

Efficient Generation of Parallel Spin-images Using Dynamic Loop Scheduling

Ahmed Eleliemy, Ali Mohammed, and Florina M. Ciorba

Department of Mathematics and Computer Science

University of Basel, Switzerland

Email: {ahmed.eleliemy, ali.mohammed, florina.ciorba}@unibas.ch

Abstract— Over the last decades, the high performance computing (HPC) systems underwent a significant increase in their processing capabilities. Modern HPC systems combine very large numbers of homogeneous and heterogeneous computing resources. Scalability is, therefore, an important aspect of scientific applications to efficiently exploit the massive parallelism and computing power of modern HPC systems. This work introduces a scalable version of the parallel spin-image algorithm (PSIA), called APSIA. The PSIA is a parallel version of the well known spin-image algorithm (SIA). The (P)SIA is used in various domains, such as 3D object recognition, categorization, and 3D face recognition. To the best of our knowledge, the scalability of the PSIA has not yet been studied. APSIA refers to the extended version of the PSIA that integrates various well known dynamic loop scheduling (DLS) techniques. Through loop scheduling and dynamic load balancing, this integration enables an *improved* and *scalable* execution of the PSIA on homogeneous and heterogeneous HPC systems. The present work: (1) Proposes APSIA, a novel flexible and scalable version of PSIA; (2) Showcases the benefits of applying DLS techniques for optimizing the performance of the PSIA; (3) Assesses the performance of the proposed APSIA by conducting several scalability experiments on more than 300 heterogeneous computing cores. The performance results are promising and show that using well known DLS techniques, the performance of the APSIA outperforms the performance of the PSIA by a factor of 1.2 and 2 for homogeneous and heterogeneous computing resources, respectively.

Index Terms—Spin-image algorithm; Static loop scheduling; Dynamic loop scheduling; Weak scalability; Strong scalability; Heterogeneous and homogeneous computing resources; Self scheduling; Guided self scheduling; Factoring

I. INTRODUCTION

Modern high performance computing (HPC) systems can be characterized along two main dimensions: a large number of computing resources and their heterogeneity. Efficiently exploiting HPC systems along these two dimensions represents a significant challenge for parallel applications. Increasing the number of computing resources assigned to a parallel application can reduce the execution time. However, such a reduction is not guaranteed due to management and communication overheads. Also, heterogeneity adds several constraints in managing the assigned computing resources. Executing parallel applications on heterogeneous resources requires that the effects of the lower performance computing resources do not dominate the performance of the other computing resources.

The scalability of parallel applications can be categorized into two main profiles: strong and weak [1]. Fixing the problem size and varying (increasing) the number of the computing resources assigned to the application is known as *strong scalability*. *Weak scalability* is defined as a linear variation of the problem size with the number of the computing resources assigned to the application. The relation between executing parallel applications with different problem sizes and on systems of different scales has a significant importance in studying the performance of these parallel applications. For instance, a specific number of computing resources R is allocated to two different applications: A and B . Application A may have a lower performance (longer execution time) than application B when they are individually executed on R resources. If the performance of application A outperforms the performance of application B , when executing on a large number of allocated computing resources R' ($R' > R$), then application A has a better strong scalability profile than application B . Therefore, from a scalability perspective, it is recommended to execute application A rather than application B on R' resources.

The parallel spin-image algorithm (PSIA) [2] is a parallel version of the well known spin-image algorithm (SIA) [3]. The SIA is widely used in different domains, such as face detection [4], object recognition [5], [6], 3D map registration [7], and 3D database retrieval systems [8]. The main limitation of the SIA is its computational time complexity and, consequently, its execution time. The PSIA [2] is introduced to overcome this limitation of the SIA. However, the PSIA [2] only employs a static load balancing technique to distribute the process of the spin-image generation among the available computing resources. Moreover, to the best of our knowledge, the scalability of the PSIA has not yet been studied.

