

DOI: 10.19797/j.cnki.1000-0852.20190201

基于 MPI 的分布式水循环模型并行计算性能研究

向东^{1,2}, 周祖昊², 袁 胜¹, 秦泽宁^{1,2}, 刘佳嘉², 朱家松³

(1.华北水利水电大学,河南 郑州 450046;

2.中国水利水电科学研究院流域水循环模拟与调控国家重点实验室,北京 100038;

3.深圳大学土木与交通工程学院,广东 深圳 518000)

摘要:为满足分布式水文模型快速模拟的需要,引入并行计算技术。基于 MPI 编程模型实现了 WEP-L 分布式水循环模型产流模块的并行编程,采取子流域任务划分和对等模式实现了模型的并行设计。为了减少进程间的通信时间,在产流计算之前,数据按随年、月、日变化以及不随时间变化分成四类,分批进行通信;产流计算完成之后,采用聚合通信方式中的数据收集,快速统计结果。模型并行化后应用于黄河流域,结果如下:(1)随着参与计算的进程数增加,并行计算的加速比呈先增加后减少的趋势,并行效率随进程数增加呈线性下降趋势。(2)模型并行性能受通信开销制约,当通信开销增量大于产流计算时间减少量时,加速比达到峰值 4.8。

关键词:分布式水循环模型;MPI;并行计算;加速比;并行效率

中图分类号:P333.9;O246

文献标识码:A

文章编号:1000-0852(2020)05-0036-05

大流域、精细化模拟是分布式水文模型发展的趋势,其发展要求大量数据信息的处理和多要素多层次的物理计算过程,因而也面临着模拟时间长,计算效率低的问题,而并行计算正是解决耗时问题的有效技术。

在 2000 年以前,各种并行语言如分布式存储并行编程环境下的 PVM、MPI 和共享存储编程环境下的 Pthreads、HPF、OpenMP 等刚刚起步,此时水文模型的并行化研究较少,并行效果不佳。如 1997 年 Apostolopoulos 等^[1]采用 Encore Parallel Fortran 实现了基于子流域的水文模型的并行计算,9 个 CPU 计算的加速比仅为 1.5。2000 年之后,水文模型的并行计算逐渐成熟。随着 OpenMP 和 MPI 等高性能并行标准的官方认定和广泛应用,基于这两种并行编程的分布式水文模型并行算法成为主流。利用 MPI 编程的研究:Cui 等^[2]以 Illinois 河流域为研究区,实现了基于子流域的分布式水文模型 r.water.fea 的并行化,11 个进程计算的加速比为 4.5;李铁键等^[3]以黄河流域支流伊洛河陆

浑、故县水库以下流域为研究区,实现了基于子流域的数字流域模型产汇流过程的并行化,20 个进程计算的加速比为 5.39,并在后续研究中使用动态并行算法进行改进^[4],13 个进程计算的加速比为 5.34;Vivoni 等^[5]以 Baron Fork 流域为研究区,实现了基于不规则三角网的 tRIBS 模型的并行化,128 个进程计算的加速比达到 60~80;申蒙蒙等^[6]实现了 SWAT 和 HIMS 模型在参数优化模块的并行化,120 个进程计算的加速比分别为 11.6 和 108.8。利用 OpenMP 编程的研究:Yu 等^[7]实现了一个基于栅格的分布式水文模型的并行计算,该方法对嘉陵江支流鹤鸣观流域进行了模拟,加速比在 4 个线程时达到 2.42;Liu 等^[8]研究基于栅格的上下游顺序相关的分布式水文模型,提出一种分层并行算法,24 个线程计算的加速比为 12.49。

近几年随着高性能计算的发展,空间分解的方式被进一步广泛应用。Liu 等^[9]将计算单元为子流域内栅格的模型分成两层,采用外层 MPI 计算、内层 OpenMP

收稿日期:2019-06-25

基金项目:国家重点研发计划课题(2016YFC0402405);江西省水利科技重大项目(KT201501);坪山河干流综合整治及水质提升工程专项课题(CSCEC-PSH-2017-03)

