# 大规模并行高阶矩量法的容错算法研究

陈 岩 林中朝* 张 玉 赵勋旺
(西安电子科技大学天线与微波技术重点实验室 西安 710071)

**摘 要**：基于超级计算机的大规模并行电磁计算对于解决实际工程中的复杂电磁难题具有重要意义，但超级计算机中由节点故障导致的进程崩溃事件的概率远远高于普通计算机。该文针对传统电磁计算难以有效应对进程崩溃的现状，提出一种高效的、适用于大规模并行高阶矩量法的容错算法。在现有并行高阶矩量法的基础上，基于“硬盘缓存”和“直接内存读取”设计高效率、高可靠性的现场保护算法，同时设计了高效的断点恢复算法。算法的有效性主要在于“固定的现场保护点”这一特点，它使得算法在有故障的情况下仍然可以正常有序地进行；而原算法每次碰到故障，则只能从头计算。数值仿真实验验证了容错算法在应对进程崩溃事件时的有效性，大幅提高了大规模并行高阶矩量法的可靠性。

**关键词**：超级计算机；并行矩量法；容错算法；现场保护；可靠性
**中图分类号**： TN820 **文献标识码**： A **文章编号**：1009-5896(2017)09-2245-07
**DOI**: 10.11999/JEIT161308

# Fault Tolerant Algorithm of Higher-order Method of Moments

CHEN Yan LIN Zhongchao ZHANG Yu ZHAO Xunwang
(*Key Laboratory of Antennas and Microwave Technology*, *Xidian University*, *Xi'an* 710071, *China*)

**Abstract**: The large scale parallel electromagnetic computation based on the supercomputer is of great significance for solving complicate electromagnetic problems in practical engineering. However, the probability of the process crash event caused by node failure in the supercomputer is much higher than that in the regular computer. Considering the incapable action for traditional electromagnetic computation to overcome the process crash event, an efficient fault-tolerance algorithm for large scale parallel high order Method of Moments (MoM) is proposed in this paper. According to the parallel higher order method of moments algorithm available, a scene protection algorithm and a scene recovery algorithm with high efficiency and reliability are designed, based on the "disk cache" and "direct memory access" technique. The efficiency of this algorithm lies on the feature of the "fixed site protection", which makes it possible for the algorithm to work normal and ordered even encountering crash failure, while the original algorithm can only restart from the beginning. The numerical simulations demonstrate the efficiency of the fault-tolerant algorithm in dealing with the process crash, which improves greatly the reliability of the large scale parallel high order MoM.

**Key words**: Super computer; Parallel Method of Moments (MoM); Fault-tolerance algorithm; Scene protection; Reliability

## 1 引言

作为电磁特性仿真中最精确的数值方法，矩量法(Method of Moments, MoM)可以有效处理各种复杂电磁问题[1]。矩量法需要建立并求解矩阵方程：$\boldsymbol{AX}=\boldsymbol{B}$，其中 $\boldsymbol{A}$ 代表阻抗矩阵，它是一个复数稠密矩阵，$\boldsymbol{X}$ 代表未知的电流向量，$\boldsymbol{B}$ 代表已知的电压向量。采用直接法求解该矩阵方程的存储复杂度和计算复杂度分别为 $O(N^2)$和 $O(N^3)$，其中 $N$ 为方程未知量，这导致矩量法的内存需求和计算量会随着电磁目标的电尺寸增大而急剧增长[2,3]，难以在实际工程中得到有效应用。随着国内超级计算机的迅猛发展，现代高性能计算技术为这一难题提供了高效的解决途径。基于超级计算机硬件平台，利用大规模并行计算技术，将大型稠密矩阵分布存储、并行

收稿日期：2016-12-08；改回日期：2017-02-26；网络出版：2017-05-11
*通信作者：林中朝 zclin@xidian.edu.cn
基金项目：国家自然科学基金(61301069)，教育部新世纪优秀人才支持计划(NCET-13-0949)，中央高校基本科研业务费(JB160218)，国家 863 计划项目(2012AA01A308)
Foundation Items: The National Natural Science Foundation of China (61301069), The Program for New Century Excellent Talents in University of China (NCET-13-0949), The Fundamental Research Funds for the Central Universities (JB160218), The National 863 Program of China (2012AA01A308)

