An Efficient and Effective Nonlinear Solver of Large Scale Petroleum Reservoir Simulation

Jianwen CAO

Collaborate with

Prof. Jiachang SUN, Feng PAN, Fuqiang LEI

Parallel Computing Laboratory

Institute of Software

the Chinese Academy of Sciences

Milestones in recent years

>>

TWO months

Seq. Commercial Code

MIPS R10K (1998)

0.94 hrs

Our Parallel Code,

LSSC-II 64CPU (2002)

Xeon 2G with Myrinet

Computation Capacity (64 CPUs case)

1998 → → → → 2002

CPU frequency improves : 5 times

Coarsely parallelization : 8 times

Parallel strategy tuning : 8 times

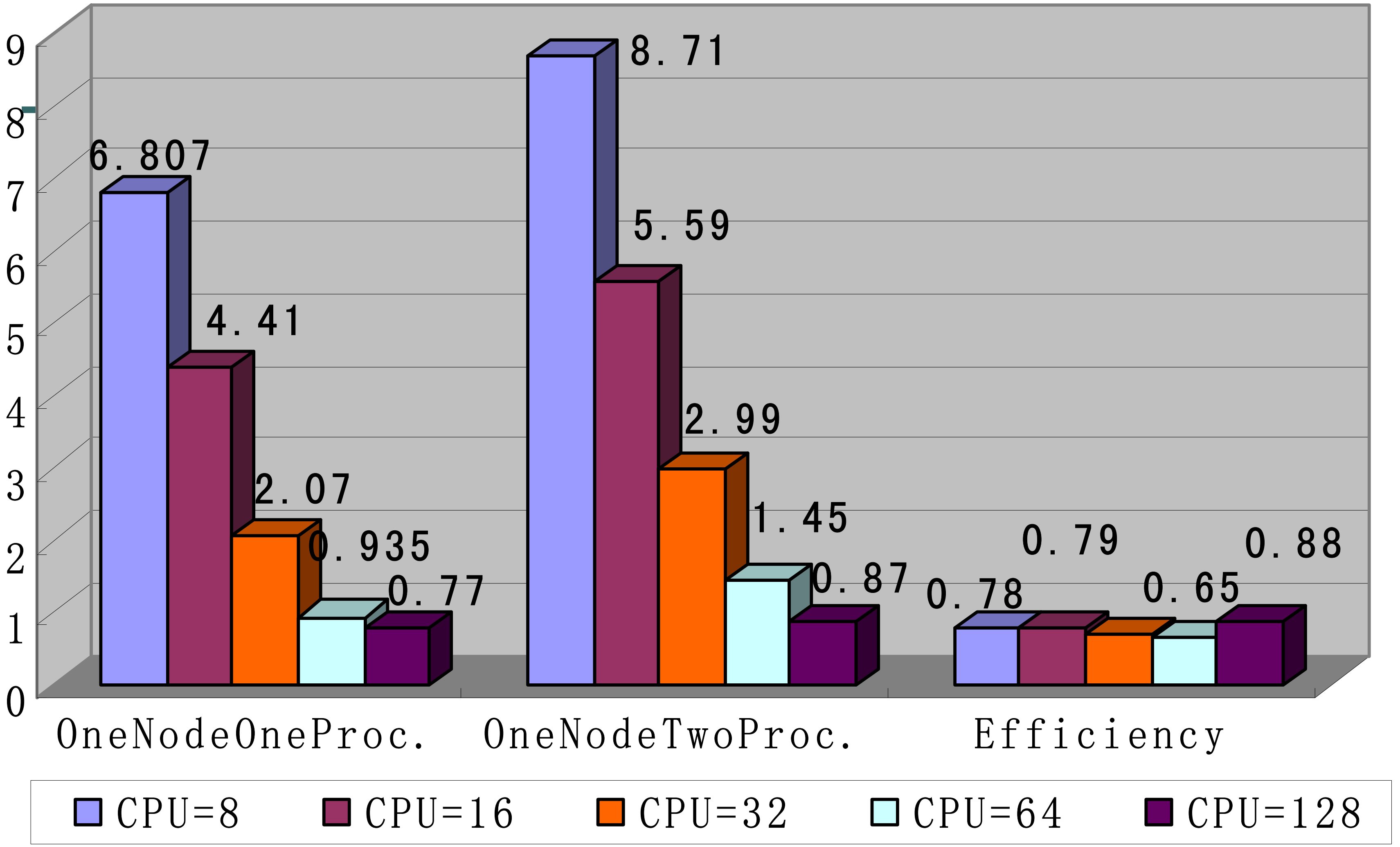
Algorithm acceleration : 5 times

TOTAL CAPACITY : 5x8x8x5=1600

Why this para code is efficient ?

- **Newton-Krylov-Schwartz solution strategy**
  - Robust solver is choosed
  - Avoid frequent global message passing hand and foot
- **MPI-based parallelism strategy**
  - Communication frequency minimization
  - Thoroughly data-based parallelization

- **Dealing with wells & faults using parallel strategy**
  - Wells are divided properly and conquered by its owned processor
  - Faults are localized hand and foot, and treated by a group of processors which own this fault
  - Group communication instead of global communication
- **File I/O process is parallel in order to avoid I/O bottleneck**



9
8
7
6
5
4
3
2
1
0
6.807
4.41
2.07
0.935
0.77
8.71
5.59
2.99
1.45
0.87
0.78
0.79
0.65
0.88
OneNodeOneProc.
OneNodeTwoProc.
Efficiency
CPU=8
CPU=16
CPU=32
CPU=64
CPU=128