作者简介:向东(1995-),男,河南信阳人,硕士研究生,研究方向为分布式水文模型并行计算。E-mail:1184985094@qq.com

通讯作者:周祖昊(1975-),男,湖北武汉人,教授级高工,博士,研究方向为流域水循环及伴生过程模拟与调控。E-mail:zhzh@iwhr.com

的计算结构,在核数为 12,每个进程的线程数为 3 的情况下,加速比达到 8。同时,为了打破空间结构并行计算的瓶颈,Wang 等^[10]同时考虑时间和空间结构的并行性,在拉萨河流域,24 个进程计算的加速比达到 15.07,突破了其空间结构最大加速比为 7.98 的瓶颈。

相比 OpenMP,MPI 并行标准适用性强,既可用于单机,又可用于计算机集群,因此在水文领域应用更为广泛。基于 MPI 的并行算法尽管取得长足的进展,但是对计算效率瓶颈问题的分析以及如何提高并行计算效率缺乏深入研究。本文拟结合广泛使用的 WEP-L 分布式水循环模型,针对该问题进行探讨。

1 计算需求

WEP-L (Water and Energy transfer Processes in Large river basins)模型是耦合模拟天然水循环过程与人工侧支水循环过程的大尺度流域分布式水循环模型,该模型在各种流域应用广泛^[11-13]。WEP-L 模型采用“子流域内的等高带”作为基本计算单元,适于空间分解的并行计算方法。

本文拟以该模型为研究对象,以黄河流域(面积 79.5×10⁴km²)为研究区,开展并行计算方法研究。模型首先基于 1km 数字高程模型和实测数字河网,将黄河流域划分为具有空间拓扑关系的 8 485 个子流域,进而按照高程划分为 38 720 个等高带;对黄河流域 1956-2017 年共 62 年长系列逐日的水循环过程进行连续模拟计算,经过不断优化发展,总的计算时间为 2.5h。使用英特尔 VTune(tm)性能分析器进行热点分析,发现产流模块占该模型总运行时间的 60%以上,是模型的计算热点。故本文针对 WEP-L 模型的产流过程,开展基于 MPI 的并行计算性能研究。

2 并行化设计与实现

2.1 MPI 计算原理

MPI 是一种消息传递编程模型,目的是服务于进程间通信。其并行计算开始时,每个 CPU 核心绑定一个进程进行实时运算。其并行程序的一般流程为:(1)启用 MPI 并行库;(2)获取当前通信域和进程信息;(3)进行任务划分,映射到计算的进程,进程间通信传递消息;(4)退出对 MPI 并行库的调用。

MPI 并行计算通常由一组运行在相同或不同计算机或计算结点上的进程构成。该算法有一个主进程(master),其余称为从进程(slave),参与计算的进程在

CPU 内核上一一对应进行计算。其中,主从模式和对等模式是 MPI 的两种基本设计模式,在主从模式中,主进程负责动态分配任务,从进程负责计算,当从进程计算完成后,主进程继续分配未处理的任务,直至所有任务计算完成,便于处理某些动态负载均衡的问题,但设计程序复杂;对等模式的各个进程地位相等,处理不同的任务对象,其功能和代码基本一致,主进程与从进程同时参与计算,便于有效使用聚合通信函数。考虑到本文设计的任务划分算法能有效减小负载均衡问题,及通信设计中所使用的聚合通信函数,选取对等模式进行应用。

2.2 WEP-L 模型并行性分析

WEP-L 分布式水循环模型已完成模块化改造,主要模块包括:输入数据读取模块、产流模块、汇流模块、输出统计模块等。各模块之间采用全局变量进行关联,减少了模块调用的参数传递,该模型基本满足并行化改造的要求。模型结构如图 1 所示。

