

Message Passing Interface

F.CS306 Parallel programming – Lecture 16

Topic Overview

- ◆ Collective мэдээлэл солилцоо
- ◆ *Barrier* синхрончлол, Broadcast, Scatter, Gather үйлдлүүд
- ◆ Глобал *Reduce* үйлдэл

Collective мэдээлэл солилцоо

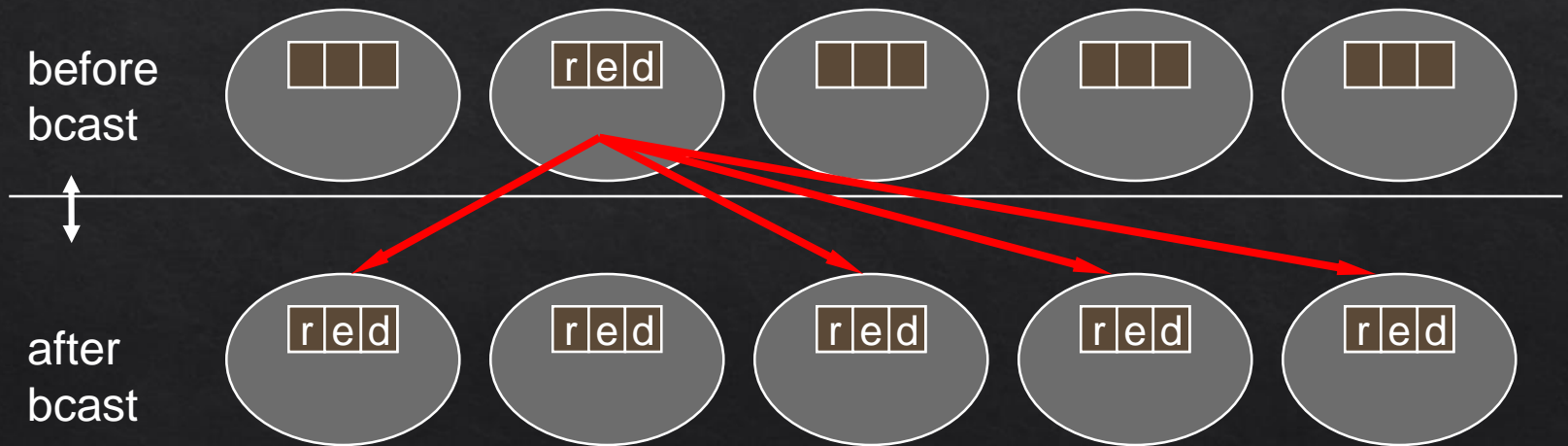
- ◇ Процесс бүлгийн үүсгэх мэдээлэл солилцоонууд.
- ◇ Communicator доторх бүх процессоор оролцох ёстой.
- ◇ Barrier синхрончлолт; Broadcast, scatter, gather; Глобал нийлбэр, глобал максимум гэх мэт.
- ◇ Синхрончлол үүсч болно, үгүй ч байж болно.
- ◇ Бүх мэдээлэл солилцоо blocking байна.
- ◇ Tag байхгүй.
- ◇ Хүлээн авах буфер нь илгээх буфертай яг ижил хэмжээтэй байх ёстой.

Barrier синхрончлол

- ◇ `int MPI_Barrier(MPI_Comm comm)`
- ◇ MPI_Barrier нь ихэвчлэн шаардлагагүй байдаг:
 - ◇ Бүх синхрончлол нь өгөгдлийн дамжуулалтаар автомат хийгдэг: Шаардлагатай өгөгдөлтэй болохоос өмнө процесс үргэлжлэх боломжгүй.
 - ◇ Хэрэв debug хийхэд хэрэглэж байгаа бол: Release хийж бүтээгдэхүүн болгохдоо устгадаг.