求解[4−6]，在解决矩量法的存储问题的同时大大加快了矩量法的求解速度，使得这一精确方法越来越具有实际工程意义。

为满足现代复杂大型电磁工程对计算资源的需求，国内超级计算机系统不断扩大和更新，其通常包含成千上万的计算节点，如国家超级计算济南中心的“神威蓝光”超级计算机[7]，包含 8700 多个计算节点，CPU 核数高达 10 万以上；国家超级计算广州中心的“天河 2 号”超级计算机[8]，包含将近 40 万 CPU 核；最近获得“戈登贝尔奖”[9]的“太湖之光”超级计算机[10]，更是包含了数百万处理器核。在如此庞大的计算机集群中开展大规模电磁计算，节点故障时有发生，以前很多小概率故障也变得经常发生。然而，传统并行电磁计算难以有效应对硬件或进程崩溃的状况，任何一个硬件错误和进程崩溃都将导致整个计算任务的失败，所有的计算都需要重新开始，这会造成计算资源的严重浪费，对于大规模并行电磁计算极其不利。前期工作中，笔者在“神威蓝光”、“天河 2 号”等超级计算机平台中成功开展了大规模并行高阶矩量法的移植、测试与仿真工作，并行规模达到 20 万 CPU 核，并取得许多工程应用成果[11−14]。本研究在总结前期工作的基础上，针对并行高阶矩量法的特征，研究其大规模并行时的容错算法，以解决计算过程中可能出现的进程崩溃问题，确保计算任务的最终完成。

## 2 并行高阶矩量法的容错算法

相对于传统的 RWG 基函数矩量法，高阶矩量法具有较大的优势。一方面，高阶面片可以在保证获得更高拟合精度的同时采用更大的面元离散目标物体表面，从而减少面元的数目，进而减少未知量的数目；另一方面，高阶基函数可以用更少的未知量来拟合目标表面的真实电流[6,15]。由于高阶面片和高阶基函数理论对大规模并行容错算法的设计关联性较小，本文并不针对这一理论进行详细展开，读者可以参考文献[6]。

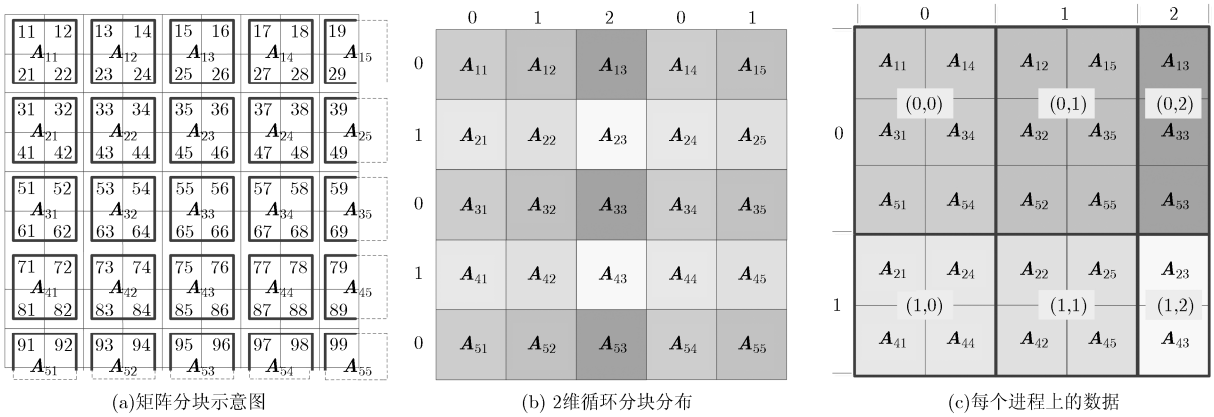
### 2.1 并行高阶矩量法

并行高阶矩量法主要分为矩阵并行填充和矩阵方程并行求解两方面，理论分析以及大量测试表明，矩阵并行填充所消耗的时间占矩量法整个运行过程的 10%以下，随着并行规模的不断扩大，这一比例甚至会低于 2%[3]。因此，并行高阶矩量法在填充过程中崩溃的概率极低，容错算法的设计主要针对矩阵方程并行求解进行。

