3D Parallel FEM (III) Parallel Visualization and ppOpen-HPC

Kengo Nakajima
Information Technology Center
The University of Tokyo

ppOpen-HPC: Overview

- Application framework with automatic tuning (AT)
  - ✓ "pp" : post-peta-scale
- Five-year project (FY.2011-2015) (since April 2011)
  - ✓ P.I.: Kengo Nakajima (ITC, The University of Tokyo)
  - ✓ Part of "Development of System Software Technologies for Post-Peta Scale High Performance Computing" funded by JST/CREST (Supervisor: Prof. Mitsuhisa Sato, RIKEN AICS)
- Target: Oakforest-PACS (Original Schedule: FY.2015)
  - ✓ could be extended to various types of platforms
- Team with 7 institutes, >50 people (5 PDs) from various fields: Co-Design
- Open Source Software
  - ✓ http://ppopenhpc.cc.u-tokyo.ac.jp/
  - ✓ https://github.com/Post-Peta-Crest/ppOpenHPC
  - ✓ English Documents, MIT License

独立行政法人 科学技術振興機構 Japan Science and Technology Agency

CREST 戦略的創造研究推進事業 Core Research for Evolutionary Science and Technology

Oakforest-PACS: OFP

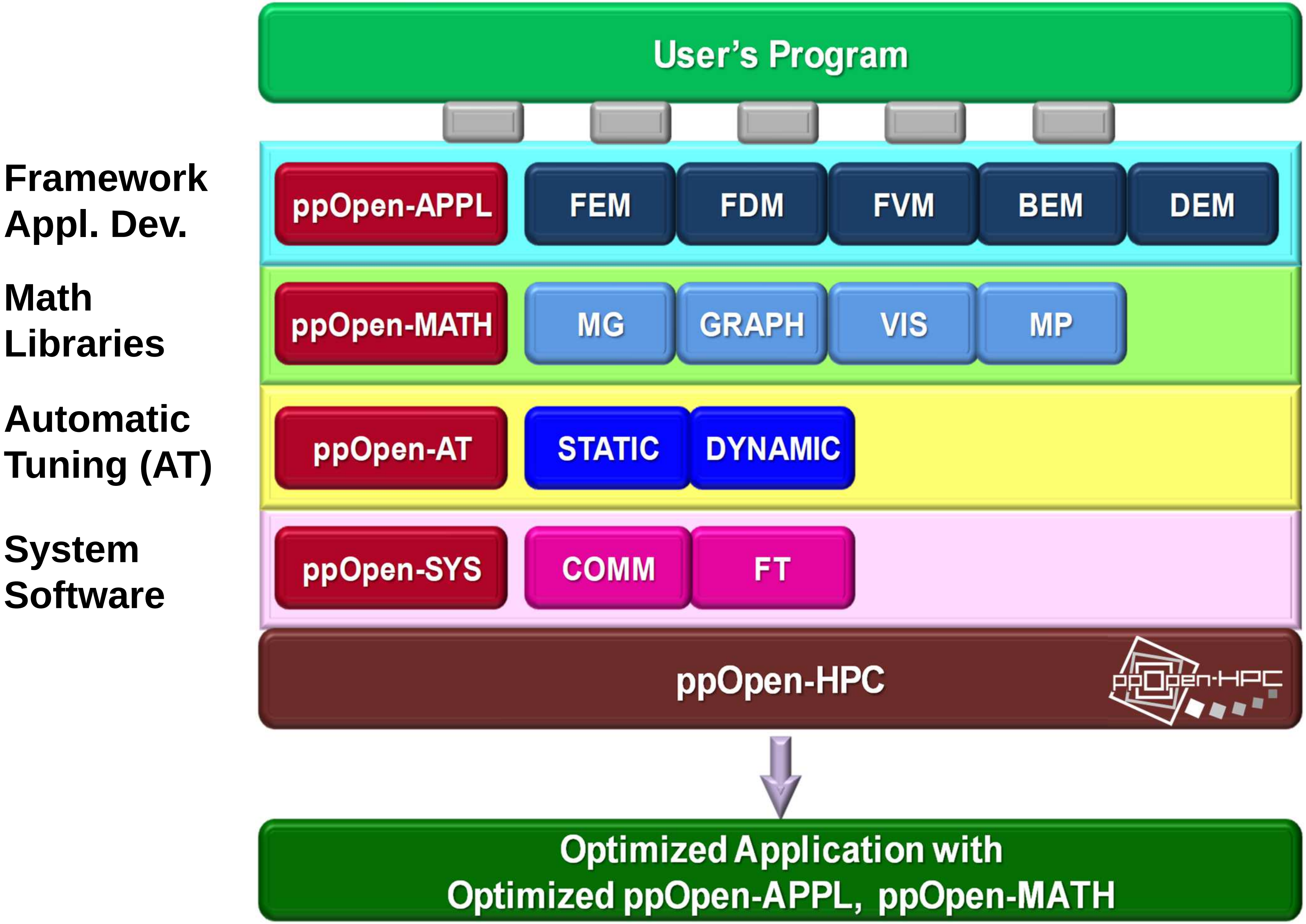
2nd Fastest System in Japan

- Full Operation started on December 1, 2016
  - After Our Post-Peta CREST ended
- 8,208 Intel Xeon/Phi (KNL), 25 PF Peak Performance
  - Fujitsu
- **TOP 500 #14 (#2 in Japan), HPCG #9 (#3) (Nov.2018)**
- **JCAHPC: Joint Center for Advanced High Performance Computing)**
  - University of Tsukuba
  - University of Tokyo
    - New system is installed at Kashiwa-no-Ha (Leaf of Oak) Campus/U.Tokyo, which is between Tokyo and Tsukuba
  - http://jcahpc.jp

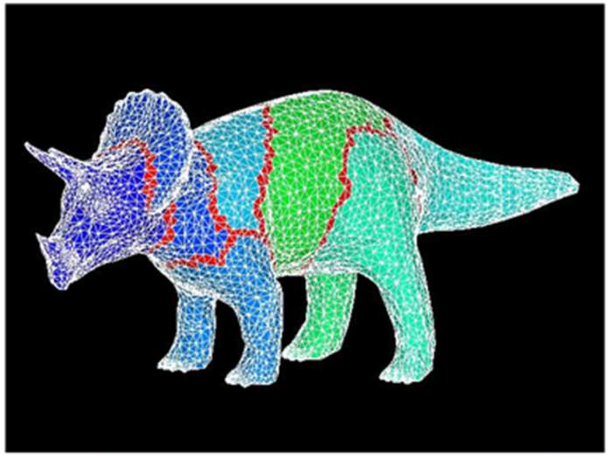
JCAHPC

東京大学 THE UNIVERSITY OF TOKYO

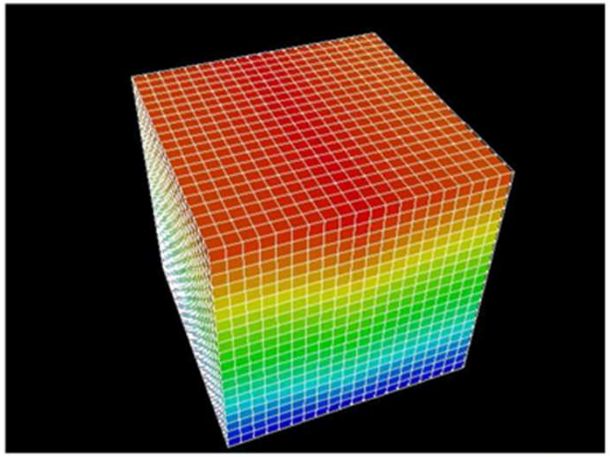
筑波大学 University of Tsukuba


User's Program
Framework Appl. Dev.
ppOpen-APPL
FEM
FDM
FVM
BEM
DEM
Math Libraries
ppOpen-MATH
MG
GRAPH
VIS
MP
Automatic Tuning (AT)
ppOpen-AT
STATIC
DYNAMIC
System Software
ppOpen-SYS
COMM
FT
ppOpen-HPC
Optimized Application with
Optimized ppOpen-APPL, ppOpen-MATH

