

面向国产异构众核架构的 CFD 非结构网格计算并行优化方法

陈鑫, 李芳, 丁海昕, 孙唯哲, 刘鑫, 陈德训, 叶跃进, 何香

引用本文

陈鑫, 李芳, 丁海昕, 孙唯哲, 刘鑫, 陈德训, 叶跃进, 何香. [面向国产异构众核架构的 CFD 非结构网格计算并行优化方法](#)[J]. 计算机科学, 2022, 49(6): 99-107.

CHEN Xin, LI Fang, DING Hai-xin, SUN Wei-ze, LIU Xin, CHEN De-xun, YE Yue-jin, HE Xiang. [Parallel Optimization Method of Unstructured-grid Computing in CFD for Domestic Heterogeneous Many-core Architecture](#)[J]. Computer Science, 2022, 49(6): 99-107.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于国产众核架构的非结构网格分区块重构预处理算法研究](#)

Study on Preprocessing Algorithm for Partition Reconnection of Unstructured-grid Based on Domestic Many-core Architecture

计算机科学, 2022, 49(6): 73-80. <https://doi.org/10.11896/jsjcx.210900045>

[OpenFoam 中多面体网格生成的 MPI + OpenMP 混合并行方法](#)

Hybrid MPI+OpenMP Parallel Method on Polyhedral Grid Generation in OpenFoam

计算机科学, 2022, 49(3): 3-10. <https://doi.org/10.11896/jsjcx.210700060>

[一种面向构件化并行应用程序的性能骨架分析方法](#)

Performance Skeleton Analysis Method Towards Component-based Parallel Applications

计算机科学, 2021, 48(6): 1-9. <https://doi.org/10.11896/jsjcx.201200115>

[基于神威平台的 Floyd 并行算法的实现和优化](#)

Implementation and Optimization of Floyd Parallel Algorithm Based on Sunway Platform

计算机科学, 2021, 48(6): 34-40. <https://doi.org/10.11896/jsjcx.201100051>

[\(n,k\)-冒泡排序网络的子网络可靠性](#)

Subnetwork Reliability of (n,k)-bubble-sort Networks

计算机科学, 2021, 48(4): 43-48. <https://doi.org/10.11896/jsjcx.201100139>

面向国产异构众核架构的 CFD 非结构网格计算并行优化方法

陈鑫¹ 李芳¹ 丁海昕² 孙唯哲¹ 刘鑫¹ 陈德训¹ 叶跃进¹ 何香¹

¹ 国家超级计算无锡中心 江苏 无锡 214000

² 中国空气动力研究与发展中心 四川 绵阳 621000

(ischen.xin@foxmail.com)

摘要 神威太湖之光在 2016—2018 年度全球超算 top500 榜单中排名第一,峰值性能为 125.4 PFlops,其计算能力主要归功于国产 SW26010 众核处理器。由于 CFD 非结构网格计算存在拓扑关系复杂、离散访存问题严重、存在强相关的线性方程求解等问题,导致 CFD 非结构网格计算一直是国产众核超级计算机移植与优化的难题。为充分发挥国产异构众核架构的计算效能,首先,提出了一种数据重构模型,提高了数据的局部性和可并行性,使得数据结构更加适应众核架构的特点;然后,针对非结构网格数据存储的无序性导致的离散访存问题,提出了一种基于信息关系预存的离散访存优化方法,将离散访存转化为连续访存;最后,对于存在强相关的线性方程求解问题,引入了从核阵列流水线并行的思想,实现了众核并行。优化后 CFD 非结构网格计算的整体性能相比原始版本提升了 4.19 倍,相比通用 CPU 提升了 1.2 倍,并扩展到 62.4 万计算核心的并行规模,能保持 64.5% 的并行效率。

关键词: 计算流体力学;异构众核;神威超级计算机;非结构网格;并行计算

中图法分类号 TP311

Parallel Optimization Method of Unstructured-grid Computing in CFD for Domestic Heterogeneous Many-core Architecture

CHEN Xin¹, LI Fang¹, DING Hai-xin², SUN Wei-ze¹, LIU Xin¹, CHEN De-xun¹, YE Yue-jin¹ and HE Xiang¹

¹ National Super Computing Center in Wuxi, Wuxi, Jiangsu 214000, China

² China Aerodynamics Research and Development Center, Mianyang, Sichuan 621000, China

Abstract Sunway TaihuLight ranked first in the global supercomputer top 500 list 2016—2018 with a peak performance of 125.4 PFlops. Its computing power is mainly attributed to the domestic SW26010 many-core RISC processor. CFD unstructured-grid computing has always been a challenge for porting and optimizing in domestic many-core supercomputer, because of its complex topology, serious discrete memory access problems, and strongly correlated linear equation solution. In order to give fully play to the computing efficiency of domestic heterogeneous multi-core architecture, firstly, a data reconstruction model is proposed to improve the locality and parallelism of data, and the data structure is more suitable for the characteristics of multi-core architecture. Secondly, aiming at the discrete memory access problem caused by the disorder of unstructured-grid data storage, a discrete memory access optimization method based on prestorage of information relation is proposed, which transforms discrete memory access into continuous memory access. Finally, the pipeline parallelism mechanism in core array is introduced to realize many-core parallelism for solving linear equations with strong correlation. Experiments show that the overall performance of unstructured-grid computing in CFD is improved by more than 4 times, and is 1.2x faster than the general CPU. The computing cores scale to 624000, and the parallelism efficiency is maintained at 64.5%.

Keywords Computational fluid dynamics, Heterogeneous many-core, Sunway supercomputer, Unstructured-grid, Parallel computing

到稿日期:2021-04-15 返修日期:2021-07-15

基金项目:国家重点研发计划(2016YFB0201100);国家科技重大专项(2017-I-0004-0004)

This work was supported by the National Key Research and Development Project of China(2016YFB0201100) and National Science and Technology Major Project (2017-I-0004-0004).

