

《大规模科学计算研究G1999032800》第三课题汇报之——

油藏数值模拟高效并行求解 技术与应用

曹建文, 赵国忠

中国科学院软件研究所

大庆油田有限责任公司勘探开发研究院

汇报提纲

- 研究背景与应用驱动
- 近几年的标志性成果
- 高效率的并行软件
- 高性能的并行求解器
- 实际数据的运行效果图示
- 在大庆油田的应用成果总结

研究背景

- 描述地下剩余原油的精细分布，需要大规模的油藏数值模拟技术
 - 石油勘探给出的油藏描述数据，网格点的数量一般在百万~千万量级
 - 油藏的开采，需要实现进行大规模的数值模拟
 - 百万网格点规模成为必需，千万网格点有需求
- 油藏数值模拟需要高效地求解
 - 黑油模型：3方程耦合的 PDE 方程组
 - 组分模型：5-10方程耦合的 PDE 方程组

-
- 本质：主要是（很）低雷诺数的 C F D 方程组，断层、井筒处的雷诺数稍高
 - 动量方程由 **Darcy's Law** 实验公式替代
 - 黑油模型分为：油、气、水三相
 - 特点：
 - （油）压力方程为非线性抛物型的，空间xyz上呈现椭圆型方程的特性
 - （油）饱和度方程主要为输运方程，呈现双曲特性
 - （气）饱和度方程为非线性的对流-扩散方程

应用驱动

- 百万网格点规模的油藏模拟问题难以高效求解
 - 全隐格式离散，Newton-Krylov-Schwarz求解
 - 计算密集、存储密集、对I/O有高要求、标量全局通讯频率高
- 实例说明：
 - 大庆百万网格数据，PowerChallenge上，98年，用某串行商业软件，模拟一次超过两个月
 - RDCPS 16处理器的机群上，99年，用并行的模拟程序，模拟一次需63.9小时，通讯时间40.2小时
 - 需要对求解算法和并行实现进行深入研究和改进

96年以来，油藏模拟专家们(李保树、马志远、陈文兰等)在不同场合多次指出，油藏数值模拟工作的主要瓶颈在于绝对计算能力太弱，以至于无法进行更多网格点规模的油藏模拟。他们提出了2000年前实现计算百万量级网格点数据的目标。大庆专家更是给出了心目中的理想计算能力：百万网格点的一次模拟，夕发朝至(16小时)时段内完成，达到朝发夕至(8小时)的能力，求之不得。

对绝对计算性能进行大幅度提升，消除油藏数值模拟工作中的最主要瓶颈，成为高性能数值计算领域的当务之急，也是大规模科学计算研究在能源领域应当承担的、具有挑战性的一个课题。

-
- 这是973第三课题组“预期目标相关内容”设立的依据
 - 以国内油田实际数据为背景：大庆或胜利
 - 利用本课题研究成果构造非线性问题并行求解器，进而应用到油藏模拟并行软件中
 - 针对百万量级的油田实际数据，使得绝对计算性能提高两个数量级
 - 其中，硬件对性能的贡献约一个数量级
 - 更重要的是，计算方法与软件实现对性能的贡献至少达到一个数量级
 - 完成核心程序，算出符合油田实际的结果

是否已经达到了预期的目标？

已经达到甚至超过了预定的指标

预期：两个数量级

结果：1600倍



百万网格点数据的
夕发朝至
目标已经达到

千万网格点数据
夕发朝至
的前景在望

这意味着：

绝对计算性能瓶颈已经消除，主要矛盾正在发生转移

百万网格点规模油藏模拟的实用化阶段即将来临

科学技术的进步超乎想象，
可以极大地提升应用部门的科学计算能力

超过了预定指标，体现在：

近几年的标志性成果……

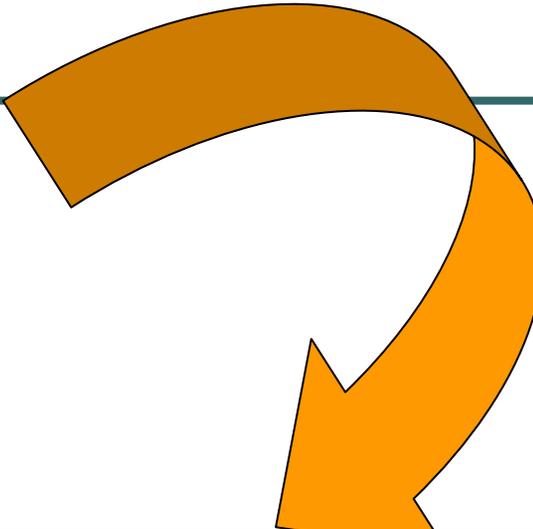
标志性成果：纵向体现

>>

TWO months

Seq. Commercial Code

MIPS R10K (1998)