Broadcast

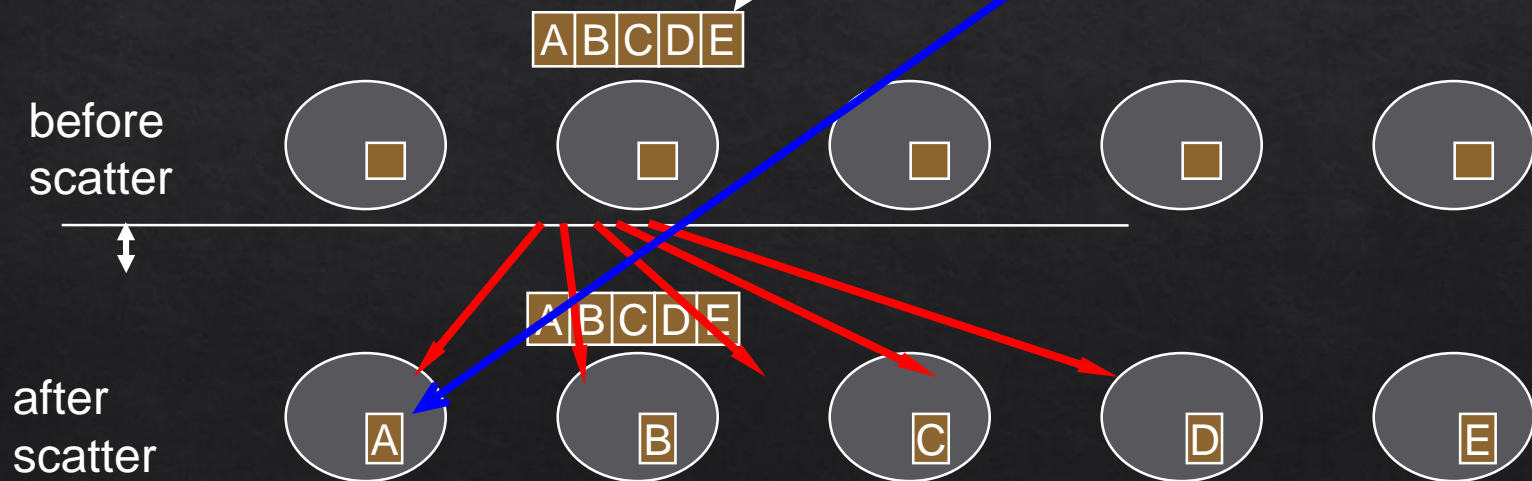
◇ `int MPI_Bcast(void *buf, int count,
MPI_Datatype datatype,
int root, MPI_Comm comm)`



