

冯新政, 张大伟, 徐海卿, 等. 基于多 GPU 数值框架的流域地表径流过程数值模拟[J]. 南水北调与水利科技(中英文), 2024, 22(1): 48-55. FENG X Z, ZHANG D W, XU H Q, et al. Numerical simulation of surface runoff process based on multi-GPU numerical framework[J]. South-to-North Water Transfers and Water Science & Technology, 2024, 22(1): 48-55. (in Chinese)

# 基于多 GPU 数值框架的流域地表径流过程数值模拟

冯新政<sup>1,2</sup>, 张大伟<sup>1</sup>, 徐海卿<sup>1</sup>, 鞠琴<sup>2,3</sup>

(1. 中国水利水电科学研究院, 北京 100038; 2. 河海大学水灾害防御全国重点实验室, 南京 210098;  
3. 水安全与水科学协同创新中心, 南京 210098)

**摘要:** 与传统概念性水文模型相比, 二维水动力模型可提供更丰富的流域地表水力要素信息, 但是计算耗时太长的  
问题限制其推广应用, 提升二维水动力模型的计算效率成为当前数字孪生流域建设工作中的关键技术难题之  
一。采用基于 Godunov 格式的有限体积法离散完整二维浅水方程组建立模型, 通过消息传递接口 (message  
passing interface, MPI) 与统一计算设备架构 (compute unified device architecture, CUDA) 相结合的技术实现了基于  
多图形处理器 (graphics processing unit, GPU) 的高性能加速计算, 采用理想算例和真实流域算例验证模型具有较  
好的数值计算精度, 其中, 理想算例中洪峰的相对误差为 0.011%, 真实流域算例中洪峰的相对误差为 2.98%。选  
取宝盖寺流域为研究对象, 分析不同单元分辨率下模型的加速效果, 结果表明: 在 5、2、1 m 分辨率下, 使用 8 张  
GPU 卡计算获得的加速比分别为 1.58、3.92、5.77, 单元分辨率越高, 即单元数越多, 多 GPU 卡的加速效果越明显。  
基于多 GPU 的水动力模型加速潜力巨大, 可为数字孪生流域建设提供有力技术支撑。

**关键词:** 水动力模型; 地表径流; 多 GPU; Godunov 格式; 加速比

**中图分类号:** TV131.2 **文献标志码:** A **DOI:** 10.13476/j.cnki.nsbdqk.2024.0007

随着信息化技术的快速发展, 以“四预”功能为目标的数字孪生流域建设成为当前水利信息化高质量发展的重要标志。在“2+N”的应用体系中, 防洪应用处于首要位置, 与传统的防洪信息化系统相比, 新一代的智慧防洪系统对水利业务模型提出了更高的要求<sup>[1]</sup>。相较于传统集总式、概念性水文模型<sup>[2]</sup>的汇流计算仅提供流域出口点的径流过程信息, 完整二维水动力模型可以提供流域空间内任一网格点的水动力要素分布信息, 可以模拟各种复杂流态下的水流运动过程, 细致考虑各种河道构筑物的影响, 因此, 在数字孪生流域建设中的应用前景更为广阔<sup>[3-7]</sup>。实际流域中产汇流过程复杂, 变源产流面积的动态变化会对地表径流运动造成影响<sup>[8-9]</sup>, 但为方便问题求解, 在实际应用时可先独立计算出净雨过程, 然后再进行二维地表径流的模拟, 张大伟等<sup>[10]</sup>采用完整二维浅水方程组构建的非结构模型, 对解家湾流域进行了数值模拟, 结果与实测值吻合

较好。Xia 等<sup>[11]</sup>的研究成果表明, 基于 Godunov 格式的二维水动力模型可以准确模拟流域地表径流的运动过程。

限制二维水动力模型实时应用的关键问题是计算效率问题。近些年, 图形处理器 (graphics processing unit, GPU) 硬件技术的快速发展, 为洪水模型的高性能计算提供了一种新的解决途径, 使得应用二维水动力模型实时模拟流域洪水过程成为了可能<sup>[12-13]</sup>。GPU 最初被设计用于图形渲染方面的工作, 随着架构的不断更新, 单块 GPU 卡内部已有数千个甚至过万个运算核心可用于执行并行计算任务, 使得其单卡计算能力远超过 CPU。目前, 已有较多国内外学者<sup>[14-18]</sup>使用 GPU 进行洪水高性能数值模拟。侯精明等<sup>[19]</sup>基于 GPU 加速技术构建的非结构流域雨洪数值模型, 在保证计算精度的同时实现流域雨洪过程的快速模拟。Kalyanapu 等<sup>[20]</sup>基于统一计算设备架构 (compute unified device