0.94 hrs

Our Parallel Code,
LSSC-II 64CPU (2002)
Xeon 2G with Myrinet

绝对计算性能的提升 (以64 进程为例)

1998 → → → → 2002

CPU主频的提升对性能的贡献 : 5 times

粗技术/自动并行对性能贡献不超过 : 8 times

通讯极小、并行调优等手段约获得 : 8 times

(非)线性求解算法的改进至少提升了 : 5 times

绝对计算性能的提升 : $5 \times 8 \times 8 \times 5 = 1600$

计算机硬件的进步带来了

40倍的绝对性能提升

非线性问题求解的NKS算法，

自适应的预处理系统的研究，

并行策略的改进、并行实现与调优

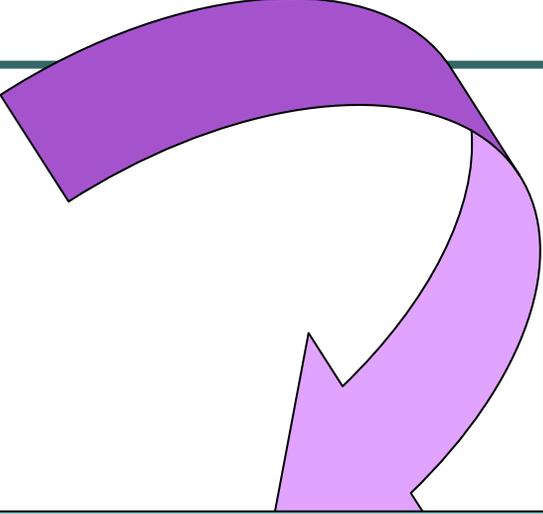
同样带来了

40倍的绝对性能提升

标志性成果：**横向体现**与Parallel VIP2K比较

> 70 hrs

Parallel VIP 2000 Code
Data from DaQing
MIPS 8CPU (2001)



6.8 hrs

Our ParaCode – PRIS v2.1
LSSC-II 8CPU (2002)
Xeon 2G with Myrinet

国外商业并行软件配国外高性能计算机

国内并行软件配PC Cluster机群系统

绝对计算性能比较：后者提高了**10.3倍**

其中，后者硬件的进步带来了**5.1倍**的绝对性能提升

后者的并行求解器，至少带来了**2倍**的绝对性能提升

更重要的是：后者的**Scalability**已经达到**128~256CPU**

我们这里强调的是绝对计算性能的提升，与973密切相关

综合优势 = 绝对计算性能优势 + 油藏功能优势
+ 软件工程设计优势 + 实时可视化显示

油藏功能优势的提升，是石油行业的主攻方向

软件工程设计优势的体现，是软件产业化的主要目标

实时可视化显示，是各部门密切合作的终极结果

绝对计算性能优势的提升，是大规模科学计算界的任务，
更大、更快、更高，是永远的追求

分析：高效率的并行软件

- 对全隐格式采用了Newton-Krylov-Schwarz 求解体系
 - 非线性问题求解器是强壮的 (robust)
 - 尽可能地避免频繁的全局通讯
- 基于MPI通讯库的并行策略
 - 标量全局通讯极小化
 - 完全意义上的数据并行, I/O采用局部文件读写方式
- 采用具有并行本性的算法重构井和断层程序

解释：高性能的并行求解器

- 精心组织的非线性问题求解器
 - 采用了Inexact Newton迭代算法
 - 带有回溯、阻尼功能
 - Inexact Newton难以收敛甚至发散时，采用BFGS构造初值
 - 实在不起作用，缩短时间步长，重新进行

-
- 经过细致调试的线性问题求解器
 - 并行的Preconditioned FGMRES
 - FGMRES(10+), 最大迭代步数限制: 88
 - 采用迭代技术构造的preconditioner:
 - ILU+GMRES(10)迭代, 允许重启动 3 次
 - 收敛标准是动态的, 与下列因素相关
 - 非线性求解迭代收敛状况的评价参数
 - 线性求解迭代的剩余向量模
 - 问题的规模

