

大数据时代的石油地震勘探系统与软件平台

谢玮¹, 刘斌², 刘鑫³, 李玉¹, 翟焱森¹

1. 中国地质大学(北京)地球物理与信息技术学院, 北京 100083
2. 中国石化石油工程地球物理有限公司胜利分公司, 东营 257086
3. 中国石化华北油气分公司采油气工程服务中心, 咸阳 712000

摘要 石油勘探开发精度的不断提高,促进了低频可控震源、宽频带、宽方位、高密度和高效采集技术的推广应用,石油地震勘探已进入了大数据时代,对质量监控、数据处理、数据安全存储和管理带来了新的挑战。本文分析了石油地震勘探大数据的特点,阐述了中国石化基于Hadoop分布式大数据处理系统研发的 π -Frame地震数据解释软件平台基本构架,举例说明了该平台在石油地震勘探大数据中的应用,对其发展前景进行展望。

关键词 石油地震勘探;大数据;Hadoop系统;地震数据解释;软件平台

随着全球数据总量的持续增长和急速膨胀,人类已经步入大数据时代^[1]。对于大数据,麦肯锡全球研究院给出的定义是:一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合,具有“4V”特征,即海量的数据规模(volume)、快速数据流转(velocity)、多样数据类型(variety)和较低价值密度(veracity)的^[2-3]。自大数据概念提出以来,掀起了信息界的一场数据革命,大数据逐渐成为所有行业中最受关注的技术话题之一,石油工业也不例外。几十年来,随着对地下地质构造了解的不断深入、具有成本效益的钻井作业的实施、高效油藏管理的实现,石油工业产生了大规模的勘探开发数据。然而,据麦肯锡2015年度报告,石油工业仅从其创造的大数据中提取了不足1%的价值^[4]。

如何从大数据中高效地获取信息,有效地深加工并最终得到有用数据是能源企业涉足大数据的目标^[5]。随着全球石油储备的逐步减少,石油勘探难度日益增大,信息化的成熟度已经成为影响行业增长幅度的首要因素^[6]。精准、快速的地质勘测成为世界能源巨头们倚重的核心竞争力之一,其中高性能计算技术和大数据技术的应用是关键因素。石油勘探中主要利用地球物理方法推断地下岩层的性质和形态信息。地球物理方法是使用现代物理方法进行地质勘探的方法,包括重力法、磁法、电法、地震波法、放射性和测井6种勘探方法,其中以地震波法即地震勘探最为重要^[7]。首先接

收由震源激发的经由地下岩层反射和折射的地震波海量数据,包括二维数据(TB级)和三维数据(PB级);然后进行大规模的计算和成像处理,以方便专家对数据进行解释,为钻井定位提供参考,从而实现最佳勘探效益。由于这些海量数据的特殊性和复杂性,对高性能计算提出了非常苛刻的要求。

Hadoop作为一个能够对大数据进行分布式并行计算的软件框架,凭借其高可靠性、高扩展性、高效性和高容错性的优点,已被许多国际知名互联网和石油公司所采用^[8]。其中,美国雪佛龙公司在其储层分析和识别的过程中使用Hadoop技术来进行地震数据分析,从而寻找能够更精确的使用更多数据的方法;Cloudera公司将Hadoop大数据技术引入Seismic UNIX地震软件平台,构建了能够存储和处理地震数据的Seismic Hadoop集群;PointCross公司利用NoSQL和Hadoop技术建立了基于数据云的大规模地震数据储存库和服务端(SDSR),使客户能够随时随地在云上存储和管理地震大数据。荷兰皇家壳牌公司在开源Hadoop框架的基础上,利用私有云计算技术对勘探开发数据进行实时采集和分析,从而提高油井开采的成功率。

Hadoop技术在石油勘探开发领域的推广应用也为地震勘探大数据的计算方法提供了新的解决思路。为了应对大数据时代全球日益复杂的地震勘探问题,中国石化集团公司于2010年启动研发了 π -Frame地震数据解释软件平台(简称 π -Frame平台)。该平台是国际地球物理业界第一个