通信作者:李芳(lifang56@163.com)

1 引言

计算流体力学(Computational Fluid Dynamics, CFD)是一门利用计算机对流体复杂流动进行数值模拟的学科^[1]。流体的运动规律可由欧拉方程、NS(Navier-Stokes)方程等基本方程组^[2]来描述,CFD则采用数值计算方法来求解这些方程,研究流体运动特性,得出流体的时空流动规律。随着CFD领域基础研究的深入,传统的结构网格^[3]已经不能完全满足复杂外形的数值模拟的精度要求。近年来,非结构网格技术^[4-5]越来越受到广大学者和研究人员的青睐,已被广泛应用于众多行业领域的研究中,如全球电磁模拟^[6]、高超声速流动模拟^[7]、海洋模式^[8]、船舶建设^[9]、地震仿真^[10]等。

CFD非结构网格的拓扑关系复杂,通常在并行计算前需要建立各进程内网格点、网格面、网格单元信息,以及各进程间的通信关系^[11]。在核心计算部分,要反复读取这些拓扑信息。由于非结构网格内部点不具有相同的毗邻单元,导致大部分拓扑信息的读取为离散读取。而对于国产众核架构而言,从核局部存储空间有限,无法将全部的拓扑信息加载到从核局存空间,而从核直接访问主存空间又会带来极高的延迟,严重影响运行效率,如何精简拓扑关系、提高数据的局部性和可并行性成为了CFD非结构网格计算众核优化的首要问题。

同时,非结构网格数据存储的无序性带来了计算过程中的离散访问问题^[12],并且网格规模越大,离散访问对计算性能的影响就越严重。在通用CPU上会严重降低访存的Cache命中率,而在进行基于国产众核超算系统的大规模计算时,会产生访存瓶颈问题,不能充分发挥国产众核强大的计算能力。目前离散访问问题已经成为制约非结构网格计算性能提升的关键因素之一。

此外,LU-SGS隐式格式被广泛用于非结构网格线化方程的求解^[13-14],其基本思想是利用通量线化假设得到特征分解后的雅可比矩阵,再利用最大特征值方法对块对角矩阵进行ILU分解。分解后采用递推法进行求解。递推法本身属于一种严格的串行执行方法,计算前后存在很强的数据相关性,这种强相关性使得众核并行成了一种挑战。

本文基于国产超级计算机神威太湖之光^[15]平台进行非结构网格计算众核优化的研究,针对国产异构众核架构的特点,提出了一种具有一定普适性的非结构网格计算众核并行方法。首先,基于国产异构众核架构的特点(从核强大的计算能力但内存空间有限),提出了一种数据重构模型,避免了拓扑关系复杂导致的反复查表的问题,同时使新的数据结构更加适应国产异构众核架构;然后,针对离散访问问题,提出了一种基于信息关系预存的离散访问优化方法,利用快速片上通信将离散访问转化成连续访存;最后引入了从核阵列流水线并行的思想,解决了LU解方程的并行难题。

2 背景介绍

2.1 SW26010众核处理器架构

SW26010众核处理器^[16]的结构组成如图1所示。该处理器集成了4个核组,每个核组包含一个主核(运算控制核心)和一个 8×8 的从核阵列(计算核心),核组之间通过片上

互联网络交互,主核与从核之间通过协处理部件相连。处理器的工作频率为1.45 GHz,理论内存带宽为130 GB/s,内存配置为32 GB,每个主核内存空间为8 GB,每个从核LDM(局存)大小为64 kB。从核可以直接离散访问主存,也可通过DMA(Direct Memory Access)的方式批量访问主存,将数据拷入LDM,同一核组内的从核可以通过寄存器的方式实现高效通信^[17]。

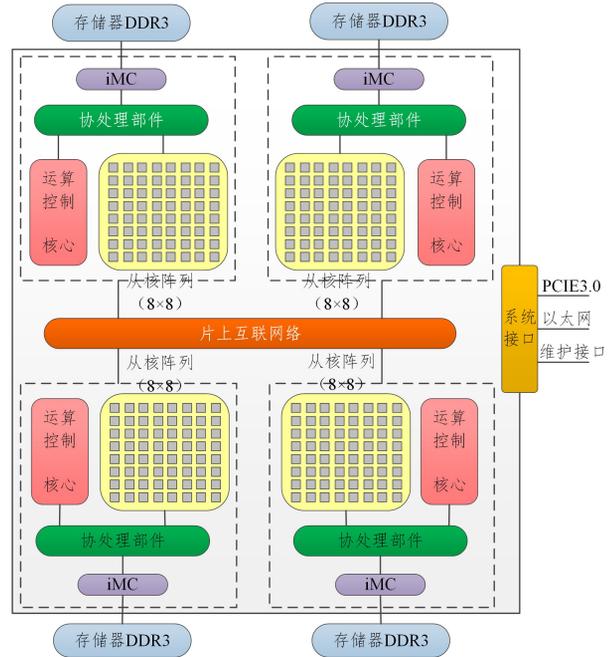


图1 SW26010众核处理器的结构

Fig. 1 Architecture of SW26010 processor

主核是一个64位RISC结构,可以在用户或系统模式下运行,进行计算、存储管理、消息收发等服务;从核也是64位RISC结构,但它主要提供强大的计算能力,其他功能有限,不支持中断功能,且只能在用户模式下运行。

主核采用两级Cache的架构,一级指令Cache容量为32 kB,采用四路组互联的结构;二级指令Cache容量为256 kB,由八路组互联构成。从核由16 kB的一级指令Cache和64 kB的LDM(局存)构成,其中从核高速存储区(LDM)由用户自己配置。显然,从结构设计上考虑,这种存储方式省去了缓存上的控制开销,并且避免了处理从核一致性导致的设计复杂度上升和性能降级^[18]。同时,若要利用好SW26010强大的计算能力,则将面临以下两个挑战:

(1)从核采用Gload/Gstore指令方式直接访问主存的访问带宽只有2.5 GB/s和1.0 GB/s,而通过DMA批量访问主存的带宽达26 GB/s,因此从核内尽量采用DMA的方式批量访问主存,可获得较高的访存效率。

(2)从核局存LDM空间有限,而非结构网格计算中的数据离散访问问题严重,核心数据的存储、冗余问题严重,因此提高LDM的利用率也是解决该类问题的关键。

2.2 相关工作

CFD非结构网格计算中面临的挑战主要体现在以下3个方面:拓扑关系复杂、离散访存问题、隐式(强相关性)方程求解。目前已有很多学者对这几个问题的并行求解进行了

研究,并取得了一定的研究成果。

(1)针对非结构网格拓扑关系复杂的问题,Liu等^[19]提出了一种非结构网格并行计算预处理方法。首先在建立各网格单元的相邻关系图时,采用了一种基于缓冲数据结构的快速搜索算法,解决了预处理内存需求过大的问题。Metis 分块之后,使用了一种新的网格重排算法,使得相邻网格单元或节点的编号尽可能接近,从而提高了后续的计算效率。但该方法只能在一定程度上提高主核的 Cache 命中率,非结构网格拓扑复杂,核心计算反复查表问题依然没有得到解决。

(2)针对非结构网格中的离散访存问题,Meng等^[20]提出了数据预排序的思想,通过提前开辟空间计算并存储排序后的预处理向量,来减少 openform 中 PCG 求解模块预处理部分的离散访存,但其预处理向量不随计算循环更新,而非结构网格计算中大部分离散访存数组是随循环更新的。Ni等^[21]提出了一种基于排序思想的众核优化算法,减少了非结构网格计算中的随机访存,相比主核获得了 10 倍左右的加速。但该算法不具有普适性,仅针对稀疏求解问题。

(3)对于非结构网格隐式算法求解问题,Xu^[22]讨论并验证了隐式 LU-SGS 算法在 GPU 上的可行性。Chen等^[23]采用 RCM 和 Metis 网格重排序方法,实现了基于非结构网格点格式的隐式 LU-SGS 算法的 GPU 并行加速。但目前在国内众核超级计算机上尚未有该求解问题的相关研究。

