

Outline Recap One-sided Comm

Dynamic

I/O PDE

Distributed Computing - MPI

Dimitar Lukarski, Jarmo Rantakokko

Division of Scientific Computing
Department of Information Technology
Uppsala Programming for Multicore Architectures Research Center
(UPMARC)
Uppsala University

Programming of Parallel Computers, Jan, 2014

500

jliji

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



Outline

Recapitulation

One-sided Communication

Recap One-sided Comm Dynamic 1/O PDE

Outline

Dynamic Process Creation

Parallel I/O

PDE Example

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

pliji



Distributed computing

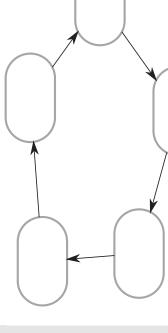




D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



Distributed computing



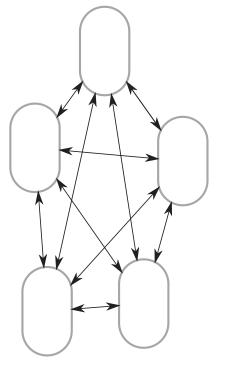
Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



Distributed computing

Outline
Recap
One-sided Comm
Dynamic
I/O
PDE



D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

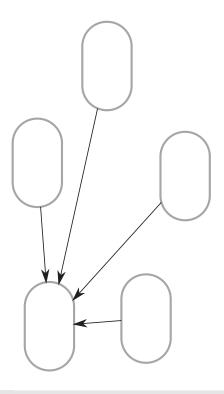
000



Distributed computing



Outline
Recap
One-sided Comm
Dynamic
I/O
PDE





Distributed computing

Outline
Recap
One-sided Comm
Dynamic

I/O PDE

You know how to communicate with MPI!

- ► Communicators
- Sending/Receiving data
- Synchronous/Blocking modes
- ► Other Point-to-Point Functions
- Global for synchronization
 - ► Global for communication

500

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

UPPSALA UNIVERSITET

One-sided Communication

Outline Recap One-sided Comm



Problems with Send/Recv

Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

What issues are expected when using Send and Recv functions?

500 D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



Problems with Send/Recv

Outline Recap One-sided Comm Dynamic I/O PDE

What issues are expected when using Send and Recv functions?

- ► Send/Recv match
- Handshaking overhead
- ► Extra buffers



One-sided Communication

Outline Recap One-sided Comm

Dynamic I/O PDE

► Direct access to the remote memory (RMA)

No hand-shaking procedure

Flexible and dynamic data distribution

Hardware optimization

► No extra buffers

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

500



One-sided Communication

Recap
One-sided Comm
Dynamic
1/O
PDE

Outline

► Create accessing window

Accessing pattern

Sync (fence)

► Reduction (MPI_Accumulate())

► Lighter syncs



One-sided Communication

Process 1

Process 2

Outline Recap One-sided Comm Dynamic

> I/O PDE

unl ock(...)

MPI _ W ndow_I ock(...)

MPI _ Get (...)

MPI _ Put (...)

MPI _ W n_ unl ock(...)

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

MPI _W n_f ence(..)

MPI _ W n_f ence(...)

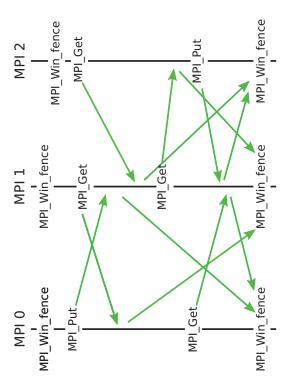


UPPSALA UNIVERSITET

Recap
One-sided Comm
Dynamic
I/O
PDE

Outline

One-sided Communication





Access Epoch

Outline Recap One-sided Comm

Dynamic I/O PDE

► Lighter/local than fence

► Epoch: time-frame when a local window may be used for RMA

Origin

- ► MPI_Win_start
- ► MPI_Win_complete

Target

- ► MPI_Win_post
- ► MPI_Win_wait

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

500



Dynamic Process Creation

Outline Recap One-sided Comm Dynamic 1/O PDE

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

php



Dynamic Process Creation

Outline Recap One-sided Comm

Dynamic

I/O

PDE

A MPI process can spawn new MPI processes at run time which starts running a new program.

- MPI_Comm_spawn(...);
- ► MPI_Comm_spawn_multiple(...) (MPMD)

500

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

Code Example

UPPSALA UNIVERSITET

```
MPI_Recv(data,count,type,0,tag,childcomm,status);
                                                                                                            MPI_Send(data,count,type,0,tag,childcomm);
                                                            MPI_Comm_spawn("childprog", arg, nproc,...,
&childcomm);
                                                                                                                                                                              MPI_Comm_free(&childcomm);
                                          MPI_Comm childcomm;
// Parent program
                                                                                                                                                                                                  MPI_Finalize();
                    MPI_Init(...);
                        Recap
One-sided Comm
Dynamic
             Outline
```