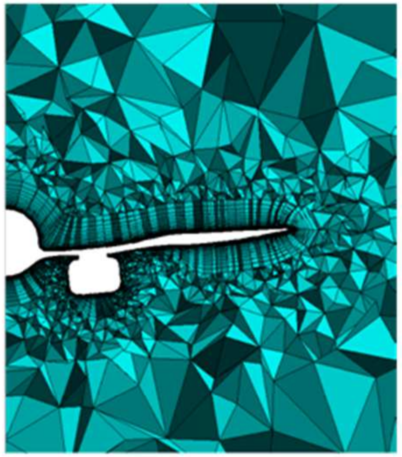
ppOpen-HPC covers …



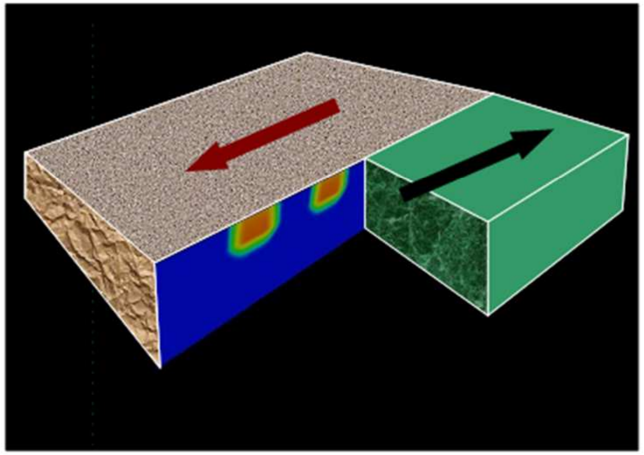
FEM
Finite Element Method



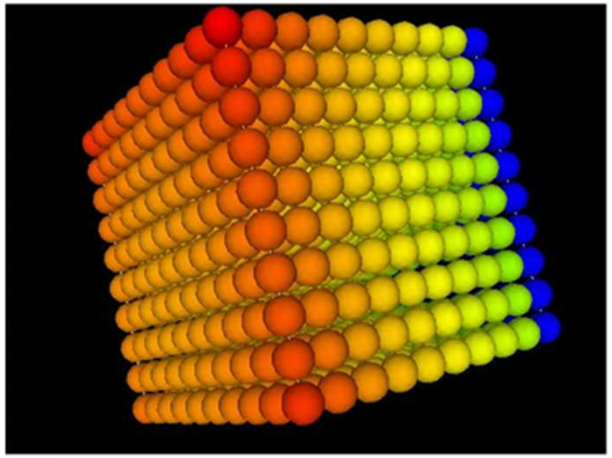
FDM
Finite Difference Method



FVM
Finite Volume Method



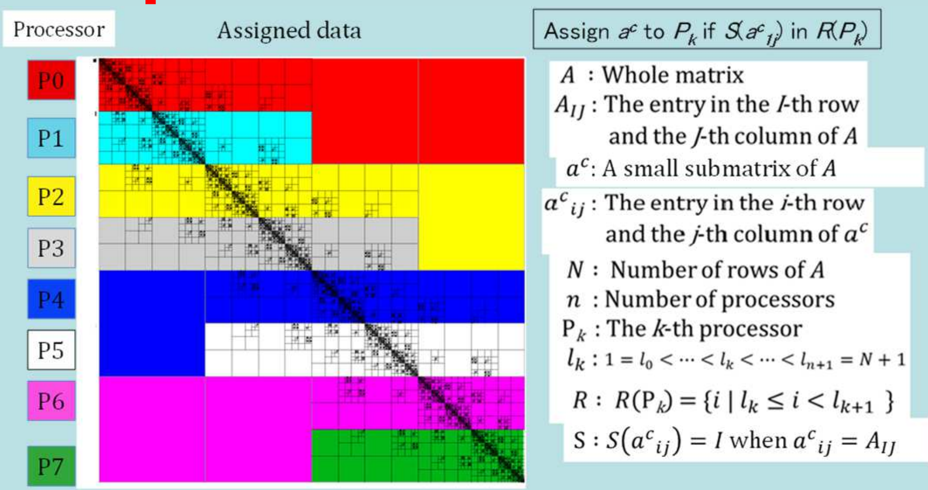
BEM
Boundary Element Method

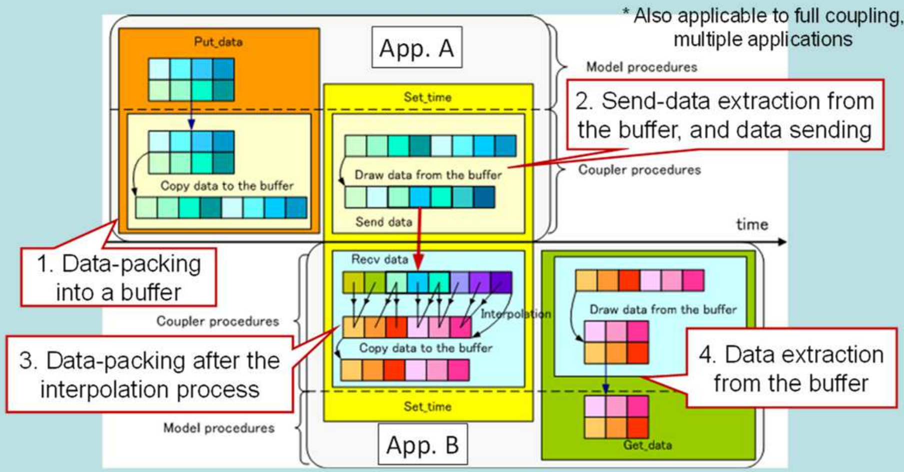


DEM
Discrete Element Method

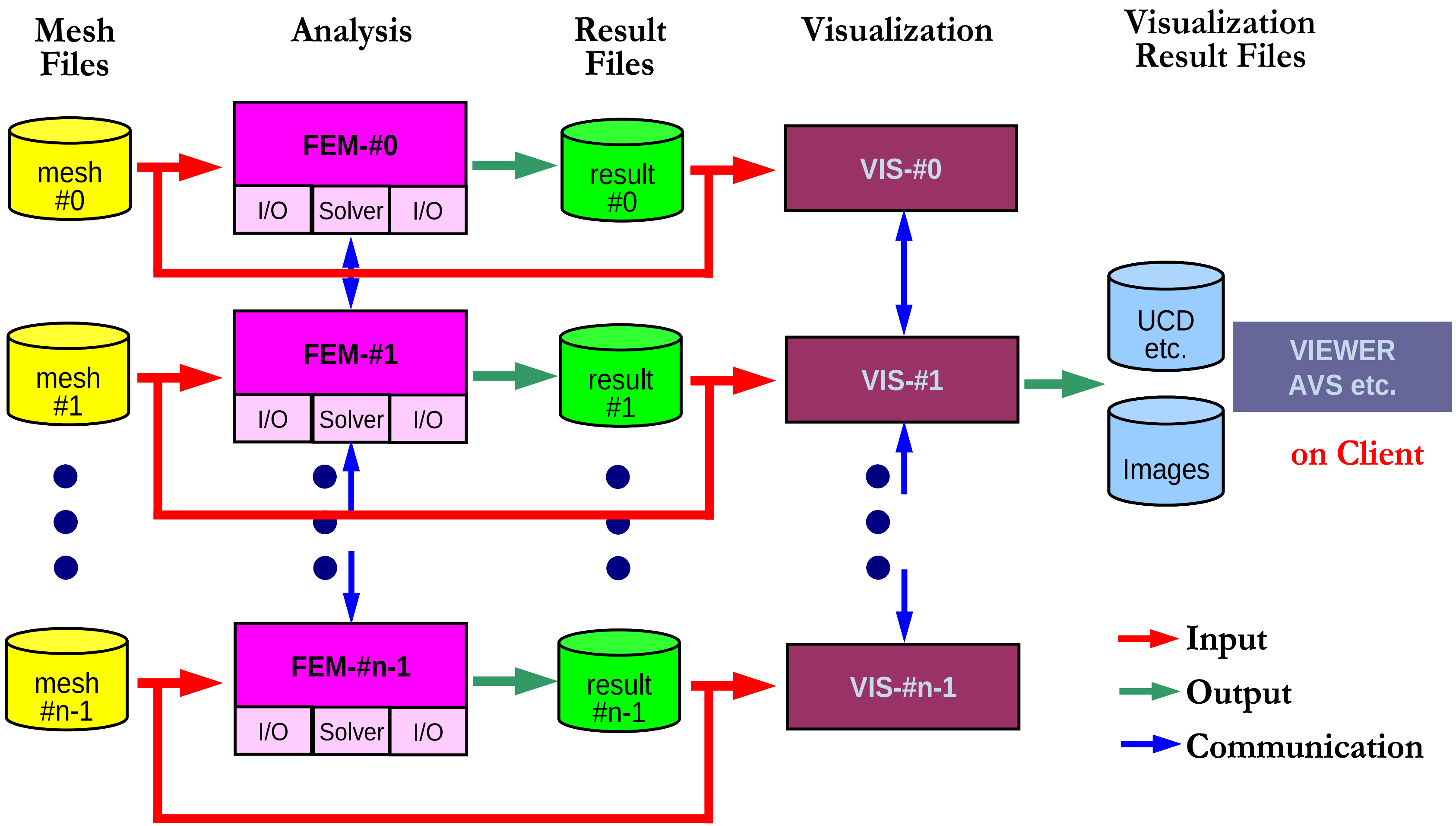
Featured Developments

- ppOpen-AT: AT Language for Loop Optimization
- HACApK library for H-matrix comp. in ppOpen-APPL/BEM (OpenMP/MPI Hybrid Version)
  - First Open Source Library by OpenMP/MPI Hybrid
- **ppOpen-MATH/MP (Coupler for Multiphysics Simulations, Loose Coupling of FEM & FDM)**
- **Sparse Linear Solvers**