Why this para code is effective?

- An elaborate nonlinear solver
  - Inexact Newton iteration is used
  - When IN don't converge satisfactorily, the iteration will be damped or backtracking
  - If IN is difficult to converge or even diverge, initial approximation will rebuild by using BFGS
  - Cutting time step by half is the third choice

- **A well-tuned linear solver**
  - Parallel Preconditioned FGMRES
  - FGMRES(10+), maximum iteration limitation :88
  - Adopt an iterative technique as its preconditioner :
    - ILU+GMRES(10) as a preconditioner : max. # restart: 3
  - Tolerance criteria is related to :
    - Nolinear iteration history evaluation
    - Residual of linear iteration vector,Problem scale

- **A distinct preconditioner**
  - It is based on a multipurpose oblique projection correction method
  - There are three types of oblique projection:
    - A → A_P   global matrix to pressure submatrix
    - A → A_Omega   DDM as a projection
    - A → A_Coarse   AMG as a projection
  - There are eight preconditioning components:
    - Relaxed ILU decomposition

- Additive Schwarz used for DDM
- AMG for initial guess & two level schwarz method
- Watts correction for anisotropy of permeability
- Constraint Residual Precond for BLK model
- CRPextended precond for compositional model
- Shur complement & Multi step Method for coupled system
- Decoupling operator used as left precond

It is adaptive, there are 10 difficult level, higher level, more powerful, and more overhead.