- 一套与众不同的预条件子系统

- 基于多角度斜交投影校正的思想构造的

- 采用了三种类型的斜交投影校正

- $A \rightarrow A_P$ 压力矩阵投影

- $A \rightarrow A_{\Omega}$ 区域分解

- $A \rightarrow A_{\text{Coarse}}$ 代数多重网格

- 该预条件子系统由8个预处理构件组成

- 该预条件子系统采用了自调节机制，共给出10个困难级。根据迭代计算的困难度，自动调节预条件子的级别，以达到收敛效果与计算代价之间的平衡。

预条件子系统的数学公式表示:

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

$$T_1 := T_{AddSchwarz} \cup T_{RelaxILU}$$

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

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

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

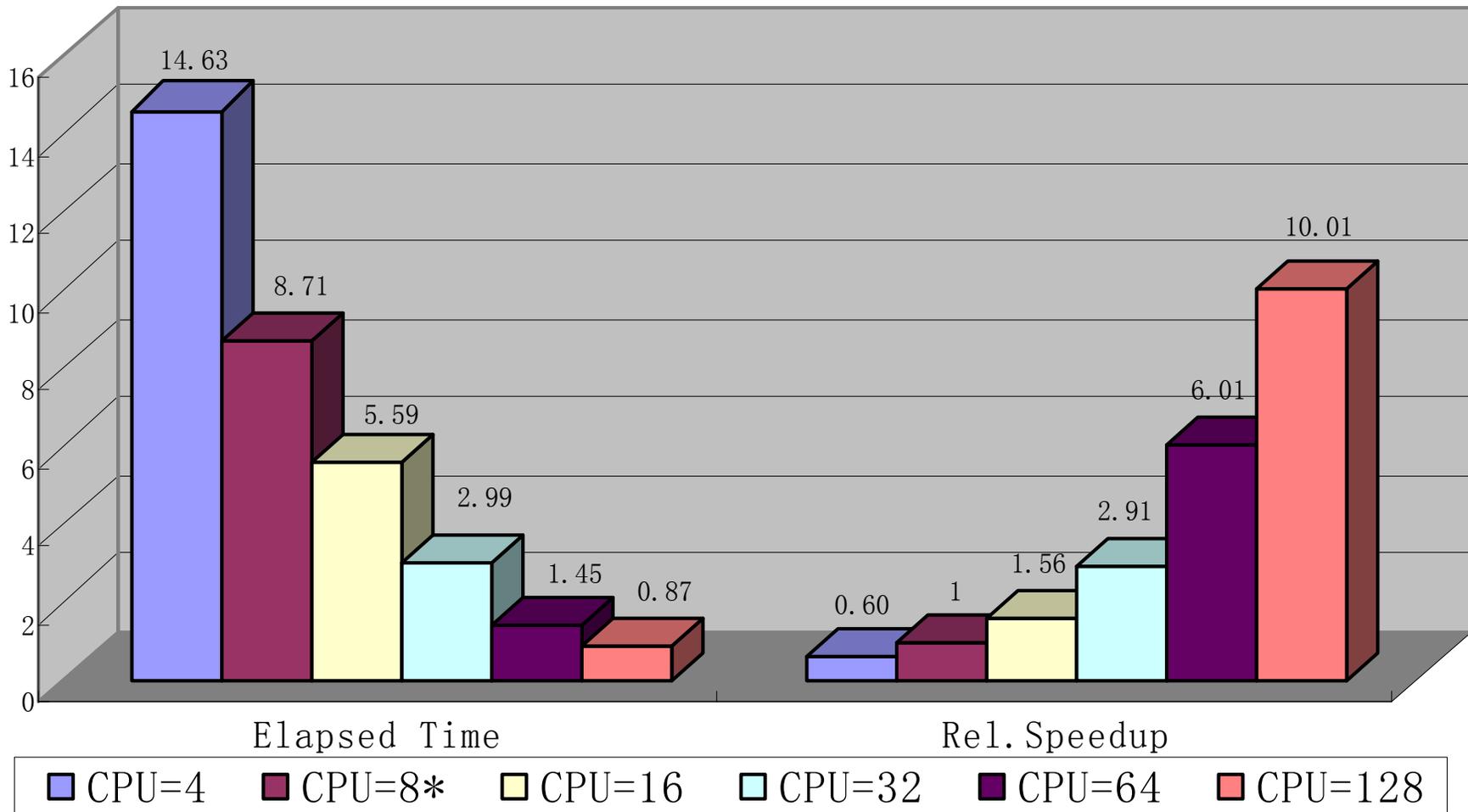
$$B = T_D \cup T_{Right}$$

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

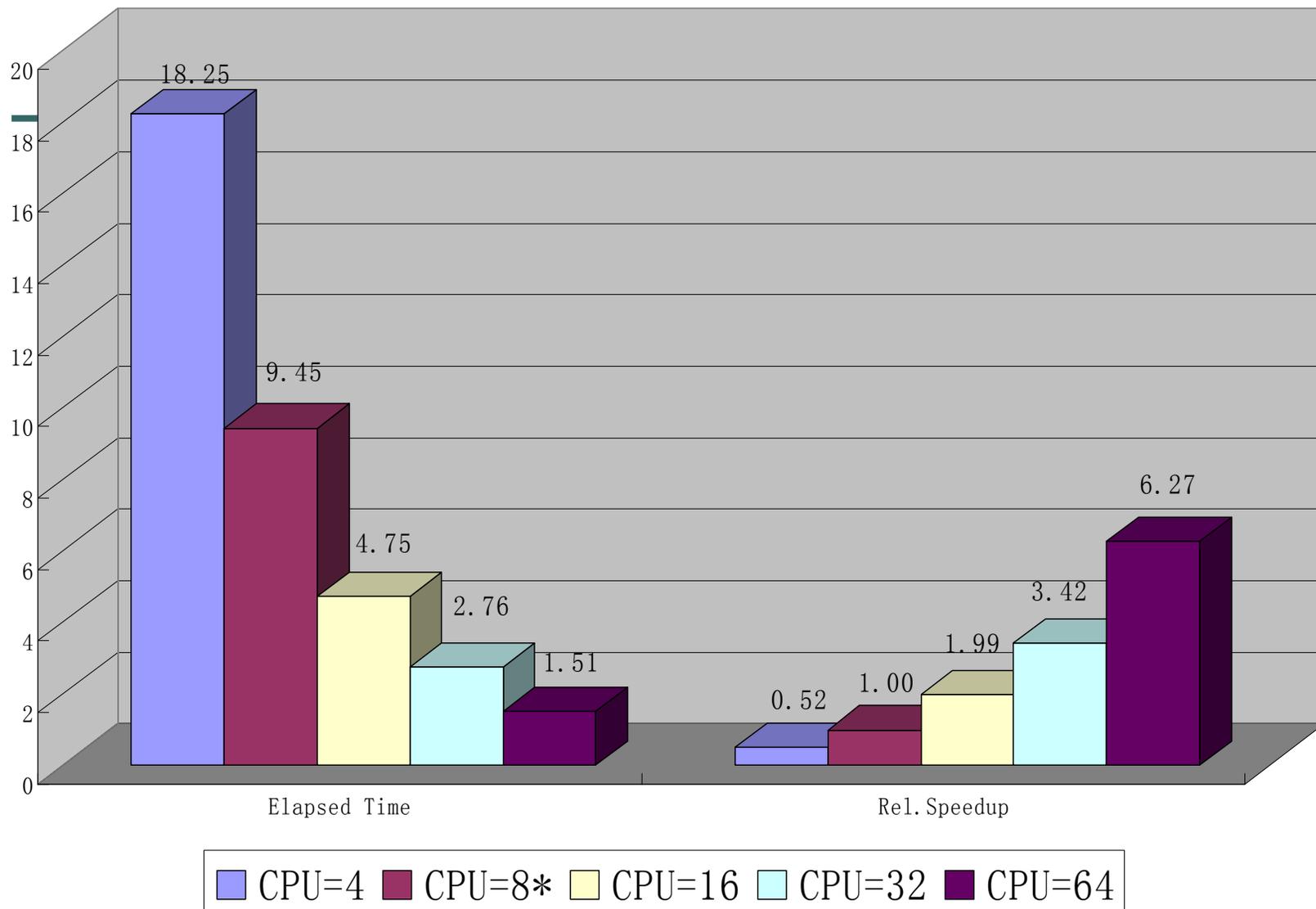
三套实际油田数据