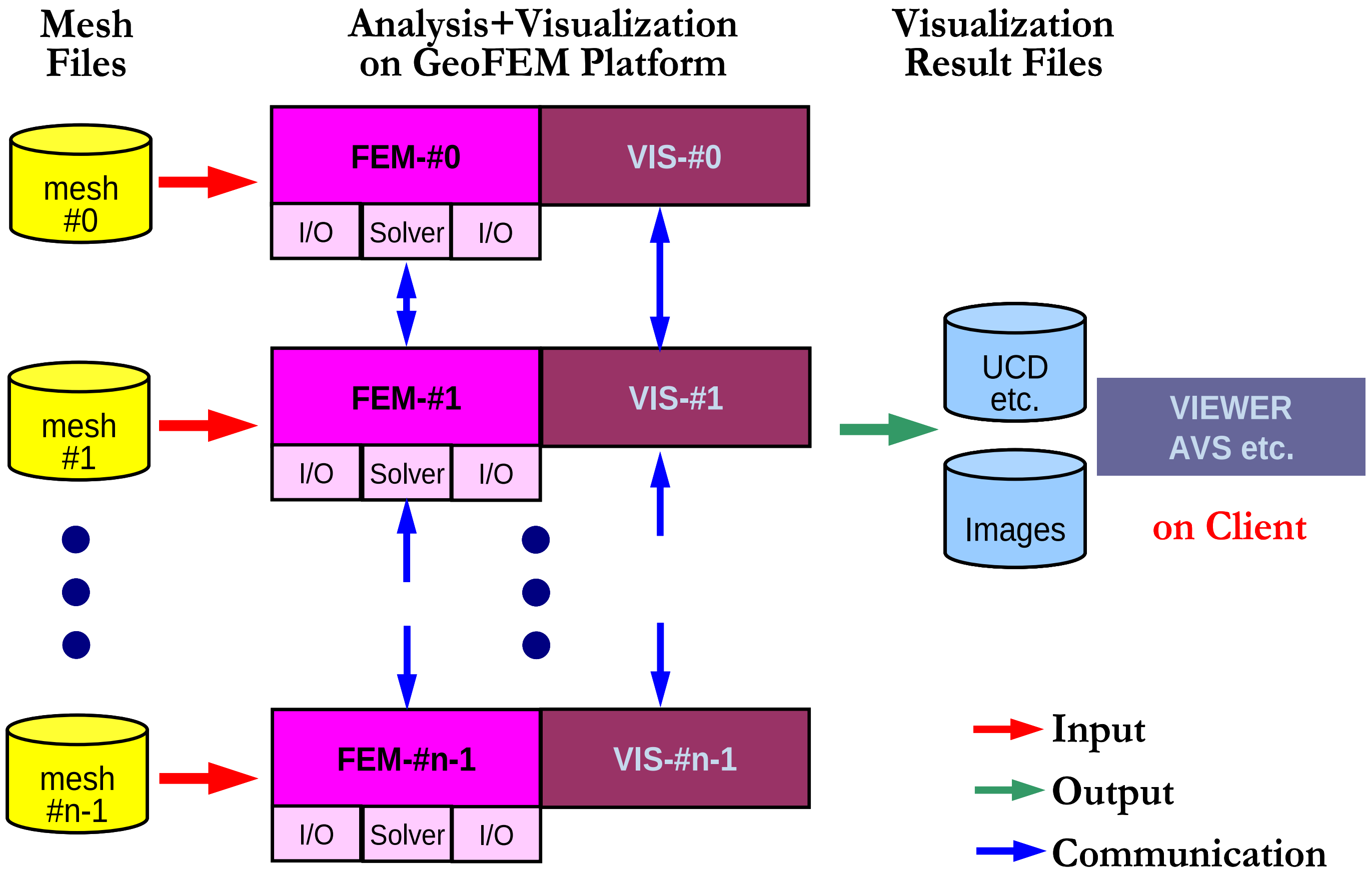




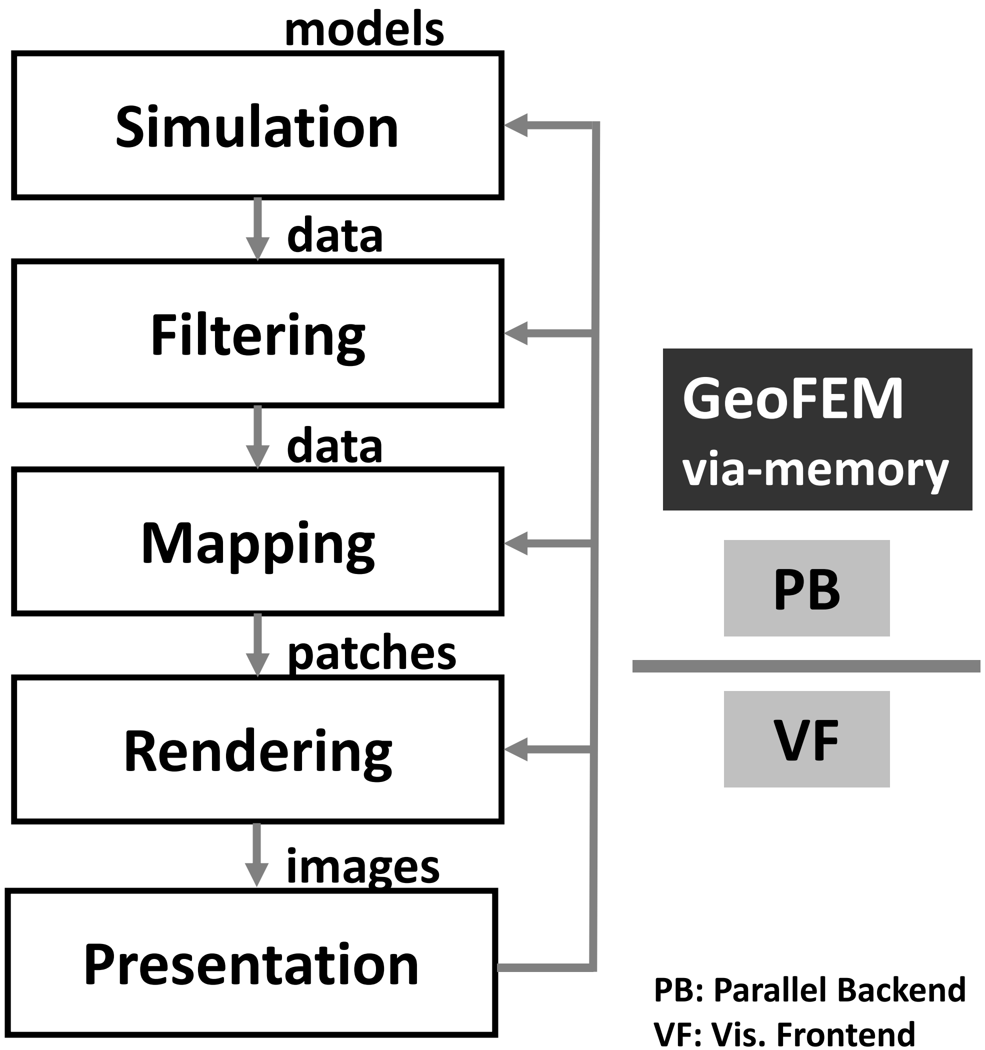
Framework for Parallel Visualization 1
Via-File



Framework for Parallel Visualization 2 Via-Memory (GeoFEM Project)



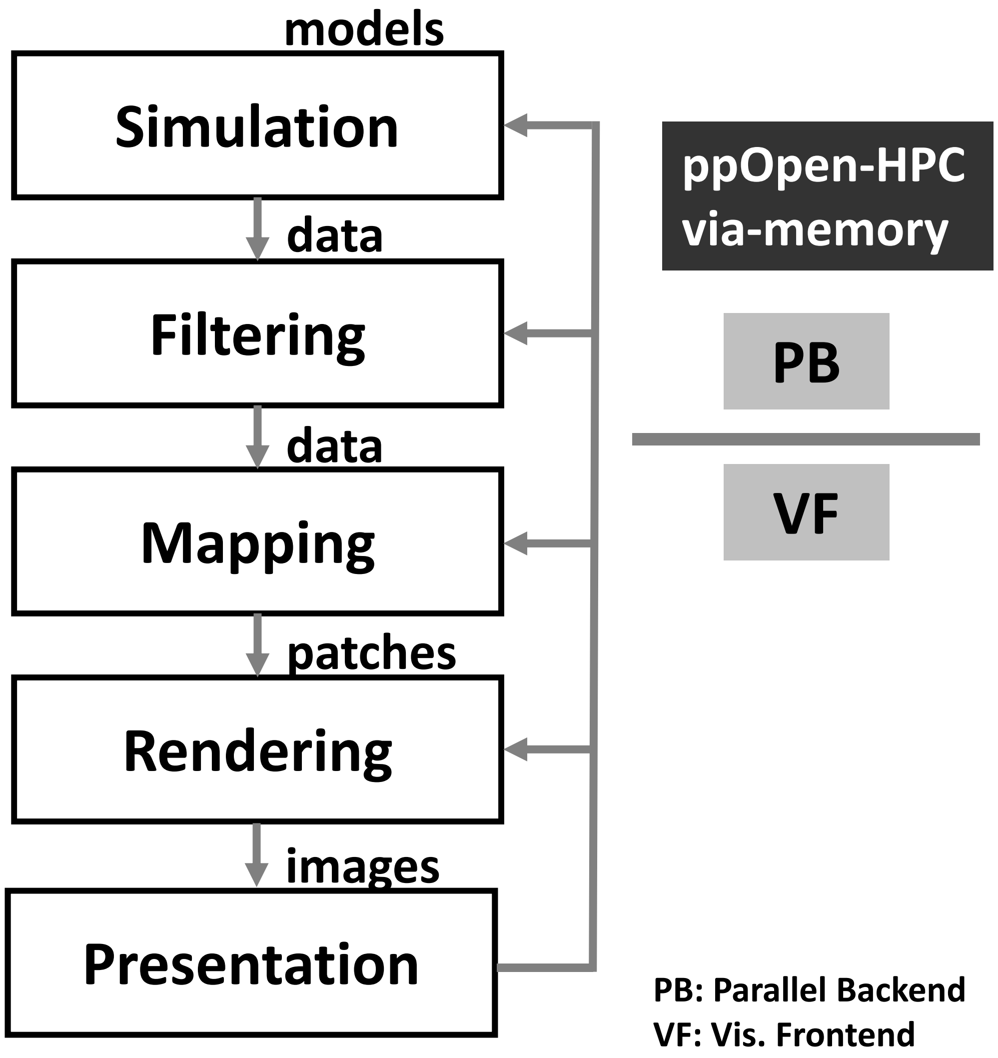
Visualization in ppOpen-HPC



PB: Parallel Backend
VF: Vis. Frontend

- Concurrent Visualization-Computation (Via Memory)
- Output files (single “self-contained（自己完結）” file) are browsed by MicroAVS & Paraview on a PC
  - ✓ In GeoFEM (previous project), only patch files were obtained.