收稿日期:2017-02-17;修回日期:2017-04-27

基金项目:国家科技重大专项(2016ZX05003-003)

作者简介:谢玮,博士研究生,研究方向为地震资料处理与解释,电子信箱:xw2008xwcs@163.com

引用格式:谢玮,刘斌,刘鑫,等. 大数据时代的石油地震勘探系统与软件平台[J]. 科技导报, 2017, 35(15): 57-62; doi: 10.3981/j.issn.1000-7857.2017.15.008

基于 Hadoop 大数据技术体系构建的大型地震勘探软件平台。本文介绍了石油地震勘探大数据的特点, π -Frame 平台的基本构架, 及 π -Frame 平台在石油地震勘探大数据中的应用, 并对其发展前景进行展望。

1 石油地球物理勘探大数据

在石油领域, 随着勘探开发的不断深入以及难度的不断增大, 对目标成像精度和油藏表征精度的要求越来越高, 石油地震勘探已朝着“两宽一高”(宽方位、宽频带、高密度)、全波场、多方法方向发展^[9-10], 使大道数、高效采集、叠前偏移、属性分析、叠前/叠后反演等技术的不断发展, 导致地震勘探数据量与计算量呈指数式增长。随着海量并行存储技术^[11]、CPU+GPU 异构协同计算技术^[12]、集成众核(many integrated core, MIC)^[13]架构技术及相关大数据处理解释软件的快速发展, 地震勘探已进入了大数据时代^[14]。

地震勘探由地震资料采集、处理和分析解释3部分组成, 地震资料的内涵是数据, 通过数据反映地质结构、地层变化, 通过数据为石油勘探开发提供依据。近年来, 地震勘探装备的不断进步, 实现了高效采集施工方法的大跨越。中国石油东方地球物理公司(BGP)利用低频可控震源和“两宽一高”采集技术在国内外地震勘探项目进行了高效作业, 采集覆盖次数由几十次增至几千次, 每炮数据量由 10 MB 增至 250 MB, 每天产生的数据量达到了 TB 级别。采用“两宽一高”技术采集的大数据能够提供更多的有用信息, 有利于地震资料成像, 而这些大数据的处理解释需要采用更先进的软件技术和信息挖掘技术。需要采用以叠前三维去噪、方位各向异性校正、共炮检距矢量片(OVT)域数据处理和基于 OVT 道集裂缝检测为核心的不同技术和流程。这些处理过程较为复杂, 在给软硬件处理能力带来巨大压力的同时, 也给数据质控、数据安全存储和管理带来新的挑战。

综上所述, 石油勘探过程中所产生的大数据具有如下“4C”特征: 1) 数据容量大(content), 每天会产生 TB 级别的生产数据, 而且在勘探的过程中还会产生大量的中间数据; 2) 数据采集来源单一(collection), 地震资料数据是由人工模拟地震波激发, 利用定点采集仪器接收和采集得到的, 数据来源和数据格式都比较单一; 3) 数据计算量大(calculation), 以 BGP 为例, 54 TB 的原始数据通过 4000 个 CPU 的计算集群处理, 需要 50 多天时间; 5) 数据处理流程复杂(complexity), 地震资料处理过程涉及复杂的处理算法、频繁的输入/输出和数据库存储过程的调用, 操作复杂。

2 π -Frame 地震数据解释一体化软件平台

面对石油勘探开发日趋复杂的需求与地震勘探采集技术进步带来的挑战, 油气地震勘探呈现出采集处理解释一体化的发展趋势, 以及海量地震数据管理和超大规模并行计算问题对石油勘探软件架构提出的新挑战, 从而对地球物理勘探软件平台提出了更新换代的要求。基于这种强大的需求