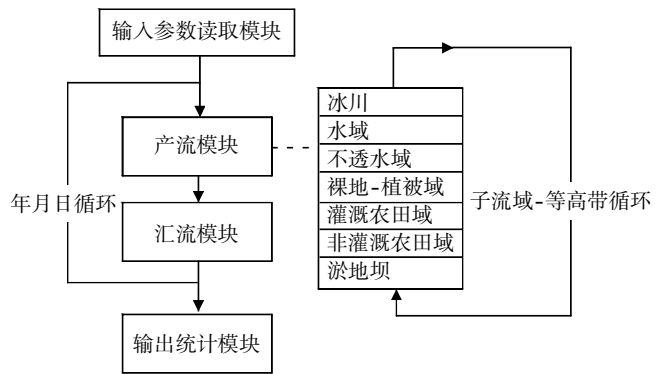


图 1 模型结构简图

Fig.1 The model structure

从模型结构上看,WEP-L 水循环模型主要有两个大循环。一是时间循环。即按年、月、日进行连续时间的模拟,涉及到后一天时段初的输入数据初始化需前一天时段末的状态值等情况,如时段末河道汇流流量、气温等,因为有时间上的顺序依赖性,所以不便按照时间进行并行计算;二是产流过程和汇流过程分空间单元的循环计算。其中,模型产流模块的物理结构为:子流域从上游到下游依次计算,子流域内等高带按高程从高到低计算,基本单元(等高带)内设置相同的物理过程。由于各个子流域的产流模块在计算上相对独立,且计算过程基本相同,而同一子流域内的等高带间涉及地下水量的交换,存在数据依赖性,故选

取相对独立的子流域作为并行计算的基本单元。汇流模块按照子流域上下游关系依次计算, 由于涉及各个子流域之间的水量交换, 数据依赖性大, 本次暂不考虑并行化计算。

2.3 WEP-L 产流模块并行化设计

2.3.1 任务划分

对 WEP-L 模型采取空间区域分割的方法, 并行化原理为: 以模型的子流域单元为基础, 尽可能地划分为 N 块计算时间均等的子区域(包含若干子流域), 之后分配给 N 个 CPU 计算, 这是并行算法设计的两个关键步骤。

考虑到对等模式的负载平衡问题, 在预处理时先将模型应用于黄河流域, 对 8 485 个子流域采用串行计算, 累积所有子流域的产流计算时间, 然后根据每个子流域的产流计算时间, 将黄河流域划分为近似均等的子区域, 找出对应子区域第一个和最后一个子流域编号, 再将不同子区域分配给不同的进程进行计算。在并行域中, 主进程和从进程地位相同, 运行代码一样, 负责黄河流域不同的子区域计算; 在并行域之外, 只有主进程运行, 主要负责输入数据的读取和输出结果的统计。子流域任务划分如图 2 所示, 选取 4 个进程做并行计算为例, 每个进程对应唯一的 CPU 内核进行计算, 根据设计的划分算法, 对 8 485 个子流域划分为计算时间近似均等的 4 个子区域, 每个子区域可能包含数量不均等的子流域。

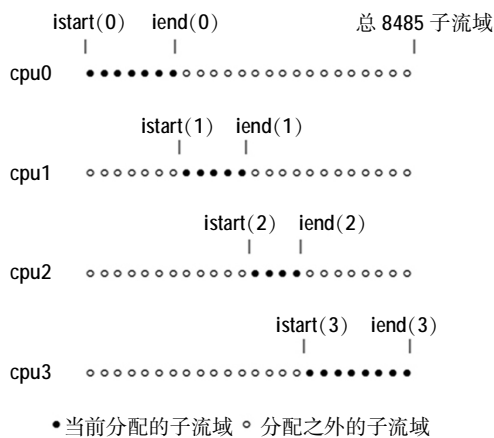


图 2 任务划分简图
Fig.2 The task division

注: cpu0 和 cpu1、cpu2、cpu3 分别代表主进程 0 和从进程 1、2、3, istart(0)、istart(1)、istart(2)、istart(3) 分别代表每个子区域第一个子流域编码, iend(0)、iend(1)、iend(2)、iend(3) 分别代表每个子区域最后一个子流域编码。

2.3.2 通信设计

采取对等并行设计模式, 主进程负责输入数据的读取和输出结果的统计, 并与从进程同时参与计算。由此设计模型并行化有两次结构性的通信: 第一次在产流计算之前, 主进程按照分配的子流域任务将输入数据通过通信传递给其它进程; 第二次在产流计算之后, 从进程将负责计算的子流域模拟结果汇总到主进程进行统计更新。