收稿日期: 2023-07-11 修回日期: 2023-11-28 网络出版时间: 2024-01-10

网络出版地址: <https://link.cnki.net/urlid/13.1430.TV.20240109.0939.002>

基金项目: 国家重点研发计划项目 (2022YFC3005501); 水利部重大科技项目 (SKS-2022007); 中国水科院科研专项 (WH0145B022021; WH0145B042021; JZ110145B0022023)

作者简介: 冯新政 (2000—), 男, 河南新乡人, 主要从事降雨径流数值模拟研究。E-mail: fengxinzheng@hhu.edu.cn

通信作者: 张大伟 (1980—), 男, 山东滕州人, 正高级工程师, 博导, 主要从事水动力模型研发和应用研究。E-mail: zhangdw@iwhr.com

architecture, CUDA) 开发的 Flood2D-GPU 模型, 模拟 30 min 溃坝导致的洪水事件耗时 2 min, 相较于 CPU 模型加速比达到了 80 多倍。Xia 等<sup>[21]</sup> 基于 Godunov 格式的有限体积法构建的二维水动力模型 HiPIMS, 在应用于模拟面积为 2 500 km<sup>2</sup> 的流域洪水过程时取得了较好的效果。中国水利水电科学研究院研发的 FASFLOOD 模型<sup>[1]</sup>, 采用其模拟蒙洼蓄滞洪区 3~5 d 的分洪过程在 40 s 内就可以完成计算。

目前, 在流域地表径流模拟方面, 基于 GPU 框架的二维地表径流水动力模型应用成果<sup>[22-23]</sup> 多是在单 GPU 环境下取得的, 针对多 GPU 加速效果的研究报道<sup>[24-25]</sup> 相对较少。随着 GPU 硬件技术的快速发展, 以 GPU 硬件为主体的云端算力会成为未来算力应用的主要形式, 因此, 研发基于多 GPU 的水动力模型具有重要的现实意义。本文采用消息传递接口 (message passing interface, MPI) 与 CUDA 相结合的计算架构实现了基于多 GPU 的水动力模型高性能加速计算, 通过 2 个算例对模型的精度进行了验证, 随后量化分析了不同数量 GPU 带来的模型加速效果。

## 1 模型原理

### 1.1 模型控制方程与数值方法

模型的控制方程采用完整的守恒型二维浅水方程:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial y} = \mathbf{S} \quad (1)$$

$$\mathbf{U} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} hu \\ hu^2 + gh^2/2 \\ huv \end{bmatrix}; \quad \mathbf{G} = \begin{bmatrix} hu \\ huv \\ hv^2 + gh^2/2 \end{bmatrix};$$

$$\mathbf{S} = \begin{bmatrix} r \\ -gh\partial Z_b/\partial x \\ -gh\partial Z_b/\partial y \end{bmatrix} + \begin{bmatrix} 0 \\ -gn^2h^{-1/3}u\sqrt{u^2+v^2} \\ -gn^2h^{-1/3}v\sqrt{u^2+v^2} \end{bmatrix} \quad (2)$$

式(1)和(2)中:  $h$  为水深, m;  $u$  和  $v$  分别为  $x$  和  $y$  方向的流速, m/s;  $t$  为时间, s;  $g$  为重力加速度, m/s<sup>2</sup>;  $\mathbf{S}$  为源项矢量, 包括净雨源项  $r$ 、底坡源项及摩阻力源项;  $Z_b$  为地面高程, m;  $n$  为糙率系数。

为便于使用流域 DEM 原始数据, 模型基于结构化网格建立, 采用基于 Godunov 格式的有限体积法离散公式(1), 通过 Roe 格式求解器求解通过各单元边的数值通量<sup>[10]</sup>, 底坡和干湿边界具体的数值方法见文献 [26]。

### 1.2 多 GPU 并行计算技术

为满足大规模并行计算任务的要求, 实现多个 GPU 同时工作, 选用 MPI 与 CUDA 相结合的技术。MPI 基于分布式存储模型实现并行进程之间的消息传递, 具有较高的可扩展性, 1 个 MPI 程序可由 1 组运行在相同或不同节点上的进程组成, 每个进程在运行时都具有唯一的进程号, 使用每个进程分别控制管理 1 个 GPU 设备, 由于不同进程的内存空间是独立的, 各个进程之间需要通过 MPI 函数进行通信。

与单 GPU 的计算流程(图 1)相比, 采用多 GPU 进行计算时, 需要将计算区域划分为多个子域, 子域的数量与 GPU 的数量相等, 每个 GPU 负责 1 个子域的计算。在子域的边界外部有一层网格用于跟相邻的子域进行通信, 这部分网格在每一时间步计算完成后从相邻子域接收数据来进行更新, 通过使用 MPI\_Isend 和 MPI\_Irecv 非阻塞通信函数来实现数据的发送与接收, 通信完成后在子域内继续进行计算, 数据传输是各个 GPU 需要同步的操作, 需要通过 MPI\_Barrier 函数来设置同步点, 该函数的作用是使程序在执行该函数时进入阻塞挂起状态, 只有当本次迭代中所有进程都执行至此, 才会继续向后执行, 因此划分子域的时候应尽可能使每个 GPU 的计算任务量相近以获得更好的加速效果。子域划分及通信示意图见图 2。