本文基于国产众核架构的特点,首先提出了一种具有普适性的 CFD 非结构网格计算并行优化方法,建立了面向众核架构的数据重构模型,精简拓扑关系,提高了数据的局部性和可并行性;然后,利用一种基于信息关系预存的离散访存优化方法,解决了非结构网格中数据存放无序性导致的离散访存问题,实现了众核并行;最后,基于从核阵列流水线并行的思想,实现了非结构网格中具有强相关性的隐式算法的众核求解。

3 面向国产异构众核架构的 CDF 非结构网格并行优化方法

CDF 非结构网格计算在神威处理器上高效运行面临的挑战主要有:1)网格规模增大导致拓扑关系非常复杂,核心解算部分(包括网格面循环、网格单元循环)需要反复查表,不适应国产众核架构的特点,因为大量与拓扑相关的访存操作会大大降低计算访存比,不能完全发挥众核架构强大的计算性能;2)非结构网格固有的节点排列无序性导致的离散访存问题(在网格面循环中查询网格单元离散,在网格单元循环中查询网格面离散)会带来极低的访存带宽利用率;3)线性化方程求解 LU-SGS 方法具有极强的数据相关性,常规化并行方法难以实现。本节将针对不同的挑战提出相应的解决方法。

3.1 面向众核架构的数据重构模型

非结构网格间的拓扑关系复杂,并行计算需要建立各进程内网格单元、网格面和网格点信息以及进程间网格单元的通信关系,每个核心段的计算都要同时读取这些拓扑信息。对于众核架构来说,从核的局部存储空间非常有限,无法放下

这么多的拓扑信息,而从核直接访问主存又会造成很大的开销,如何解决拓扑关系复杂与从核局存内存空间有限之间的矛盾成为了非结构网格计算众核优化的首要问题。

图 2 为网格单元与网格面的拓扑关系示意图,很多核心计算中需要这些拓扑的相互映射,例如当对网格面循环的过程中求相应网格单元面数时(伪代码见算法 1),原程序需要先查表 FaceIndex 得到该网格面相邻的两个网格单元编号 CellNo,再查表 CellIndex 得到相应网格单元的类型,根据网格类型查表 CellnumofType 得到网格单元面数。从中可以发现,第二、三次查表本质上是一种离散访存,同时 FaceIndex, CellIndex 和 CellnumofType 都是与拓扑相关的数组。由于从核局存空间有限(64kB),显然在一般情况下无法将全部数组拷入 LDM,而直接访问主存又会导致很大的访问延迟;此外,反复查表会导致计算访存比降低,进而影响程序的运行效率。

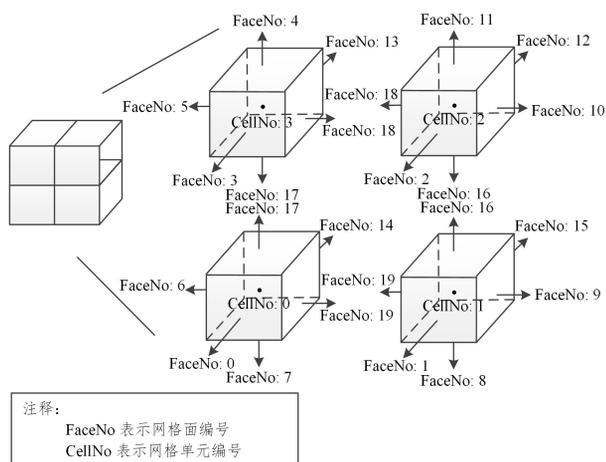


图 2 网格单元与网格面的拓扑关系示意图

Fig. 2 Diagram of topology between grid cells and surfaces

算法 1 面循环连续,网格单元变量访问离散

```

1. for i=is Face to ieFace do
2.   for im=0 to 1 do
3.     CellNo←FaceIndex[i], CellNo(im)
4.     CellType←CellIndex[cellNo], CellType
5.     Cellnum←CellnumofType[cellType]
6.   end for
7. end for

```

为了解决这一问题,本文建立了一个面向众核架构的数据重构模型,在预处理阶段重新建立一套新的数据结构,这些数据结构仅存计算所需的数据。具体模型如下:

(1)遍历各个核心计算部分,找到所有与拓扑相关的有效变量,如网格单元/面编号、网格单元/面类型、网格面面积、网格单元体积等;

(2)针对网格单元循环、网格面循环分别建立新的数据结构,分别取名为 SWFace 和 SWCell,结构成员包含步骤 1 中核心计算中与拓扑相关的有效变量;

(3)在预处理阶段,根据原拓扑关系,将相应的数据重新映射至新的数据结构中对应的成员变量。

(4)在核心计算部分用新的数据结构代替旧的数据结构。显而易见,经过数据重构,避免了反复查拓扑表,获取

网格单元面数仅需查一次表 SWFace 即可(如算法 2 所示),且原始查表次数越多,优化效果就越明显。更重要的是,通过数据重构,将与拓扑相关的离散访存转换成了连续访存。在后续众核并行时,仅需将本从核任务划分后(针对网格面/单元循环划分)对应的数据段拷至 LDM 即可,这样就解决了原拓扑关系复杂与从核局存空间有限的矛盾,且避免了高延迟的访主存操作。

算法 2 数据结构重构后,网格单元变量访问变连续

```

1. for i=isFace to ieFace do
2.   for im=0 to 1 do
3.     Cellnum←SWFace[i].CellnumofType[im]
4.   end for
5. end for
    
```

值得注意的是,本节提出的数据重构模型不仅有利于后续的众核并行,其本身也是一种纯主核的性能提升,且对各平台均适用,在 3.2 节将试验该模型性能提升的效率。

3.2 基于信息关系预存的离散访存优化方法

一些与拓扑关系相关的变量,已经通过 2.1 节中的方法将离散访存转化为连续访存,但还有一些是循环过程中计算出来的离散访存变量,不能通过数据重构来解决,以求网格面心密度 QyFace 为例,具体步骤如算法 3 所示。具体实现过程如下:在网格面循环中,通过查表 FaceIndex 得到网格面相邻的两个网格单元编号 CellNo,再利用 CellNo 得到两个网格单元的密度,利用这两个网格单元的密度求得网格面心密度 QyFace。在该过程中,网格单元密度 Qyx 对网格单元 CellNo 的查表属于离散访存,因为 CellNo 具有无序性。针对 Qyx 这一类型的离散访存数组,我们无法利用 2.1 节提出的数据重构的方法来解决,因为这些变量会随着迭代步进行更新,而不同于拓扑相关的变量,只要网格确定了,它们的值就确定了。本节提出了一种基于信息关系预存的离散访存优化方法,提高了数据的局部性,避免了高延迟的访主存操作。

