie de probleme diverse de convecție a mantalei pământului, de la convecția simplă pur termică la

mai complexa problemă a convecției termo - chimice. Toate simulările numerice au fost efectuate în domenii de analiză sferice sau regiuni ale unui domeniu sferic. Am folosit diferite densități ale rețelei de discretizare și un număr diferit de nuclee de calcul; cu cât mai mare densitatea rețelei de discretizare, cu atât mai mare numărul de nuclee utilizate. Fiecare test a fost efectuat de mai multe ori pentru a ne asigura că am obținut rezultate consistente, repetabile și precise.

2.3. REZULTATELE TESTELOR DE REFERINȚĂ.

Mai jos vom prezenta în detaliu rezultatele de referință obținute pe ambele sisteme HPCC. Având în vedere limitarea la 300 de nuclee pentru HPCC Horus, rezultatele sunt comparabile doar la simulările cu 192 nuclee de calcul pentru modelele care folosesc un domeniu de analiza sferic și 256 nuclee de calcul pentru modele regionale. Pe ambele sisteme HPC am efectuat o serie de teste de referință, folosind trei scenarii diferite de simulare FEM, două probleme simple, de convecție termică, unul ca model regional și altul ca model sferic, și o simulare mult mai complexă a convecției termo-chimice pe un domeniu de analiză sferic.

În cazul simulărilor de convecție pur termică, pe ambele sisteme HPC, am obținut performanțe similare pentru rezoluții ale rețelei de discretizare limitate la 129x129x65 noduri (figurile 1 și 2). Crescând dimensiunea rețelei de discretizare și a numărului de nuclee de calcul, CyberDyn HPCC devine mai rapid decât Horus HPCC, fenomen datorat infrastructurii de rețea de latență scăzută QDR dedicată exclusiv traficului MPI.

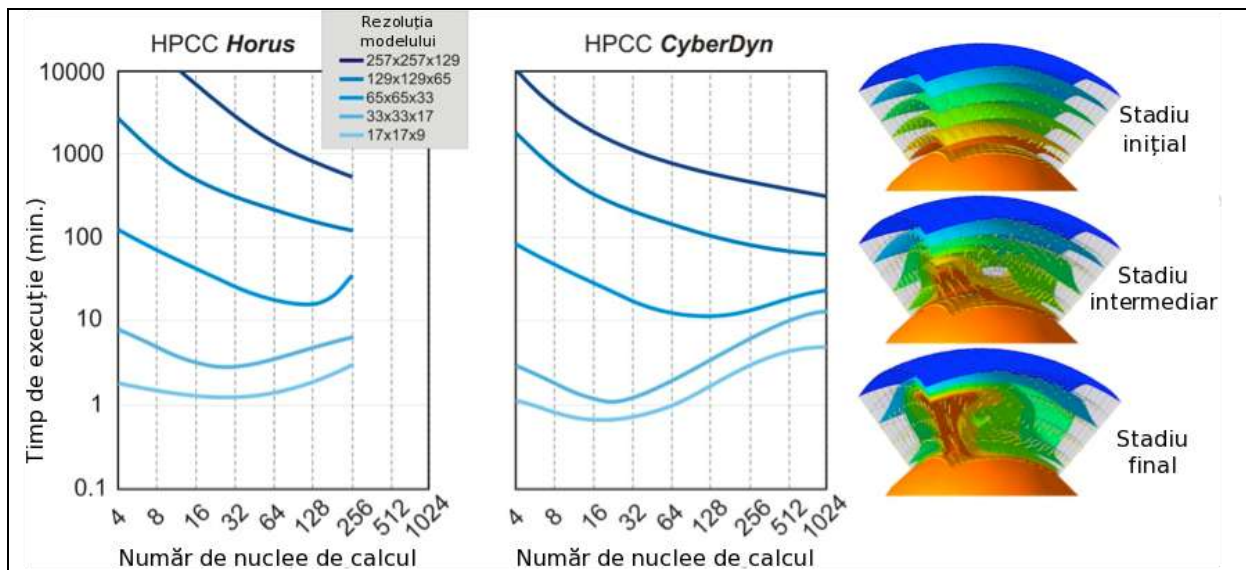


Figura 1. Comparație între rezultatele testelor de referință pe HPCC Horus și HPCC CyberDyn pentru o simulare de convecție pur termică FEM într-un model regional, care arată influența dimensiunii rețelei de discretizare asupra timpului de execuție în funcție de numărul de procesoare. În dreapta sunt prezentate, prin izosuprafețe de temperatură, trei etape din evoluția simulării de convecție termică (vizualizări efectuate cu OpenDX - software open source). Culoarele calde corespund temperaturilor ridicate, și culoarele reci reprezintă temperaturi mai scăzute în interiorul mantalei Pământului. Sfera portocalie de la etapa inițială reprezintă nucleul de fier al Pământului.