与信息技术发展带来的机遇, 中国石化集团公司组织研发了面向未来业务需求、开放的地震数据解释软件平台, 即 π -Frame 平台。

2.1 π -Frame 平台基本构架

作为新一代地震勘探软件平台, π -Frame 基于先进的 Hadoop 大数据技术平台架构, 采用面向海量地震数据高效管理、超大规模并行计算和地震处理解释一体化三大设计理念, 实现了海量数据高效管理与快速存取、异构并行计算框架、大规模并行计算作业及工作流引擎、二维与三维图形交互应用框架等功能, 支持勘探开发多学科数据的一体化管理与应用, 为地球物理应用软件提供了先进的集成开发环境与生产应用环境。

π -Frame 平台由开发平台、应用系统和生态系统三部分组成(图 1)。其中开发平台提供了功能丰富、使用方便、开放集成的软件开发包(SDK)和集成开发环境(IDE), 支持大型地震软件的开发和方法技术的软件集成; 应用系统包括各类 UI 插件、批处理模块、地震处理解释系统引擎及 Hadoop 分布式大数据计算系统, 为海量地震数据管理、大规模并行计算、大规模项目作业提供了强有力的支撑; 生态系统以价值分享与共赢发展为理念, 形成开放的地震专业软件社区, 实现集技术支持、技术交流、成果分享、需求发布和技术交易于一体的用户互动。

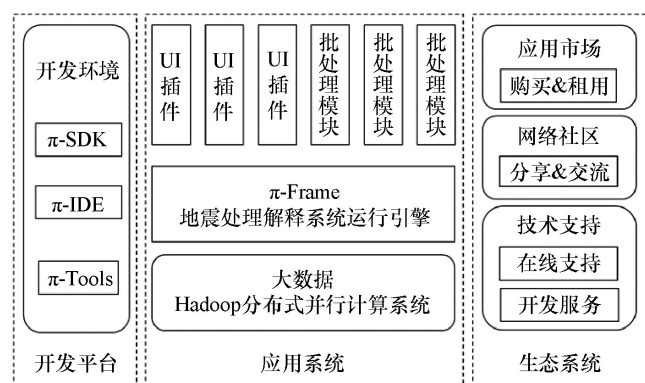


图 1 π -Frame 平台的三大组件

Fig. 1 Three components of π -Frame platform

2.2 Hadoop 大数据技术体系

Hadoop 系统框架由 Apache 软件基金会基于 Java 开发, 是当下主流的大数据标准开源软件^[15]。Hadoop 架构为应用程序提供了一组稳定可靠的 API 接口, 从而实现具有高可靠性和良好扩展性的分布式系统, 运行于由大量廉价设备组成的集群系统上, 能够进行大数据的分布式并行计算。Hadoop 包括 HDFS(Hadoop distributed file system) 和 MapReduce 两个核心构件, 以及 HBase、Hive 和 ZooKeeper 等其他构件, 如图 2 所示。

HDFS 是一种具有高可靠性的分布式文件系统, 能够存储 PB 级别的大规模数据, 该技术的应用使得 π -Frame 平台



图2 Hadoop 分布式系统架构示意

Fig. 2 Architecture diagram of Hadoop distributed system

的数据存取访问性能、稳定性及容错能力达到国际领先水平。HDFS 采用主从模式搭建集群系统,通常由一个名字节点(namenode)和多个数据节点(datanode)构成,如图3所示。其中,namenode 节点通常由高性能的服务器充当,主要负责文件命名空间和调节客户端文件访问服务的管理工作; datanode 节点在集群环境中通常由廉价PC构成,主要负责节点数据的存储和处理,以及响应客户端的读写请求。HDFS 的工作原理是将单个文件分割后的多个 block(默认值为 64 MB)冗余存储在多个 datanode 节点上,并通过 namenode 节点来保存整个文件系统的目录信息、文件信息以及相应的分块信息。这种数据分块以及数据块的备份和分散存储策略,保证了 HDFS 的高度容错性,降低了集群系统对硬件的要求。