算法 3 求网格面心密度,网格单元密度离散且不断更新

```

1. for i=isFace to ieFace do
2.   QyFace[i]←0
3.   for im=0 to 1 do
4.     CellNo←FaceIndex[i].CellNo[im]
5.     QyFace[i]←QyFace[i]+Qyx[CellNo][im]
6.   end for
7.   QyFace[i]←QyFace[i]/2
8. end for
    
```

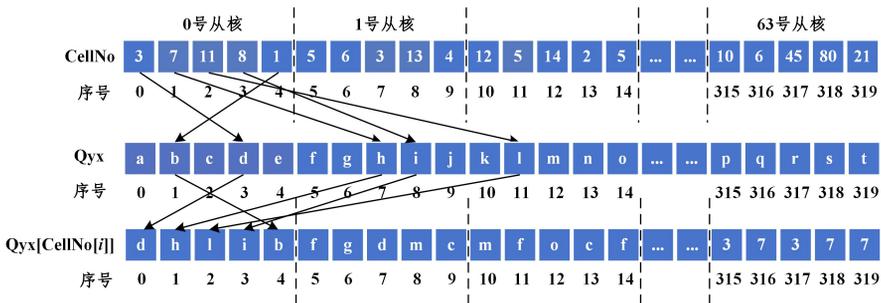


图 5 核组分布式存储示意图

Fig. 5 Diagram of distributed storage in core group

3.2.1 众核并行任务划分

对最外层的网格面循环进行众核并行任务划分,划分示意图如图 3 所示,采取均匀划分的方式以保证各众核之间计算任务的负载均衡。



图 3 众核任务划分示意图

Fig. 3 Diagram of many-core task partition

3.2.2 信息关系预存

进一步研究发现,这类问题有一个特点:虽然网格单元编号 CellNo 是离散的,但是这些信息关系,例如网格面周围存在的网格单元是确定的。即只要网格确定了, Qyx 的下标 CellNo(网格单元编号)的值就确定了。因此,该问题的实质是预知信息关系的离散访存问题。

基于此,本文通过信息关系预存,一些预先可知的拓扑关系,如本从核处理的网格面周围的网格单元、通信关系(包括后续需要从从核取得的数据等信息),以及重排序所需的映射关系,可以通过预处理得到并被放入主存。信息关系预存示意图如图 4 所示。

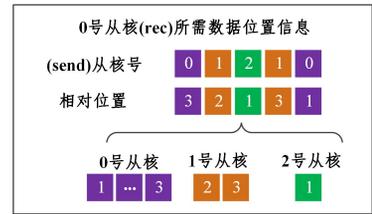


图 4 信息关系预存示意图

Fig. 4 Diagram of pre-storage of information relations

3.2.3 核组分布式存储

为了充分利用从核的内存空间,采用核组分布式的存储方式,将目标数组分布式地存储于各从核 LDM 上,如图 5 所示。这样相当于从核空间扩大了 64 倍,理论上可以存放一个 4 MB 大小的数组(LDM 空间为 64 kB)。假设要计算的变量为 30 个,数据类型为 Double,通过计算,理论上可以处理的网格大小约为 0.96n 万(n 为进程数),基本能够满足实际工程需求。与此同时,还有一个问题亟需解决,即对离散访存数组进行取值计算时,需从各个从核离散访问数据(从核间通信),如此细粒度的从核间通信也会严重影响加速性能。

3.2.4 基于快速片上通信的数组重排

国产异构众核架构在体系结构上支持从核内部寄存器通信^[16],在此基础上,利用 2.2.1 节中预存的信息关系,各个从核通信前将数据进行打包发送,这样将从核间细粒度通信转

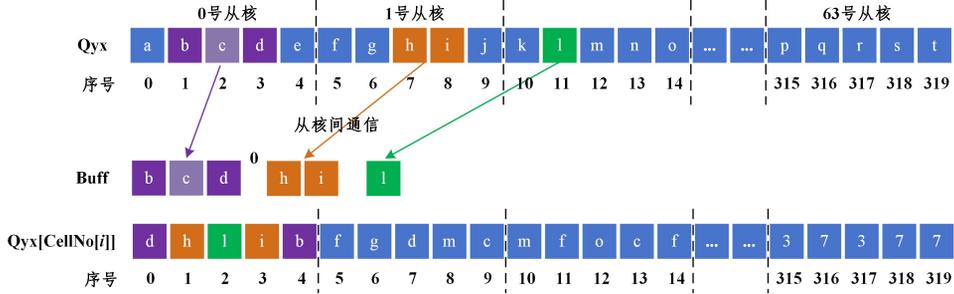


图 6 基于快速片上通信的数组重排示意图

Fig. 6 Diagram of data rearrangement based on fast on-chip communication

注释 1: 针对任意网格规模问题,若网格规模继续增大(大于 0.96n 万),超过了核组分布式存储的上限,则需要根据实际问题离散程度的高低采取不同的策略:1)若离散程度低,可以通过循环分段、分批调用的方式来解决;2)若离散程度高,则只能用 DMA 代替核组内的寄存器通信,LDM 只存计算所需数据,但由于 DMA 带宽低于寄存器通信带宽,因此会牺牲部分性能。

3.3 从核阵列流水线并行方法求解 LU-SGS 方程

由于非结构网格体系中网格排列是无序的,一般都采用有限体积法^[24]。通过对 Navier-Stokes 方程组(纳维-斯托斯方程)进行半离散处理,再对无粘项、粘性项和化学非平衡源项进行线性化,得到隐式基本方程为:

$$\begin{aligned} & [1 + \Delta \tilde{t} \sum_{i_F \in i_C} (\frac{1}{2} \rho (\tilde{A}_p)_{i_C} + (\lambda_{P_V})_{i_F \in i_C}) + \Delta \tilde{t} \mathbf{D}^n] \cdot \delta \mathbf{Q}_{i_C}^{n+1} + \\ & \Delta \tilde{t} \sum_{i_F \in i_C} (\tilde{A}_p^+)_{i_C n < i_C} \delta \mathbf{Q}_{i_C n < i_C}^{n+1} + \Delta \tilde{t} \sum_{i_F \in i_C} (\tilde{A}_p^-)_{i_C n > i_C} \delta \mathbf{Q}_{i_C n > i_C}^{n+1} = \Delta \tilde{t} \cdot \\ & \mathbf{RHS}_{i_C}^n \end{aligned} \quad (1)$$