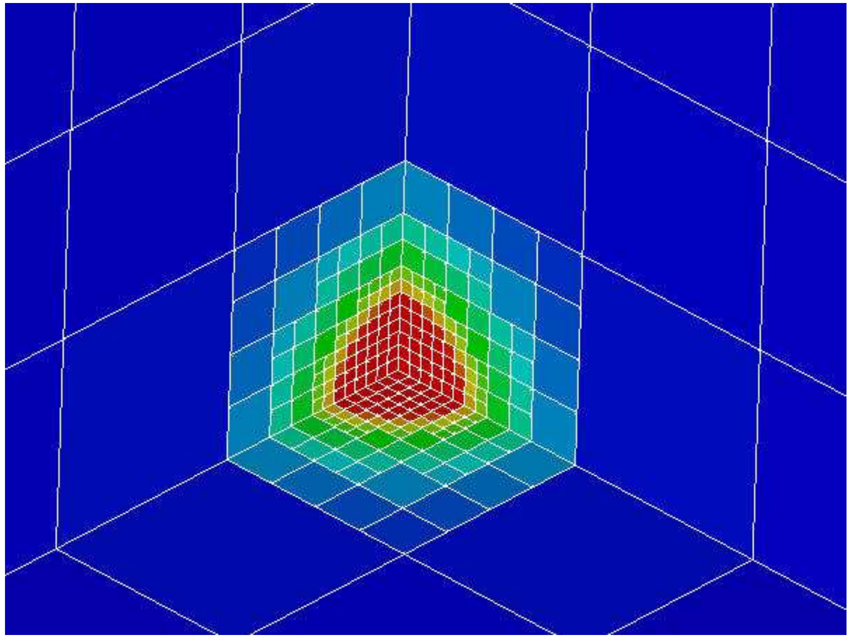
Visualization in ppOpen-HPC



- Concurrent Visualization-Computation (Via Memory)
- Output files (single “self-contained（自己完結）” file) are browsed by MicroAVS & Paraview on a PC
  - ✓ In GeoFEM (previous project), only patch files were obtained.
- Not detailed visualization.
- Just for understanding MIN-MAX, and peaks
- Detailed geometry is preferable

ppOpen-MATH/VIS

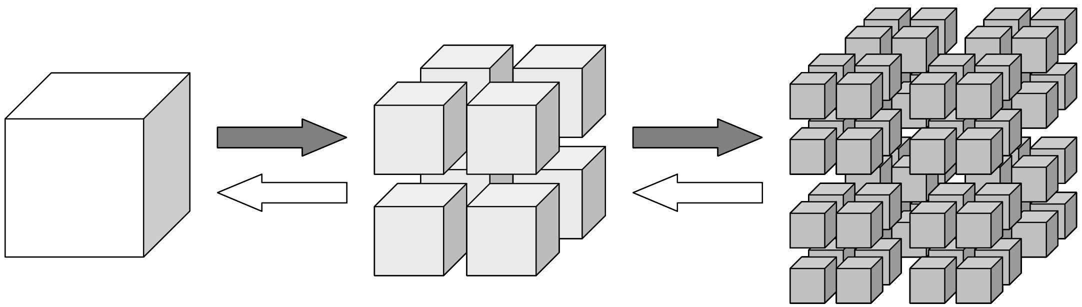
- Parallel Visualization using Information of Background Voxels [Nakajima & Chen 2006]
  - FDM version is released: ppOpen-MATH/VIS-FDM3D
- UCD single file

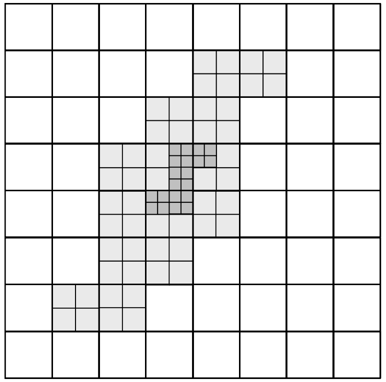


```
[Refine]
AvailableMemory = 2.0     Available memory size (GB), not available in this version.
MaxVoxelCount   = 500     Maximum number of voxels
MaxRefineLevel  = 20      Maximum number of refinement levels
```

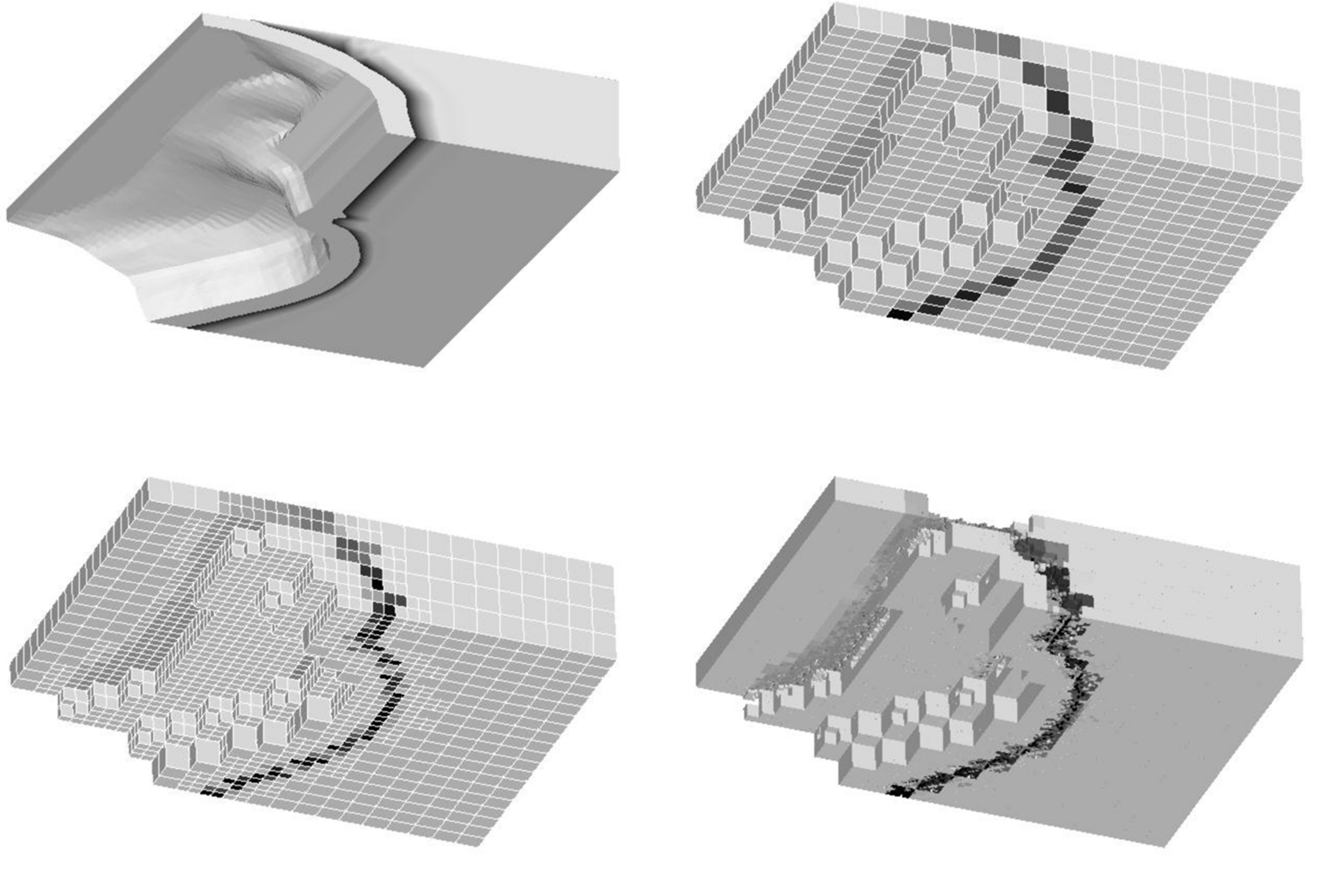
Simplified Parallel Visualization using Background Voxels