影响并行算法效率的一个重要因素是负载均衡度。高阶矩量法矩阵方程的求解过程是一个计算、数据和通信三重密集的过程，为确保计算负载均衡、存储负载均衡和通信负载均衡的协调统一，需要将矩阵以“2 维循环分块分布”[3]的方式分布到各个进程上，所有进程并发执行，最终完成矩阵方程的求解。2 维循环分块分布是一种均衡策略的矩阵分布方式。设 $P_r\times P_c$ 个进程参与计算，形成的进程网格为 $P_r$ 行与 $P_c$ 列。矩阵大小为 $N\times N$，先将矩阵分为若干子阵 $\boldsymbol{A}_{ij}$，分块大小为 $m_b\times n_b$，则子阵 $\boldsymbol{A}_{ij}$ 被分布到行坐标为$\{(i\text{-}1) \bmod P_r\}$、列坐标为$\{(j\text{-}1) \bmod P_c\}$的进程中，其中 mod 表示取余数。举例而言，考虑把一个 9×9 的矩阵 $\boldsymbol{A}$ 采用 2×2 的分块大小分布到 2×3 的进程网格中，如图 1 所示。

图 1(a)中每一个子阵 $\boldsymbol{A}_{ij}$ 由实线圈出，最外面的虚线表示的是这些子阵未被填满。图 1(b)为矩阵在进程网格上的分布，矩阵左侧一列数字与上侧一行数字分别代表进程行坐标与进程列坐标，两个维度坐标的相应组合便确定了一个进程。子阵 $\boldsymbol{A}_{ij}$ 便分布在其所在的行和列所对应的数字确定的进程上。图 1(c)为分布后每个进程上拥有的数据。

在将矩阵按照 2 维循环分块分布的方式分布到




(a)矩阵分块示意图 (b) 2维循环分块分布 (c)每个进程上的数据

图 1 矩阵2维循环分块分布示意图

各进程上之后，施以必要的进程间通信，便可以对矩阵方程进行并行求解。矩阵方程并行求解主要分为 4 个步骤：列块分解，行交换，行块更新和余项更新。其中，列块是处于同一列上的分块矩阵，如图 2(a)中阴影部分所示；行块是处于同一行上的分块矩阵，如图 2(c)中阴影部分所示；余项矩阵是处于矩阵右下角的子矩阵，如图 2(d)中阴影部分所示。矩阵方程并行求解按照列块和行块逐步推进，每一步都由上一步的余项矩阵中取出新的列块进行计算，直到将余项矩阵完全消除。图 2 给出了矩阵方程并行求解算法推进到第 $k$ 步时的示意图。图中给出了每一过程的执行公式，此处不再展开讨论[3]。

### 2.2 并行容错算法研究

大规模并行矩量法容错技术主要涉及两个方面的内容，一是现场保护技术，一是断点恢复技术。所谓现场保护技术，是指将矩量法当前运行的状态、内存中的数据以文件的形式保存到硬盘中，并确认保存成功。所谓断点恢复技术，是指在矩量法由于集群故障停止后，程序由断点重新启动，恢复到停止前的状态，并确认恢复成功。容错算法可以充分利用已经完成的部分计算结果，大大减少由进程崩溃造成的时间浪费。

**2.2.1 现场保护技术**　现场保护技术涉及到如下几个方面的问题：(1)确定现场保护的时机，即何时对矩量法程序运行状态以及数据进行保存；(2)确定需要保存的信息，即为了能完全恢复矩量法程序在断点之前的运行状态和数据，最少需要保存哪些信息；(3)现场保护对程序整体性能的影响要尽量小；(4)如何确认现场保护成功。

对于现场保护的时机，一种直接的方案是：何时发生故障便何时进行现场保护。即机器发生故障后触发中断信号，中断处理在极短时间内将现场进行保护。这一过程如图 3(a)所示。考虑到矩量法巨大的数据量及其复杂的状态信息，除去程序设计复杂这一因素之外，中断程序几乎不可能在极短时间内完成现场保护，也没有有效的手段确认现场保护是否成功完成。另外一种合理的方案是：设定固定的现场保护点。即无论发生故障与否，当程序执行到特定的步骤之时都会进行现场保护；无论发生故障与否，在程序执行到特定的步骤之前都不会进行现场保护。由于这两个特点的存在，保护点既不能选的太多，也不能选的太少。选的太多，会影响程序的整体性能；选的太少，则两个保护点之间执行时间过长，保护点不能起到有效保护数据和程序状态的作用。本文选择固定保护点的方案。为了方便描述，本文假设保护点选在某次列块分解之前，这一过程如图 3(b)所示。