其中, $\Delta \tilde{t} = \Delta t / Vol$, Δt 是局部时间步长, Vol 是控制体的体积; \mathbf{Q} 代表解向量; 下标 i_C 表示网格单元中心点的编号, 下标 $i_C n$ 代表与网格单元 i_C 相邻网格单元的中心点, $i_F \in i_C$ 表示属于网格单元 i_C 的网格面; $(\tilde{A}_p)_{i_C} = S_F (n_x \frac{\partial F}{\partial Q} + n_y \frac{\partial G}{\partial Q} + n_z \frac{\partial E}{\partial Q})$, 其中 S_F 是网格控制体 i_C 与其相邻控制体 $i_C n$ 共有控制面的面积, F, G, E 为 NS 方程中的 3 个通量项^[25], n_x, n_y, n_z 分别代表共有控制面的外法向单位量的 3 个分量; $\rho(\tilde{A}_p)$ 为矩阵 \tilde{A}_p 的谱半径; $(\lambda_{P_V})_{i_F \in i_C}$ 为粘性项特征值; $(\tilde{A}_p^+)_{i_C n < i_C}$, $(\tilde{A}_p^-)_{i_C n > i_C}$ 是特征分解后的 Jacobian 矩阵; $\mathbf{RHS}_{i_C}^n$ 为所有粘性通量、无粘通量、各源项的和; $\delta \mathbf{Q}_{i_C n < i_C}^{n+1}$ 表示第 $n+1$ 个迭代步的解向量的增量; \mathbf{D}^n 是关于源项的对角矩阵。

LU-SGS 分解算法由于收敛速度快、稳定性好,被广泛应用于工程流体计算中^[26]。首先,利用最大特征值方法把矩阵分解成上、下三角矩阵,将基本方程式(1)写成矩阵形式。

$$(\mathbf{D} + \mathbf{U} - \mathbf{L}) \delta \mathbf{Q} = \hat{\mathbf{R}} \quad (2)$$

其中, \mathbf{D} 代表由 $\delta \mathbf{Q}_{i_C}^{n+1}$ 的系数组成的对角矩阵, \mathbf{U} 表示由 $(\tilde{A}_p^+)_{i_C n > i_C}$ 项组成的上三角矩阵, \mathbf{L} 表示由 $(\tilde{A}_p^-)_{i_C n < i_C}$ 项组成的

化为粗粒度的通信,提高通信效率。然后,利用预处理得到的映射关系,将离散目标数组按照计算序列进行重排,如图 6 所示。在后续计算时,利用重排后的数组代替原来的数组,离散访存也就转化为了连续访存。

下三角矩阵, $\hat{\mathbf{Q}}, \hat{\mathbf{R}}$ 分别表示 $\delta \mathbf{Q}_{i_C}^{n+1}, \Delta \tilde{t} \cdot \mathbf{RHS}_{i_C}^n$ 组成的一维向量。

由于

$$\begin{cases} \mathbf{D} + \mathbf{U} - \mathbf{L} = (\mathbf{D} - \mathbf{L}) \mathbf{D}^{-1} (\mathbf{D} + \mathbf{U}) + \mathbf{L} \mathbf{D}^{-1} \mathbf{U} \\ \mathbf{D} + \mathbf{U} - \mathbf{L} = (\mathbf{D} + \mathbf{U}) \mathbf{D}^{-1} (\mathbf{D} - \mathbf{L}) + \mathbf{U} \mathbf{D}^{-1} \mathbf{L} \end{cases} \quad (3)$$

其中, $\mathbf{D} \approx \mathbf{O}(1), \mathbf{U} \approx \mathbf{O}(\Delta t), \mathbf{L} \approx \mathbf{O}(\Delta t)$, 将式(3)代入式(2)并忽略二阶无穷小得:

$$(\mathbf{D} - \mathbf{L}) \mathbf{D}^{-1} (\mathbf{D} + \mathbf{U}) \delta \mathbf{Q} = \hat{\mathbf{R}} \quad (4)$$

或

$$(\mathbf{D} + \mathbf{U}) \mathbf{D}^{-1} (\mathbf{D} - \mathbf{L}) \delta \mathbf{Q} = \hat{\mathbf{R}} \quad (5)$$

式(4)、式(5)均为线性化方程式(2)的不完全 LU 分解,两者在关于 Δt 的二阶精度是等价的,而且都可以分解成多步格式,但求解时空间推进方向相反。将其分解为两步格式:

$$\begin{aligned} \delta \mathbf{Q}^{n+1/2} &= \hat{\mathbf{R}} + \mathbf{L} \mathbf{D}^{-1} \delta \mathbf{Q}^{n+1/2} \\ \delta \mathbf{Q}^{n+1} &= \mathbf{D}^{-1} (\delta \mathbf{Q}^{n+1/2} - \mathbf{U} \delta \mathbf{Q}^{n+1}) \end{aligned} \quad (6)$$

式(6)属于递归方程,给定边界条件后,采用递推的方法对其进行求解,可以看出 LU-SGS 方法避免了复杂的矩阵求逆运算,极大地提高了计算效率。然而,从式(6)可以看出,该递推方法存在强相关性,某网格单元的解变量 $\delta \mathbf{Q}^n$ 与之前的分量的解相关(累加关系),难以对其进行并行优化。

下文结合国产异构众核架构的特点——从核阵列之间支持寄存器通信,提出了一种针对非结构网格 LU-SGS 强相关问题的并行求解方法,具体步骤如下:

(1) 将无相关性且计算量大的求解 Jacobian 矩阵的部分分离出来,然后利用常规方法放入众核进行并行加速。

(2) 其余具有强相关计算部分(主要是关于 $\hat{\mathbf{Q}}$ 的递推部分),采用从核阵列流水线并行的方法实现众核加速(从核心间的通信关系如图 7 所示)。

1) 将网格单元循环映射到从核阵列,每个从核接收来自同行及同列的结算结果,并完成本从核对应的网格单元 $\delta \mathbf{Q}$ 的计算,然后将计算结果用寄存器通信的方式发送给同行以及同列的下一个从核。

2) 第 63 号从核在完成本从核对应的网格单元 $\delta \mathbf{Q}$ 的计算

后,将计算结果发送给 0 号从核,以此类推,直至遍历所有的网格单元。

3)最后一个网格单元对应的从核将计算结果发送给主核。

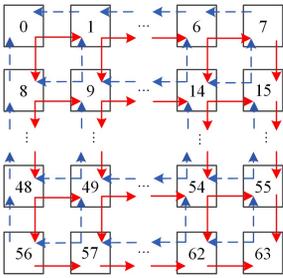


图 7 从核阵列的流水线并行

Fig. 7 Pipeline parallelism in core group

4 实验结果与分析

本文实验采用神威众核处理器和 Intel(R) Xeon(R) CPU E5-2695 v3 处理器进行测试,具体的软硬件配置如表 1 所列。

表 1 测试环境(软硬件配置)

Table 1 Test environment (software and hardware configuration)

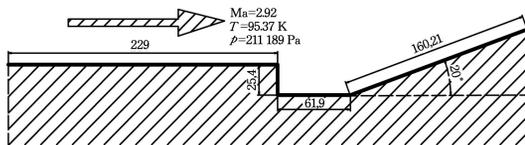
	SW26010 众核处理器	Intel(R) Xeon(R) CPU E5-2695 v3
主频(主核)/GHz	1.45	2.3
主频(从核)/GHz	1.45	—
内存/GB	8	8
页表大小/kB	8	4
编译器	swg++/sw5cc	GCC/ICC
众核编程语言	athread	—

