Une brève introduction à MPI

Destinée à l 'usage des utilisateurs de CIMENT

Laurent Desbat

juin 2002

Plan

- ***MPI, qu'est-ce que c'est ?***
- MPI, comment l'utilise-t'on ?
- Notion de communications globales/collectives
- Un exemple réel d'imagerie médicale SPECT

MPI, qu'est-ce que c'est ?
MPI : Message Passing Interface

- Le modèle de programmation est celui de processus communicant, coopérant à l'exécution d'une tâche.
  - Chaque processus a ses propres données
  - Les processus peuvent communiquer entre eux
    - MPI_Send, MPI_Recv, MPI_Bcast, ...
  - Et se synchroniser
    - MPI_Barrier, ...

MPI, qu'est-ce que c'est ?

- MPI permet de gérer
  - l 'environnement d 'exécution
  - les communications point à point (Send, Recv)
  - les communications collectives (Bcast, Ata, Ato,...)
  - les groupes de processus et des communicateurs
  - la topologie d'inter-connexion des processus (grilles, tore, arbre,…=> notion de voisinage)

Plan

- MPI, qu'est-ce que c'est ?
- ***MPI, comment l'utilise-t'on ?***
- Notion de communications globales/collectives
- Un exemple réel d'imagerie médicale SPECT

MPI, comment l'utilise-t'on ?

- Initialisation de l'environnement (début)
  - C/C++ : MPI_Init(&argc,&argv);
  - FORTRAN : MPI_INIT(ierr)
- Et destruction (fin)
  - MPI_Finalize();
  - FORTRAN : MPI_FINALIZE(ierr)

MPI, comment l'utilise-t'on ?

- Un processus est identifié par son groupe et son rang (dans le groupe).
  - int MPI_Comm_rank(MPI_comm comm, int *rank)
  - int MPI_Comm_size(MPI_comm comm, int *rank)
  - exemple d'utilisation :
    - MPI_Comm_rank(MPI_COMM_WORLD, &myrank)
      - MPI_COMM_WORLD est le communicateur global (tous les processus du programme MPI courant en sont élément). Il est créé par MPI_Init.

MPI, comment l'utilise-t'on ?

- MPI permet de gérer les communicateurs
  - Sont définis par un groupe et un contexte
  - Ils définissent la portée des communications
    - Le contexte d 'un communicateur détermine certaines propriétés (mais l 'utilisateur ne peut agir directement dessus)
- Exemple : MPI_COMM_WORLD créé à l 'initalisation comprends tous les processus de votre exécution parallèle

MPI, comment l'utilise-t'on ?

- MPI permet de gérer les communications
- Communication point à point
  - l'enveloppe d'un message est constituée de
    - le rang du processus émetteur
    - le rang du processus récepteur
    - une étiquette de message
    - le nom du communicateur
  - Les données sont typées
    - MPI_CHAR, MPI_SHORT, MPI_INT, MPI_LONG, MPI_UNSIGNED_CHAR, MPI_UNSIGNED_SHORT, MPI_UNSIGNED, MPI_UNSIGNED_LONG, MPI_FLOAT, MPI_DOUBLE, MPI_LONG_DOUBLE

MPI, comment l'utilise-t'on ?

- Communication bloquantes
  - int MPI_Send(void* buf, int count,MPI_Datatype datatype, int dest, int tag, MPI_Comm comm)
  - int MPI_Recv(void* buf,int count,MPI_Datatype datatype, int source, int tag, MPI_Comm comm, MPI_Status *status)
    - MPI_ANY_SOURCE et/ou MPI_ANY_TAG peuvent être utilisés, MPI_PROC_NULL (proc. fictif=sans effet)
    - le status est une structure qui contient en sortie la source et le tag du message reçu
      - status.MPI_SOURCE ; status.MPI_TAG

MPI, comment l'utilise-t'on ?

- Mon premier programme en C
- Mon premier programme en FORTRAN

Plan

- MPI, qu'est-ce que c'est ?
- MPI, comment l'utilise-t'on ?
- ***Notion de communications globales/collectives***
- Un exemple réel d'imagerie médicale SPECT