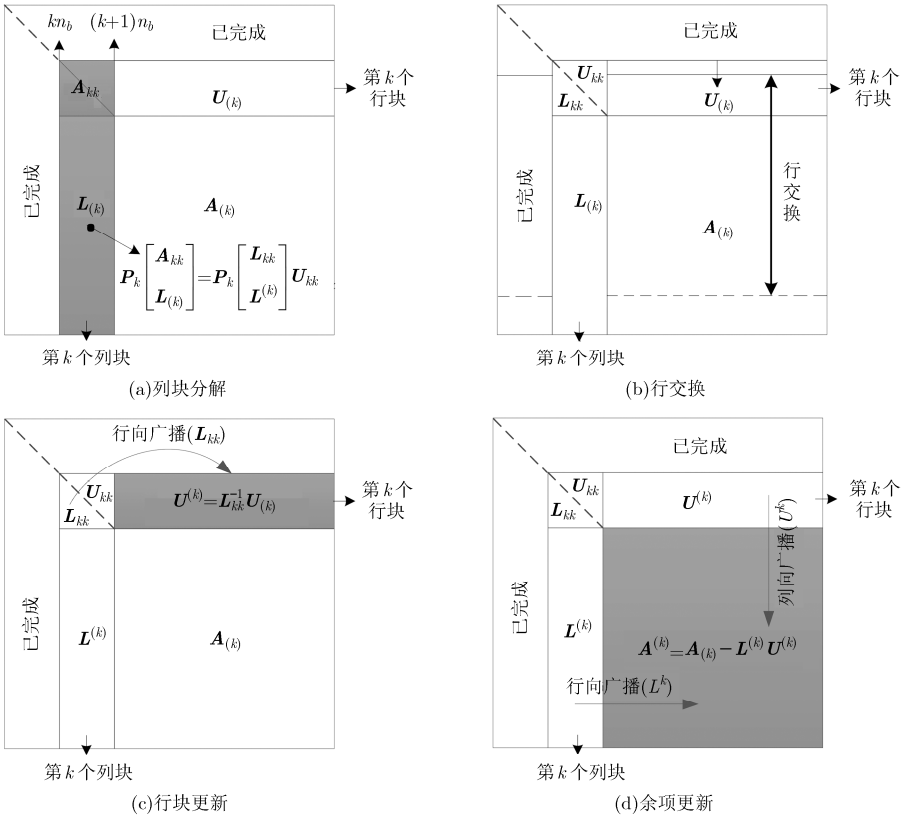
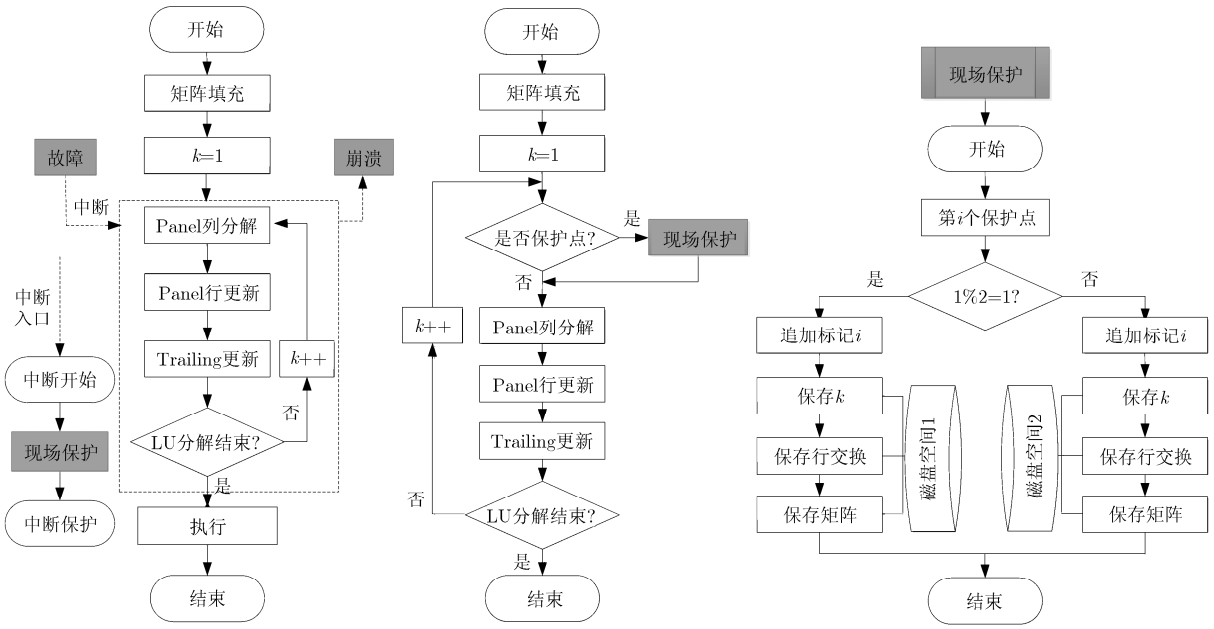