4.1 正确性验证

本文的计算软件采用空气动力学研究中心自主研发的非结构网格计算软件 AHL3D-UNS^[27],选取了以下 3 个算例进行验证,分别代表了小、中、大 3 种问题规模。

4.1.1 Settles 二维超声速凹槽

Settles^[28]二维超声速凹槽实验模型如图 8 所示,实验是在 Princeton 大学的 20 cm * 20 cm 的高雷诺数风洞里进行的。实验精度较高,对压强和温度的校准显示精度在 1% 以内。因此,该模型被众多学者作为流体力学软件的验证算例。



注:图中的距离单位为 mm

图 8 超声速凹槽实验模型

Fig. 8 Experimental model of supersonic groove

入口指定为超声速入流,出口指定为超声速出流,壁面指定为无滑移绝热壁,上表面指定为远场边界条件。控制方程采用可压的 Navier-Stokes 方程,无粘项计算采用低耗散通量分裂格式(LDFSS),限制器格式采用 VanAlbada。网格规模为 5 万,运用本文提出的非结构网格并行优化方法对 AHL3D-UNS 进行优化。进行 20 万步后收敛,图 9 给出了超声速凹槽网格,图 10 给出了上壁面压力计算结果。

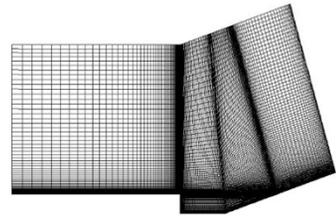


图 9 超声速凹槽计算网格

Fig. 9 Computing mesh of supersonic grooves

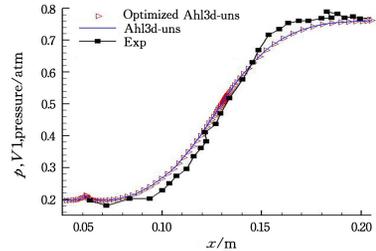


图 10 上壁面压力分布图

Fig. 10 Pressure distribution of upper wall

从图 10 可以看出,优化前后的计算结果都与实验结果吻合较好,从而验证了本文提出的非结构网格并行计算方法可成功应用于实际非结构网格的计算。

4.1.2 Hifire 直连式发动机

本节采用的模型为 NASA/Langley 研究中心的直连式发动机实验模型^[29],图 11 为该模型在流线方向上的二维截面图,其分为 4 个部分:设备过渡段、设备喷管、隔离段、燃烧室。所使用的喷管模拟飞行马赫数为 5.84,使用的燃料喷嘴 P1, S1 的位置如图 11 所示。燃料喷嘴共有 4 个,其中心点距离壁面为 1.27 cm,与发动机中轴线的距离为 2.54 cm,并且位于燃烧室突扩段的下游。S1 燃料喷嘴共有 4 个,直径为 0.2388 cm,其在 y 方向上等距排列,位于燃烧室收缩段的下游。来流条件如表 2 所列。

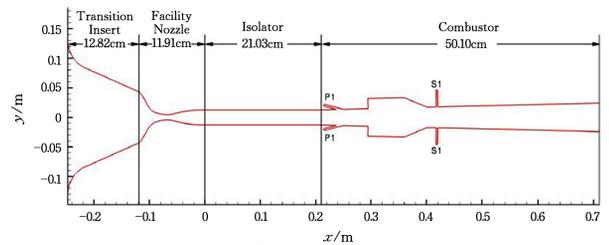


图 11 Hifire 直连式发动机模型示意图

Fig. 11 Diagram of Hifire direct-connected engine model

表 2 Hifire 发动机的来流条件

Table 2 Incoming flow conditions of Hifire engine

马赫数	2.51
总温/K	1550
总压/Mpa	1.48
静温/K	754.8
静压/kpa	81.5

控制方程、无粘项计算格式、限制器格式均与 3.1.1 节保持不变,湍流模型采用 k-omega 两方程模式,湍流 Prandtl 数取 0.89,湍流 Schmidt 数取 0.6,网格量共计 227 万。分别使用原始版本和用本文提出的非结构网格并行优化方法优化后的

软件版本对该算例进行计算,反应流 10 万步时收敛,计算结果如图 12 所示。

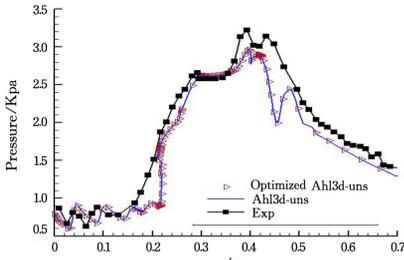


图 12 反应流壁面压力分布

Fig. 12 Pressure distribution of wall with reflecting flow

从计算结果可以看出,两种版本计算得到的壁面压力分布与实验值趋势一致,能够较好地吻合,从而验证了本文提出的非结构网格并行计算方法的正确性。

4.1.3 某超燃冲压发动机机燃烧室

超燃冲压发动机燃烧室的构型如图 13 所示,该燃烧室由等直段、扩张段和尾部喷管组成,上壁面含两个凹槽,下壁面有一个凹槽,4 个交错支板插入在燃烧室中截面上。来流条件如表 3 所列。

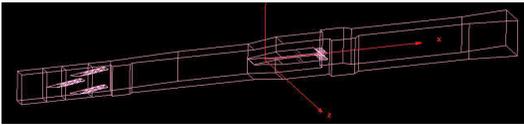


图 13 某超燃冲压发动机燃烧室计算网格

Fig. 13 Computational grid of a scramjet combustor

表 3 某超燃冲压发动机来流条件

Table 3 Incoming flow conditions of a scramjet combustor

马赫数	3.0
总温/K	1658
总压/Mpa	1.629
静温/K	663
静压/kpa	83

控制方程、无粘项计算格式、限制器格式仍与 3.1.1 节保持不变,湍流采用 TNT 模型,通量计算采用 AUSMPW+ 格式,时间推进采用 LU-SGS 方法,网格量共计 1.02 亿。分别使用原始版本和用本文提出的非结构网格并行优化方法优化后的软件版本对该算例进行计算,反应流 30 万步时收敛。图 14 给出了燃烧流场壁面压力沿程分布的情况。

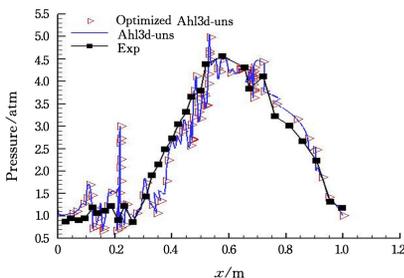


图 14 燃烧流场沿壁面压力分布

Fig. 14 Pressure distribution of combustion flow field along

the wall

从计算结果来看,优化前后的计算结果都与实验结果吻

合较好,进一步验证了本文提出的非结构网格并行计算方法的正确性。而在燃烧室入口,计算出的预测压力分布较实验值偏高,这是因为此处有 3 个支板压缩来流,导致激波/激波干扰、激波/边界层干扰现象较为显著,若采用更精细的网格配合合适的数值格式和湍流模型便能防止这一现象的出现。