Similar to most scientific applications, the main source of parallelism in the PSIA is a loop, which for PSIA consists of independent iterations. Efficient loop scheduling techniques are, therefore, needed for parallelizing and executing this loop on parallel computing systems. Among the loop scheduling techniques, the dynamic loop scheduling (DLS) techniques have been shown to be most effective in optimizing the execution of parallel loop iterations by scheduling them during execution [9]–[12].

The DLS techniques aim to balance the computational workload between the computing resources to decrease the execution time of the application. This work proposes a novel

version of PSIA, namely APSIA, that integrates a number of DLS techniques, such as self scheduling SS [13], guided self scheduling GSS [14], and factoring FAC [15].

APSIA is generic enough to integrate different DLS techniques. The performance of the PSIA and the performance of the APSIA are evaluated via different scalability experiments and on different homogeneous and heterogeneous computing resources. The achieved performance of the APSIA under different DLS techniques is studied.

The remainder of this work is organized as follows. In Section II, a background of the PSIA and the DLS techniques used in this work is provided. The most relevant work in the literature, concerning the performance optimizations of the SIA, is also reviewed in Section II. In Section III, the proposed APSIA that can use different DLS techniques is introduced. The experimental setup and the information needed to reproduce this work is presented in Section IV. In Section V, the results of executing the proposed APSIA are compared with the results of executing the PSIA on homogeneous and heterogeneous computing resources to derive their weak and strong scalability profiles, respectively. In Section VI, the conclusion of this work and the potential future work are outlined.

II. BACKGROUND AND RELATED WORK

A. Parallel Spin-Image Algorithm

The spin-image algorithm (SIA) was originally introduced in 1997 by Johnson [3]. It converts a 3D object to a set of 2D images which are considered as a shape descriptor for that 3D object. The crux of the SIA is the process of generating the 2D images. The spin-image generation process can be explained as a process of spinning a sheet of paper through a 3D object. When a sheet of paper spins around a certain oriented point through a 3D object, other oriented points of that object are pasted onto that sheet of paper. After a complete cycle around the oriented point, the spinning sheet of paper represents a spin-image generated at that oriented point. In Fig. 1, taken from [3], the spin-image generation process is illustrated using eight animated frames.

Three parameters characterize the SIA: W , B , and S , and are included in Table I. W denotes the number of pixels in a row or column of the generated spin-image and is similar to the width of the spinning sheet of paper. The SIA assumes square spin-images with equal widths and heights. B is a factor of the 3D mesh resolution, used to determine the storage capacitance of each cell on the spinning sheet of paper. Increasing B means that many oriented points will be pasted to the same cell on the spinning sheet of paper. Consequently, the effect of individual oriented points on the generated spin-image will be reduced. S is a constraint for the spin-image generation process. If the angle between np_i and np_j , the normal vectors of the two oriented points P_i and P_j , respectively, is greater than S , then the oriented point P_j does not contribute to the generated spin-image at P_i . In Fig. 2, θ is the angle between the two normal vectors np_i and np_j .



Fig. 1. Eight animation frames, taken from [3], represent an analogy of the spin-image generation process

TABLE I
GLOSSARY OF NOTATION

Symbol	Description
M	Number of oriented points
N	Number of spin-images $1 \leq N \leq M$
P_i	An oriented point with a known normal vector where a spin-image can be generated, $0 \leq i < M$
OP	Set of all oriented points that belong to a 3D object $\{P_i \mid 0 \leq i < M\}$
np_i	A 3D vector that represents the normal vector of P_i
θ	The angle between two normal vectors np_i and np_j , $0 \leq i, j < M$
W	The number of pixels in a row or column of the generated spin-image where the generated spin-image is assumed to be a square matrix
B	A factor of the 3D mesh resolution that is used to determine the storage capacitance of the generated spin-image, $0 < B \leq 10$
S	Maximum allowed angle θ between P_i and P_j , where P_j contributes to the generated spin-image at P_i
WO	The number of workers used to generate the spin-images
wo_k	A worker that represents an MPI rank, pinned to a certain computing resource (core), $1 \leq k \leq WO$