图 2 矩阵方程并行求解算法推进到第 $k$ 步时的基本过程




(a)中断方式的现场保护 (b)固定保护点方式的现场保护 (c)基于Buffer和DMA的现场保护子程序

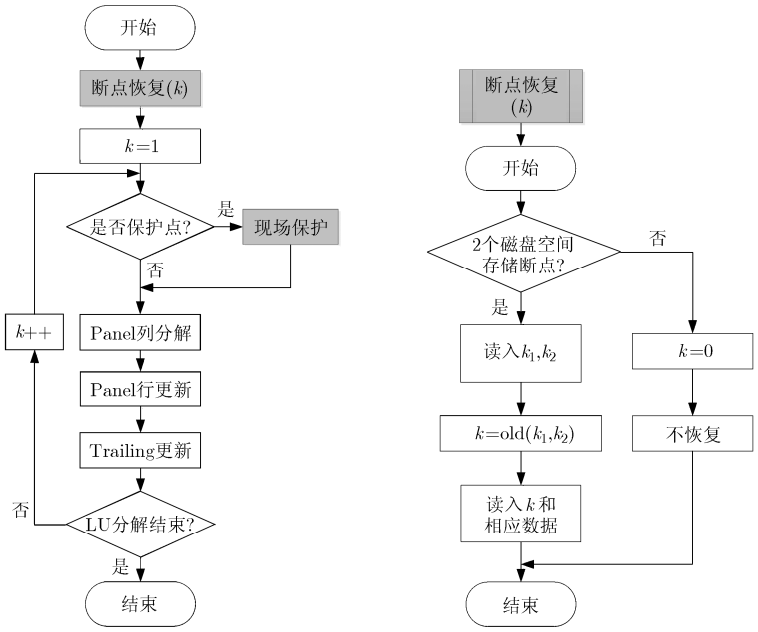
图 3 容错算法的现场保护技术

在选定现场保护点后，便可以确定需要保存的信息。由于保护点选在列块分解之前，因此需要保存的状态信息和数据有：当前矩阵方程求解算法推进的步数 $k$、整个阻抗矩阵和行交换信息(置换矩阵)。考虑到阻抗矩阵有一部分数据是已经分解完成的，这一部分子矩阵被保存之后便不会被再次改变，因此每次现场保护的时候这一部分矩阵不需要重新保存。这使得随着 $k$ 的递增，现场保护需要保存的数据量不断递减。为了节省硬盘空间，每一个新的现场保护点并非开辟新的存储空间，而是直接覆盖上一次现场保护的数据。

现场保护实质上是硬盘读写操作，这是一种低效率的操作，保护点的引入会降低程序的整体性能。为了能尽量减少这一不利影响，本文引入内存缓冲区(Buffer)以及直接内存访问(DMA)设计方案，将CPU 由硬盘读写操作中释放出来，合理设计程序流程，加速现场保护的过程。

引入内存缓冲区和直接内存访问后，确定现场保护的可靠性变得非常困难：不采用二者时，只要程序由现场保护子程序返回，则可确定现场保护成功完成；采用二者后，程序由现场保护子程序返回并不意味着现场保护过程完成。本文由以下两点确保现场保护的可靠性：(1)在矩阵的第 1 个元素之前和最后一个元素之后追加一个标记，此标记只与保护点相关，若此两个标记都被输出到硬盘上，则认为本次现场保护成功。(2)开辟两个相同的硬盘空间；在第 1 次现场保护时将数据保存在磁盘空间 1 上；第 2 次时保存在磁盘空间 2 上；第 3 次时又保存在磁盘空间 1 上，覆盖第 1 次的数据；依次类推。这样既能防止算法在现场保护时出现故障，从而破坏掉已保存的数据；同时双磁盘空间使得每一时刻磁盘上都保存有最新的 2 次现场保护的数据，且其中至少有一组(较早的那一次现场保护)一定是完整的。这一过程如图 3(c)所示。