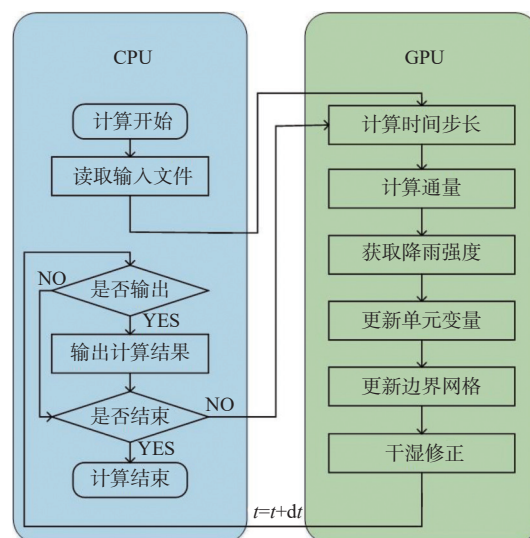


图 1 单 GPU 计算流程

Fig. 1 Calculation flowchart using single GPU

本文旨在探讨不同数量的 GPU 对于二维水动力模型的加速效果, 计算任务均在阿里云服务器上进行。计算所采用的配置硬件参数见表 1。

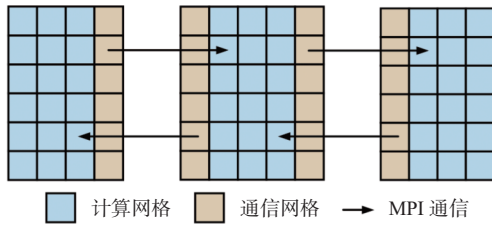


图 2 多 GPU 计算区域划分

Fig. 2 Domain decomposition on multi-GPU framework

表 1 硬件参数

Tab. 1 Hardware parameters

CPU型号	CPU主频/ GHz	GPU型号	CUDA核心 数量/个	显存容 量/GB	显存带宽/ (GB·s <sup>-1</sup> )
Intel Xeon Platinum 8163	2.5	NVIDIA Tesla V100	5 120	16	900

## 2 模型精度验证

### 2.1 理想降雨径流算例

该降雨径流算例最早是由 Gottardi 等<sup>[27]</sup>提出,具有解析解,用来检验模型的数值计算结果精度。该算例中,整个坡面长度为 200 m,坡度为 0.001,糙率系数为 0.030;降雨强度为 60 mm/h,持续时间( $t$ ) 1 h 后停止,总模拟时间为 2 h。分别采用 5.0、1.0、0.5 m 共 3 种大小的网格对计算区域进行划分,该坡面的汇流时间为 31.6 min,因降雨时间大于汇流时间,流量过程会达到一段稳定状态,此时坡面末端的单宽流量为  $3.333\ 33 \times 10^{-3} \text{ m}^2/\text{s}$ ,水深为 0.032 m。图 3 给出了 3 种网格大小情况下模型计算得到的坡面末端出流过程与解析解的对比结果。可以看出数值解与解析解整体上吻合较好:分辨率为 5.0 m 时,坡面末端单宽流量( $q$ )稳定后的数值解为  $3.332\ 86 \times 10^{-3} \text{ m}^2/\text{s}$ ,相对误差为 0.014%;分辨率为 0.5 m 时, $q$  稳定后的数值解为  $3.332\ 98 \times 10^{-3} \text{ m}^2/\text{s}$ ,相对误差为 0.011%,网格分辨率越高,数值解与解析解的吻合程度越好。通过该算例说明模型对于降雨径流过程的模拟结果是可靠的。

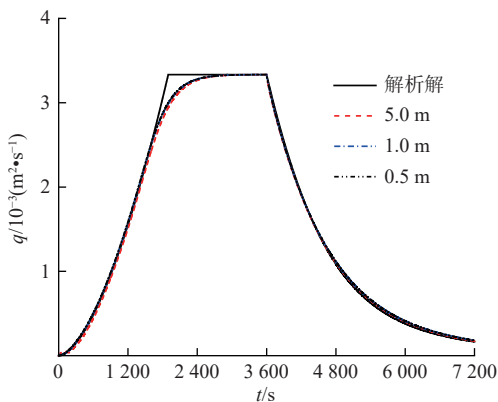


图 3 模拟流量过程与解析解对比

Fig. 3 Simulation flow process and comparison of analytical solutions

### 2.2 解家湾流域雨洪过程模拟

#### 2.2.1 流域概况

选取解家湾流域作为真实流域算例,研究区位于嘉陵江中下游的四川省遂宁市安居区,靠近遂宁水土保持试验站。该区位于  $105^{\circ}4'E \sim 105^{\circ}43'E$ 、 $30^{\circ}10'N \sim 30^{\circ}34'N$ ,处于亚热带湿润季风气候区,年平均降水量约为 932.7 mm。解家湾流域总面积  $65\ 827 \text{ m}^2$ ,高程最低 275 m,最高 330 m,流域平均坡度为 0.27,属于典型的盆中丘陵地貌,流域高程见图 4。研究区地形分辨率为 0.5 m,共计 263 276 个计算单元。由于谢家湾流域面积较小,水土流失严重,草木植被稀少,耕地和荒坡占流域面积的 90% 以上<sup>[28]</sup>,故对研究区的糙率系数取值为 0.035。模型设置下游边界为自由出流边界,其余为固壁边界。