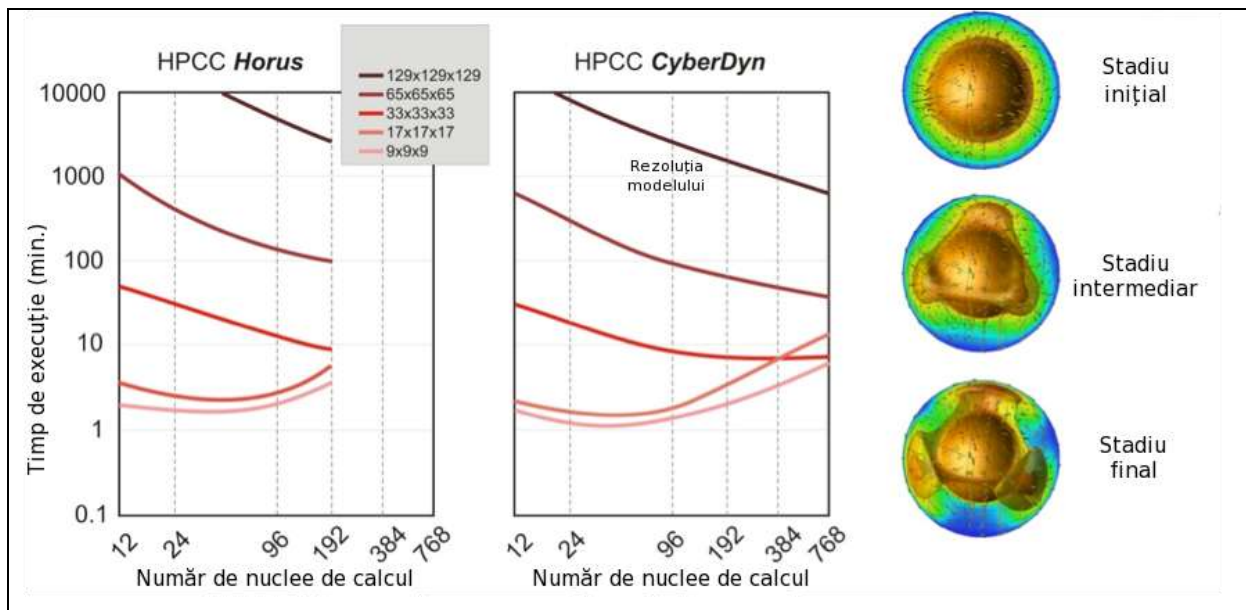


Figura 2. Comparație între rezultatele testelor de referință pe HPCC Horus și HPCC CyberDyn pentru o simulare de convecție pur termică FEM într-un model sferic, care arată influența dimensiunii rețelei de discretizare asupra timpului de execuție în funcție de numărul de procesoare. În dreapta sunt prezentate, prin izosuprafețe de temperatură, trei etape din evoluția simulării de convecție termică (vizualizări efectuate cu OpenDX - software open source). Culorile calde corespund temperaturilor ridicate, și culorile reci reprezintă temperaturi mai scăzute în interiorul mantalei Pământului. Sfera portocalie de la etapa inițială reprezintă nucleul de fier al Pământului.

O simulare termo-chimică efectuată într-un domeniu de analiză sferic reprezintă cel de-al treilea test comparativ între Horus HPCC și CyberDyn HPCC. Aceste simulări implică particule - marcatori pentru a urmări modificările termo-chimice din interiorul modelului. Particulele marcatori sunt generate pseudo-aleator, având în vedere să se pastreze un număr egal de marcatori pe element finit. În simulările noastre am folosit un număr de 20 de marcatori pe element, prin urmare, am variat numărul total de marcatori de la 1 milion la 500 de milioane, în funcție de rezoluția modelului. Rezultatele de referință arată că pentru simulări complexe numerice HPCC CyberDyn se comportă mai bine ca Horus, din cauza latenței scăzute a infrastructurii de rețea InfiniBand QDR și a dispozitivului de stocare de înaltă performanță Panasas. Am constatat, de asemenea că pentru modele de foarte mare rezoluție numărul maxim de nuclee de calcul, care oferă timp de execuție minim este în jurul valorii de 384 (a se vedea figura 3). La Horus, deși este mai lent

decât CyberDyn pentru aceste simulări FEM, putem observa o continuă scădere a timpului de execuție pentru aproape toate rezoluțiile modelelor și numărul nucleeleor de calcul. Acest rezultat demonstrează că Horus HPCC are încă un potențial real de a se extinde prin adăugarea de nuclee de calcul în viitor, rețeaua fiind capabilă să deservească un număr de nuclee ce depășește probabil 500.