ЖНЬ: root=1

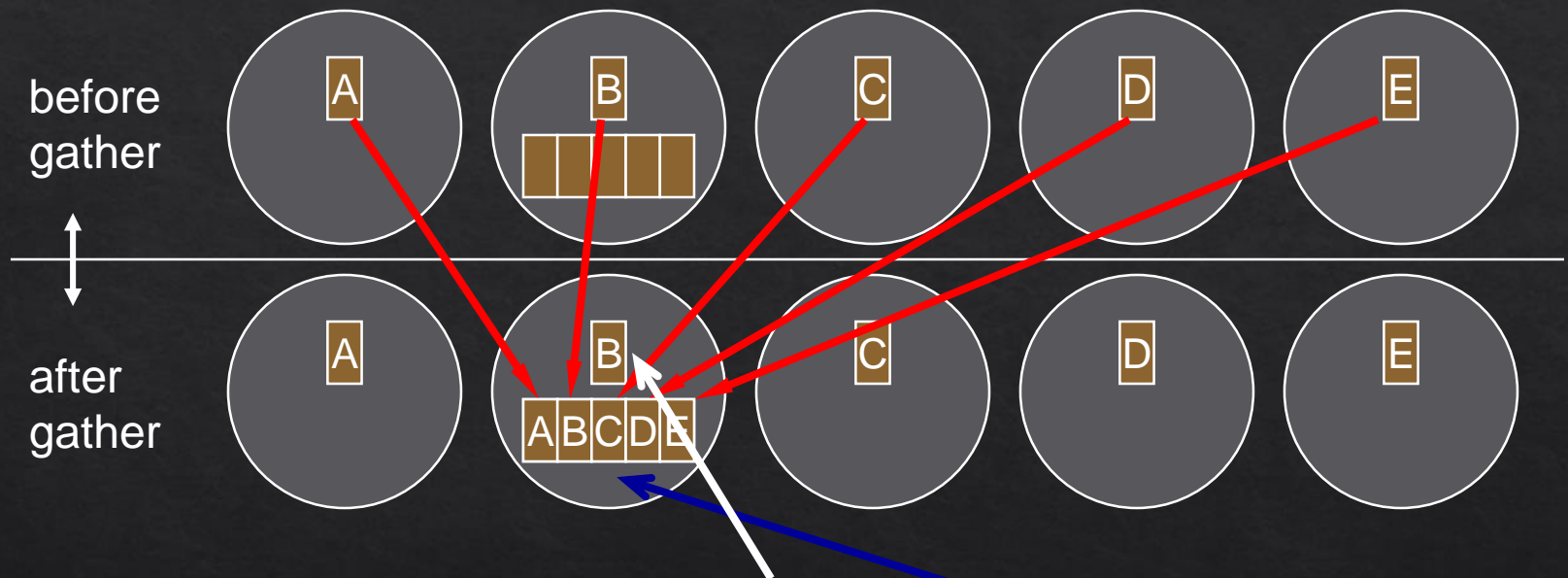
Scatter

◇ `int MPI_Scatter(void *sendbuf, int sendcount, MPI_Datatype sendtype, void *recvbuf, int recvcount, MPI_Datatype recvtype, int root, MPI_Comm comm)`



Gather

ЖНБ: root=1



```
◇ int MPI_Gather(void *sendbuf, int sendcount,  
MPI_Datatype sendtype, void *recvbuf, int  
recvcount, MPI_Datatype recvtype,  
int root, MPI_Comm comm)
```

Global Reduction Operations

- ◇ Глобал *Reduce* үйлдэл нь группын бүх гишүүн дээр биелдэг.
- ◇ $d_0 \circ d_1 \circ d_2 \circ d_3 \circ \dots \circ d_{s-2} \circ d_{s-1}$
 - ◇ $d_i = i$ *rank*-тай процесс дахь өгөгдөл, single хувьсагч, эсвэл vector
 - ◇ \circ = холбогч үйлдэл
 - ◇ Жишээ: глобал нийлбэр, үржвэл; глобал хамгийн их эсвэл бага олох; глобал хэрэглэгчийн тодорхойлсон үйлдэл
- ◇ floating point rounding may depend on usage of associative law:
 - ◇ $[(d_0 \circ d_1) \circ (d_2 \circ d_3)] \circ [\dots \circ (d_{s-2} \circ d_{s-1})]$
 - ◇ $(((((d_0 \circ d_1) \circ d_2) \circ d_3) \circ \dots) \circ d_{s-2}) \circ d_{s-1})$

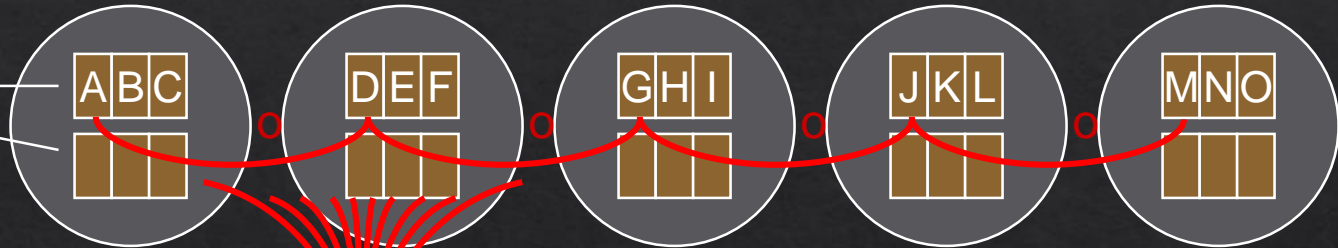
Predefined Reduction Operation Handles

Predefined operation handle	Function
MPI_MAX	Maximum
MPI_MIN	Minimum
MPI_SUM	Sum
MPI_PROD	Product
MPI LAND	Logical AND
MPI_BAND	Bitwise AND
MPI_LOR	Logical OR
MPI_BOR	Bitwise OR
MPI_LXOR	Logical exclusive OR
MPI_BXOR	Bitwise exclusive OR
MPI_MAXLOC	Maximum and location of the maximum
MPI_MINLOC	Minimum and location of the minimum