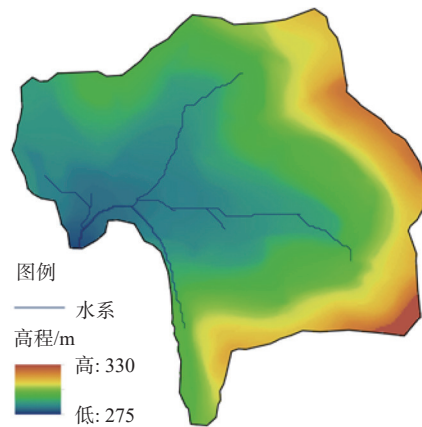


图 4 解家湾流域数字高程

Fig. 4 Digital elevation map of Xiejiaowan watershed

#### 2.2.2 流域 S 曲线验证

选取的单位时段长为 10 min,单位时段内的净雨量为 10 mm,由理论公式计算可得该流域的 S 曲线最大值为  $1.097 \text{ m}^3/\text{s}$ 。图 5 为模型计算所得该流域的 S 曲线过程,出口流量( $Q$ )在经历较短的涨水过程之后达到稳定,稳定后的值同样为  $1.097 \text{ m}^3/\text{s}$ ,与理论值相符。图 6 为计算过程中涨水阶段与达到稳定状态时,径流深大于 1 cm 区域的水深分布。

#### 2.2.3 流域实测降雨径流过程模拟

选取解家湾流域 2010 年 8 月 21 日 11 点至 8 月 22 日 11 点的一场实测降雨数据进行模拟,共历时 24 h,由新安江三水源模型计算得到净雨过程并将其作为模型的输入条件<sup>[10]</sup>,模拟结果见图 7。从图 7 可以看出,模拟结果与实测流量过程吻合较好,模拟流量过程线在上升过程略有延迟:第一个洪峰出现时间较实测流量过程延后 20 min 左右,模拟洪峰流量为 30.21 L/s,实测值为 35.81 L/s,相对误差约

为 15.64%; 第二个洪峰出现时间较实测流量过程延后 15 min 左右, 模拟洪峰流量为 63.77 L/s, 实测值为 65.73 L/s, 相对误差约为 2.98%。通过该算例的检验说明, 模型在处理实际流域复杂地形时可获得较高的数值计算精度。