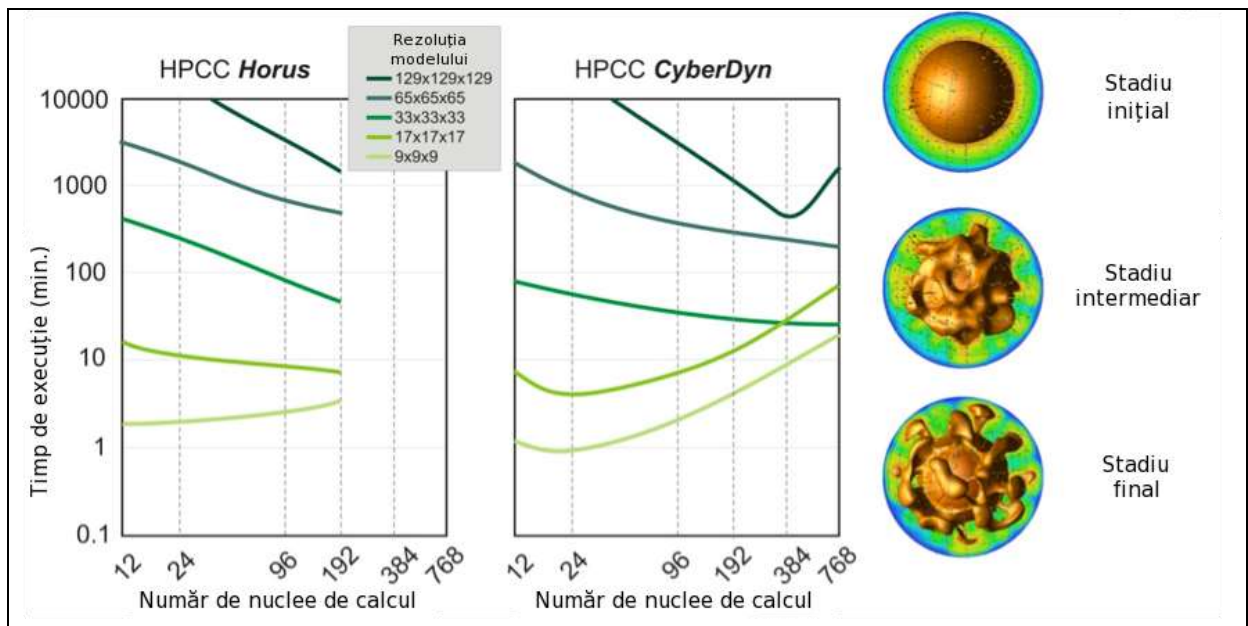


Figura 3. Comparație între rezultatele testelor de referință pe HPCC Horus și HPCC CyberDyn pentru o simulare de convecție termo-chimică FEM într-un model sferic, care arată influența dimensiunii rețelei de discretizare asupra timpului de execuție în funcție de numărul de procesoare. În dreapta sunt prezentate, prin izosuprafețe de temperatură, trei etape din evoluția simulării de convecție termo-chimică (vizualizări efectuate cu OpenDX - software open source). Culorile calde corespund temperaturilor ridicate, și culorile reci reprezintă temperaturi mai scăzute în interiorul mantalei Pământului. Sfera portocalie de la etapa inițială reprezintă nucleul de fier al Pământului.

2.4. DISCUȚIE ȘI CONCLUZII

Infrastructura de rețea de mare viteză Infiniband oferă posibilitatea de a exploata întregul potențial al supercalculatoarelor de dimensiuni mari, și reprezintă o componentă-cheie care influențează pozitiv atât scalabilitatea cât și performanța pe sistemele HPC mari. Testând un cod de simulare a dinamicii fluidelor, CFD (computational fluid dynamics) de înaltă rezoluție specific pentru științele pământului, sistemul HPC Horus bazat pe infrastructura de rețea Gigabit Ethernet s-a comportat remarcabil de bine în comparație cu omologul său CyberDyn HPCC care se bazează pe infrastructura de rețea InfiniBand QDR. Cu toate acestea, sistemele HPC bazate pe Gigabit Ethernet sunt încă destul de populare și reprezintă o alegere eficientă din punctul de vedere al costurilor, dar potrivită pentru calculatoare paralele de dimensiuni mici, sau eventual mijlocii proiectate pentru a rula coduri CFD specifice pentru științele pământului. Pe de altă parte, pentru calculatoare paralele medii și mari proiectate pentru a rula coduri similare, infrastructura cu latență mică și lățime de bandă mare ca cea Infiniband este foarte recomandată. Din moment ce în prezent simulările numerice folosesc rezoluții din ce în ce mai mari pentru predicții ale fenomenelor geodinamice de aceeași scară cu observațiile (de la kilometri la mii de kilometri), infrastructura HPC utilizată în științele pământului ar trebui să beneficieze în viitor de investiții pentru realizarea de noi sisteme bazate pe rețele de mare viteză.

3. UN COD PARALEL DE URMĂRIRE A PARTICULELOR PENTRU SIMULĂRILE DE DINAMICA FLUIDELOR ÎN ȘTIINȚELE PĂMÂNTULUI: IMPLEMENTARE SI TESTARE PE HPCC CYBERDYN.

Problema convecției într-un fluid cu vâscozitate mare reprezintă o direcție de cercetare des întâlnită în domeniul științelor Pământului. Pentru studiul mișcării convective a fluidului, se folosește o sursă de particule pasive (sau marcatori), care urmăresc viteza locală de deplasare și nu afectează modelul. Această metodă de lucru este frecvent utilizată.

Se prezintă un cod paralel care utilizează particule pasive fără greutate a căror poziție este calculată pe baza deplasării lor într-un interval scurt de timp, conform cu o viteză calculată anterior pentru anumite puncte, în spațiu și timp. Acest cod paralel cu marcatori este integrat în CitcomS un software care este utilizat pe

scară largă în comunitatea oamenilor se știință care se ocupă de studiul geodinamicii (www.geodynamics.org). Am testat codul cu marcatori pe o mașină de calcul de ultimă generație, CyberDyn, un cluster de performanță ridicată de calcul (High Performance Computing Cluster - HPCC), cu 1344 nuclee de calcul, instalat la Institutul de Geodinamică al Academiei Române. Testele de referință sunt efectuate folosind o serie de modele 3D geodinamice în care am introdus grupuri diferite de marcatori în locuri diferite. Cu toate că s-au folosit mai multe milioane de particule marcatori, testele de referință arată că performanțele codului folosit sunt bune, numărul optim de nuclee de calcul situându-se între 32 și 64. Din cauza volumului mare de date care se transferă între nucleele de calcul, simulările de dinamica fluidelor, CFD (computational fluid dynamics) de înaltă rezoluție folosite pentru predicții geodinamice care necesită zeci de milioane, sau chiar miliarde de marcatori pentru a urmări cu precizie mișcările din manta, vor beneficia foarte mult de sistemele bazate pe HPCC cu viteză mare de transfer între nucleele de calcul.

3.1. INTRODUCERE

În ultimii ani, puterea de convingere a simulărilor numerice de mecanica fluidelor ca un instrument pentru explorarea științei, în general, și a științelor pământului, în special, a atins noi nivele. Această direcție de cercetare reprezintă un viitor promițător în oferirea de alternative utile experimentelor de laborator, care sunt greu sau chiar imposibil de realizat în parametri realiști în cazul științelor pământului. Noile dezvoltări în domeniul calculului numeric de înaltă performanță (HPC), atât în termeni de componente hardware și de rețea cu lățime de bandă suficientă și latență redusă, precum și în termen de algoritmi avansați de calcul capabili să simuleze probleme complexe din domeniul științelor pământului, fac astăzi posibilă simularea cu exactitate a deplasării rocilor în interiorul globului terestru. În această lucrare, se prezintă un cod paralel pentru urmărirea curgerii materiei vâscoase în diferite scenarii de simulare numerică specifice pentru științele pământului. De asemenea, scopul acestei cercetări este de a oferi date de referință privind timpii de execuție obținuți de acest cod paralel cu particule - marcatori care este utilizat pe scară largă în cadrul modelelor de advecție a rocilor, specifice pentru științele pământului. Advecția materiei în interiorul Pământului, mișcarea materiei ce se comportă ca un lichid foarte vâscos, este un proces fizic fundamental care stă la baza multor coduri

de simulare. Au fost luate în considerare o serie de simulări numerice 3D (rezolvate cu metoda elementelor finite) care tratează deplasări ale unui lichid incompresibil cu un comportament ce se modifică în timp, în interiorul mantalei Pământului. Codul paralel cu marcatori prezentat în această lucrare este integrat în pachetul de programe CitcomS, un software open source, utilizat pe scară largă în comunitatea oamenilor de știință care studiază fenomenele geodinamicii, pentru simularea fenomenelor termo - chimice de scară largă (la nivel planetar) ce se petrec în interiorul Pământului (www.geodynamics.org).