$$T_C := T_{CoarseGrid} := T_{Watts} \bigcup T_{AMG}$$

$$T_1 := T_{AddSchwarz} \bigcup T_{\mathrm{Re}\ laxILU}$$

$$T_2 = P_1 (P_1^T A P_1)^{-1} P_1^T$$

$$T_3 = P_2 (P_2^T A P_2)^{-1} P_2^T$$

$$T_D A T_{Right} (T_{Right}^{-1} x) = T_D b$$

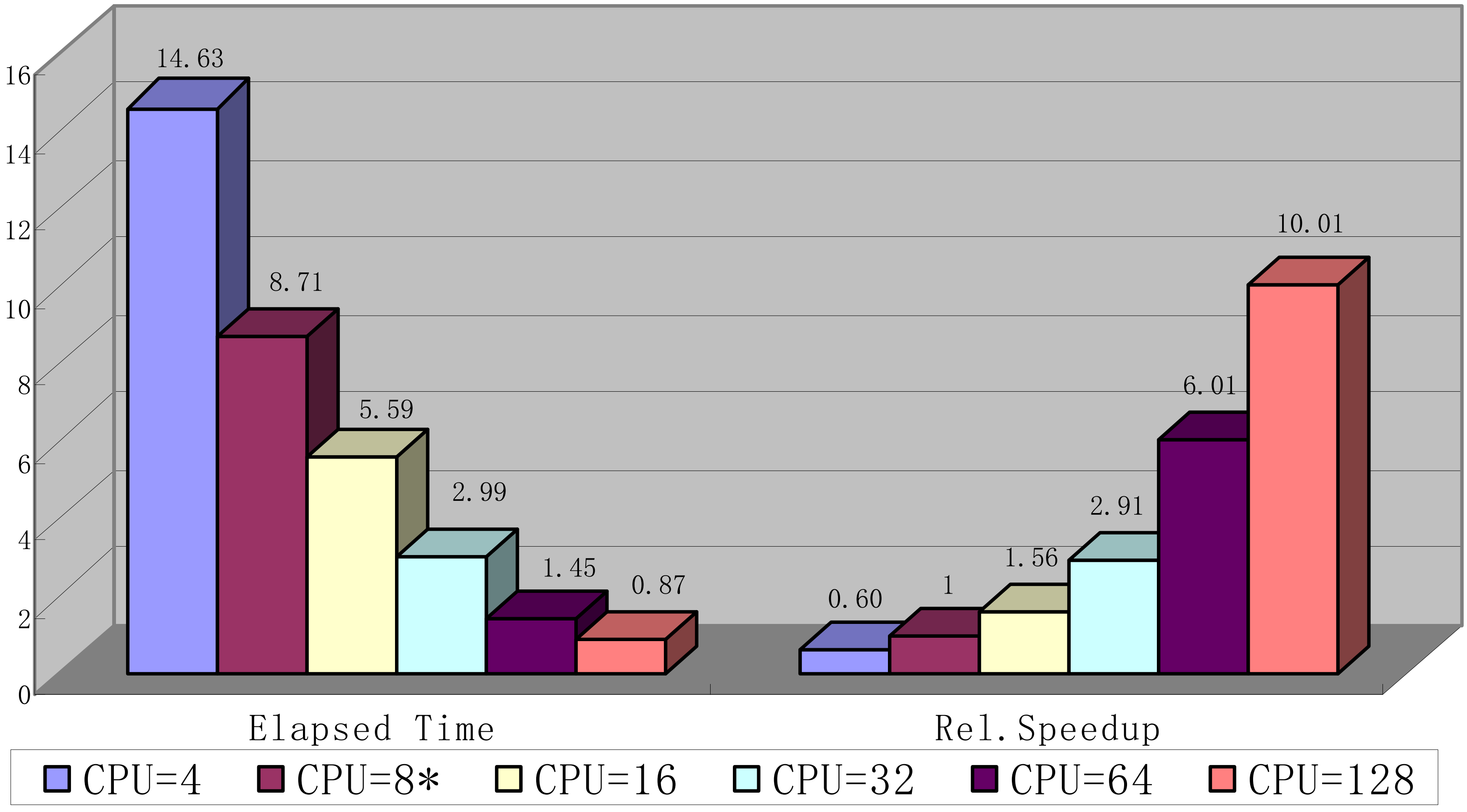
$$B = T_D \bigcup T_{Right}$$

$$I - A T_{Right} = (I - T_3)(I - T_C)(I - T_2)(I - T_1)$$

Practical test cases and results

- **1159K scale BLK data from DaQing**
  - 291 wells, 31.5 years history matching
- **1382K scale BLK data from ShengLi**
  - 326 wells, 14 years history matching
- **5529K scale BLK data from ShengLi**
  - 326 wells, 14 years history matching
- **430K scale BLK data from CNOOC**
  - 29 wells, 22 years history matching

1159K scale BLK data from DaQing



4 x 6 x 8 = 192
4 x 6 x 8 x 12 = 2304

- One time step needs 4 Newton iteration steps
- One Newton step needs 6 FGMRES Steps
- One FGMRES step needs 8 linear iterations
- One linear iteration needs 12 precond. iters.