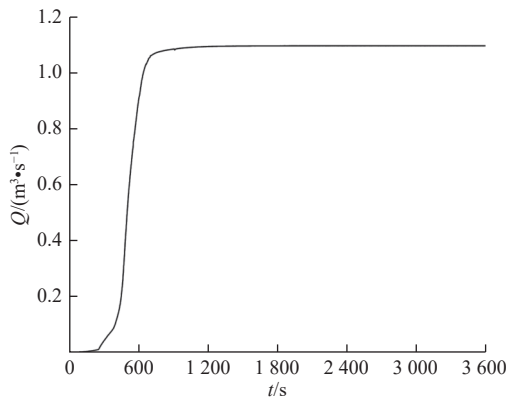


图 5 解家湾流域 S 曲线

Fig. 5 S curve of Xiejiawan watershed

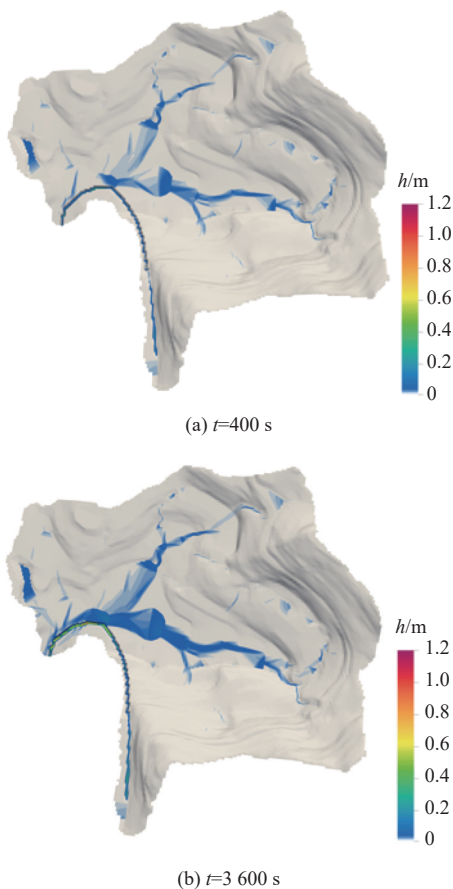


图 6 不同时刻水深分布

Fig. 6 Water depth distribution at different times

### 3 模型加速效果验证

模型采用 GPU 并行加速技术提高计算效率, 但由于解家湾流域面积较小, 用来验证模型的加速效

果时并不明显, 所以选取流域面积相对较大的湖南省宝盖寺流域对模型的多 GPU 加速效果进行验证。

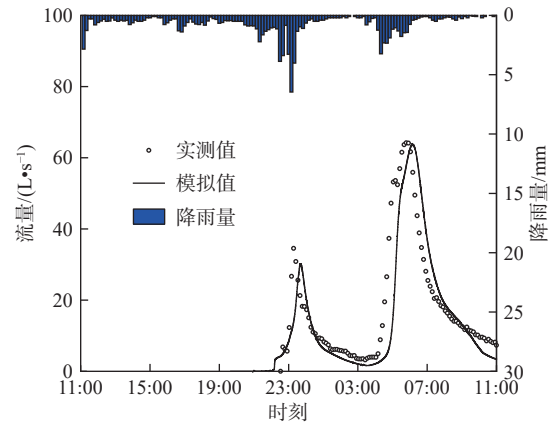


图 7 谢家湾流域模拟流量过程及实测值对比

Fig. 7 Simulated flow process and comparison with measured flow process in Xiejiawan watershed

### 3.1 流域概况及参数设置

宝盖寺流域集水面积约为 22.1 km<sup>2</sup>, 河道长约 9.1 km<sup>[29]</sup>, 流域高程图见图 8。计算采用的降雨强度恒为 10 mm/h, 降雨时长为 6 h。为了对比不同分辨率地形下的加速效果, 分别采用 1、2、5、10 m 大小的网格进行计算, 网格数分别为 21 539 061、5 384 807、861 605、215 387。糙率系数统一取为 0.035, 模型设置下游边界为自由出流边界, 其余为固壁边界, 模拟结果仅在计算结束时进行输出, 计算过程中 CFL 数取值为 0.5。

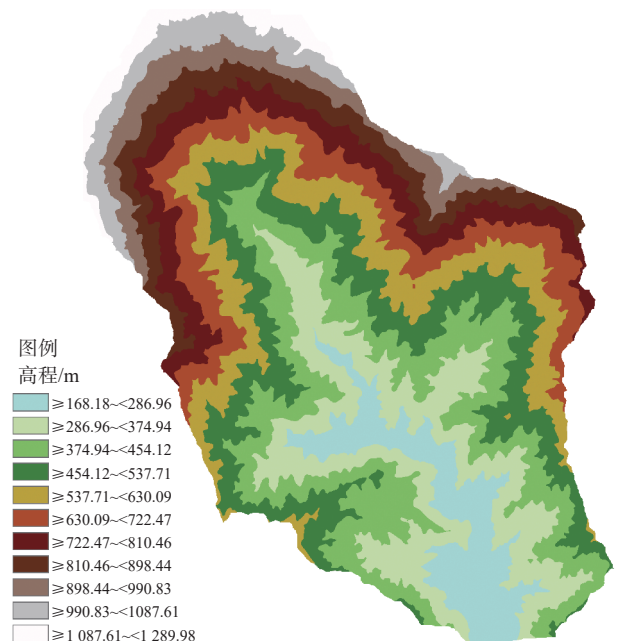


图 8 宝盖寺流域地形

Fig. 8 Topographic map of Baogaisi watershed

### 3.2 模型加速效果

采用加速比来量化评价多 GPU 较单 GPU 数值

方法的加速效果,表达式为

$$J = \frac{T_{GPU}}{T_{GPU_n}} \quad (3)$$

式中:  $J$  为加速比;  $T_{GPU}$  和  $T_{GPU_n}$  分别为模型在单个 GPU 和  $n$  个 GPU 上的运行时间, min。

各种模拟方案加速效果见表 2 和图 9, 可以看出, 加速比与 GPU 数量之间并不是线性关系, 分辨率越高, 越能体现地形的复杂程度及水流运动细节, 模型的计算量更大, 每张 GPU 所承担的计算任务也随之增加, GPU 的利用率变得更高, 加速比也更大,

如在 1 m 分辨率地形下, 8 张 GPU 的加速比可达 5.77 倍, 但对于 10 m 分辨率地形来说, 由于网格总数较少, 使用单 GPU 已经能在 30 s 内完成计算, 随着 GPU 数量的增加, 经过划分之后分配到每个 GPU 上的网格数量更少, 由于 Tesla V100 型号的 GPU 性能较为强大, 在网格数较少的情况下, 通信时间和计算时间相比占据份额更大, 所以, 在网格总数较少的情况下, 会出现多 GPU 计算用时大于单 GPU 计算用时的情况。