- Octree-based AMR
- AMR applied to the region where gradient of field values are large
  - stress concentration, shock wave, separation etc.
- If the number of voxels are controlled, a single file with 10^5 meshes is possible, even though entire problem size is 10^9 with distributed data sets.

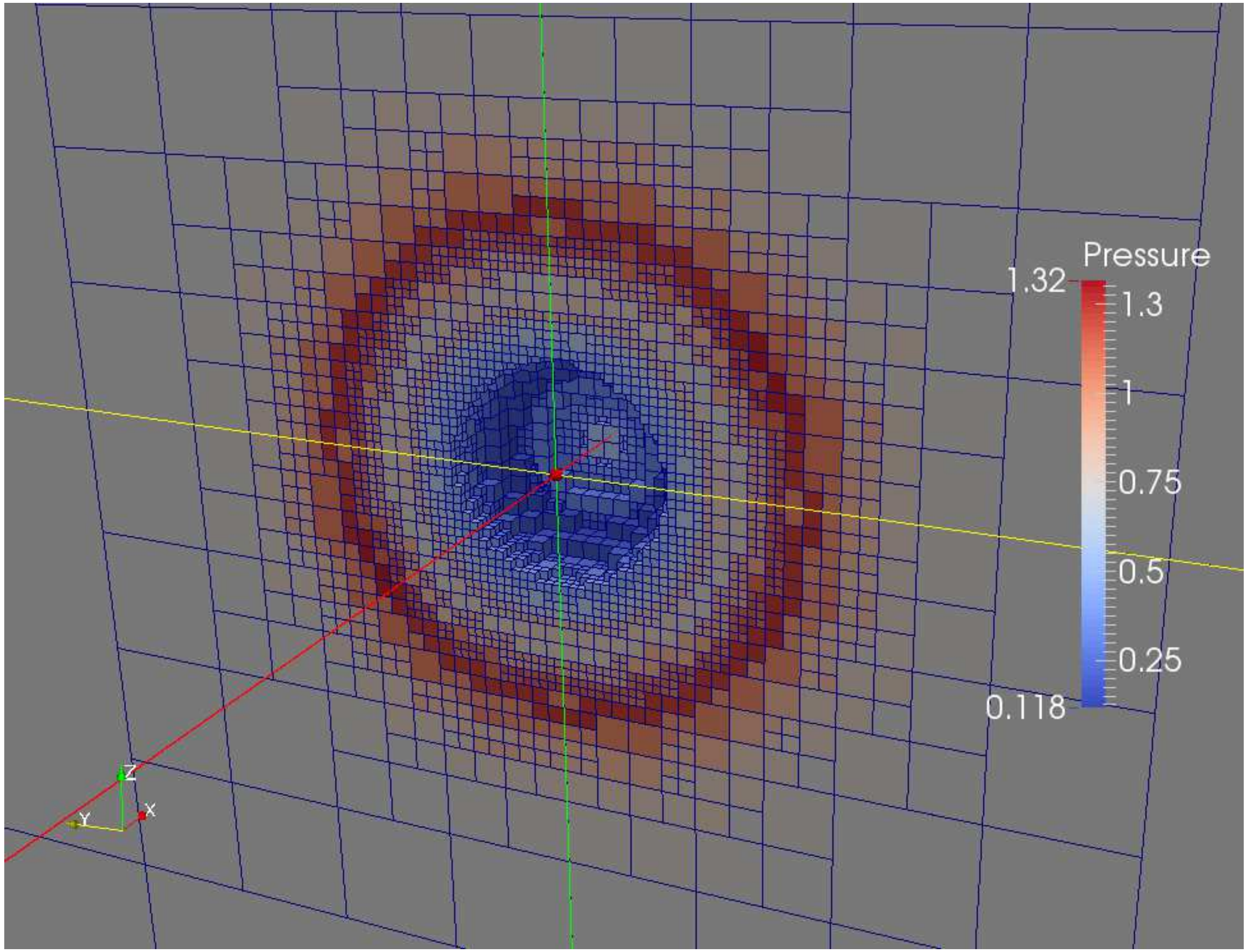




Voxel Mesh (adapted)



Flow around a sphere



pFEM3D + ppOpen-MATH/VIS

```
>$ cd /lustre/gt18/t18XXX/pFEM/pfem3d/srcV
>$ make
>$ cd ../run
>$ ls solv
      solv
```

Makefile（Fortran）

```
include Makefile.in

FFLAGSL   = -I/lustre/gt00/z30088/class_eps/include
FLDFLAGSL = -L/lustre/gt00/z30088/class_eps/lib
LIBSL     = -lfppohvispfem3d -lppohvispfem3d

.SUFFIXES:
.SUFFIXES: .o .f90 .f

.f.o:
	$(FC) -c $(FFLAGS) $(FFLAGSL) $< -o $@
.f90.o:
	$(F90) -c $(F90FLAGS) $(FFLAGSL) $< -o $@
TARGET = ../run/solv
OBJS = ¥
	pfem_util.o …

all: $(TARGET)

$(TARGET): $(OBJS)
	$(F90) -o $(TARGET) $(F90FLAGS) $(FFLAGSL) $(OBJS) 
$(LDFLAGSL) $(LIBS) $(LIBSL) $(FLDFLAGSL)

clean:
	rm -f *.o *.mod $(TARGET)
distclean:
	rm -f *.o *.mod $(TARGET)
```