Pentru efectuarea calculelor paralele a fost folosit HPCC CyberDyn, un supercomputer de 1344 de nuclee de calcul aflat la Institutul de Geodinamică al Academiei Române.

3.2. FOLOSIREA PARTICULELOR MARCATORI ÎN CONTEXTUL ȘTIINȚELOR PĂMÂNTULUI

Folosirea particulelor - marcatori fără masă oferă o mai bună înțelegere a fenomenelor complexe care au loc în interiorul Pământului. Această tehnică face posibilă analiza unor factori care se crede ca joacă un rol cheie în generarea vulcanismului sau a seismicității din adâncul Pământului. Aceste particule lipsite de masă, care joacă rolul de marcatori, se deplasează simultan în tot domeniul de calcul pe măsură ce soluția evoluează, și sunt des utilizate pentru a preleva eșantioane reprezentative de date, și pentru a vizualiza mișcarea materiei în timp și spațiu. De exemplu, de un interes deosebit în simulările științelor pământului este timpul de rezidență, o mărime care are legătură cu particulele implicate în studierea mișcării mantalei, și practic măsoară timpul petrecut într-o anumită regiune de analiză, ajutând astfel la caracterizarea diverselor regiuni de studiu ca regiuni de stagnare sau regiuni în care convecția se petrece mai rapid. Uneori, de aceste particule se atașează diverse proprietăți, cum ar fi compoziția chimică, ori ele pot reprezenta diferite materiale, care suportă o mișcare de advecție în interiorul Pământului, ca de exemplu crusta continentală sau cea oceanică. Particulele ne ajută, de asemenea, să înțelegem traiectoria complexă a mișcării mantalei în interiorul Pământului, în anumite locații, ca de exemplu sub un arc vulcanic activ, o regiune în care se presupune în mod normal că se generează magma la suprafață. Acest context de

fapt motivează dezvoltarea unei astfel de modalități de rezolvare a problemelor geodinamice care folosește aceste particule - marcatori care oferă informații suplimentare prin posibilitatea de a urmări traiectoria mișcării fluidului, un instrument de cercetare care nu este frecvent întâlnit în software-ul open source construit să ruleze în mod paralel. Am dezvoltat un cod care folosește particule - marcatori, ca parte integrantă a unui cod matur și robust de simulare cu ajutorul FEM (metoda elementelor finite) pentru simulările de mecanica fluidelor - CFD (computational fluid dynamics) în interiorul Pământului, CitcomS. În studiul științelor Pământului avem de-a face cu modele de 4D (3D + timp), iar din simulările numerice rezultă o cantitate mare de date. În scopul urmăririi cu cât mai mare acuratețe a schimbărilor de direcție a mișcării mantalei, numărul acestor particule poate varia de la câteva zeci de mii la câteva sute de milioane sau chiar miliarde. Codurile care folosesc un număr atât de mare de marcatori ar trebui să fie și eficiente, acest lucru presupunând implementarea lor ca și coduri paralele.

3.3. ALGORITM DE URMĂRIRE A MARCATORILOR.

În această lucrare am testat codul dezvoltat de noi într-o serie de scenarii 3D de mecanica fluidelor, în care s-a considerat că fluidul este incompresibil. Vom prezenta alcătuirea unei soluții numerice care folosește particulele marcatori, dezvoltate având în vedere eficiența din punctul de vedere al folosirii puterii de calcul. Datele de intrare folosite sunt reprezentate de câmpul de viteze rezultat din simulările numerice ale mișcărilor mantalei efectuate cu CitcomS. Soluția noastră numerică modifică poziția particulelor marcatori care participă în mod pasiv la mișcarea ce se desfășoară conform câmpului de viteze cunoscut, iar pentru calcularea noilor poziții se folosesc funcții de interpolare. Codul numeric cu marcatori este scris în limbajul de programare C, și în scopul de a beneficia de numărul mare de nuclee de calcul disponibile pentru testare pe HPCC CyberDyn și Horus, folosim pentru paralelizare MPI (Message Passing Interface). MPI este un standard de transmitere a mesajelor între nodurile de calcul, care oferă posibilitatea de a scrie și rula aplicații pe calculatoare paralele. Standardul MPI a fost definitivat în 1994 (Ulterior au fost dezvoltate noi standarde: MPI1.2 și MPI2, a se vedea <http://www.mcs.anl.gov/mpi>). Acum, mulți producători respectă acest standard, și există mai multe implementări de MPI în domeniul public. În această lucrare, folosim

varianta MPICH de la Argonne National Laboratory (Using MPI, 2nd. Ed. by W. Gropp, E. Lusk, and A. Skjellum MIT Press, Cambridge, 1999).