**2.2.2 断点恢复技术** 断点恢复技术关键在于两点，(1)确定断点；(2)将任务恢复到断点时的状态。在任务崩溃并重启后，程序首先监测两个磁盘空间上是否存在断点，如果只有一个磁盘空间存储了断点数据，说明任务没有运行到第 2 断点便已经崩溃，此时不能完全保证第 1 断点可用，程序重新执行；如果两个磁盘空间都存储了断点数据，说明任务至少运行到第 2 断点，此时首先判断哪一个磁盘空间存储的是最新的断点数据，然后舍弃此断点(最新的断点不一定可靠)，使用另一磁盘空间的断点数据(较旧的断点可靠)恢复任务。为了简化描述，将磁盘空间 1 中存储的信息“$k$”记为 $k_1$，磁盘空间 2 中存储的数据“$k$”记为 $k_2$，显然可以利用 $k_1$ 和 $k_2$ 的大小判断哪一个断点是最新断点：若 $k_1 > k_2$，则磁盘空间 1 中存储的是最新断点；反之，则磁盘空间 2 中存储最新断点。在确认断点之后，程序首先读入 $k$(实际为 $k_1$ 或 $k_2$ 之一)，其次读入行交换信息，最后读入整个矩阵，当整个矩阵读入完成后，断点恢复成功，之后程序按照正常流程继续执行即可。图 4 给出了容错算法的断点恢复技术流程图。




(a)现场恢复方式　　(b)断点恢复子程序

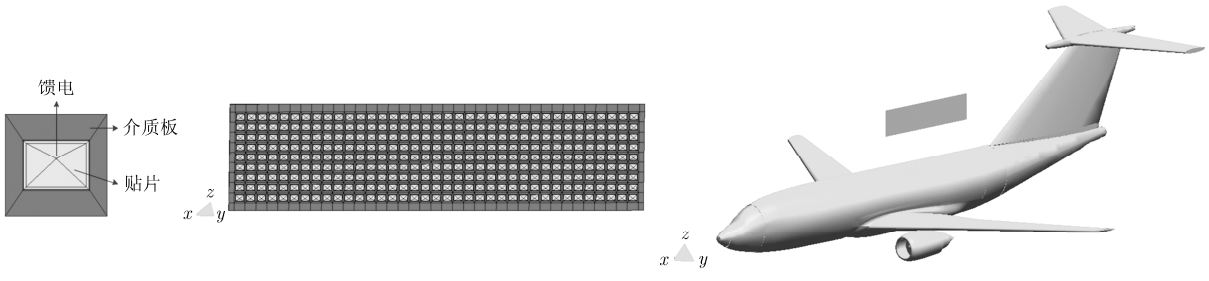
图 4 容错算法的断点恢复技术

由于程序是并行的，在现场保护时，所有进程都必须根据需要备份的数据和信息确定自身需要保存的数据和信息。进程之间是相对独立的，每个进程可以独立存储各自的数据和信息。在任务重新启动时，另外，断点恢复只需要在任务启动的时候执行一次即可，而现场保护在任务运行过程中会执行多次。

## 3　数值算例

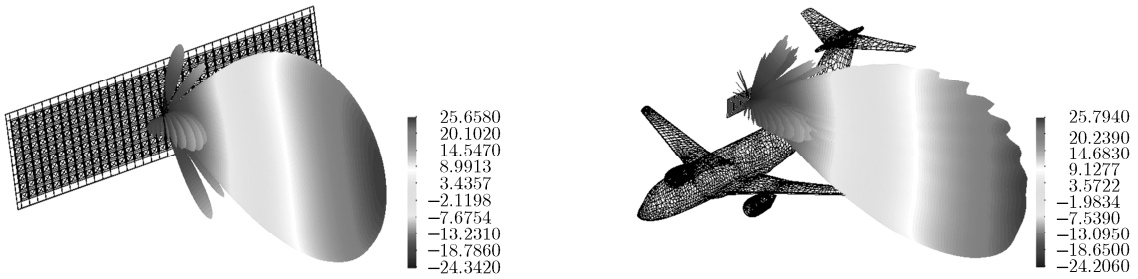
本例对机载微带相控阵进行电磁仿真，分析飞机平台对相控阵辐射特性的影响。微带阵列工作频率为 440 MHz，微带介质基板相对介电常数为 $\varepsilon_r = 4.5$，厚度为 18 mm。阵列单元模型和阵列模型如图 5(a), 5(b)所示，阵列单元数为 37×9。将天线阵列架设到飞机平台上，机身长 49.5 m，翼展 50 m，高 13.8 m，如图 5(c)所示，飞机机头沿+$x$ 方向，机翼平行于 $xoy$ 面。本文采用的计算平台共有 170 个计算节点，每个节点配置两路 Xeon E5-2692V2 2.2 GHz 12 核的 CPU, 64 GB 内存，节点之间使用 Mellanox FDR 56 Gb/s InfiniBand 实现互连。本例使用其中的 150 个节点，即 3600 CPU 核，进行仿真。为了考察容错算法应对进程崩溃的有效性和高效性，在测试过程中人为对节点进行宕机，比较并分析原算法和容错算法的运行特征。测试分为 6 组，分别设置 0, 1, 2, 3, 4, 5 个故障，其中容错算法每隔约 300 s 进行一次现场保护。程序完整运行时间大约为 1800 s，对于非容错算法，故障点都设置在程序运行后大约 700～1100 s 处；对于容错算法，故障点的设置相对随意，对运行时间影响较小。测试结果如表 1 所示。图 6 给出了微带天线阵列的辐射方向图，以及将阵列安装到机载平台之后的辐射方向图。从图中可以看出机载平台对阵列方向图的干扰。

