Multi-Scale Discrete Simulation on Multi-Scale HPC System

Group of Complex System and Multi-scale Simulation
Institute of Process Engineering, Chinese Academy of Sciences
Presented by Wei Ge
Outline

Challenges and approaches
Software and hardware development
Applications in different areas
Summary and prospects
Multi-scale simulation in process engineering

Can we find an accurate, efficiency and general way?

Too... costly!

Agent

Flow sheet

MD

MC / DPD

DEM

PBM

TFM

Mm

year

environment

molecule

molecular cluster

particle

molecular cluster

particle cluster

reactor

process

Angstrom

fs

nm

ps

μm

μs

mm

ms

m

s

km

hour

fs

ps

μs

s

hour

year

m

mm

Mm

km

m

ms

μs

nm

ps

Angstrom

MD
Generality vs Efficiency

- Fully general
- General Algorithmic Framework for Special Architecture
- Fully special
- Exist? → discrete method
Traditional parallelization

Multi-scale world

Global communication, Mono-scale parallelization
Multi-scale parallelization

Multi-scale world

Local communication, Multi-scale parallelization
Hierarchy of discrete approaches for complex flows

Micro-scale: fluctuating, conservative
MD, DSMC, LGA, PPM, …

Meso-scale: fluctuating, dissipative
DPD, FPM, DSPH, LBM, …

Macro-scale: smooth, dissipative
SPH, MPS, DEM, MaPPM, …
Consistency: Physics, Algorithm, Architecture

- **Macro**
- **Meso**

Long-Range Correlation

Parameter Exchanges

Local Interaction

Software organization

Switch

CPU

GPU

Hardware architecture
Outline

Challenges and approaches
Software and hardware development
Applications in different areas
Summary and prospects
General Platform for Discrete Simulation

- Particle Method: MD, PPM, SPH, DPD, DEM, MaPPM
- Data Partition: Uniform Domain, Uniform Load
- Preprocess: Boundary Disposal, Particle Generation
- Configuration: CAD Drawing Conversion
- Algorithm: Space Decomposition, Link Cell + Neighbor List, Dynamic Load Balance, Communication Scheme
- Data Structure: Particle, Boundary, Potential, Communicator, Organizer, Assistant
- Computation and Communication: MPI, STL, Loki

C&CE, 2005, 29:1543-1553; Ge et al., Sci. in China, 2005
100Tflops GPU system (2008.2.18)

R_peak : 127Teraflops SP
Nodes : 126×HP8600
CPU : 252×Intel 2.66GHz
GPU cards : 200×NV Tesla C870 + 20×NV GeForce 9800×2
Network : Gigabit Ethernet (mesh+tree)
Switch : H3C 7506R
OS : RedHat Linux 5.2
R_real : 20 ~ 40Tflops

6.023e(23-9.7) flops
Mole-9.7 (<70kW)
China’s first HPC system with 1.0 Petaflops peak performance in single precision (2009.3.19)

2x48x GE SW

Top + Middle
2x4C+HD4870x2
201.6T

2x48x IB-DDR SW

288x GE SW

Middle
4C+2xC1060
64T

Bottom
4C+3xG280
140.4T

288x IB-DDR SW

Bottom
4C+2xHD4870x2
153.6T

Top + Bottom
2x4C+2xG295
460.8T

200T(IPE/Dawning )

200T(IPE/Lenovo)

150T(IPE)

450T(IPE)

Mole-8.7 (<300kW)
### Real performance in Couette-cavity flow

<table>
<thead>
<tr>
<th>Scale</th>
<th>Tflops(sp)</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>464GT200</td>
<td>163.3 / 432.9</td>
<td>39.4%</td>
</tr>
<tr>
<td></td>
<td>(288.6)</td>
<td>(59.1%)</td>
</tr>
<tr>
<td>120RV770</td>
<td>24.9 / 144</td>
<td>17.3%</td>
</tr>
<tr>
<td>680GT200+</td>
<td>118 / 963</td>
<td>12.5%</td>
</tr>
<tr>
<td>274RV770</td>
<td>(306)</td>
<td>(31.7%)</td>
</tr>
</tbody>
</table>

D2Q9 with Kahan ~ D3Q19 > D2Q9

Grooved micro-channel 1024x1024
**First Fermi-based GPU supercomputing system**

2010.04.24

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rpeak SP</td>
<td>2 Petaflops</td>
</tr>
<tr>
<td>Rpeak DP</td>
<td>1 Petaflops</td>
</tr>
<tr>
<td>Total RAM</td>
<td>17.2 TB</td>
</tr>
<tr>
<td>Total GRAM</td>
<td>6.6 TB</td>
</tr>
<tr>
<td>Total HD</td>
<td>360 TB</td>
</tr>
<tr>
<td>Data Comm.</td>
<td>Mellanox QDR InfiniBand</td>
</tr>
<tr>
<td>Inst. Comm.</td>
<td>H3C Gigabit Ethernet</td>
</tr>
<tr>
<td>Occupied area</td>
<td>150 m$^2$ (with internal cooling)</td>
</tr>
<tr>
<td>Weight</td>
<td>12.6 T (with internal cooling)</td>
</tr>
<tr>
<td>Max Power</td>
<td>600 kW (computing) + 200 kW (cooling)</td>
</tr>
<tr>
<td>System</td>
<td>CentOS 5.4, PBS</td>
</tr>
<tr>
<td>Monitor</td>
<td>Ganglia, GPU monitor</td>
</tr>
<tr>
<td>Languages</td>
<td>C, C++, Fortran, CUDA, OpenCL</td>
</tr>
</tbody>
</table>
Collective capacity:
4.907 Petaflops SP
1.300 Petaflops DP