Schema logică a codului nostru, împreună cu porțiuni de cod sunt prezentate și discutate în detaliu mai jos. Codul este împărțit în trei părți principale. Prima parte importantă a codului este utilizată pentru a seta proprietățile inițiale ale marcătorilor și să distribuie datele către procesoare (Figura 4). Aici vom folosi doar două structuri MPI, și anume MPI_Comm și MPI_Status. Structura MPI_Status este folosită de funcțiile de primire a mesajelor pentru returnarea de informații despre un mesaj, iar MPI_Comm este dispecerul implicit de mesaje. Practic avem două sisteme de coordonate, unul la nivel global, pentru domeniul de calcul, și tot mai multe sisteme de coordonate locale, în funcție de numărul de procesoare. În versiunea actuală, datele de intrare ale acestui sistem includ informații cu privire la poziția inițială a marcătorilor (x, y, z), identificarea grupului de particule, ambele specificate în sistemul de coordonate global. Folosim descompunerea de domeniu, în scopul de a atribui procesoare în fiecare direcție a domeniului de calcul. Cu alte cuvinte, fiecărei direcții geometrice îi corespunde un anumit număr de procesoare, sau nuclee de calcul.

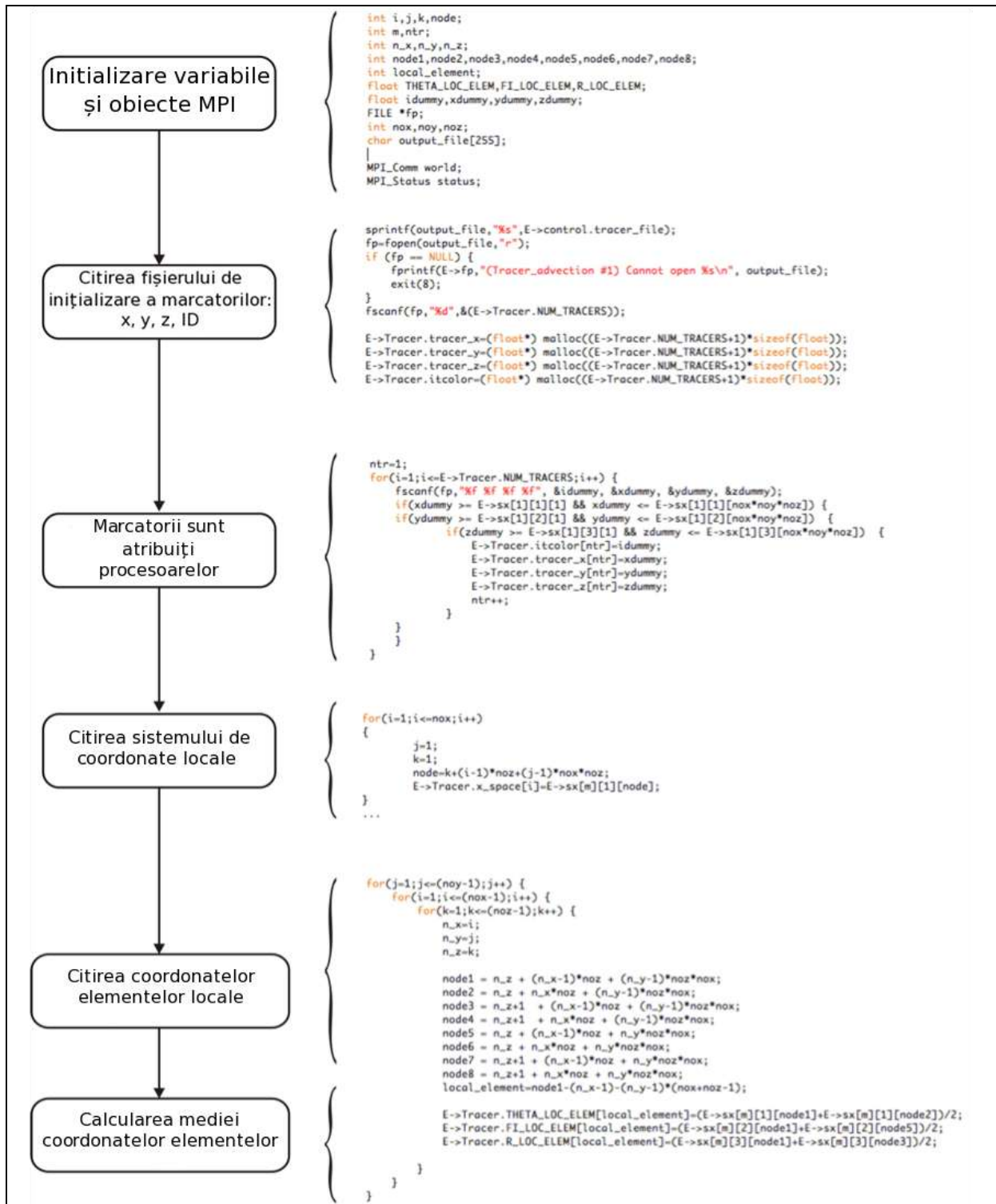


Figura 4. Un extras din codul paralel de urmarire a marcatilor, în care fiecare procesor controlează un anumit număr de marcatiori și rețeaua globală de discretizare este descompusă în rețele locale, conform hărții procesoarelor.

Al doilea pas din prima parte a codului este împărțirea marcatorilor în conformitate cu harta procesarelor, fiecărui nucleu de calcul revenindu-i sarcina sa controleze un anumit număr de marcatori. CitcomS utilizează elemente de forma paralelipipedică, și prin urmare, fiecare element finit are opt noduri. Folosim aceste informații pentru a calcula centrul fiecărui element în sistemul local de coordonate.

În a doua parte a codului nostru, poziția marcatorilor este modificată cu ajutorul nucleului de calcul care controlează marcatorul respectiv. Aici s-a folosit o funcție suplimentară MPI, MPI_Request, care este utilă în trimiterea și primirea cererilor. Aici, câmpul de viteză provenit din CitcomS este importat în codul de urmărire a marcatorilor, și este folosit pentru a modifica poziția marcatorilor folosind funcții de interpolare de ordin 2. Practic, pentru fiecare procesor se itereaza toți marcatorii controlați și se identifică elementul finit în care se află fiecare marcator. Pentru a modifica poziția unui marcator, avem nevoie să știm viteza marcatorului situat în interiorul unui element finit. În acest scop, se interpolează vitezele din cele opt colțuri ale elementului finit în poziția marcatorului cu ajutorul funcțiilor de pondere (Figura 5).