由表 1 可以看出，当没有进程崩溃时，由于容错算法会执行多次现场保护子程序，因此造成了约 2%的性能损失。当有进程崩溃时，原算法只能从头




(a)微带天线单元　　(b)阵列仿真模型　　(c)平台-阵列一体化仿真模型

图 5 机载平台微带天线阵列仿真模型




(a)阵列本身 (c)平台-阵列一体化

图 6 机载平台微带天线阵列受扰方向图

**表 1 容错算法和原算法测试结果**

| 测试组 | 故障个数 | 采用算法 | 现场保护时间(s) | 断点恢复时间(s) | 总计算时间(s) | 效率提升(%) |
|---|---|---|---|---|---|---|
| 1 | 0 | 原算法 | 0.00 | 0.00 | 1795.25 | – |
| | | 容错算法 | 35.52 | 0.00 | 1832.94 | −2.10 |
| 2 | 1 | 原算法 | 0.00 | 0.00 | 2695.88 | – |
| | | 容错算法 | 36.48 | 8.72 | 1845.34 | 31.55 |
| 3 | 2 | 原算法 | 0.00 | 0.00 | 2872.41 | – |
| | | 容错算法 | 38.22 | 17.11 | 1852.23 | 35.52 |
| 4 | 3 | 原算法 | 0.00 | 0.00 | 3590.51 | – |
| | | 容错算法 | 34.69 | 27.70 | 1861.72 | 48.15 |
| 5 | 4 | 原算法 | 0.00 | 0.00 | 4308.78 | – |
| | | 容错算法 | 38.63 | 39.36 | 1875.11 | 56.48 |
| 6 | 5 | 原算法 | 0.00 | 0.00 | 5463.81 | – |
| | | 容错算法 | 36.43 | 44.53 | 1881.27 | 65.57 |

开始计算，而容错算法可以由保护点继续计算，即便是考虑现场保护和断点恢复的开销，容错算法带来的效率提升仍然十分可观。随着崩溃点的增多，容错算法比原算法的优势体现的更为明显。值得注意的是，在并行规模更加庞大的超级计算机上，如果系统的平均稳定运行时间小于电磁计算的所需时间，则仿真基本上不可能完成；而利用具有容错算法的大规模并行高阶矩量法，只要现场保护点的时间间隔小于系统平均稳定运行时间，则可以顺利完成电磁场仿真运算。

## 4 结束语

本文针对大规模并行高阶矩量法难以有效应对超计算机系统中经常发生的进程崩溃问题，提出了一种高效、可靠的并行容错算法。该算法首先根据高阶矩量法并行求解的特点设置固定的现场保护点，并采用内存缓冲区和直接内存访问提高现场保护的效率，采用双磁盘空间和双标记确保现场保护成功；然后根据基于双磁盘空间实现可靠的断点恢复算法；最后考虑进程的并发执行，实现现场保护、断点恢复和并行算法的高效融合。利用机载微带天线阵列辐射方向图的仿真对算法进行了验证，实验结果表明，本文所提出的大规模并行高阶矩量法容错算法可以有效地应对进程崩溃问题，大大提高并行高阶矩量法的可靠性和高效性，进一步提升了并行高阶矩量法在实际工程应用中的价值。

## 参 考 文 献

[1] HARRINGTON R F. Field Computation by Moment Methods[M]. New York: IEEE Press, 1993.

[2] 王长清. 现代计算电磁学基础[M]. 北京: 北京大学出版社, 2005: 116–157.
WANG C. Computational Advanced Electromagnetics[M]. Beijing: Peking University Press, 2005: 116–157.