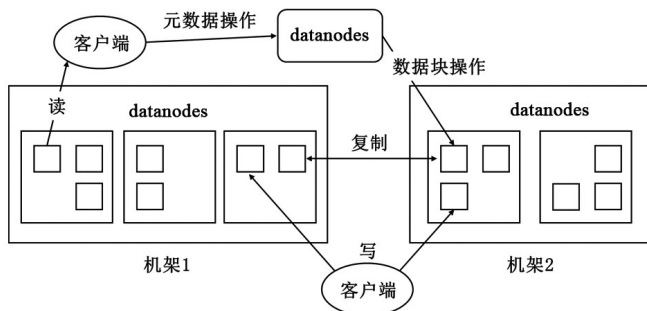


图3 HDFS 分布式文件系统结构示意

Fig. 3 Structure diagram of Hadoop distributed file system

MapReduce 是一种以 HDFS 为基础的大数据分布式并行计算模型,通过对大数据进行分割处理,再将分割后的小数据集交给不同的计算节点进行处理,从而实现大数据的并行计算。其计算流程如图4所示,具体过程为:1) 对输入目录中的大数据文件按照相同的大小进行分割,得到一组小数据块,并将这些数据块看做是键值对(key-value pair); 2) Map 函数接收一个键值对,映射成一组新的中间键值对; 3) 将 Map 函数产生的中间键值对里键相同的值传递给一个 Reduce 函数; 4) Reduce 函数接收一个键,以及对应的一组值,对这组值进行归约后产生一组规模更小的值,并将结果写入到输出目录的文件中。 π -Frame 平台利用 MapReduce 技术实现了地震数据成像处理和所有批处理模块的全并行化,具备了超万核级大规模并行计算部署与生产应用能力,具有良好的并行效率和可扩展性,显著提高了地震数据处理的效率和能力。

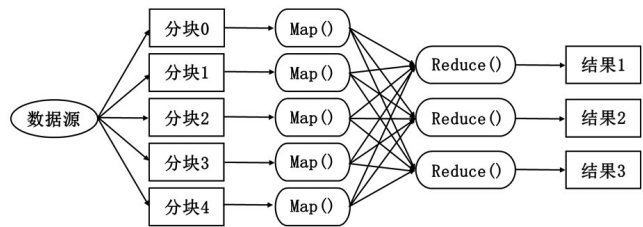


图4 MapReduce 分布式并行计算模型结构示意图

Fig. 4 Structure diagram of MapReduce distributed parallel computing model

3 π -Frame 平台应用

3.1 Kirchhoff 叠前时间偏移

作为常规地震数据处理流程的叠前时间偏移,由于其具有较高的计算效率、很强的适应能力、适应野外不规则观测系统等优点,特别是对速度模型精度要求较低,从而成为广泛应用的偏移方法^[6]。随着处理数据规模的不断增长,面对三维海量地震数据的处理,基于常规 MPI 并行策略的叠前时间偏移已无法满足地震资料快速处理的需求。为此,亢永敢等^[7]通过采用输入数据的两次分块及成像空间的分块计算等策略,研究了 π -Frame 平台基于 Hadoop 框架的 Kirchhoff 叠前时间偏移并行算法,并在 67 节点、2×10 核 CPU/节点、64 GB 内存/节点的刀片集群上对比其与 MPI 框架 Kirchhoff 叠前时间偏移并行算法的计算效率。

首先,对比分析 Hadoop 与 MPI 框架 Kirchhoff 叠前时间偏移对海量数据的计算效率。选取某工区 748 km² 的实际资料为测试数据,计算输入数据量为 1.9 TB,在两种框架偏移参数相同的前提下,输出数据量为 65 GB。其中,Hadoop 框架的偏移程序耗时 5.68 h,而 MPI 框架的偏移程序耗时 6.09 h。从中可以看出,Hadoop 框架偏移程序的计算效率比 MPI 框架高 6.7%,这得益于 HDFS 分布式文件系统的高 I/O 吞吐能力。