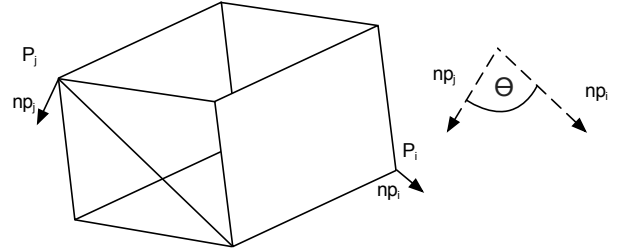


Fig. 2. A 3D object in which P_j does not contribute to the generated spin-image at P_i because θ is greater than S

The time complexity of the SIA is $O(NM)$. If N approximately equals M , 3D objects with more than $100K$ oriented points represent a significant challenge for the SIA in terms of its execution time. PSIA [2] exploits the inherent parallelism within the SIA where the calculation of each individual spin-image is independent from other spin-image calculations. The steps of the PSIA are listed in Algorithm 1.

There are two experimental setups for executing any implementation of Algorithm 1. The first setup is when the number of parallel computing resources, WO , used in the experiment equals N . In such a setup, each worker generates exactly one spin-image, i.e., according to Algorithm 1, it executes the code between Lines 5-20 *only once*. In practice, it is not always feasible for the number of workers WO to equal N , especially when N approximately equals M . The second setup is when WO is smaller than N . Each worker generates a certain number of spin-images proportional to the ratio of N

Algorithm 1: Parallel spin-image algorithm (PSIA) described in [2]

```

1 calculateSpinImages (W, B, S, OP, N)
  Input parameters : W: image width, B: bin size,
                      S: support angle,
                      OP: list of oriented points,
                      N: number of generated spin-images
  Output parameter: spinImages: list of generated
                      spin-images
2 spinImages = createSpinImagesList(N)
3 M = getLength(OP)
4 Parallel for  $i = 0 \rightarrow N$  do
5   tempSpinImage[W, W]
6   init(tempSpinImage)
7   P = OP[i]
8   for  $j = 0 \rightarrow M$  do
9     X = OP[j]
10     $np_i = \text{getNormal}(P)$ 
11     $np_j = \text{getNormal}(X)$ 
12    if  $\text{acos}(np_i \cdot np_j) \leq S$  then
13       $k = \left\lceil \frac{W/2 - np_i \cdot (X - P)}{B} \right\rceil$ 
14       $l = \left\lceil \frac{\sqrt{\|X - P\|^2 - (np_i \cdot (X - P))^2}}{B} \right\rceil$ 
15      if  $0 \leq k < W$  and  $0 \leq l < W$  then
16        tempSpinImage[k, l]++
17      end
18    end
19  end
20  add(spinImages, tempSpinImage)
21 end

```

divided by WO . This means that each worker executes the code between Lines 5-20 of Algorithm 1 *more than once*. In both experimental setups, the performance of the algorithm is dominated by the performance of the slowest worker. A worker can be the slowest performing worker in two cases: (1) It has a larger amount of computations than others and/or (2) It has lower processing capabilities than others.

B. Dynamic Loop Scheduling

In scientific applications, loops are, in general, one of the main sources of parallelism. Parallel loops are categorized as DOALL and DOACROSS loops [16]. The DOALL loops have no dependencies between their iterations while DOACROSS loops consist of iterations that are data-dependent. As shown in Algorithm 1, there are no dependencies between the iterations of the outer loop (Lines 4-21). Therefore, the PSIA is an example of a DOALL loop. In this section, the most common and successful dynamic loop scheduling (DLS) techniques for the DOALL loops are discussed. The DLS techniques are used to schedule loops with no dependencies between their iterations, or loops where most dependencies between

iterations can be eliminated via various loop transformations. Using DLS, the scheduling decisions are performed during the application execution time [17].