表 2 降雨径流过程不同数量 GPU 的模拟时间及对比

Tab. 2 Simulation time and comparison of different number of GPUs in rainfall runoff process

运算条件	网格尺寸/m	网格数量/个	计算时间/min	加速比/倍
NVIDIA Tesla V100	1	21 539 061	194.69	1.00
	2	5 384 807	32.62	1.00
	5	861 605	1.79	1.00
	10	215 387	0.36	1.00
NVIDIA Tesla V100*2	1	21 539 061	101.97	1.91
	2	5 384 807	18.78	1.74
	5	861 605	1.47	1.22
	10	215 387	0.46	0.78
NVIDIA Tesla V100*4	1	21 539 061	59.54	3.27
	2	5 384 807	11.87	2.75
	5	861 605	1.23	1.46
	10	215 387	0.46	0.78
NVIDIA Tesla V100*6	1	21 539 061	42.68	4.56
	2	5 384 807	9.49	3.44
	5	861 605	1.17	1.53
	10	215 387	0.48	0.75
NVIDIA Tesla V100*8	1	21 539 061	33.75	5.77
	2	5 384 807	8.32	3.92
	5	861 605	1.13	1.58
	10	215 387	0.50	0.72

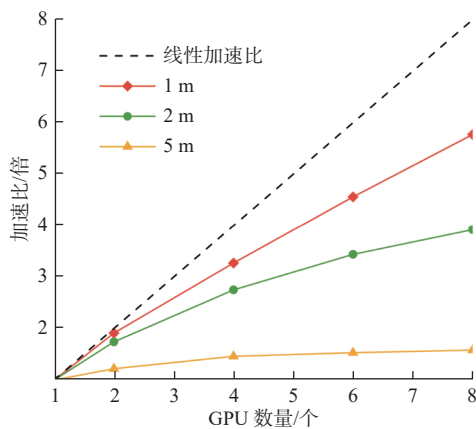


图 9 不同分辨率地形下的加速比

Fig. 9 Speedup ratio of topography with different resolutions

#### 4 结论与展望

采用基于 Godunov 格式的有限体积法离散二维浅水方程组, 同时使用 CUDA 与 MPI 相结合的技术, 成功实现基于多 GPU 数值框架的二维地表径流水动力过程的高性能模拟, 通过理想解析解算例及实际流域算例验证模型具有较高的数值计算精度。本文方法与传统的单位线等汇流计算方法相比, 可以摆脱对水文实测历史数据的依赖, 同时可提供流域空间内更丰富的水力要素信息, 而不仅仅是流域出口的流量过程信息。

采用宝盖寺流域算例对多 GPU 的加速效果进

行测试,结果表明,当网格总数量达到一定规模后,采用多 GPU 加速技术可获得满意的加速效果,当流域网格分辨率分别为 5、2、1 m 时,对应的网格单元数量为 861 605、5 384 807、21 539 061,使用 8 张 Tesla V100 型号 GPU 卡获得的加速比分别为 1.58、3.92、5.77,随着网格单元的增多,多 GPU 加速技术展现出了巨大的加速潜力,可为防洪“四预”业务的秒级计算需求提供一种可行的技术解决方案。

限于测试条件限制,仅测试了 8 块 GPU 的加速情况,测试的网格最大数量为 21 539 061 个。接下来的工作是继续研究在充分 GPU 算力资源的支撑下,拥有千万级或亿级数量网格单元的算例实现秒级或分钟级计算的可能性。另外,文中采用的实际流域面积相对较小,当模拟流域面积较大时,为考虑流域的空间差异性,可以对不同子流域对应的计算网格单元设置不同的净雨输入过程,也可以在每个网格单元上采用具有物理机制的产流方法直接计算每个网格单元的净雨量,汇流计算时可根据不同的土地利用类型数据给计算网格单元赋值不同的糙率值。