然后,测试 Hadoop 框架 Kirchhoff 叠前时间偏移的稳定性。对 1.9 TB 的实际数据进行全偏移计算,共耗时 70.58 h,计算输出数据量为 1.15 TB。对每个 Map 运行时间进行统计,如图5所示,图中有 12 个 Map 运行时间异常。对系统运行日志进行分析,第 27 号节点的网络异常中断导致运行在该节点上的 Map 作业不能正常完成,为此 Hadoop 框架的容错机制进行了异常 Map 作业的迁移,从而保证了作业的顺利完成。从中可以看出 Hadoop 框架能够在节点出现异常的情况下,保证作业正确完成,验证了 Hadoop 系统的高容错性和稳定性。

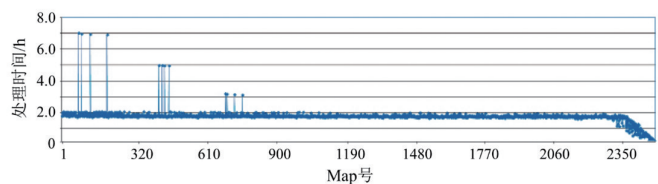


图5 Map 运行时间统计

Fig. 5 Statistical chart of Map run time

最后,验证Hadoop框架偏移程序计算结果的精确性,对测试数据某一条测线的Hadoop和MPI框架程序偏移结果成像剖面进行对比,如图6所示,其中图6(c)为两个偏移结果相减得到

的误差剖面,可以看出,Hadoop框架程序的计算结果与MPI框架程序高度吻合,二者的均方误差为 1.3×10^{-4} 。由此可知,Hadoop框架程序能够满足Kirchhoff叠前时间偏移的精度要求。