- **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

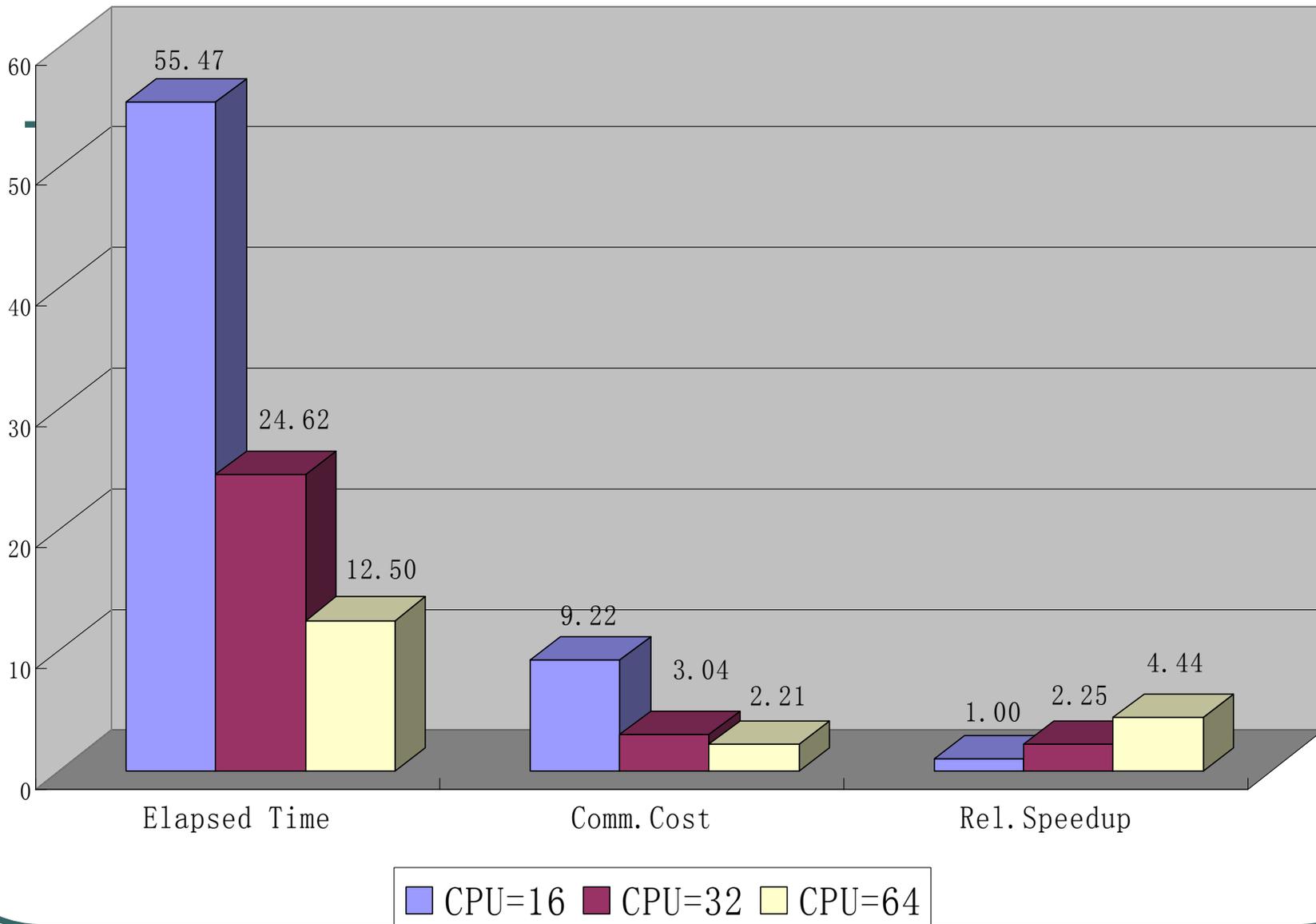
1159K数据的效果



1382K数据的效果



5529K数据的效果



在大庆油田的应用成果总结一

Cluster引入油田，实现了在用软件并行化

- 在先期试验基础上，构建大庆油田第一套用于并行油藏模拟的微机机群，以此为开发环境，应用通过高效并行解法器研究获得的求解思想，实现了油田在用串行黑油模拟软件的并行化。
- 应用结果显示，实现的全套机群并行模拟技术的模拟规模已超过百万节点，模拟速度也高于引进的Origin2000 + 的PVIP组成的成套并行模拟技术，对于百万节点以上规模的模拟，速度至少高出40%。