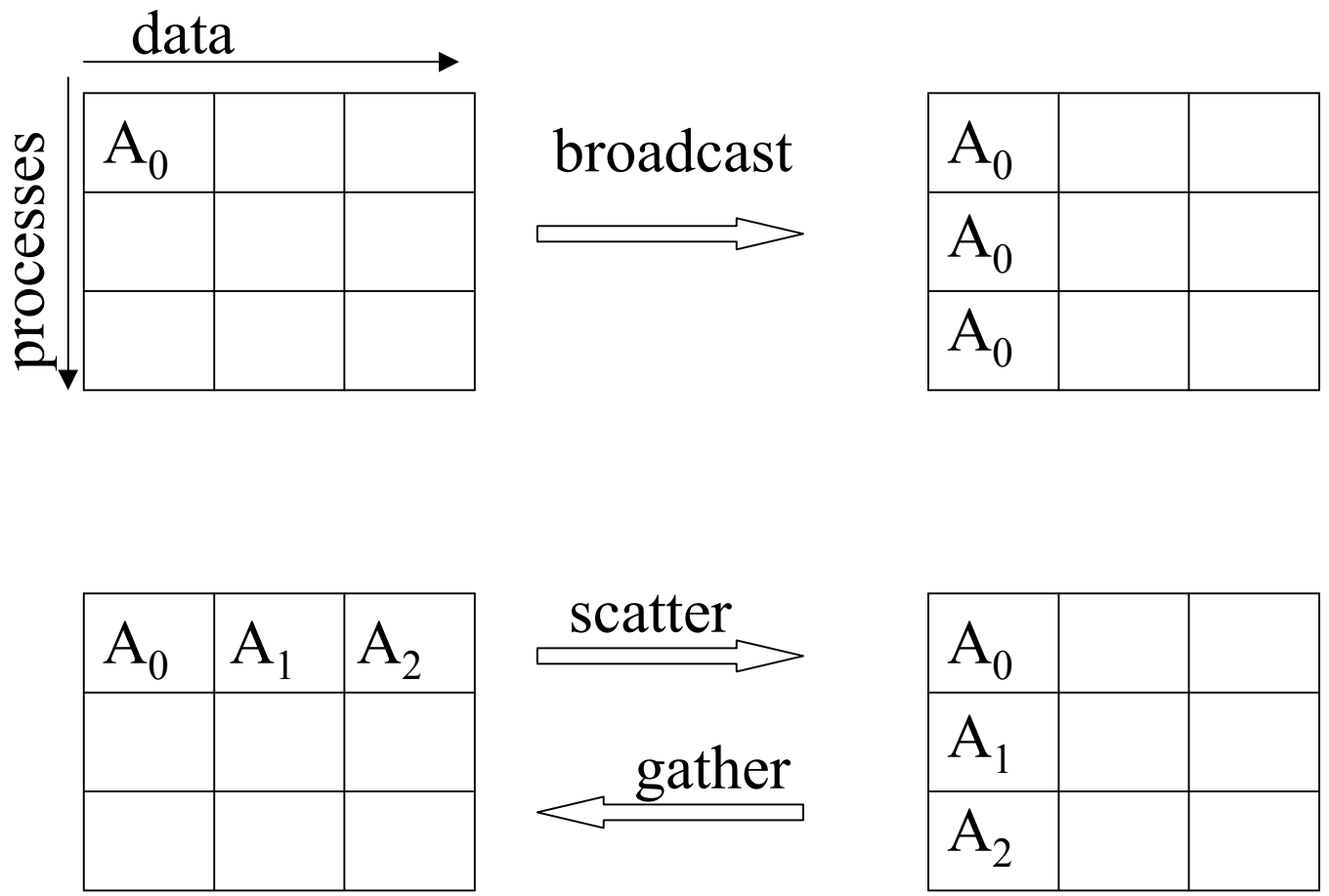
Communications collectives

- Elles impliquent un groupe de processeurs
- Contraintes
  - la quantité d ’information envoyée doit être identique à la quantité d ’info reçue
  - le même communicateurs doit être invoqué pqr tous les processus participant
  - pas de TAG