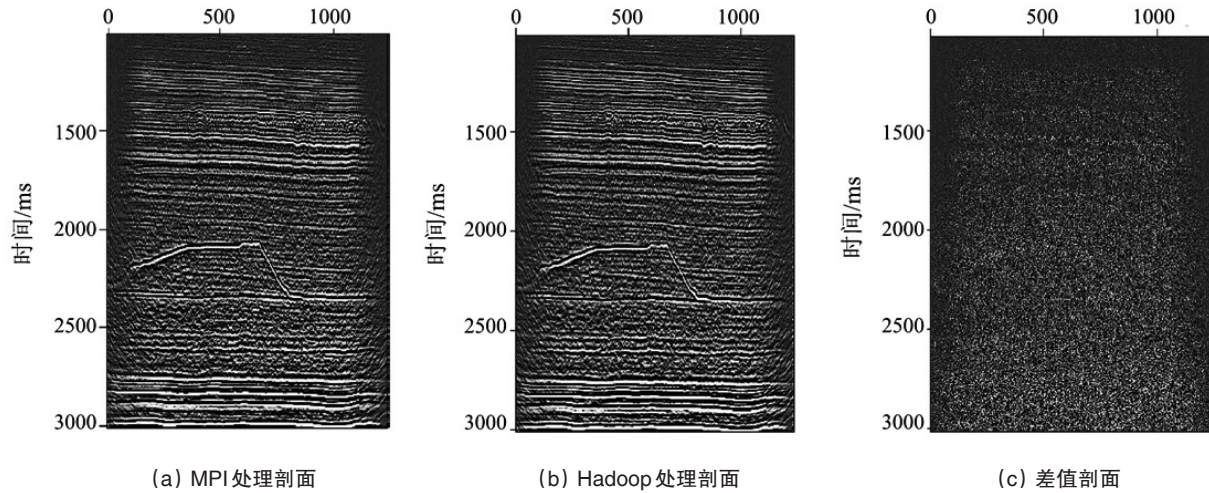


图6 不同并行程序偏移结果对比

Fig. 6 Comparison of results computed by different parallel migration program

3.2 批处理及多任务性能测试

选取“道头更新”、“面波压制”及“地表一致性振幅补偿”等6个主要模块对 π -Frame平台进行实际资料批处理性能测试^[18]。测试环境为57个节点的刀片集群,每个节点包含2个10核20线程的CPU,64 GB内存,各节点之间通过10 Gbit/s网络连接;实际测试数据为18 TB。测试结果如表1所示,6个主要模块的批处理速度均值达到了8.79 TB/h。

选取“道头更新”、“一致性反褶积”和“选排”3个模块对

π -Frame平台进行实际资料多任务性能测试^[18]。测试环境为472个节点的刀片集群,每个节点包含2个10核20线程的CPU,64 GB内存,各节点之间通过10 Gbit/s网络连接;测试数据大小约为40 TB。测试结果如表2所示,其中“一致性反褶积”和“选排”2个模块的多任务处理速度约为11.72 TB/h,“道头更新”的多任务处理速度达到了76.17 TB/h。综上所述, π -Frame平台能够为多任务、多模块同时进行的大数据批处理提供高速、全面的实际生产应用环境。

表1 π -Frame平台批处理性能测试

Table 1 Performance test of batch modules on π -Frame

批处理模块	数据量/TB	处理时间/h	处理速度/(TB·h ⁻¹)
道头更新	18	1.52	11.48
面波压制	18	1.38	12.66
地表一致性振幅补偿	18	2.60	6.74
最大能量剩余静校正	18	2.48	7.03
选排	18	4.56	3.81
初至自动拾取	18	1.24	14.12

表2 π -Frame平台多任务性能测试

Table 2 Multi-job performance test of batch modules on π -Frame

批处理模块名称	数据量/TB	处理时间/h	处理速度/(TB·h ⁻¹)	任务数
道头更新	42.0	0.53	79.72	14
一致性反褶积	47.5	4.58	10.37	15
选排	42.0	2.47	16.97	14

4 结论

石油地震勘探已经进入了大数据时代,总结了石油地震勘探大数据的数据容量大(content)、数据采集来源单一(collection)、数据计算量大(calculation)和数据处理流程复杂(complexity)的“4C”特点,并介绍了基于Hadoop大数据技术体系的“ π -Frame地震数据解释一体化软件平台”基本构架。对比分析了 π -Frame平台Hadoop框架与MPI框架两种Kirchhoff叠前时间偏移并行算法的计算效率,Hadoop版本与MPI版本相比获得了6.7%的效率提升。进行了 π -Frame平台对“道头更新”、“一致性反褶积”和“选排”批处理模块的多任务性能测试,结果表明 π -Frame平台能够为多任务、多模块同时进行的大数据批处理提供全面的实际生产应用环境。

地震勘探流程除了资料采集、处理,还包括解释和反演。其中在解释环节,需要通过统计分析、多属性交汇、多属性融合等技术从地震属性大数据中提取出能够反映储层含油气性的特征参数(如振幅、频率、相位、能量、波形以及这些参数的梯度变化)^[19];反演环节的大数据主要体现在多次正演与迭代拟合,特别地如全波形反演方法对内存的需求能够达到几百GB甚至TB级^[20],对计算资源的要求非常之高,制约了其发展。因此基于Hadoop框架的 π -Frame平台凭借其高稳定性、高容错性以及海量数据的处理能力,在地震勘探大数据的处理、解释和反演3个环节都有着广泛的应用前景。

石油地震勘探技术的综合研究成果蕴含在大数据之中。大数据对石油地震勘探行业既是挑战又是机遇,驾驭大数据,建立大数据质控体系,强化 π -Frame平台的高效计算能力建设和数据标准化建设,从中挖掘有效信息,提高决策的准确性,从而更有力地掌控资源和市场。