数据通信开销对并行算法的性能有着重要影响, 本文中, 将两次通信进行优化操作。通信设计如图 3 所示, 对于产流计算之前的通信, 如果将所有的输入数据都按照每日一次进行通信, 通信时间太长, 本文先将输入数据按照随年、月、日变化和随时间变化划分为四类, 根据类别分别按照一次或者一日一次、一月一次或者一年一次的频率进行通信, 从而达到减少数据冗余、降低通信次数的效果。对于产流计算完成之后的通信, 当所有从进程向主进程汇总任务结果时, 若主进程缓存区达到一定规模的通信数据量, 则会产生严重的通信延迟, 当数据等待接收的时间过长, 甚至会导致通信失败。故将原来每个从进程依次给主进程传递结果的“点对点”通信模式改为主进程直接对通信域中的计算结果进行收集的聚合通信模式, 这样避免数据在缓存区排队等待问题, 保证了数据安全。

针对 WEP-L 模型的输入和输出数据, 将图 3 通信内容做以下 5 类简要说明: 通信 1 为河床材质渗透系数、三层土壤厚度、含水层厚度、洼地储留深等; 通信 2 为土地利用百分比、各下垫面占计算单元面积比例等; 通信 3 为植被覆盖度、叶面积指数、植被高度、

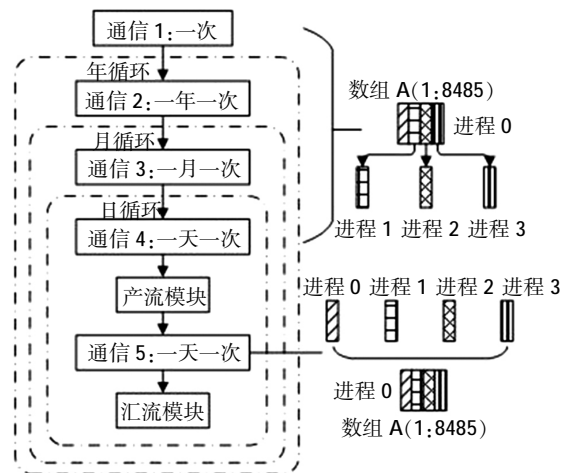


图 3 通信设计简图

Fig.3 The telecommunication design

植被根深等;通信 4 为降雨、气温、日照、时段末河道汇流流量等;通信 5 为各下垫面蒸发量、植物蒸腾量、净辐射量、显热通量等。以上数据是模型概化的物理参数,随实际相比,某些物理量随时间的变化可能归类不同。图 3 以数组 A 表示某个物理量为例,数组空间按子流域个数设置 8 485,模拟 4 个进程参与计算。将数组 A 分成 4 块子区域的数据,产流之前的通信为“点对点”模式,主进程把数据依次传递给从进程;产流之后的通信为聚合通信模式,所有进程同时进行通信,主进程收集通信域内所有进程的计算结果。

3 测试结果与分析

测试的目的主要是验证设计的 WEP-L 水循环模型产流模块并行算法的合理性和效果。测试环境:WEP-L 模型采用 Fortran 语言编写,编写软件为 Microsoft Visual Studio, MPI 编程模型实现的工具为 MPICH2,硬件环境为服务器 Intel(r) Xeon(r) CPU E5-2630 v4 @2.2GHz,系统为 Windows Server 2016。并行化后的模型可用于单机和集群,本文选用 CPU 核数较多及通信稳定的服务器平台测试。

设置进程数为 1-12,对黄河流域 8485 个子流域 1956-2017 年长系列的水循环过程进行连续模拟计算,得到产流过程计算时间、各进程间的通信时间(包括通信延迟和数据传输时间)、产流模块计算总时间(前两项之和)。根据公式(1)和(2)计算不同进程数条件下产流模块的加速比和并行效率,如表 1 和图 4-5 所示。

表1 并程序测试结果

Table1 The test results of the parallel program

参与计算 进程数	产流过程 计算时间/s	通信时间 /s	产流模块 计算总时间/s	加速比	并行 效率
1	5 874	0	5 874	1.0	100%
2	3 083	335	3 418	1.7	86%
3	2 262	379	2 641	2.2	74%
4	1 665	397	2 062	2.8	71%
5	1 331	372	1 703	3.4	69%
6	1 118	418	1 536	3.8	64%
7	996	511	1 507	3.9	56%
8	842	512	1 354	4.3	54%
9	765	474	1 239	4.7	53%
10	701	530	1 231	4.8	48%
11	670	649	1 319	4.5	41%
12	625	715	1 340	4.4	37%