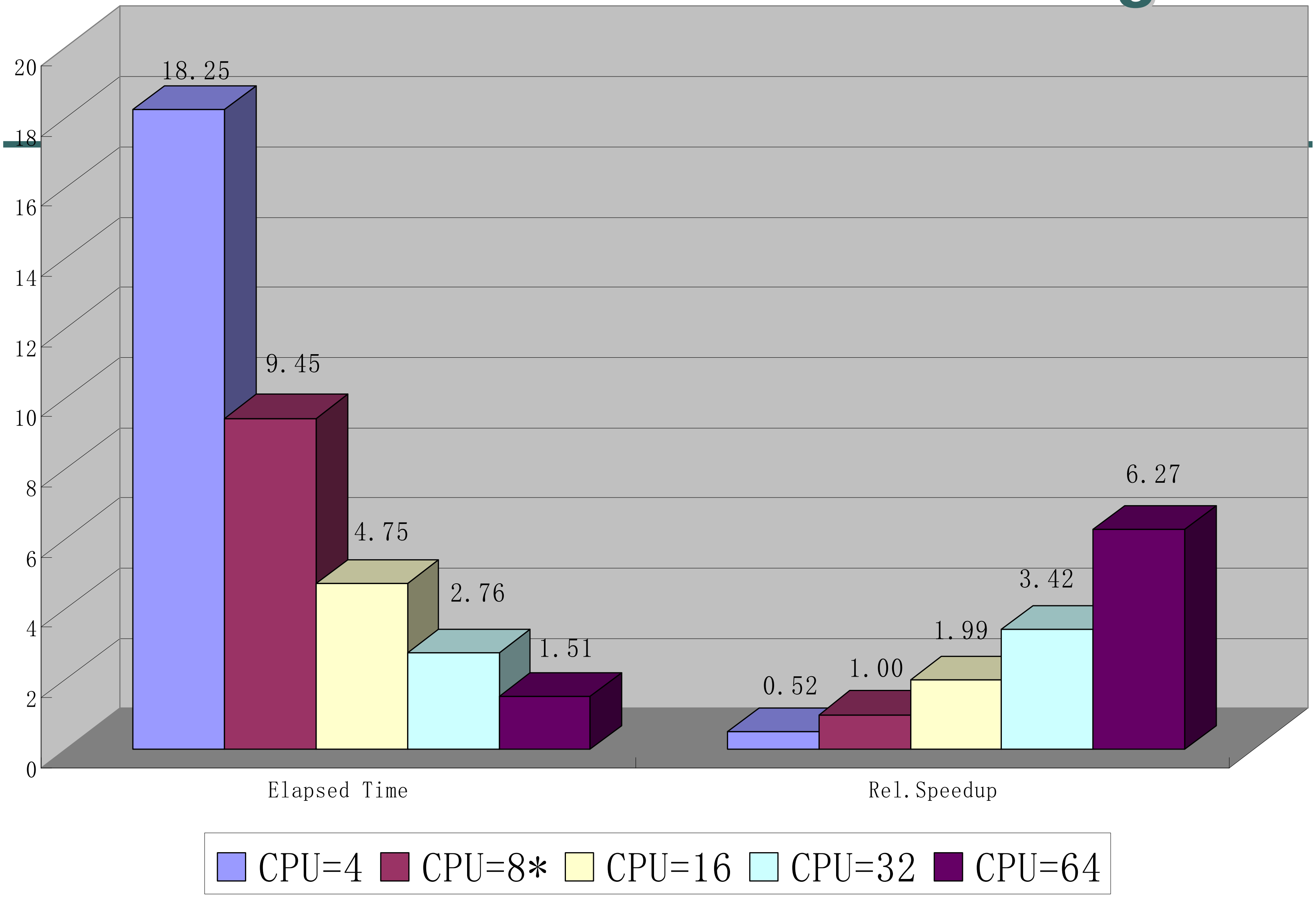
- One Newton step needs 3000 scale reductions
- One FGMRES step needs 500 scale reductions
- One linear iteration needs 65 scale reductions
- One precond.iteration needs 6 scale reductions