#### 参考文献:

- [1] 吕娟,张大伟.智慧防洪对水利业务模型的功能需求与技术实现构想[J].*中国水利*,2022(8):65-66. DOI: 10.3969/j.issn.1000-1123.2022.08.029.
- [2] 张润庆,何蒙,姚成,等.新安江-海河模型在烟台市典型流域的径流模拟[J].*南水北调与水利科技(中英文)*,2020,18(6):78-84. DOI: 10.13476/j.cnki.nsb-dqk.2020.0117.
- [3] MOGHIM S, GHAREHTORAGH M A, SAFAIE A. Performance of the flood models in different topographies[J]. *Journal of Hydrology*, 2023, 620: 129446. DOI: 10.1016/j.jhydrol.2023.129446.
- [4] 余富强,鱼京善,蒋卫威,等.基于水文水动力耦合模型的洪水淹没模拟[J].*南水北调与水利科技*,2019,17(5):37-43. DOI: 10.13476/j.cnki.nsb-dqk.2019.0107.
- [5] 申言霞,周琦,段艳华,等.基于多重网格的地表水文与二维水动力动态双向耦合模型研究[J].*水利学报*,2023,54(3):302-310. DOI: 10.13243/j.cnki.slxb.20220645.
- [6] 苑希民,贾帅静,田福昌,等.洪水风险快速分析技术方法研究进展[J].*水利水电技术*,2018,49(7):62-70. DOI: 10.13928/j.cnki.wrahe.2018.07.008.
- [7] 陈海梅,徐金英,赖锡军.二维坡面雨洪水动力模型研究[J].*水动力学研究与进展(A辑)*,2017,32(4):484-490. DOI: 10.16076/j.cnki.cjhd.2017.04.012.
- [8] 张博美,高红凯,余敦先,等.以地形为基础的HSC产流模型在辽西半干旱丘陵流域的适用性[J].*南水北调与水利科技(中英文)*,2021,19(6):1072-1082. DOI: 10.13476/j.cnki.nsb-dqk.2021.0110.
- [9] GAO H, BIRKEL C, HRACHOWITZ M, et al. A simple topography-driven and calibration-free runoff generation module[J]. *Hydrology and Earth System Sciences*, 2019, 23(2): 787-809. DOI: 10.5194/hess-23-787-2019.
- [10] 张大伟,权锦,马建明,等.基于Godunov格式的流域地表径流二维数值模拟[J].*水利学报*,2018,49(7):787-794,802. DOI: 10.13243/j.cnki.slxb.20180388.
- [11] XIA X L, LIANG Q H, MING X D, et al. An efficient and stable hydrodynamic model with novel source term discretization schemes for overland flow and flood simulations[J]. *Water Resources Research*, 2017, 53(5): 3730-3759. DOI: 10.1002/2017WR021563.
- [12] 李健,张大伟,姜晓明,等.并行化洪水演进模拟研究综述[J].*计算机工程与应用*,2021,57(13):1-7. DOI: 10.3778/j.issn.1002-8331.2011-0376.
- [13] 覃金帛,曾志强,梁藉,等.GPU并行优化技术在水利计算中的应用综述[J].*计算机工程与应用*,2018,54(3):23-29,63. DOI: 10.3778/j.issn.1002-8331.1711-0336.
- [14] SHAW J, KESSERWANI G, NEAL J, et al. LIS-FLOOD-FP 8.0: the new discontinuous Galerkin shallow water solver for multi-core CPUs and GPUs[J]. *Geoscientific Model Development*, 2021, 14(6): 3577-3602. DOI: 10.5194/gmd-14-3577-2021.
- [15] LIANG Q H, XIA X L, HOU J M. Catchment-scale high-resolution flash flood simulation using the GPU-based technology[J]. *Procedia Engineering*, 2016, 154: 975-981. DOI: 10.1016/j.proeng.2016.07.585.
- [16] 许栋,徐彬, PAYET D, 等.基于GPU并行计算的浅水波运动数值模拟[J].*计算力学学报*,2016,33(1):114-121. DOI: 10.7511/jslx201601018.
- [17] 侯精明,王润,李国栋,等.基于动力波法的高效高分辨率城市雨洪过程数值模型[J].*水力发电学报*,2018,37(3):40-49. DOI: 10.11660/slf-dxb.20180305.
- [18] 赵旭东,赵杨,孙家文,等.基于GPU加速的潮流模型及其在岛群二维水动力数值模拟中的应用[J].