推进了油田大规模油藏模拟的应用，为剩余油的精细预测和高效挖潜提供了技术保障，潜在经济效益巨大

-
- 为实现油藏模拟与精细地质研究的技术衔接，开发了沉积相带图的数字化技术和实用相控地质建模软件

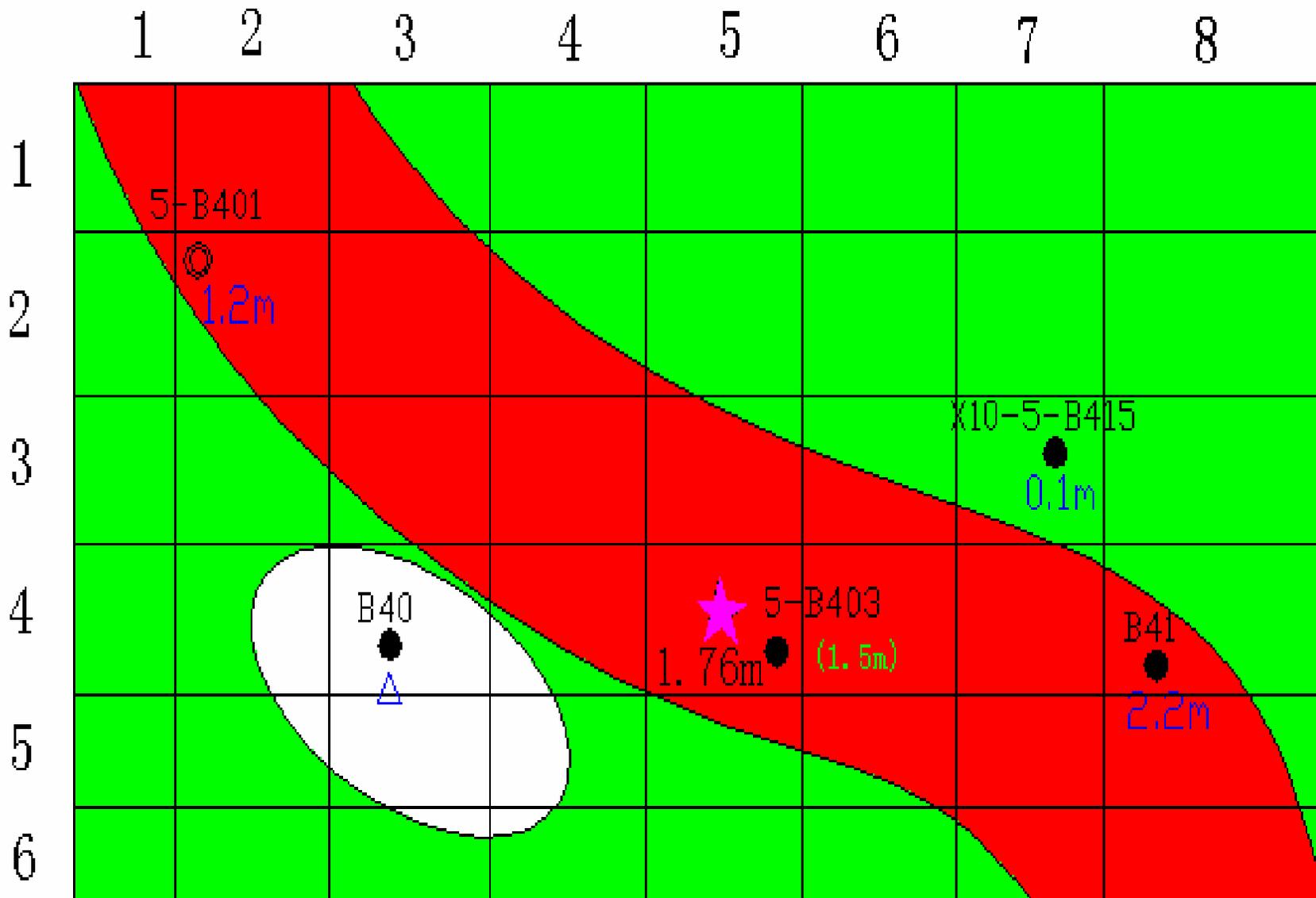


图1相控建模结果示意图

-
- 实现了与机群并行模拟技术的综合集成，形成了大庆油田高含水后期多学科油藏研究的工作环境和平台

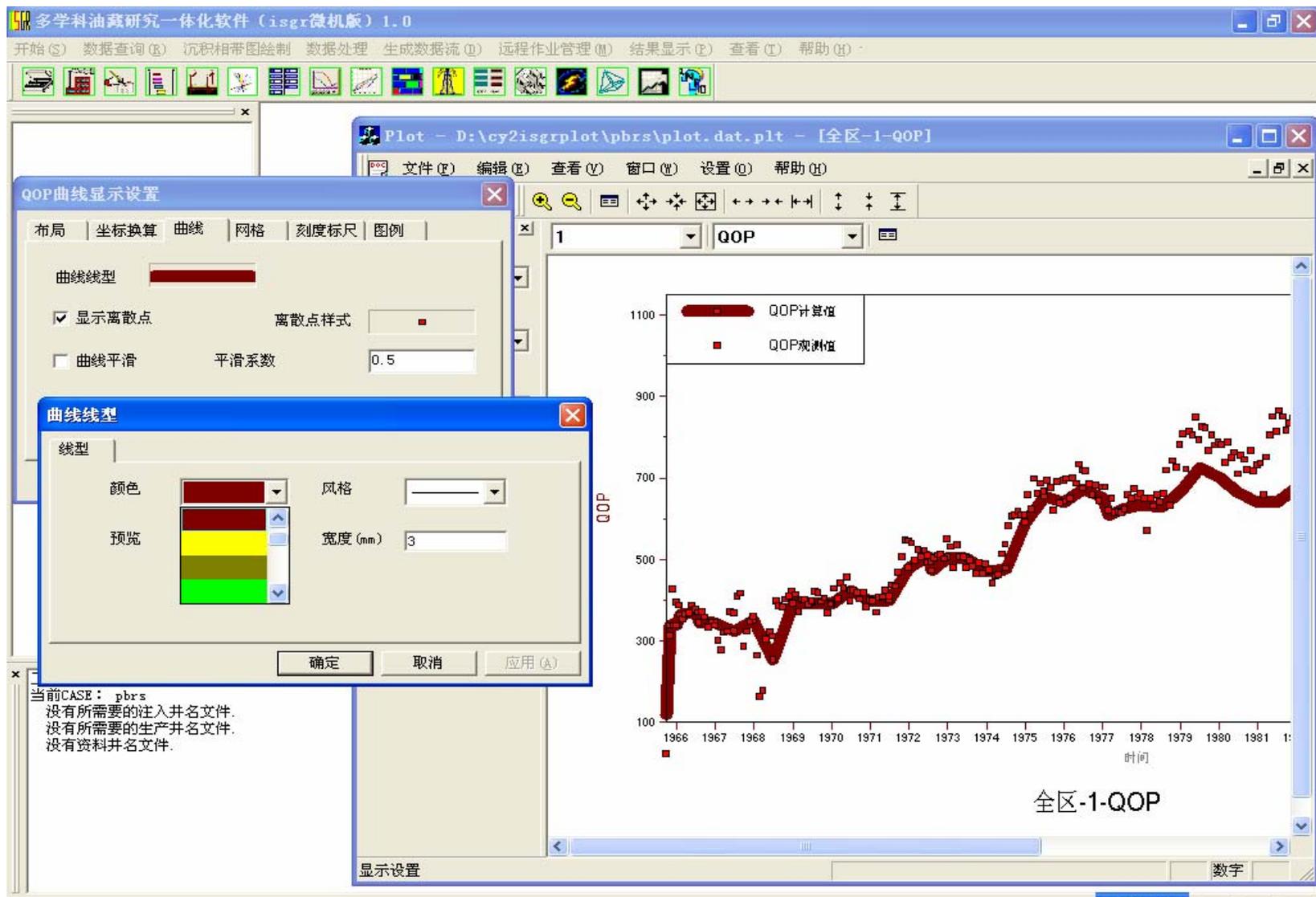


图2 建模数模一体化软件ISGR示意

-