MPI_REDUCE

before MPI_REDUCE

- inbuf
- result



after



root=1

AoDoGoJoM

Variants of Reduction Operations

- ◆ MPI_ALLREDUCE

- ◆ no root,
- ◆ returns the result in all processes

- ◆ MPI_REDUCE_SCATTER

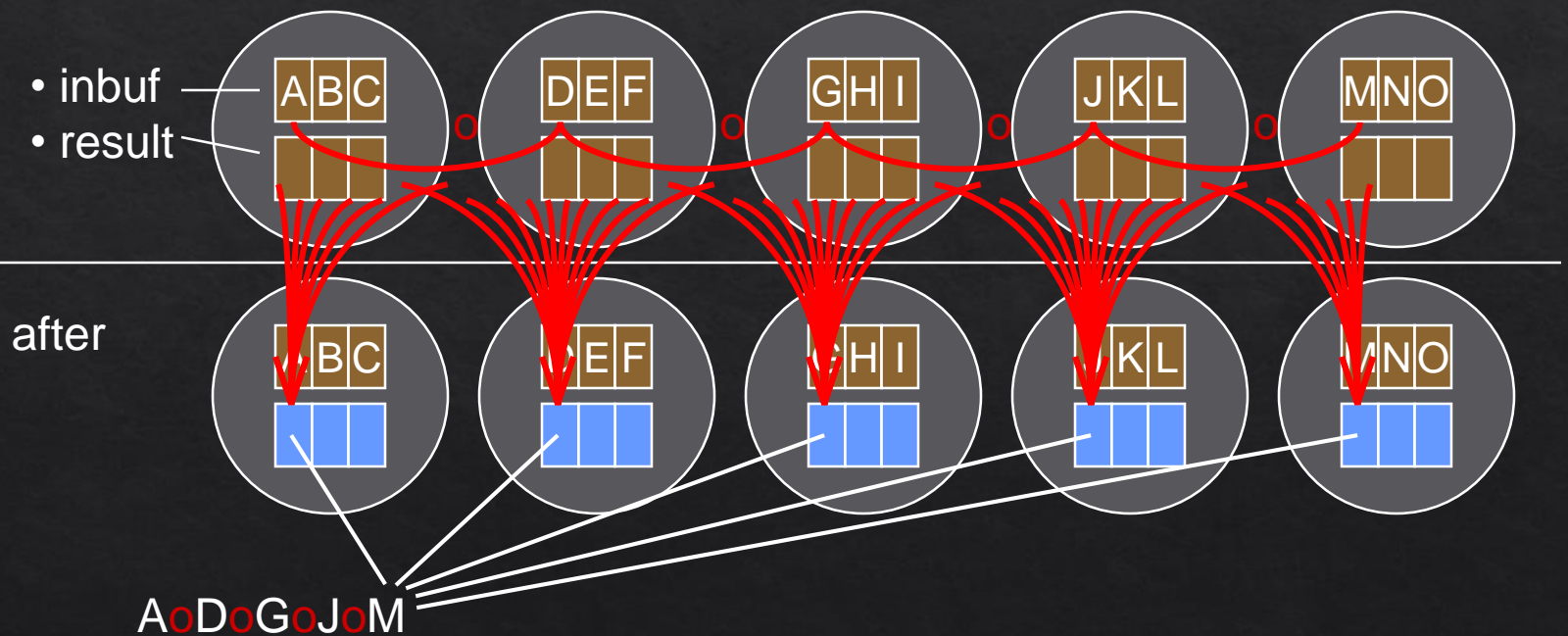
- ◆ result vector of the reduction operation
is scattered to the processes into the real result buffers

- ◆ MPI_SCAN

- ◆ prefix reduction
- ◆ result at process with rank i :=
reduction of inbuf-values from rank 0 to rank i