- 海洋环境科学, 2017, 36(5): 781-790. DOI: 10.13634/j.cnki.mes.2017.05.022.
- [19] 侯精明, 张兆安, 马利平, 等. 基于GPU加速技术的非结构流域雨洪数值模型[J]. 水科学进展, 2021, 32(4): 567-576. DOI: 10.14042/j.cnki.32.1309.2021.04.008.
- [20] KALYANAPU A J, SHANKAR S, PARDYJAK E R, et al. Assessment of GPU computational enhancement to a 2D flood model[J]. *Environmental Modelling and Software*, 2011, 26(8): 1009-1016. DOI: 10.1016/j.envsoft.2011.02.014.
- [21] XIA X L, LIANG Q H, MING X D. A full-scale fluvial flood modelling framework based on a high-performance integrated hydrodynamic modelling system (HiPIMS)[J]. *Advances in Water Resources*, 2019, 132: 103392. DOI: 10.1016/j.advwatres.2019.103392.
- [22] BUTTINGER-KREUZHUBER A, KONEV A, HORVÁTH Z, et al. An integrated GPU-accelerated modeling framework for high-resolution simulations of rural and urban flash floods[J]. *Environmental Modelling and Software*, 2022, 156: 105480. DOI: 10.1016/j.envsoft.2022.105480.
- [23] CARLOTTO T, CHAFFE P L B, SANTOS C I D, et al. SW2D-GPU: A two-dimensional shallow water model accelerated by GPGPU[J]. *Environmental Modelling and Software*, 2021, 145: 105205. DOI: 10.1016/j.envsoft.2021.105205.
- [24] CAVIEDES-VOULLIÈME D, MORALES-HERNÁNDEZ M, NORMAN M R, et al. SERGHEI (SERGHEI-SWE) v1.0: A performance-portable high-performance parallel-computing shallow-water solver for hydrology and environmental hydraulics[J]. *Geoscientific Model Development*, 2023, 16(3): 977-1008. DOI: 10.5194/gmd-16-977-2023.
- [25] MORALES-HERNÁNDEZ M, SHARIF M B, KALYANAPU A, et al. TRITON: A multi-GPU open source 2D hydrodynamic flood model[J]. *Environmental Modelling and Software*, 2021, 141: 105034. DOI: 10.1016/j.envsoft.2021.105034.
- [26] MURILLO J, GARCÍA-NAVARRO P. Weak solutions for partial differential equations with source terms: Application to the shallow water equations[J]. *Journal of Computational Physics*, 2010, 229(11): 4327-4368. DOI: 10.1016/j.jcp.2010.02.016.
- [27] GOTTARDI G, VENUTELLI M. An accurate time integration method for simplified overland flow models[J]. *Advances in Water Resources*, 2008, 31(1): 173-180. DOI: 10.1016/j.advwatres.2007.08.004.
- [28] 付瀚, 郑江坤, 任雨之, 等. 降雨与植被变化对川中丘陵区典型小流域侵蚀产沙的影响[J]. *中国水土保持科学*, 2019, 17(4): 67-74. DOI: 10.16843/j.sswc.2019.04.009.
- [29] 刘昌军, 文磊, 周剑, 等. 小流域暴雨山洪水文模型与水动力学方法计算比较分析[J]. *中国水利水电科学研究院学报*, 2019, 17(4): 262-270, 278. DOI: 10.13244/j.cnki.jiwhr.2019.04.003.

## Numerical simulation of surface runoff process based on multi-GPU numerical framework

FENG Xinzhen<sup>1,2</sup>, ZHANG Dawei<sup>1</sup>, XU Haiqing<sup>1</sup>, JU Qin<sup>2,3</sup>

( 1. China Institute of Water Resources and Hydropower Research, Beijing 100038, China; 2. The National Key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210098, China; 3. National Cooperative Innovation Center for Water Safety & Hydro Science, Nanjing 210098, China )

**Abstract:** Compared to the traditional conceptual hydrological models, two-dimensional hydrodynamic models can provide more comprehensive hydraulic information of watershed surfaces, but the issue of long computational time restricts its widespread application. Exploring ways to improve the computational efficiency of two-dimensional hydraulic models has become one of the hot points and key technological challenges in current digital twin watershed development. The rapid advancement of GPU hardware technology has enabled the utilization of two-dimensional hydraulic models for the purpose of simulating watershed flood processes in real time. The potential applications of this technology in the construction of digital twin watersheds appear bright.

The model is established based on a structured grid and adopts the Godunov scheme based on the finite volume

method to discretize the complete two-dimensional shallow water equations. A high-performance accelerated calculation based on multiple GPUs is realized by combining MPI and CUDA computing architecture to meet the requirements of large-scale parallel computing tasks and realize the simultaneous work of multiple GPUs. MPI implements message passing between parallel processes based on the distributed storage model. Each process has a unique process rank at runtime and controls a GPU device. When using multiple GPUs for computation, the computational domain needs to be divided into multiple subdomains, and each GPU is assigned to compute a specific subdomain. Each subdomain is surrounded by an additional layer of grid cells that is used to communicate with adjacent subdomains. This outer layer of grid cells receives data from the adjacent subdomains to perform updates. Once the communication is completed, the computation continues within each subdomain.

The model's numerical accuracy has been verified using ideal and real watershed cases, with a relative error of 0.011% for the peak discharge in the ideal case and 2.98% for the peak discharge in the real watershed case. The acceleration effect of the model under different cell resolutions was analyzed in the Baogaisi watershed. The results showed that when the total number of grids reaches a certain scale, the multi-GPU acceleration technology can obtain a satisfactory acceleration effect. When the grid resolutions of the watershed are 5 m, 2 m, and 1 m, the corresponding grid units are 861,605, 5384,807, and 21,539,061. The speedup ratios obtained by 8 Tesla V100 GPUs are 1.58, 3.92, and 5.77, respectively. Higher cell resolutions lead to more significant acceleration effects with multiple GPUs. The hydrodynamic model based on multi-GPU has great potential for acceleration and can provide strong technical support for the construction of digital twin river basins.

**Key words:** hydrodynamic model; surface runoff; multi-GPU; Godunov scheme; speedup ratio