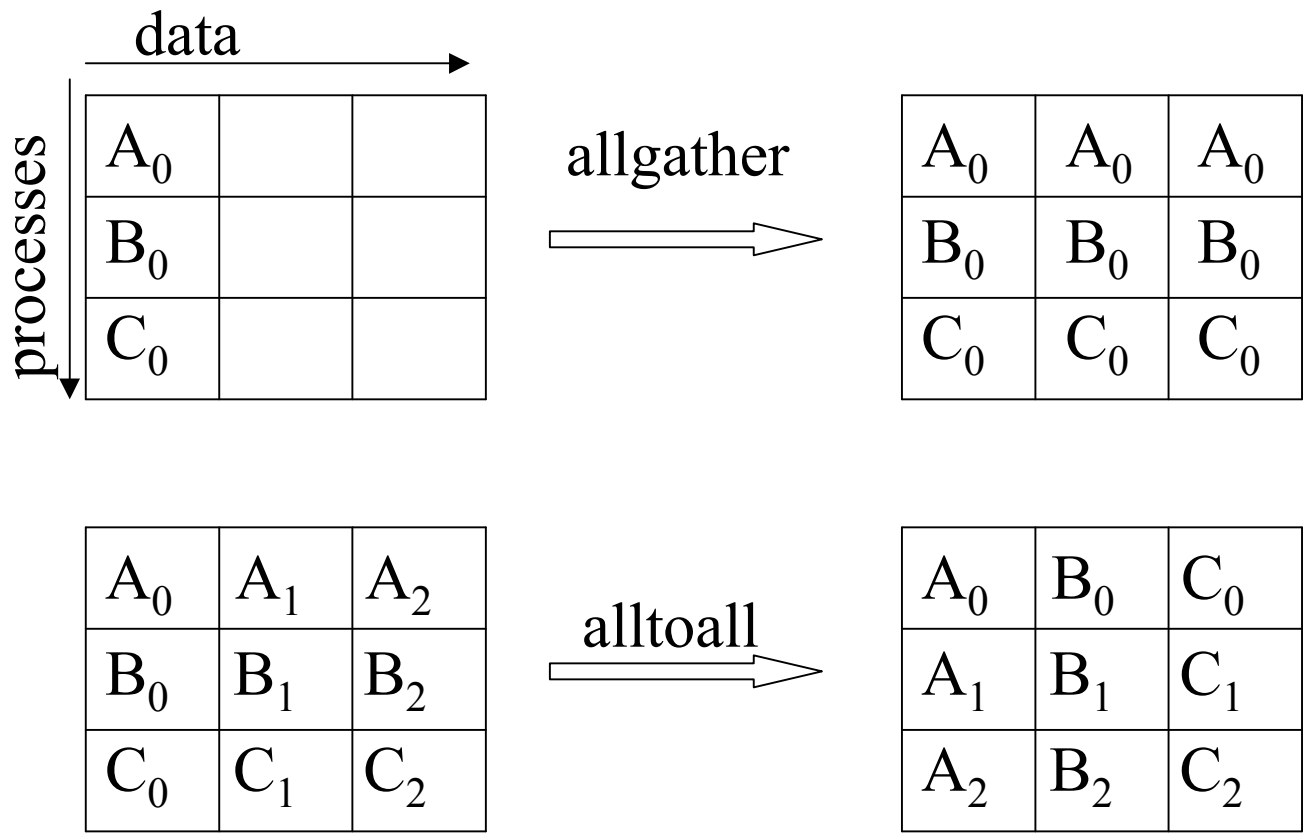
Communications collectives

- Synchronisation
  - int MPI_Barrier(MPI_Comm com)
- Transfert de données : MPI_Bcast, MPI_Scatter, MPI_Gather, MPI_Allgather,MPI_Alltoall
  - int MPI_Bcast(void* buffer, int count, MPI_Datatype datatype, int root, MPI_Comm comm)
- Transfert de données et calcul, puis eventuellement transfert du résultat : MPI_Reduce, MPI_Allreduce, MPI_Reduce_Scatter, MPI_Scan
  - int MPI_Reduce(sendbuf,recvbuf,count,datatype,op,root,comm)
  - op est un MPI_Op (MPI_MAX, MPI_SUM, MPI_PROD), voir aussi MPI_Op_create pour la création de nouveaux op

Communications collectives



Communications collectives



Plan

- MPI, qu'est-ce que c'est ?
- MPI, comment l'utilise-t'on ?
- Notion de communications globales/collectives
- ***Un exemple réel d'imagerie médicale SPECT***

Lien vers tomo3D.ppt

Lien vers les sources séquentielles

Lien vers TP

Lien vers solution

Conclusion

- Très peu de connaissances sont nécessaires pour démarrer sous MPI
- Primitives de base
  - MPI_Init et MPI_Finalize
  - MPI_Comm_rank, MPI_Comm_size
  - MPI_Send et MPI_Recv
  - MPI_Barrier

Quelques Liens

- Sur eclair (machine alpha)
  - /usr/opt/MPI190/mpi_guide.ps
- Bibliographie
  - Marc Snir *et al*. MPI, the complete reference. Second edition, MIT Press, 1998. Vol1, The MPI core, Vol2, MPI-2.
  - William Gropp, Ewing Lusk et Anthony Skjellum. Usining MPI : Portable parallel Programming with the Message Passing Interface. MIT Press, 1994.
  - Peter S. pacheco. Parallel Programming with MPI. Morgan Kaufman Ed., 1997.

Quelques Liens

- Distributions (domaine public)
  - http://www.lam-mpi.org/
  - http://www-unix.mcs.anl.gov/mpi/mpich
- Tutoriaux
  - http://webct.ncsa.uiuc.edu:8900/webct/public/home.pl
  - http://www-unix.mcs.anl.gov/mpi/tutorial
  - http://www.idris.fr/

Plan d ’un cours avancé

- Notion de communications non-bloquantes
- Notion de types et types dérivés
- Notion de communicateurs
- Notion de topologie