Makefile（C）

```
include Makefile.in

CFLAGSL  = -I/lustre/gt00/z30088/class_eps/include
LDFLAGSL = -L/lustre/gt00/z30088/class_eps/lib
LIBSL    = -lppohvispfem3d

.SUFFIXES:
.SUFFIXES: .o .c

.c.o:
	$(CC) -c $(CFLAGS) $(CFLAGSL) $< -o $@

TARGET = ../run/solv

OBJS = test1.o …

all: $(TARGET)

$(TARGET): $(OBJS)
	$(CC) -o $(TARGET) $(CFLAGS) $(CFLAGSL) $(OBJS)
$(LDFLAGSL) $(LIBS) $(LIBSL)
clean:
	rm -f *.o *.mod $(TARGET)

distclean:
	rm -f *.o *.mod $(TARGET)
```

Fortran/main (1/2)

```
      use solver11
      use pfem_util
      use ppohvis_pfem3d_util

      implicit REAL*8(A-H,O-Z)
      type(ppohVIS_BASE_stControl)                  :: pControl
      type(ppohVIS_BASE_stResultCollection)         :: pNodeResult
      type(ppohVIS_BASE_stResultCollection)         :: pElemResult
      character(len=PPOHVIS_BASE_FILE_NAME_LEN)     :: CtrlName
      character(len=PPOHVIS_BASE_FILE_NAME_LEN)     :: VisName
      character(len=PPOHVIS_BASE_LABEL_LEN)         :: ValLabel
      integer(kind=4)                                :: iErr

      CtrlName = ""
      CtrlName = "vis.cnt"

      VisName = ""
      VisName = "vis"

      ValLabel = ""
      ValLabel = "temp"

      call PFEM_INIT

      call ppohVIS_PFEM3D_Init(MPI_COMM_WORLD, iErr)
      call ppohVIS_PFEM3D_GetControl(CtrlName, pControl, iErr);
      call INPUT_CNTL
      call INPUT_GRID

      call ppohVIS_PFEM3D_SETMESHEX(                                  &
     &       NP,       N,            NODE_ID, XYZ,                    &
     &       ICELTOT, ICELTOT_INT, ELEM_ID, ICELNOD,                  &
     &       NEIBPETOT, NEIBPE, IMPORT_INDEX, IMPORT_ITEM,            &
     &                        EXPORT_INDEX, EXPORT_ITEM, iErr)
```

Fortran/main (2/2)

```
       call MAT_ASS_MAIN
       call MAT_ASS_BC

       call SOLVE11

       pNodeResult%ListCount = 1
       pElemResult%ListCount = 0
       allocate(pNodeResult%Results(1))

       call ppohVIS_PFEM3D_ConvResultNodeItem1N(                         &
      &       NP, ValLabel, X, pNodeResult%Results(1), iErr)

       call ppohVIS_PFEM3D_Visualize(pNodeResult, pElemResult, pControl, &
      &                              VisName, 1, iErr)

       call PFEM_FINALIZE

       end program heat3Dp
```

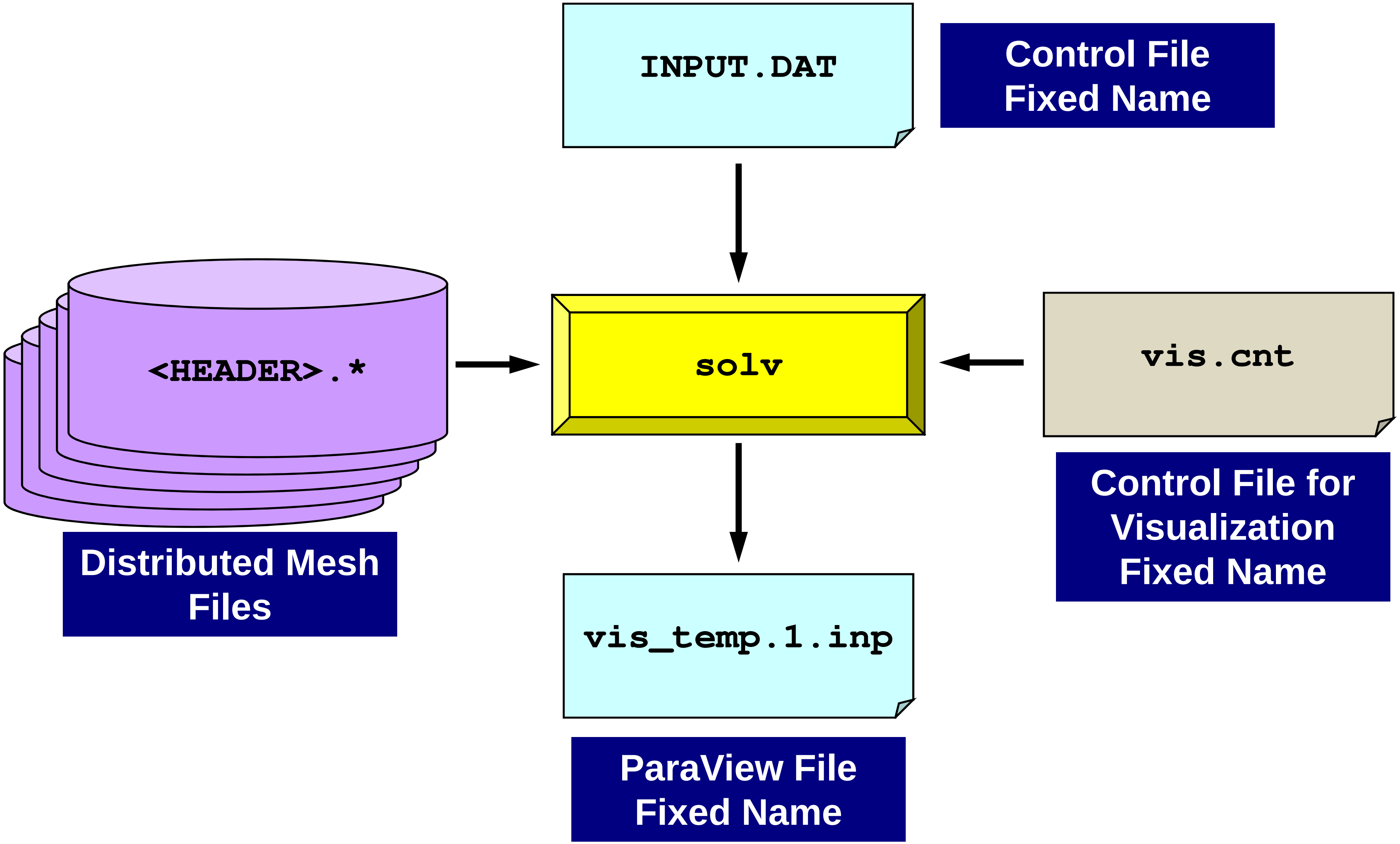
C/main (1/2)

```
#include <stdio.h>
#include <stdlib.h>
FILE* fp_log;
#define GLOBAL_VALUE_DEFINE
#include "pfem_util.h"
#include "ppohVIS_PFEM3D_Util.h"
extern void PFEM_INIT(int,char**);
extern void INPUT_CNTL();
extern void INPUT_GRID();
extern void MAT_CON0();
extern void MAT_CON1();
extern void MAT_ASS_MAIN();
extern void MAT_ASS_BC();
extern void SOLVE11();
extern void OUTPUT_UCD();
extern void PFEM_FINALIZE();
int main(int argc,char* argv[])
{
  double START_TIME,END_TIME;
  struct ppohVIS_FDM3D_stControl *pControl = NULL;
  struct ppohVIS_FDM3D_stResultCollection *pNodeResult = NULL;

  PFEM_INIT(argc,argv);
  ppohVIS_PFEM3D_Init(MPI_COMM_WORLD);
  pControl = ppohVIS_FDM3D_GetControl("vis.cnt");

  INPUT_CNTL();
  INPUT_GRID();

  if(ppohVIS_PFEM3D_SetMeshEx(
      NP,N,NODE_ID,XYZ,
      ICELTOT,ICELTOT_INT,ELEM_ID,ICELNOD,
      NEIBPETOT,NEIBPE,IMPORT_INDEX,IMPORT_ITEM,EXPORT_INDEX,EXPORT_ITEM)) {
            ppohVIS_BASE_PrintError(stderr);
            MPI_Abort(MPI_COMM_WORLD,errno);
  };
```