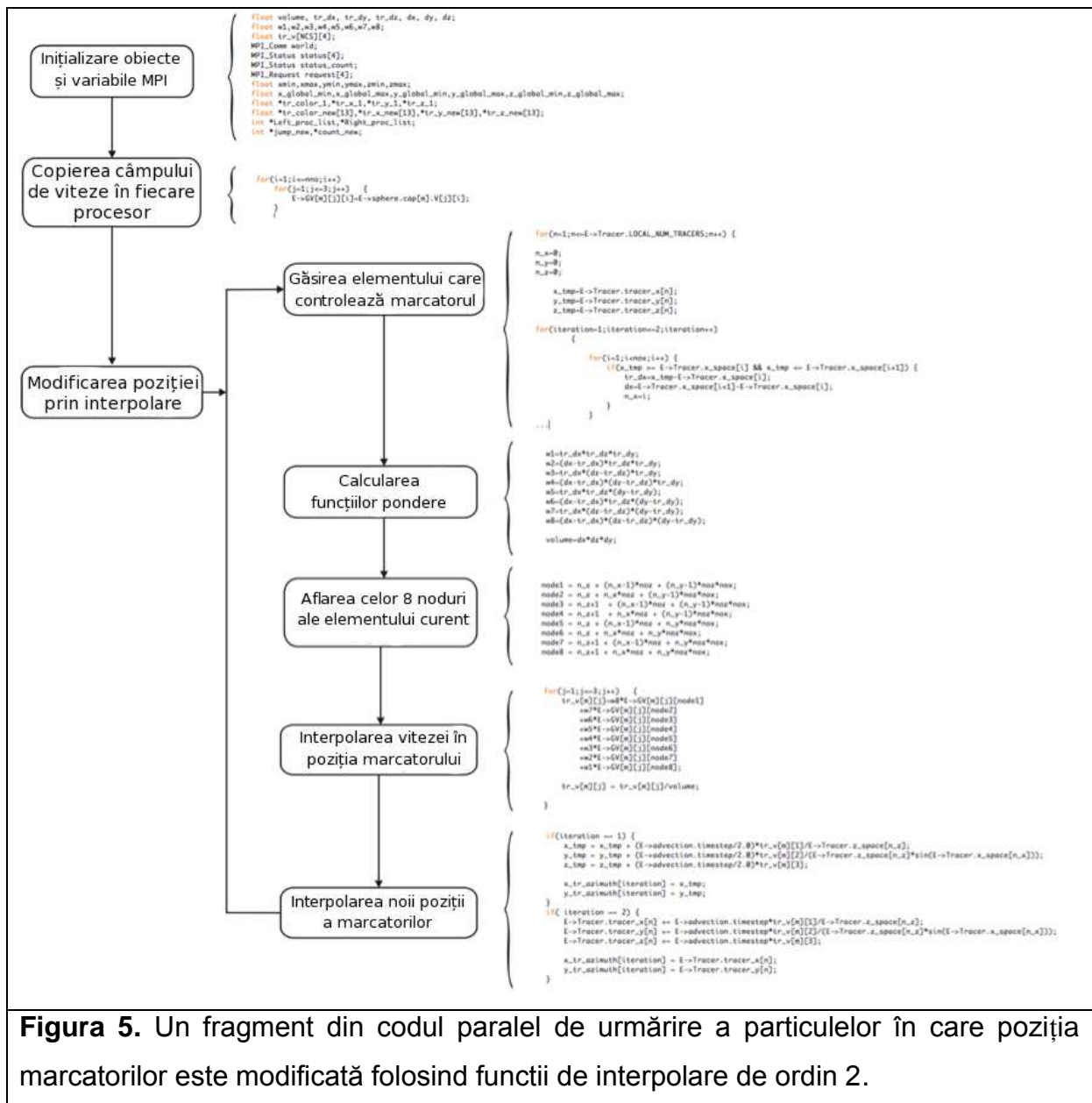


Figura 5. Un fragment din codul paralel de urmărire a particulelor în care poziția marculatorilor este modificată folosind funcții de interpolare de ordin 2.

Ultima parte a codului dezvoltat este dedicat în întregime calculului paralel. Se începe cu definirea hărții procesoarelor, practic se afla care sunt vecinii (procesoare) fiecărui procesor, informație utilizată pentru împărțirea domeniului de calcul. Regula generală este de a găsi care procesor este la stânga/dreapta, față/spate și sus/jos pentru procesorul curent (Figura 6).

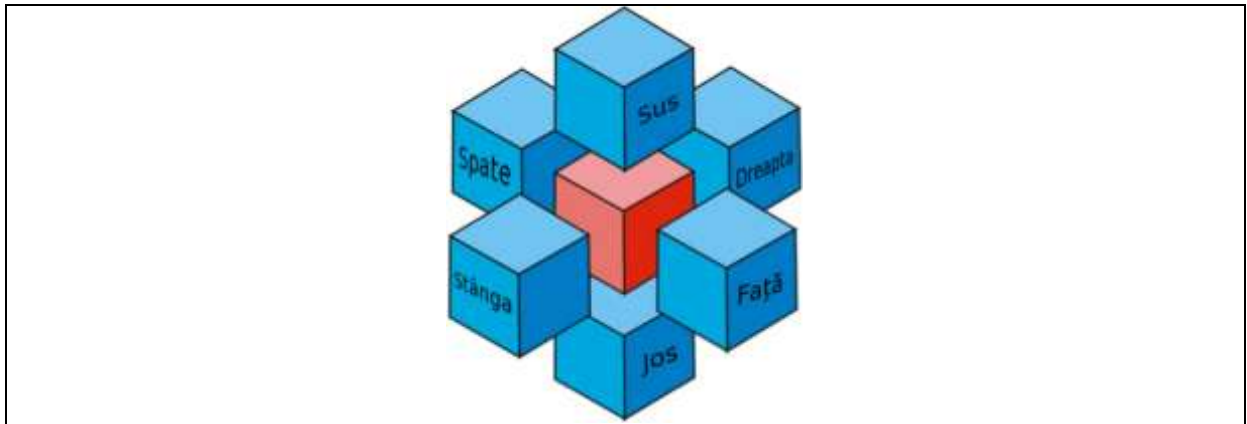


Figura 6. Harta procesoarelor folosită pentru paralelizarea codului de urmărire a marcătorilor prin MPI. Cuburile reprezintă nuclee de calcul. Cubul roșu reprezintă procesorul curent și arată poziția celor șase procesoare vecine.