- Scale / Broadcast = 91.06%
- Vector / Broadcast = 8.93%

- Neighbor Exchanges / Broadcast = 24.52%

- Average timestep : 90 days
- For CASE1, Delta(t) is larger, so the needed linear iterations & precond. iterations is larger than CASE2.

1382K scale BLK data from ShengLi



5 x 7 x 5 = 175

5 x 7 x 5 x 6 = 1050

- One time step needs 5 Newton iteration steps
- One Newton step needs 7 FGMRES Steps
- One FGMRES step needs 5 linear iterations
- One linear iteration needs 6 precond. iters.

- One Newton step needs 1500 scale reductions
- One FGMRES step needs 220 scale reductions
- One linear iteration needs 44 scale reductions
- One precond.iteration needs 9 scale reductions

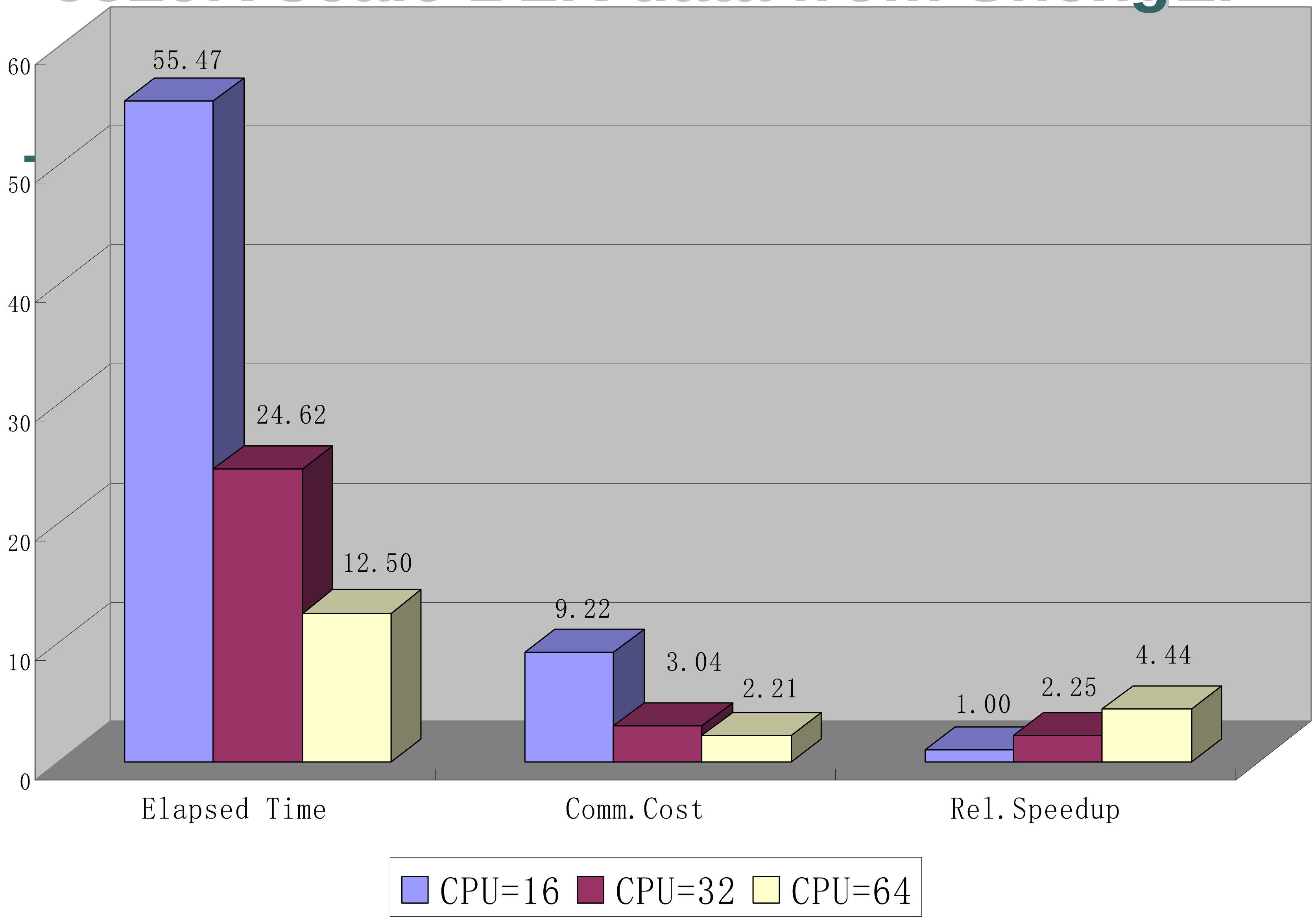
- Scale / Broadcast = 84.54%
- Vector / Broadcast = 15.45%