参考文献(References)

- [1] 李金诺. 浅谈石油行业大数据的发展趋势[J]. 价值工程, 2013, 32(29): 172-174.
Li Jinnuo. The development tendency of big data in petroleum industry [J]. Value Engineering, 2013, 32(29): 172-174.
- [2] Biesdorf S, Court D, Willmott P. Big data: What's your plan? [J]. McKinsey Quarterly, 2013(2): 40-51.
- [3] 方欣友. 试析油气勘探大数据化面临的挑战与机遇[J]. 环球人文地理, 2015(10): 102.
Fang Xinyou. Challenges and opportunities for big data in oil and gas exploration [J]. Cultural Geography, 2015(10): 102.
- [4] Manyika J, Chui M, Bisson P, et al. The Internet of things: Mapping the value beyond the hype[R/OL]. [2017-04-27]. http://www.ifuture.org/documentacion/Unlocking_the_potential_of_the_Internet_of_Things_Full_report.pdf.
- [5] 姜睿. 大油气需要大数据[J]. 中国石油石化, 2012(24): 29.
Jiang Rui. Big oil and gas need big data[J]. China Petrochem, 2012(24): 29.
- [6] 张建. 大数据点金大油气[J]. 中国石油企业, 2014(4): 69-71.
Zhang Jian. Big data highlights the big oil and gas[J]. China Petroleum Enterprise, 2014(4): 69-71.
- [7] 曹建鹏, 陈松涛, 李志萍. 云计算在石油勘探领域的应用[J]. 信息系
统工程, 2012(6): 89-89.
Cao Jianpeng, Chen Songtao, Li Zhiping. The application of Cloud computing in petroleum exploration[J]. China Chief Information Officer News, 2012(6): 89-89.
- [8] 王凯, 曹建成, 王乃生, 等. Hadoop支持下的地理信息大数据处理技术初探[J]. 测绘通报, 2015(10): 114-117.
Wang Kai, Cao Jiancheng, Wang Naisheng, et al. Research on GIS big data computing technologies based on Hadoop[J]. Bulletin of Surveying and Mapping, 2015(10): 114-117.
- [9] 谭晔, 马永, 刘长军. 大数据:海量信息中“掘金”[N]. 中国石油报, 2014-05-13(03).
Tan Ye, Ma Yong, Liu Changjun. Big data: panning for gold in mass information[N]. China Petroleum Daily, 2014-05-13(03).
- [10] 撒利明, 甘利灯, 黄旭日, 等. 中国石油集团油藏地球物理技术现状与发展方向[J]. 石油地球物理勘探, 2014, 49(3): 611-626.
Sa Liming, Gan Lideng, Huang Xuri, et al. The status and development direction of reservoir geophysics in CNPC[J]. Oil Geophysical Prospecting, 2014, 49(3): 611-626.
- [11] 李存琛. 海量数据分布式存储技术的研究与应用[D]. 北京: 北京邮电大学, 2012.
Li Cunchen. Research on distributed storage technology based on mass data[D]. Beijing: Beijing University of Posts and Telecommunications, 2012.
- [12] 马召贵, 赵改善, 武港山, 等. 基于CPU+GPU联合计算真地表叠前时间偏移实用化研究[J]. 石油地球物理勘探, 2014, 49(3): 530-537.
Ma Zhaogui, Zhao Gaishan, Wu Gangshan, et al. Kirchhoff pre-stack time migration of complex surface data on CPU + GPU[J]. Oil Geophysical Prospecting, 2014, 49(3): 530-537.
- [13] 车永刚, 张理论, 王勇献, 等. Intel多核与集成众核上CFD程序的OpenMP性能分析[J]. 计算机科学与探索, 2015, 9(10): 1153-1162.
Che Yonggang, Zhang Lilun, Wang Yongxian, et al. OpenMP performance analysis of CFD application on Intel multicore and manycore architectures [J]. Journal of frontiers of computer science and technology, 2015, 9(10): 1153-1162.
- [14] 王喜双, 赵邦六, 董世泰, 等. 油气工业地震勘探大数据面临的挑战及对策[J]. 中国石油勘探, 2014, 19(4): 43-47.
Wang Xishuang, Zhao Bangliu, Dong Shitai, et al. Challenges and strategies for large seismic exploration data of oil and gas industry [J]. China petroleum exploration, 2014, 19(4): 43-47.
- [15] 程学旗, 靳小龙, 杨婧, 等. 大数据技术进展与发展趋势[J]. 科技导报, 2016, 34(14): 49-59.
Cheng Xueqi, Jin Xiaolong, Yang Jing, et al. Technological progress and trends of big data[J]. Science & Technology Review, 2016, 34(14): 49-59.
- [16] 刘国峰, 刘洪, 王秀闽, 等. Kirchhoff积分时间偏移的两种走时计算及并行算法[J]. 地球物理学进展, 2009, 24(1): 131-136.
Liu Guofeng, Liu Hong, Wang Xiumin, et al. Two kinds of traveling time computation and parallel computing methods of Kirchhoff migration[J]. Progress in Geophysics, 2009, 24(1): 131-136.
- [17] 亢永敢, 赵改善, 魏嘉, 等. 基于Hadoop的Kirchhoff叠前时间偏移并行算法[J]. 石油地球物理勘探, 2015, 50(6): 1213-1218.
Kang Yonggan, Zhao Gaishan, Wei Jia, et al. Parallel algorithms of Kirchhoff prestack time migration based on Hadoop[J]. Oil Geophysical Prospecting, 2015, 50(6): 1213-1218.
- [18] Zhao G. Keynote: Seismic processing and interpretation platform in the big data era[C]. Workshop: High Performance Computing, Beijing,