Folosim patru obiecte MPI: `MPI_Irecv`, `MPI_Send`, `MPI_Waitall` și `MPI_Get_count`. `MPI_Irecv` începe o recepționare continuă, ceea ce înseamnă că sistemul poate începe scrierea datelor în buffer-ul de primire (Figura 7). Acesta primește de fapt o serie de marcători, care traversează granița dintre două procesoare (sau nuclee de calcul) adiacente. Numărul de marcători care a rămas în procesorul (sau nucleul de calcul) expeditor (`MPI_Send`) trebuie să fie redus, iar numărul de marcători în procesorul (sau nucleul de calcul) primitor trebuie să fie crescut. La fiecare pas de timp, fiecare procesor trebuie să-și reactualizeze numărul de marcători, în funcție de numărul de marcători primiți și trimiși între procesoare. Structura `MPI_Waitall` impune restricția ca toate procesoarele să finalizeze comunicațiile pentru un pas de timp. Cu `MPI_Get_count` se poate urmări cu exactitate numărul de marcători transmiși de la un procesor la altul.

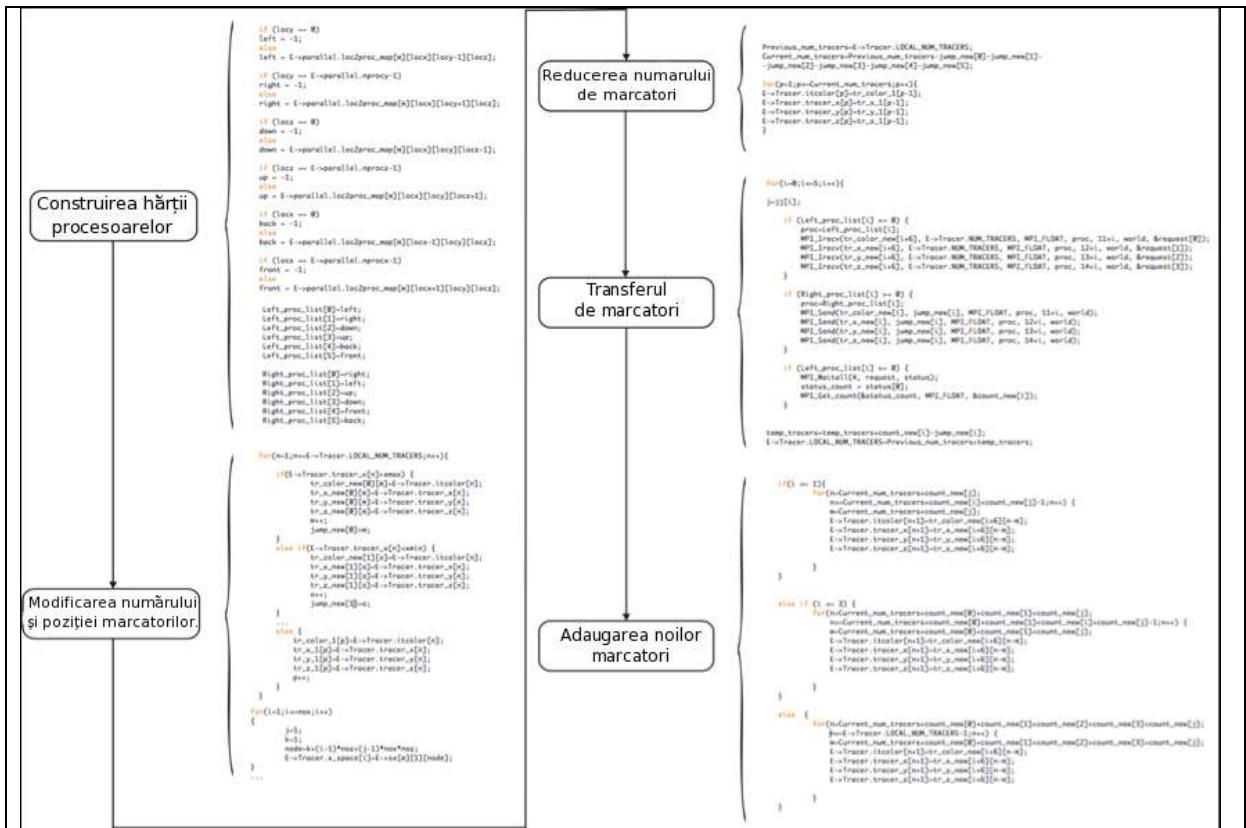


Figura 7. Schema logică și porțiunea respectivă a codului paralel, în care marcatorii sunt trimiși și primiți între procesoare, folosind funcțiile MPI.