MPI_ALLREDUCE

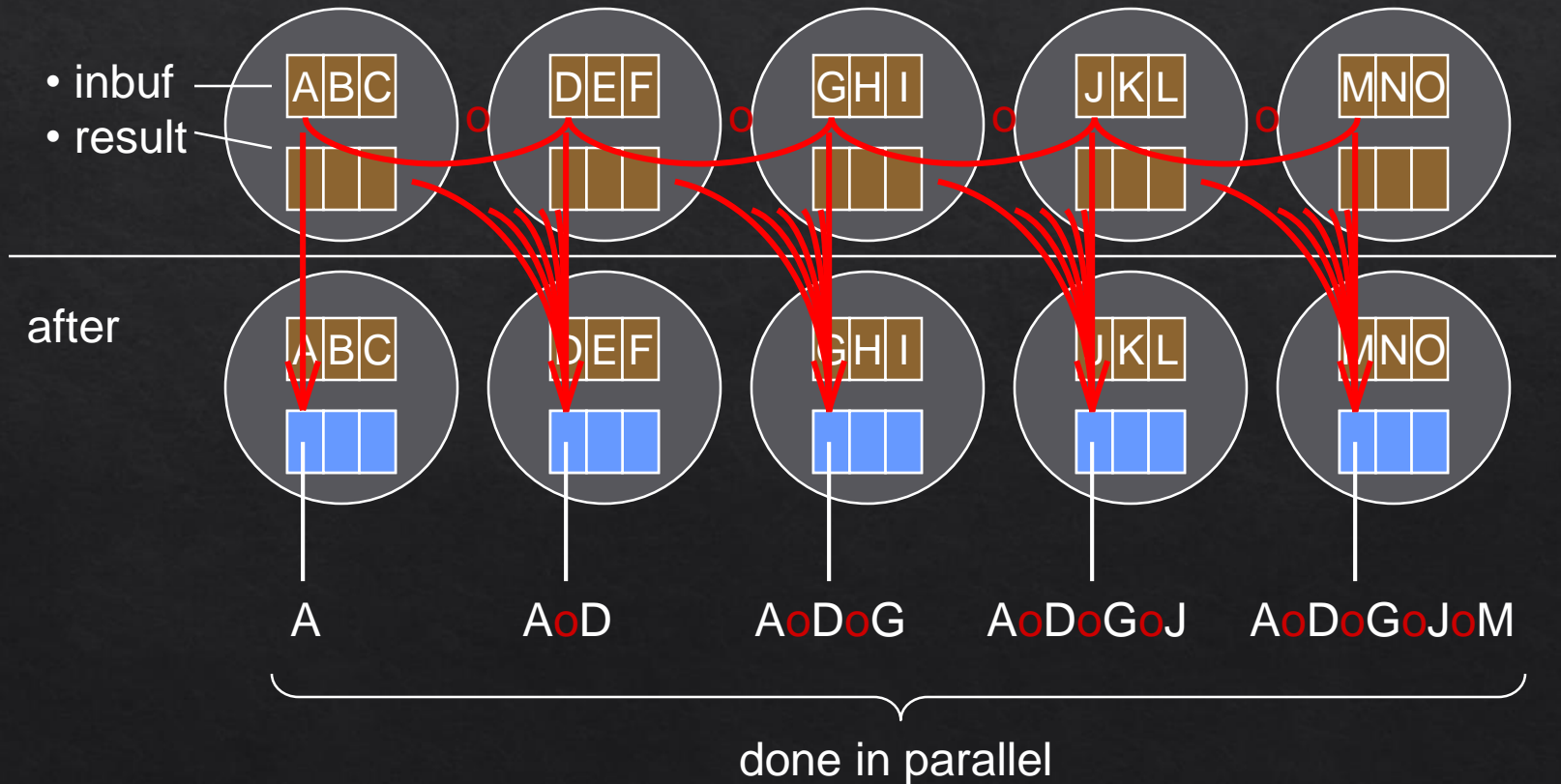
before MPI_ALLREDUCE



MPI_SCAN

before MPI_SCAN

- inbuf
- result



Thanks.