- 大庆油田研究院和采油厂构建了8套油藏模拟专用微机机群，并安装了并行化黑油油藏模拟器**PBRS2.1**

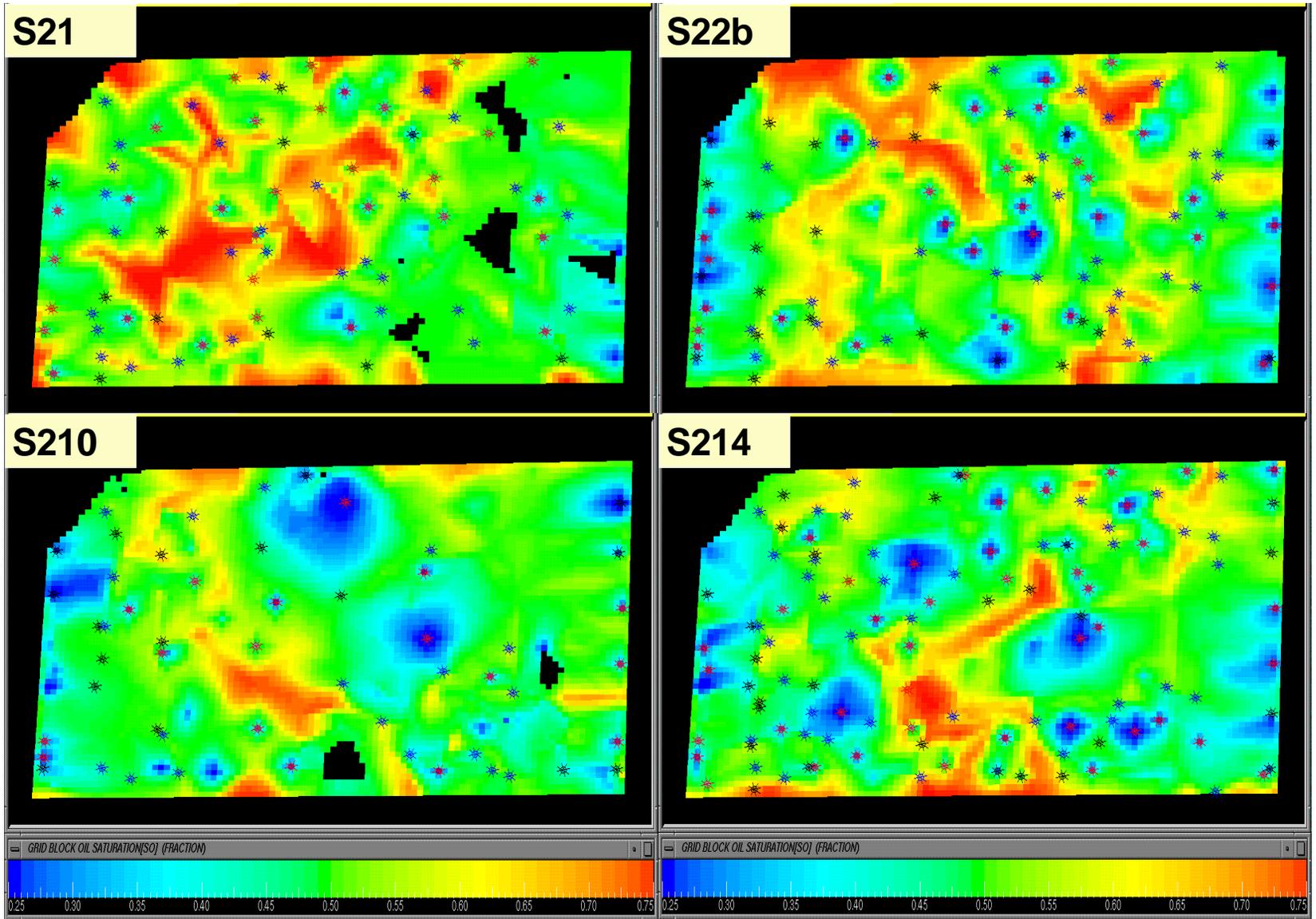


图3 大庆油田萨区剩余油分布示例

以油田大规模模拟应用为主要技术的高含水后期多学科油藏研究技术,在大庆长垣推广应用几年来,研究的总面积达到139.72 km²、总井数为4846口、总地质储量为6.3×10⁸t。

通过这些研究，深化了对储层地质和井况的认识，细化量化了剩余油分布，实现了区块综合调整方案的优化。如ZD2-15井通过补孔挖掘剩余油，设计优选了SII1和SII21两个补孔层位，实施后日产液增加19t，日产油增加11t，含水下降了28.1个百分点，与预测指标符合很好，取得了较好的增油降水效果。

目前, 先期投入研究的油藏区块已经步入生产动态跟踪模拟, 提高了工作效率, 降低了操作成本。油田大规模油藏数值模拟, 已成为剩余油精细预测和高效挖潜的关键技术之一, 开始发挥重要的保障作用。综合研究认为, 基于大规模油藏数值模拟技术的应用来预测和挖潜剩余油, 可以进一步提高油田采收率0.5~1.0个百分点。那么, 对于已经开展研究的这些油藏区块, 按提高采收率0.5%计算, 将多产原油 3.14×10^6 t, 可见潜在的经济效益是巨大的。

The End !