4.2 加速性能测试

采用与 4.1.2 节中 227 万网格规模的 Hifre 直连式发动机模拟作为本小节的测试算例,依次采用本文提出的数据重构、离散访存优化、流水线并行等优化策略,得到的运行时间及性能提升如图 15 所示。

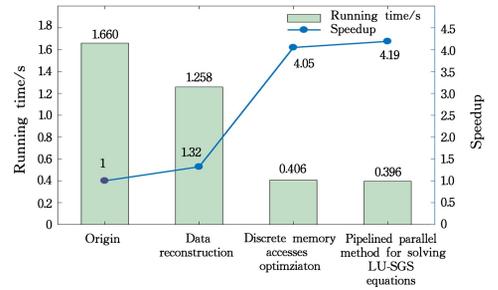


图 15 依次应用不同优化方法后的运行时间与加速比

Fig. 15 Running time and speedup after applying different optimization methods in sequence

首先本小节完成了对重构模型的验证(主核),由实验结果可知,通过针对性的数据结构重组,程序运行效率提高了 30%(这里的效率提升是纯主核的提升,不仅对众核有效,而且对各平台均适用)。此外,本文提出的数据重构模型减少了不必要的离散查表,更重要的是将离散访存转换成了连续访存,为后续的众核优化打下了基础。

接着,采用离散访存优化与众核并行后,非结构网格计算的整体运行时间缩短到 0.406 s,整体加速比达到 4.05 倍;实验结果验证了本文提出的基于信息关系预存的离散访存优化方法的高效性。

最后,采用流水线并行后,解方程部分的性能提升了 2.7 倍,非结构网格计算的整体加速比达 4.19 倍。而解方程部分对整体程序的加速贡献较小,这是因为在本算例中解方程部分的占比较小,关于流水线并行方法性能评估将在 4.3 节进行详细阐述。

4.3 流水线并行求解 LU-SGS 方程的性能评估

本节将针对 4.1 节中 3 个不同的算例对流水线并行优化求解 LU-SGS 方程的优化性能进行测试,实验结果如表 4 所列。

表 4 不同算例下流水线并行优化方法的性能对比

Table 4 Performance comparison of pipeline parallel optimization method with different examples

测试算例	前两种策略优化后解方程部分的占比/%	加速比	整体性能提升
Settles	18.72	2.6	7.20
Hifre	6.64	2.7	2.46
某燃烧室	42.50	3.2	13.28

从实验结果可以看出,在不同算例中,LU-SGS 解方程部分的占比不同,因此对整体性能的提升程度有所区别。在

这3个算例中,整体性能提升为2.46%~13%。尤其对于该部分占比高的算例(例如某超燃冲压发动机燃烧室算例,占比为42.5%),优化后整体性能能够提升13.28%,具有较好的优化效果。此外,该部分的平均加速比为2.83,流水线并行加速效果有限,这是非结构网格中求解LU-SGS方程部分前后的相关性太强、流水效率有限所致,这将是我们下一步的工作。

4.4 Intel CPU与SW26010运行时间的对比

本节将非结构网格程序在SW26010的主核、众核以及在通用CPU上的运行结果进行性能对比,实验结果如表5所列。从实验结果可以看出,sw26010的众核版相比Intel CPU性能提升了1.2倍。

表5 Intel CPU与SW26010运行时间的对比

Hifire直连式 发动机模拟算例	版本	运行时间/s	加速比
算例 网格规模 (227万)	SW26010主核版	1.660	1.000
	SW26010众核版	0.396	4.190
	Intel(R) Xeon(R) CPU E5-2695 v3	0.475	3.492

虽然众核性能高于通用的Intel CPU,但非结构网格计算在从核上的计算效率低于Intel CPU(计算核心数大于Intel CPU)。通过测试分析,从核性能主要受以下两个方面的限制:

(1)SW26010从核空间大小(64 kB)远小于Intel CPU的cache大小,非结构网格数据需要反复导入;

(2)SW26010的DMA带宽低于Intel CPU的访存带宽,多次DMA操作拉低了众核计算的计算/访存比,使得整个非结构网格计算性能受限于访存。

4.5 可扩展性分析

本节对本文提出的面向国产异构众核架构的非结构网格并行优化方法进行了可扩展性分析。本小节采用4.1.3小节中描述的1亿网格规模的超燃冲压发动机燃烧室模拟作为测试用例,考虑到1亿网格规模的内存需求,实测得到最小运行规模为19500计算核心(300个核组,每个核组包含1个主核和64个从核),实验在最小运行规模(19500计算核心数)的基础上依次翻倍扩展至624000计算核心,实验结果如图16所示。

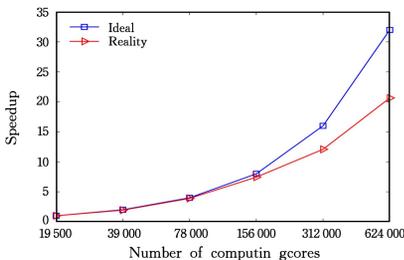


图16 加速比随计算核心数变化曲线图

Fig. 16 Speedup of unstructured-grid computing versus the number of computing cores

从实验结果可知,对于1亿网格规模的非结构网格计算,在应用本文提出的非结构网格并行优化方法后,计算核心数由1.95万扩展到62.4万,加速比达169.55,扩展效率达64.5%。

结束语 面对非结构网格计算在国产异构众核架构下移植与优化的众多挑战,本文结合国产众核体系结构的特点,提出了一种面向国产异构众核架构的非结构网格并行优化方法。首先,利用数据重构模型,打破了非结构网格众核化的壁垒;接着采用一种信息关系预存的离散访存优化方法,解决了非结构网格固有的离散访存问题,实现了大部分核心计算的众核并行;最后运用从核阵列流水线并行的思想解决了强相关线性化方程的众核求解。

基于神威太湖之光平台,实现了非结构网格计算的MPI+众核两级并行。实验结果表明,相比主核版本,整体获得了4.19倍的加速;相比通用的Intel CPU,性能提升了1.2倍。计算核心扩展至16384时加速比为169.55,扩展效率达64.6%。

同时,本文提出的非结构网格优化方法具有一定的普适性,可适用于其他类似应用课题,如分子动力学、多体运动等。

此外,还应注意,从核阵列流水线并行方法求解LU-SGS方程部分还没有完全发挥出众核高效的计算能力,其性能瓶颈主要在于数据强依赖导致这部分程序的并行度低,非结构网格LU-SGS隐式求解固有的计算访存比低,从而导致流水效率不高。因此,下一步工作将从以下两个方面进行深入研究:(1)通过层级划分充分挖掘程序中隐藏的并行性;(2)通过网格重编号提高数据的局部性,提高DMA数据的传输效率。