- China, November 14–16, 2016.
- [19] 李婷婷, 王钊, 马世忠, 等. 地震属性融合方法综述[J]. 地球物理学进展, 2015(1): 378–385.
- Li Tingting, Wang Zhao, Ma Shizhong, et al. Summary of seismic attributes fusion method[J]. Progress in Geophysics, 2015(1): 378–385.
- [20] 尹力, 方伍宝, 李振春, 等. 3D声波全波形反演计算策略研究[C]// 2014年中国地球科学联合学术年会论文集. 北京: 中国地球物理学会, 2014.
- Yin Li, Fang Wubao, Li Zhenchun, et al. Research on computing strategies of 3D acoustic full waveform inversion[C]// Proceedings of the joint annual meeting of the Earth Sciences in China in 2014. Beijing: Chinese Geophysical Society, 2014.

System and software platform of oil seismic exploration in big data era

XIE Wei¹, LIU Bin², LIU Xin³, LI Yu¹, ZHAI Yansen¹

1. School of Geophysics and Information Technology, China University of Geosciences (Beijing), Beijing 100083, China
2. Geophysical Co., Ltd, Shengli Branch, Sinopec, Dongying 257086, China
3. Service Center of Oil and Gas Engineering, North China Branch, Sinopec, Xianyang 712000, China

Abstract Low-frequency vibrator, wide-band, wide-azimuth and high-density, high-efficiency acquisition technologies have been put into wide application thanks to the continual accuracy requirement of oil exploration and development. At the same time, oil seismic exploration has entered an era of big data, leading to a series of challenges of quality control, data processing, safety storage and management. For this reason, Sinopec has developed a "π-Frame seismic data processing and interpretation integration software platform" based on the Hadoop big data architecture. In this paper, we first introduce the characteristics of big data in oil seismic exploration. Then we describe the basic framework of the π-Frame platform. Finally we illustrate the application of big data in oil seismic exploration and prospect its potential developments.

Keywords oil seismic exploration; big data; Hadoop; seismic data processing and interpretation; software platform

(责任编辑 刘志远)