It is higher than case1 because of more slant wells increase the frequency of vector comm.

- Neighbor Exchanges / Broadcast = 40.89%

- Average time step : 30 days
- For CASE2, Delta(t) is smaller, so the needed linear iteration & precond. iteration are less than that of CASE1

5529K scale BLK data from ShengLi



8 x 8 x 6 = 384

8 x 8 x 6 x 6 = 2304

- One time step needs 8 Newton iteration steps
- One Newton step needs 8 FGMRES Steps
- One FGMRES step needs 6 linear iterations
- One linear iteration needs 6 precond. iters.

- One Newton step needs 2200 scale reductions
- One FGMRES step needs 290 scale reductions
- One linear iteration needs 47 scale reductions
- One precond.iteration needs 9 scale reductions

- Scale / Broadcast = 89.18%
- Vector / Broadcast = 10.81%

- Neighbor Exchanges / Broadcast = 38.44%

- Average time step : 30 days

430K scale BLK data from CNOOC

- Grids : 101x91x47 = 431977
- Faults : 36
- Wells : 7+22
- Elapsed time :
  - 2p : 7382s 4p : 3222s
  - 8p : 2157s 16p : 1204s
  - 32p : 1376s

- Average time step : 320 days
- For CASE4, Delta(t) is largest, so the needed linear iterations & precond. iterations is largest among these CASEs

- Scale / Broadcast = 76.80%
- Vector / Broadcast = 23.19%
- Neighbor Exchanges / Broadcast = 15.63%

It is higher than others because 36 faults increase the frequency of vector comm.

Behind these results ……

- **If the number of grids improves 4 times, then the elapsed time will improve about 9 times, where**
  - **the frequency of global reduction improves about 3 times**
  - **the number of Newton step improves about 2 times**
  - **the computation complexity of linear iteration improves about 4 times**

- For one million grid scale problem, every 100 time step of discrete parabolic equations will need
  - 300-500 newton steps,
  - 2400-3500 FGMRES steps,
  - 14K-20K linear iterations,
  - 0.5M-1M global communications,
- and it can be simulated in **one hour** on Beowulf clusters **(128 Nodes).**

- For five million grid scale problem, every 100 time step of discrete parabolic equations will need
  - 600-1000 newton steps,
  - 5000-8000 FGMRES steps,
  - 30K-40K linear iterations,
  - 1M-2M global communications,
- and it can be simulated in **three hours** on Beowulf clusters **(256 Nodes).**

- **Global Communication**
  - The ratio of scale reduction exceeds 80%
  - The ratio of vector reduction is less than 20%
  - The frequency of neighbor's exchanging messages is less than 50% of global communication
  - So, the communication process of scale reduction is a bottleneck of simulation
  - Algorithm & parallelism strategy should be adopted in order to satisfy minimization of scale reduction