3.4. TESTE DE REFERINȚĂ ALE CODULUI

Testele de referință sunt efectuate folosind o serie de scenarii geodinamice 3D, în care am introdus diverse grupuri de marcatori dispuse în straturi, la limita inferioară a modelelor. Am efectuat două simulări cu diferite anomalii de temperatură impuse la limita inferioară a modelului. Modelele sunt realizate într-o regiune 3D decupată din mantaua Pamantului, cu diferite tipuri de anomalii termice specificate ca și condiții inițiale. Testele au fost efectuate pe supercomputerul CyberDyn HPCC, cu 1344 nuclee de calcul, rețea de interconectare QDR Infiniband și $R_{max} \sim 12$ teraflopi (R_{max} reprezintă performanța maximă obținută). Software-ul de distribuire a sarcinilor pe nuclee de calcul folosit este SGE (Sun Grid Engine), configurat ca "round-robin", unde fiecare nod de calcul este umplut secvențial. Ne concentrăm asupra modului în care numărul de marcatori și de nuclee de calcul influențează timpul de execuție până la finalizarea soluției scenariului geodinamic. Am folosit un număr de 0 (fără marcatori), 10^4 , 10^5 , 10^6 și 10^7 marcatori, și de la 4 la 1024 nuclee

de calcul. Rezultatele arată că numărul optim de nuclee de calcul pentru simularea fără marcatori este situat între 32 și 64. Folosind un număr mai mare de nuclee (128 sau chiar 1024) duce la o creștere semnificativă a timpului de calcul, și acest timp crește mai mult atunci când se crește numărul de nuclee utilizat. Folosind un număr tot mai mare de marcatori ($10^4 - 10^7$), am obținut același număr optim de nuclee de calcul: 32-64 pentru un timp de execuție minim.

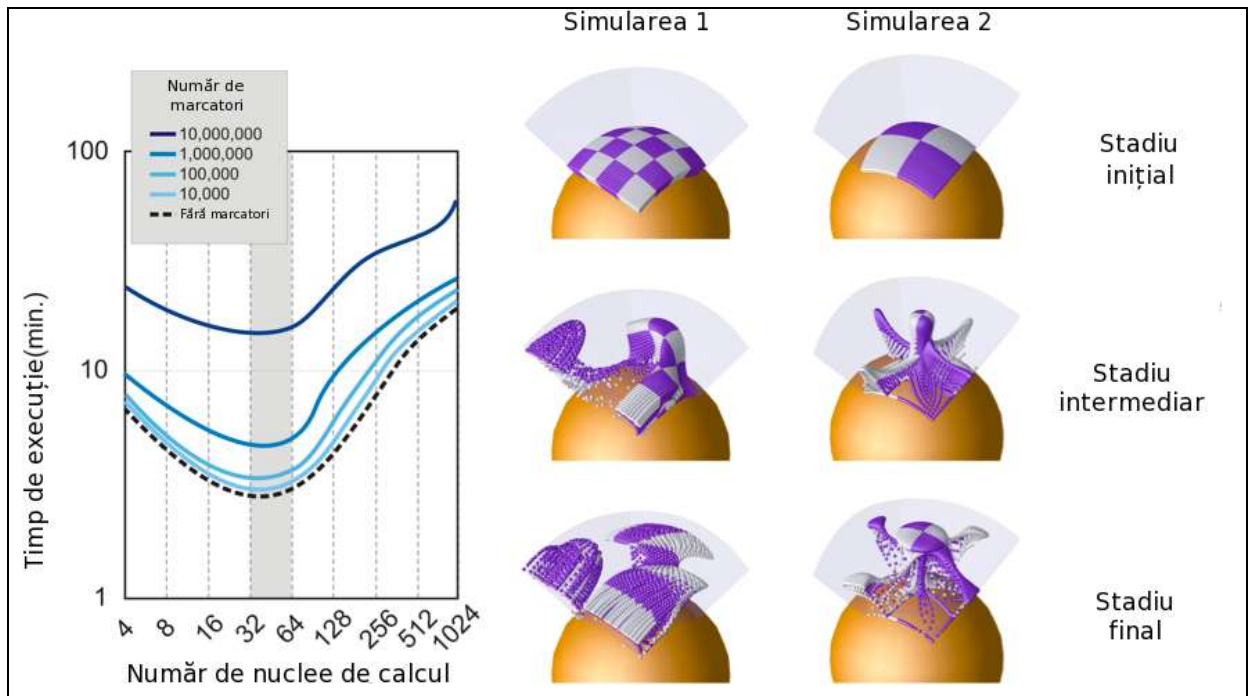


Figura 8. Rezultatul testelor de referință a codului paralel de urmărire a marcatorilor pe CyberDyn HPCC pentru simulările de convecție termice FEM (metoda elementelor finite) într-un model regional, care arată influența numărului de marcatori asupra timpului de rulaj în funcție de numărul de nuclee de calcul. În dreapta este prezentată evoluția marcatorilor pentru două simulări, în trei stadii diferite de evoluție ale simulării (vizualizare realizată cu software-ul open source OpenDX). Sfera portocalie reprezintă nucleul de fier al Pământului.

3.5. CONCLUZII

Cu cât calculele numerice folosite în studiul dinamicii fluidelor în științele pământului cresc în dimensiune, avem nevoie de algoritmi paraleli și de particule de urmărire (marcatori) pentru a analiza în detaliu deplasările diverselor porțiuni de manta. Mai mult decât atât, posibilitatea de a folosi particule marcatori aduce un plus de valoare tuturor codurilor moderne de simulare a mecanicii fluidelor. Cu toate acestea, construirea de algoritmi paraleli eficienți utilizabili pe sisteme mari HPCC ale căror cerințe de comunicare controlează timpii de execuție reprezintă în continuare o provocare pentru oamenii de știință. În această lucrare am prezentat tehnica de a paraleliza eficient un cod pentru urmărirea marcatorelor, utilizând MPI. Testele de performanță arată că acest cod dezvoltat de noi funcționează bine cu un număr mare de nuclee de calcul și, deși am folosit un număr mare de particule (până la 10^7 marcatori), care implică o cantitate imensă de date transferate între nodurile de calcul, nu a fost observată nici o întârziere în timpul de execuție (Figura 8). Acesta este un indiciu că acest cod paralel de urmărire a marcatorelor este bine integrat în codul paralel principal CitcomS, și că infrastructura de rețea de mare viteză QDR (40 Gb/sec.) utilizată de CyberDyn HPCC nu generează întârzieri în traficul MPI implicat în deplasarea marcatorelor.