$$\text{Speedup} = t_1/t_N \tag{1}$$

$$\text{Efficiency} = \text{Speedup}/N \tag{2}$$

式中:Speedup 为加速比; t_1 、 t_N 分别为 1 个进程和 N 个进程对应的运行时间;N 为参与计算的进程数;Efficiency 为并行效率。

从表 1 和图 4 可见,随着参与计算的进程数增多,产流过程的计算时间逐渐减小,进程间的通信时间逐渐增加,产流模块计算总时间先减少后增加。

进程间的通信开销是制约并行计算性能的关键因素,参与计算的进程数越多,进程间的消息收发越频繁,导致通信次数越多,通信时间延迟越长。由图 4 可见,进程数与通信时间成正相关关系,当通信时间的增量大于产流过程计算减少的时间时,所设计的并行算法的效率达到最高值。在本次测试过程中,当参与计算的进程数为 10 时,产流模块计算的总时间最少,为 1231s。图 5 为衡量并行性能的两个指标——加速比和并行效率与进程数的关系,可以看出,随进程数增多,加速比增长趋势减慢,当参与计算的进程数为 10 时,加速比达到峰值 4.8,之后由于通信开销过大,导致加速比下滑。

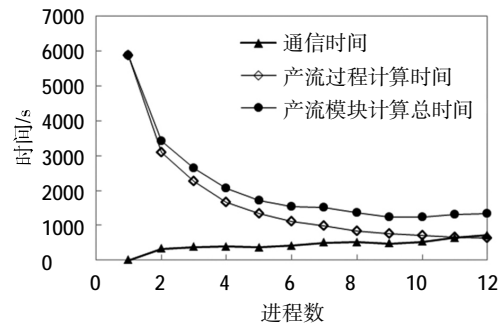


图4 时间开销与进程数的关系

Fig.4 The relationship between the telecommunication time and numbers of process

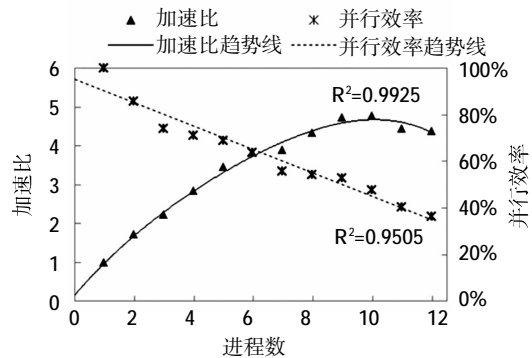


图5 并行性能与进程数的关系

Fig.5 The relationship between the parallel performance and numbers of process

由图 5 可见,产流模块的并行效率随进程数增多呈线性递减,在最大加速比为 4.8 的情况下,并行效率减小到 48%。并行效率随进程数增加而减少的原因,除了因为进程间的通信时间增加以外,主要是因为随着参与计算的进程数增加时,计算进程的闲置等待时间过长,各进程间负载不均衡的问题越来越突出,因而产流过程的计算时间随进程数增加而减小的速度逐渐减缓,而且减小的幅度越来越小。

4 结 语

本文将并行计算技术应用到 WEP-L 水循环模型,设计一种基于 MPI 的并行算法,采用子流域划分和对等模式实现模型的并行化处理。以模型划分的黄河流域的 8485 个子流域为例,在多核服务器平台进行了对比实验,实验结果表明:

(1)并行化设计可有效缩短模型的计算时间,显著提高模型的计算效率。随着参与计算的进程数增加,并行计算的加速比呈先增加后减少的趋势。当进程数为 10 时,黄河流域产流模块计算的总时间达到最小,由串行的 5 874s 缩短至 1 231s。

(2)模型并行性能受通信开销制约,随着进程数增加,通信开销逐渐增加,当通信开销增量大于产流计算时间减少量时,计算效率达到峰值。本文中当进程数为 10 时,加速比峰值为 4.8。

(3)模型并行效率随进程数增加呈线性下降趋势,当产流模块计算的总时间达到最小时,并行效率减小到 48%。若想进一步提高计算性能,需充分考虑进程间的负载平衡问题,根据各进程计算完成情况动态调整计算任务,减小计算进程的闲置等待时间,寻找最优负载方案。