IPE Mole-8.5
3 Petaflops SP, 1 Petaflops DP

Distributed GPU-Supercomputing in China

过程所
NV+AMD
1P SP
100T DP

高能所
AMD
200T SP
40T DP

金属所
NV
183T SP
15T DP

深圳先进院
NV
173T SP
14T DP

国家天文台
NV
160T SP
13T DP

地质地球所
NV+AMD
200T SP
17T DP

网络中心
NV+AMD
300T SP
39T DP

近代物理所
NV
202T SP
17T DP

紫金山台
NV
183T SP
15T DP

电工所
NV
101T SP
8T DP

中科大
NV
205T SP
17T DP
Distributed GPU computing for oil recovery

Main cluster:
Overall optimization and control
Outline

Challenges and approaches
Multi-scale: from method to hardware
Applications in different systems
Summary and prospects
Case study: Oil refining
1.4Mt/a MIP FCC process producing 1/3 gasoline in China
Simulation of gas solid flow on multi-scales

Reactor: 9*40m 3D EMMS

Section: 3*10m 2D CFD+ EMMS

Cell: 2*10cm 2D DNS in MaPPM

100M grids 432 GPUs ~3s ~100x speedup

1.2M cells 96 GPUs **Realtime** ~50x speedup

120k solids ~ 1G fluids 144 GPUs 20~30x speedup
Approach: Particle-fluid flow $\rightarrow$ particle-particle flow

Particle
- Velocity difference $\rightarrow$ tangential stress
- Density difference $\rightarrow$ normal stress

Continuum
- Viscosity
- Pressure

Fixed particle (boundary)
- Bundled particle (solids)
- Free particle (fluid)
A straightforward formulation

\[ \nabla f \bigg|_a = D \sum_i \frac{f_{ia}}{r_{ai}} r_{ai} \frac{m_i}{\rho_i} W_{ai} \]

\[ \Delta f \bigg|_a = 2D \sum_i \frac{f_{ia}}{r_{ai}}^2 m_i W_{ai} \]

Ge & Li, 2003, Powder Tech. 137:99
Animation Challenge:
9600x2400 → 1200x300 pixels
1000 → 17 frames
Oil recovery: fracture-cave type oil fields

Physical Experiment
0.5x0.8m days

Geological Structure
20x20km years

Simulation
80 GPUs
500x150m
Month (in hours)
Metallurgy: new process "COREX" 30x9m

Flow in porous media: realtime simulation
Particle handling: real time simulation of a rotating drum
(9.6 million particles, 13.5*1.5 meter)
Micro-/Nano-fluidics:

3D simulation: bubble-particle in liquid
0.1*0.1*0.15µm, bubble mean velocity 3m/s
LJ/PP fluid at 60K, NVT ensemble
7M particles, 2 GPUs
15~50x speedup

3D simulation: gas-liquid phase transition
0.1*0.05*0.1µm,
LJ/PP fluid at 60K, NVT ensemble
1M particles, 2GPU
Material: multi-scale structure in solar cells

43.2nm x 48.7nm x 5.4nm
572,800 atoms
1GPU, 50x speedup, for force
Biochemistry:

Polymer dynamics
- 1200 polyethylene chains
- Chain length: 300 CH2
- NVT ensemble
- 30x speedup

Vesicle formation
- 1392656 water
- 3375 dipalmityl phosphatidyl choline
- NPT ensemble
- 20x speedup
Data processing:

Image reconstruction for industrial CT

Scanning

>80x speedup

off-line

realtime
Prospect of virtual process engineering

hardware + software → application

CPU

CPU+GPU

CPU+XPU+GPU

macro

meso

micro

average

multi-scale discrete

pre-/post-processing

engineering design

CAD

down

realtime

down

virtual process
Summary

• Similarity between problem, software and hardware is key to real HPC.

• Multi-scale discrete simulation is natural and advantageous way for the simulation of a wide variety of complex systems.

• GPU computing provides an effective way to realize multi-scale discrete simulation with commercial components.
Thank you for your attention!

Visit us: www.ipe.ac.cn/csms (to be updated)

This work is supported by:
NSFC MOF MOST CAS
PetroChina SinoPec Baosteel
BHPbilliton CSIRO ALSTOM NVIDIA