C/main (2/2)

```
  MAT_CON0();
  MAT_CON1();

  MAT_ASS_MAIN();
  MAT_ASS_BC()   ;

  SOLVE11();

  OUTPUT_UCD();

  pNodeResult=ppohVIS_BASE_AllocateResultCollection();
        if(pNodeResult == NULL) {
                ppohVIS_BASE_PrintError(stderr);
                MPI_Abort(MPI_COMM_WORLD,errno);
        };
        if(ppohVIS_BASE_InitResultCollection(pNodeResult, 1)) {
                ppohVIS_BASE_PrintError(stderr);
                MPI_Abort(MPI_COMM_WORLD,errno);
        };

        pNodeResult->Results[0] =

  ppohVIS_PFEM3D_ConvResultNodeItemPart(NP,1,0,"temp",X);

  START_TIME= MPI_Wtime();
        if(ppohVIS_PFEM3D_Visualize(pNodeResult,NULL,pControl,"vis",1)) {
                ppohVIS_BASE_PrintError(stderr);
                MPI_Abort(MPI_COMM_WORLD,errno);
        };

  ppohVIS_PFEM3D_Finalize();

  PFEM_FINALIZE() ;
}
```

pFEM3D + ppOpen-MATH/VIS



Preparing Distributed Mesh Files

```
>$ cd /lustre/gt18/t18XXX/pFEM/pfem3d/pmesh
(mesh.inp, mg.sh)

>$ qsub mg.sh
```

mesh.inp

```
256 256 256
  4   4   4
pcube
```

256^3 nodes into 4×4×4=64 partitions
16,777,216 nodes
16,581,375 elements

Each MPI process has 64^3 nodes

mg.sh

```
#!/bin/sh
#PBS -q u-lecture8
#PBS -N mesh
#PBS -l select=2:mpiprocs=32

#PBS -Wgroup_list=gt18
#PBS -l walltime=00:05:00
#PBS -e err
#PBS -o mg.lst

cd $PBS_O_WORKDIR
. /etc/profile.d/modules.sh
export I_MPI_PIN_DOMAIN=socket
export I_MPI_PERHOST=32
mpirun ./impimap.sh ./pmesh
rm wk*
```

Computation + Visualization

```
>$ cd /lustre/gt18/t18XXX/pFEM/pfem3d/run
(INPUT.DAT, gv.sh)

>$ qsub gv.sh
```

INPUT.DAT

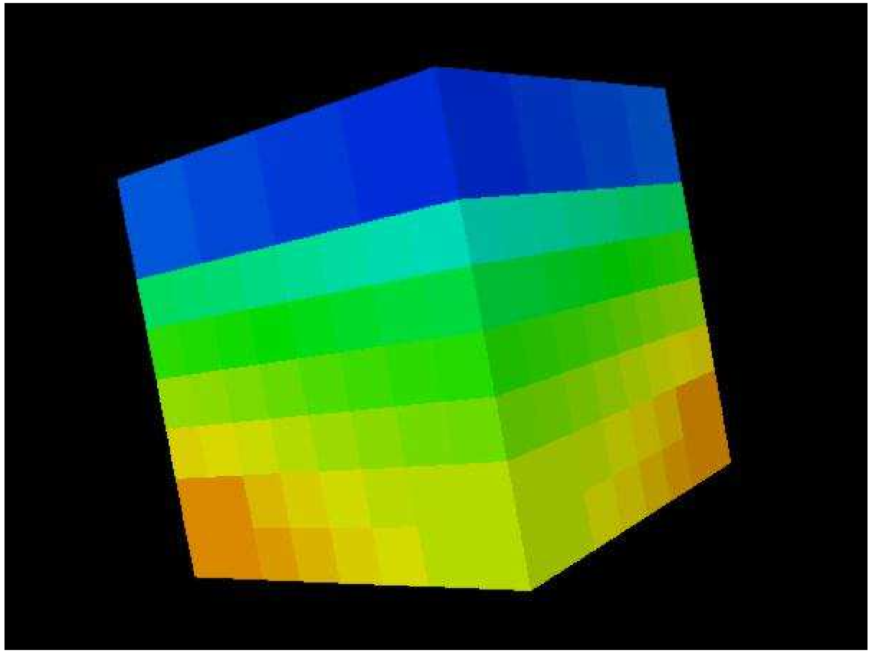
```
../pmesh/pcube
2000
1.0 1.0
1.0e-08
```

gv.sh

```
#!/bin/sh
#PBS -q u-lecture8
#PBS -N run+vis
#PBS -l select=2:mpiprocs=32

#PBS -Wgroup_list=gt18
#PBS -l walltime=00:05:00
#PBS -e err
#PBS -o test.lst

cd $PBS_O_WORKDIR
. /etc/profile.d/modules.sh
export I_MPI_PIN_DOMAIN=socket
export I_MPI_PERHOST=32
mpirun ./impimap.sh ./solv
```

vis.cnt

```
[Refine]                    Control Info. for Refinement
AvailableMemory = 2.0       Available Memory (GB) not in use
MaxVoxelCount = 1000        Max Voxel #
MaxRefineLevel = 20         Max Voxel Refinement Level
[Simple]                    Control Info. for Simplification
ReductionRate = 0.0         Reduction Rate of Surface Patches
[Output]                    Output Format
FileFormat     = 2          =1:MicroAVS, =2:ParaView
```



Values at Cell Ctr.

16,777,216 nodes
16,581,375 elem's, 64 MPI proc's

↓

vis_temp.1.inp
1,436 nodes
1,002 elements

COPY the File to Your PC

REEDBUSH

```
>$ cd /lustre/gt18/t18XXX/pFEM/pfem3d/run
>$ cp vis_temp.1.inp ~/.
```

PC

```
>$ scp t18XXX@reedbush.cc.u-tokyo.ac.jp:~/vis_temp.1.inp .
```