[3] 张玉, 赵勋旺, 陈岩, 等. 计算电磁学中的大规模并行矩量法[M]. 西安: 西安电子科技大学出版社, 2016: 112–171.
ZHANG Y, ZHAO X, CHEN Y, *et al*. Massively Parallel Method of Moment in Computational Electromagnetics[M]. Xi'an: Xidian University Press, 2016: 112–171.

[4] 张玉, 王萌, 梁昌洪, 等. PC 集群系统中 MPI 并行矩量法研究[J]. 电子与信息学报, 2005, 27(4): 647–650.

ZHANG Y, WANG M, LIANG C H, *et al*. Study of parallel MoM on PC clusters[J]. *Journal of Electronics & Information Technology*, 2005, 27(4): 647–650.

[5] 徐晓飞, 曹祥玉, 高军, 等. 基于矩量法的电大目标RCS核外并行计算[J]. 电子与信息学报, 2011, 33(3): 758–762. doi: 10.3724/SP.J.1146.2010.00519.
XU X F, CAO X Y, GAO J, *et al*. Parallel out-of-core calculation of electrically large objects' RCS based on MoM [J]. *Journal of Electronics & Information Technology*, 2011, 33(3): 758–762. doi: 10.3724/SP.J.1146.2010.00519.

[6] Zhang Y and Sarkar T K. Parallel Solution of Integral Equation Based EM Problems in the Frequency Domain[M]. Hoboken, NJ: Wiley-IEEE, 2009: 107–136. doi: 10.1002/9780470495094.

[7] 国家超级计算济南中心: 中心概况/Center Overview[OL]. http://www.nsccjn.cn/node/42.jspx. 2016.11.

[8] 国家超级计算广州中心: 产品中心[OL]. http://www.nscc-gz.cn/Product/HighPerformanceComputingService/ServiceCharacteristics.html?. 2016.11.

[9] Association for computing machinery: ACM gordon bell prize [OL]. http://awards.acm.org/bell/year.cfm. 2016.11.

[10] 国家超级计算无锡中心: 硬件资源[OL]. http://www.nsccwx.cn/soft1.php?word=soft&i=46. 2016.11.

[11] 林中朝, 陈岩, 张玉, 等. 国产CPU平台中并行高阶矩量法研究[J]. 西安电子科技大学学报, 2015, 42(3): 43–47. doi: 10.3969/j.issn.1001-2400.2015.03.008.
LIN Z, CHEN Y, ZHANG Y, *et al*. Study of the parallel higher-order MoM on a domestically-made CPU platform[J]. *Journal of Xidian University*, 2015, 42(3): 43–47. doi: 10.3969/j.issn.1001-2400.2015.03.008.

[12] ZHANG Y, LIN Z, ZHAO X, *et al*. Performance of a massively parallel higher-order method of moment code using thousands of CPUs and its applications[J]. *IEEE Transactions on Antennas and Propagation*, 2014, 62(12): 6317–6324. doi: 10.1109/TAP.2014.2361135.

[13] 林中朝, 陈岩, 张玉, 等. 高阶矩量法的超级电磁计算研究[J]. 科研信息化技术与应用, 2015, 6(4): 20–28. doi: 10.11871/j.issn.1674-9480.2015.04.003.
LIN Z, CHEN Y, ZHANG Y, *et al*. Study of super electromagnetic computing for higher-order MoM[J]. *e-Science Technology & Application*, 2015, 6(4): 20–28. doi: 10.11871/j.issn.1674-9480.2015.04.003.

[14] CHEN Y, ZHANG Y, ZHANG G, *et al*. Hybrid MIC/CPU parallel implementation of MoM on MIC cluster for electromagnetic problems[J]. *IEICE Transactions on Electronics*, 2016, 99(7): 735–743. doi: 10.1587/transele.E99.C.735.

[15] 王少刚, 关鑫璞, 王党卫, 等. 求解电场积分方程的高阶矩量法[J]. 电子与信息学报, 2007, 29(9): 2265–2268.
Wang S, Guan X, Wang D, *et al*. Solution of the electric field integral equation using higher-order method of moments[J]. *Journal of Electronics & Information Technology*, 2007, 29(9): 2265–2268.

陈　岩：　男，1990 年生，博士，研究方向为计算电磁学、大规模并行算法.

林中朝：　男，1988 年生，讲师，研究方向为计算电磁学.

张　玉：　男，1978 年生，教授，研究方向为计算电磁学、大规模并行算法.

赵勋旺：　男，1983 年生，副教授，研究方向为大型机载天线阵列分析.