 $\| f \|_1$



Code Example

```
// Child program (started by the parent program)
                                                                                                If (rank==0) MPI_Recv(data,count,type,0,tag,
                                                                                                                                                          MPI_Get_parent(&parentcomm);
MPI_Get_rank(MPI_COMM_WORLD,&rank);
                                                                                                                      parentcomm, status);
                                      MPI_Comm parentcomm;
                                                                                                                                                                                                 MPI_Finalize();
                 MPI_Init(...);
                       Outline
Recap
One-sided Comm
```

Dynamic

I/O PDE

500

jliji

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



Outline

Recap One-sided Comm Dynamic I/O PDE

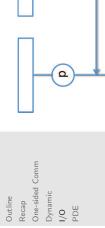
Parallel I/O

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

 $\| f \|_1$



I/O - Approach 1



Q ۵ d

500

jliji

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

UPPSALA UNIVERSITET

Outline Recap One-sided Comm Dynamic I/O PDE

I/O - Approach 1

Gather all data to process 0 and process 0 writes to file. Serial operation, lacks scalability and performance.

000

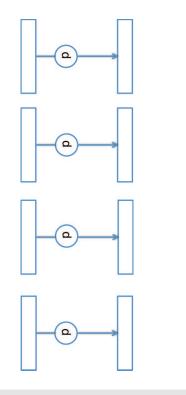
php



I/O - Approach 2



Outline Recap One-sided Comm Dynamic I/0 PDE



500

php

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



I/O - Approach 2

Recap One-sided Comm Dynamic I/O PDE Outline

but lots of small files to manage. Not a practical solution, what if the number of MPI tasks changes? Each process writes to a separate file. Good performance

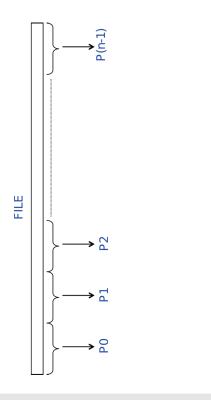
php



I/O - Approach 3

UPPSALA UNIVERSITET

Outline
Recap
One-sided Comm
Dynamic
I/O
PDE



500

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala



I/O - Approach 3

Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

MPI allows for several processes to read and write to the same file in parallel

► MPI_File_open(...)

► MPI_File_read(...)

► MPI_File_write(...)

► etc (more than 60 functions)

Portable, scalable and efficient solution!

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

 $\| f \|_1$



Outline Recap One-sided Comm Dynamic I/O PDE

PDE Example

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

500

PDE

UPPSALA UNIVERSITET

Outline Recap One-sided Comm Dynamic 1/0 PDE

1D Heat problem

$$\partial_t u - \nu \partial_x^2 u = 0$$
 in $(0, T] \times (0, 0)$

$$u(0,\cdot)=u_0(\cdot)$$
 in $(0,T]\times(0,1)$ $u(0,\cdot)=u_0(\cdot)$ in $(0,1)$ (Initial condition) $u(\cdot,1)=0$ in $(0,T]$ (Boundary condition)

$$u(0,\cdot) = u_0(\cdot)$$
 in $(0,1)$
 $u(\cdot,0) = u(\cdot,1) = 0$ in $(0,T)$

where $\nu>0$.

Solve it in numerically and in parallel!



Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

Math

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

000



Head Problem

UPPSALA UNIVERSITET

Outline
Recap
One-sided Comm
Dynamic
1/O
PDE

► Check compatibility condition

Math



Outline Recap One-sided Comm Dynamic I/O PDE

Head Problem

Math

- Check compatibility conditionCreate space grid (def space step)

Head Problem

500

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

UPPSALA UNIVERSITET

Outline Recap One-sided Comm Dynamic I/O PDE

Math

- ► Check compatibility condition
- ► Create space grid (def space step)
- ► Create time grid (def time step)



Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

Math

- ► Check compatibility condition
- ► Create space grid (def space step)
- Create time grid (def time step)
 - ▶ Discretize the problem (FD)

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

500



Head Problem

Recap One-sided Comm Dynamic 1/O PDE Outline

Math

- ► Check compatibility condition
- ► Create space grid (def space step)
- Create time grid (def time step)
- ► Discretize the problem (FD)
- ► Explicit method (no LS)



UPPSALA UNIVERSITET

Outline
Recap
One-sided Comm
Dynamic
I/O
PDE

Math

- ► Check compatibility condition
- ► Create space grid (def space step)
- Create time grid (def time step)
- Discretize the problem (FD)
- Explicit method (no LS)
- Check Stability condition

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

2000



Head Problem

UPPSALA UNIVERSITET

Outline

Recap One-sided Comm Dynamic 1/O PDE

Math

- ► Check compatibility condition
- ► Create space grid (def space step)
- Create time grid (def time step)
- Discretize the problem (FD)
- Explicit method (no LS)
- Check Stability condition $2\nu\tau \le h^2$



Outline Recap One-sided Comm Dynamic I/O PDE

Math

- ► Check compatibility condition
- Create space grid (def space step)
 - ► Create time grid (def time step)
 - ▶ Discretize the problem (FD)
- ► Explicit method (no LS)
- lacktriangle Check Stability condition $2\nu\tau \le \hbar^2$

Black-board writting...

D. Lukarski, J.Rantakokko, Jan, 2014, Uppsala

500