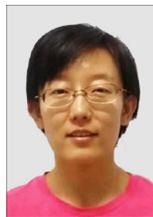
参考文献

- [1] LIN C L, TAWHAI M H, MCLENNAN G, et al. Computational fluid dynamics[J]. IEEE Engineering in Medicine & Biology Magazine, 2009, 28(3): 25-33.
- [2] XU K, MATHEMATICS D O. Direct modeling for computational fluid dynamics[J]. Acta Mechanica Sinica, 2015, 1(1): 303-318.
- [3] XU C F, DENG X G, ZHANG L L, et al. Parallelizing a High-Order CFD Software for 3D, Multi-block, Structural Grids on the TianHe-1A Supercomputer[C]// International Supercomputing Conference. Berlin, Heidelberg, 2013: 26-39.
- [4] CORRIGAN A, CAMELLI F, LOHNER R, et al. Running unstructured grid based CFD solvers on modern graphics hardware[J]. International Journal for Numerical Methods in Fluids, 2011, 66(2): 221-229.
- [5] ABBRUZZESE G, GÓMEZ M, CORDERO-GRACIA M, et al. Unstructured 2D grid generation using overset-mesh cutting and single-mesh reconstruction[J]. Aerospace Science & Technology, 2018, 78: 637-647.
- [6] JAHANDARI H, BIHLO A. Forward modelling of geophysical electromagnetic data on unstructured grids using an adaptive mimetic finite-difference method[J]. Computational Geosciences, 2021, 25: 1083-1104.
- [7] CHEN S S, HUA Y, CAI F J, et al. Multi-dimensional dissipation strategy within advection upstream splitting methods in hypersonic flows[J]. Journal of Physics: Conference Series, 2021, 1786(1): 012050.
- [8] DLA C, MP A, RL A, et al. Tracer transport within an unstruc-

- tured grid ocean model using characteristic discontinuous Galerkin advection -ScienceDirect [J]. *Computers & Mathematics with Applications*,2019,78(2):611-622.
- [9] CAI X,ZHANG Y J,SHEN J,et al. A Numerical Study of Hypoxia in Chesapeake Bay Using an Unstructured Grid Model: Validation and Sensitivity to Bathymetry Representation [J]. *JAWRA Journal of the American Water Resources Association*, 2020,10:1-24.
- [10] FUJITA K,HORIKOSHI M,ICHIMURA T,et al. Development of Element-by-Element Kernel Algorithms in Unstructured Finite-Element Solvers for Many-Core Wide-SIMD CPUs: Application to Earthquake Simulation [J]. *Journal of Computational Science*,2020,45:1-11.
- [11] SHARMA V,ESWARAN V,CHAKRABORTY D,et al. Determination of optimal spacing between transverse jets in a SCRAMJET engine [J]. *Aerospace Science and Technology*, 2020,96:1-12.
- [12] LI F,LI Z H,XU J X. Research on Adaptation of CFD Software Based on Many-core Architecture of 100P Domestic Supercomputing System [J]. *Computer Science*,2020,47(1):24-30.
- [13] LI R,WANG X,ZHAO W B. A Multigrid Block LU-SGS Algorithm for Euler Equations on Unstructured Grids [J]. *Numerical Mathematics Theory Methods & Applications*,2008,1(1):1-25.
- [14] LI W,LUO L S. An implicit block LU-SGS finite-volume lattice-Boltzmann scheme for steady flows on arbitrary unstructured meshes [J]. *Journal of Computational Physics*,2016,20(2):503-518.
- [15] FU H H,LIAO J F,YANG J Z,et al. The Sunway Taihu Light supercomputer: system and applications [J]. *Science China (Information Sciences)*,2016,59(7):113-128.
- [16] LIN H,TANG X,YU B,et al. Scalable Graph Traversal on Sunway TaihuLight with Ten Million Cores [C] // 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS). IEEE,2017.
- [17] LIN J,XU Z,NUKADA A. Optimizations of Two Compute bound Scientific Kernel on SW26010 Manycore Processor [C] // Proceedings of the 46th International Conference on Parallel Processing. IEEE,2017.
- [18] DONGARR J. Sunway TaihuLight supercomputer makes its appearance [J]. *National Science Review*,2016,3(3):265-266.
- [19] LIU X,LU L S,CHEN D X,et al. Research on Pre-processing Methods of Unstructured Grids [J]. *Computer Science*,2012,39(3):308-311.
- [20] MENG D L,WEN M H,WEI J W,et al. Porting and Optimizing OpenFOAM on Sunway TaihuLight System [J]. *Computer Science*,2017,44(10):64-70.
- [21] NI H,LIU X. Unstructured grid many-core optimization technology based on Sunway • Taihulight [J]. *Computer Engineering*,2019,45(6):45-51.
- [22] XU T H. GPU implementation of compressible viscous flow numerical method based on unstructured mesh [D]. Nanjing: Nanjing University of Aeronautics and Astronautics,2016.
- [23] CHEN L,XU T H,TIAN S L. Research on GPU Acceleration of Implicit Schemes Based on Unstructured Grids [J]. *Computer System Application*,2018,27(5):238-243.
- [24] SINGH M,SINGH R,SINGH S,et al. Discrete Finite Volume Approach for Multidimensional Agglomeration Population Balance Equation on Unstructured Grid [J]. *Powder Technology*, 2020,376:229-240.
- [25] ZHOU S,WEI W,GUO X. Notice of Retraction Unstructured grid finite volume method for NS equation [C] // International Conference on Computer Application & System Modeling. IEEE,2010.
- [26] BOCHAROV A N,EVSTIGNEEV N M,RYABKOV O I. Fully implicit multiple graphics processing units' schemes for hypersonic flows with lower upper symmetric Gauss-Seidel preconditioner on unstructured non-orthogonal grids [J]. *Journal of Physics: Conference Series*,2020,1698(1):1-13.
- [27] WANG L. Parallel Numerical Simulations of the Whole Scramjet Engine Flowfields on Unstructured grids [D]. Mianyan: China Aerodynamics Research and Development Center,2007.
- [28] HORSTMAN C,SETTLES G S,WILLIAMS D R,et al. A Re-attaching Free Shear Layer in Compressible Turbulent Flow [J]. *AIAA Journal*,1982,20(1):79-85.
- [29] BYNUM M,BAURLE R. A Design of Experiments Study for the HiFiRE Flight 2 Ground Test Computational Fluid Dynamics Results [C] // 17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference. 2013.



CHEN Xin, born in 1994, master, research assistant. His main research interests include computational fluid dynamics and high-performance parallel computation and application.



LI Fang, born in 1980, Ph.D, associate researcher. Her main research interests include computational fluid dynamics and high-performance parallel computation and application.

(责任编辑:柯颖)