参考文献:

- [1] Apostolopoulos T K, Georgakakos K P. Parallel computation for streamflow prediction with distributed hydrologic models [J]. *Journal of Hydrology*, 1997,197(1-4):1-24.
- [2] CUI Z T, Vieux B E, Neeman H, et al. Parallelisation of a distributed hydrologic model [J]. *Computer Applications in Technology*, 2005,22(1):42-52.
- [3] 李铁键,刘家宏,和杨,等. 集群计算在数字流域模型中的应用[J]. *水科学进展*, 2006,17(6):841-846. (LI Tiejian, LIU Jiahong, HE Yang, et al. Application of cluster computing in the digital watershed model [J]. *Advances in Water Science*, 2006,17(6):841-846. (in Chinese))
- [4] Li T J, Wang G Q, Chen J, et al. Dynamic parallelization of hydrological model simulations [J]. *Environmental Modelling & Software*, 2011,26(12):1736-1746.
- [5] Vivoni E R, Mascaro G, Mniszewski S, et al. Real-world hydrologic assessment of a fully-distributed hydrological model in a parallel computing environment [J]. *Journal of Hydrology*, 2011,409:483-496.
- [6] 申蒙蒙. 水文模拟中并行参数优化算法 [J]. *计算机工程与设计*, 2017,38(4):1002-1007. (SHEN Mengmeng. Parallel parameter optimization algorithm in hydrological simulation [J]. *Computer Engineering and Design*, 2017,38(4):1002-1007. (in Chinese))
- [7] Xu R, Huang X X, Luo L, et al. A new grid-associated algorithm in the distributed hydrological model simulations [J]. *Science China: Technological Sciences*, 2010,53(1):235-241.
- [8] Liu J Z, Zhu A X, Liu Y B, et al. A layered approach to parallel computing for spatially distributed hydrological modeling [J]. *Environmental Modelling & Software*, 2014,51:221-227.
- [9] Liu J Z, Zhu A X, Qin C Z, et al. A two-level parallelization method for distributed hydrological models [J]. *Environmental Modelling & Software*, 2016,80:175-184.
- [10] Wang H, Fu X D, Wang Y J, et al. A High-performance temporal-spatial discretization method for the parallel computing of river basins [J]. *Computers & Geosciences*, 2013,58:62-68.
- [11] Jia Y W, Wang H, Zhou Z H, et al. Development of the WEP-L distributed hydrological model and dynamic assessment of water resources in the Yellow River basin [J]. *Journal of Hydrology*, 2006,331(3-4):606-629.
- [12] Zhou Z H, Jia Y W, Qiu Y Q, et al. Simulation of dualistic hydrological processes affected by intensive human activities based on distributed hydrological model [J]. *Journal of Water Resources Planning and Management*, 2018,144(12).
- [13] Li J, Zhou Z H, Wang H, et al. Development of WEP-COR model to simulate land surface water and energy budgets in a cold region [J]. *Hydrology Research*, 2019,50(1):606-629.
- [14] 王敏,王明,杨明,等. 小浪底水库三维数学模型并行计算研究[J]. *人民黄河*, 2012,34(5):25-27. (WANG Min, WANG Ming, YANG Ming, et al. Parallel computing research of Xiaolangdi reservoir three-dimensional mathematical model [J]. *Yellow River*, 2012,34(5):25-27. (in Chinese))
- [15] 卢浩,王少华,李绍俊,等. 基于 OpenMP 的并行化水文分析算法研究与实现 [J]. *测绘与空间地理信息*, 2013,38:7-10. (LU Hao, WANG Shaohua, LI Shaojun, et al. Research and implementation of hydrologic analysis parallel algorithm based on OpenMP [J]. *Geomatics & Spatial Information Technology*, 2013,38:7-10. (in Chinese))
- [16] 苏丹阳. 基于物理概念的水文模型 InHM 机群并行计算研究[D]. 杭州:浙江大学, 2012. (SU Danyang. Study on the Parallelization of a Physics-based Hydrologic Model InHM Utilizing Computer Cluster [D]. Hangzhou: Zhejiang University, 2012. (in Chinese))

(下转第 27 页)