The DLS techniques considered in this work include SS [13], GSS [14], and FAC [15]. SS assigns a single loop iteration each time to a requesting computing resource. The main advantage of SS over other DLS techniques is its ability to achieve an optimized load balance between all processing elements. However, this advantage comes at a very high overhead. GSS divides the total number of loop iterations into variable size chunks of loop iterations. In each scheduling step, GSS divides the remaining loop iterations by the total number of processing elements. GSS is considered as a compromise between SS and STATIC, providing an acceptable load balance at an acceptable scheduling overhead. GSS has the disadvantage of overloading the first free and requesting computing resource with the first and largest chunk of iterations. The remaining loop iterations may not be sufficient to ensure a balanced execution among the computing resources. FAC was designed to handle iterations of variable execution time. It schedules the loop iterations in batches of P of equal sized chunks where P is the total number of computing resources.

The reason for selecting the three above-mentioned DLS techniques and STATIC for the present work is to cover a broad spectrum of the performance of the PSIA using the loop scheduling techniques. SS and STATIC represent the two extreme cases of the DLS techniques. STATIC has the lowest communication overhead and the lowest ability to balance the execution of the loop iterations among the workers. SS has the highest communication overhead and the highest ability to balance the execution of the loop iterations among the workers. The expected performance of GSS and FAC represent intermediate points between STATIC and SS. Further work is needed and planned as future work to include other more complex DLS techniques.

C. Related Work

In [2], an empirical approach was used to achieve the best performance of PSIA executing on a heterogeneous computing system that consisted of an Intel CPU and an Intel Knights Corner (KNC) co-processor. The main goal of the work in [2] was to achieve a load balanced execution of the algorithm between the 24 cores CPU and the 64 cores KNC. The approach taken in [2] *statically* divides the workload (the generation of spin-images) unequally in such a way that guarantees that the CPU cores and the KNC cores finish the execution at the same time. To perform such a static division of the generation of the spin-images, certain information regarding the time to generate each spin-image is required in practice. This information was obtained by generating each spin-image on the two available computing architectures. However, the obtained information was only valid for specific computing architectures and for the input data used. Motivated by the work in [2], the present work demonstrates the need for using *dynamic loop scheduling* within PSIA and extends it into APSIA. APSIA employs

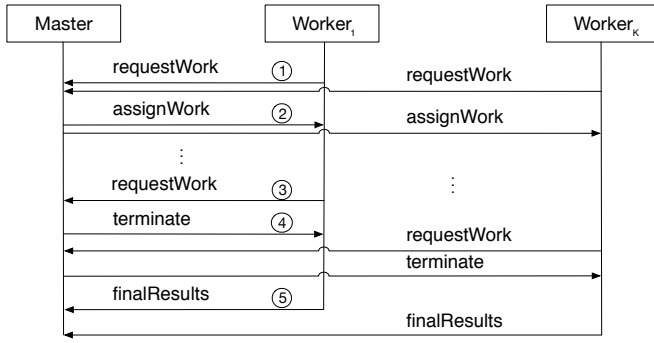


Fig. 3. Communication protocol between master and workers

dynamic loop scheduling to execute efficiently *both* on heterogeneous as well as homogeneous computing resources.

For distributed memory architectures (similar to the ones used in this work), the DLS techniques were integrated within a master-worker execution model [11], [12]. Without loss of generality, the present work differs from [11], [12] as follows: (1) The master is a dedicated resource and performs the DLS-based chunk calculations and the work assignment; (2) There is no communication or work reassignment among the workers; (3) The input data is initially replicated in the main memory of all workers; (4) The workers only send the results of calculating all chunks *after* they receive the termination signals from the master.

III. THE PROPOSED SCALABLE PSIA

The scalable version of PSIA, proposed in this work and denoted APSIA, is introduced next. The APSIA employs a master-worker execution model. As shown in Fig. 3, the master-worker communication protocol consists of five steps: (1) A free worker *requests* an amount of work (chunk of loop iterations); (2) The master *calculates* (according to the selected DLS technique) and *assigns* a chunk of loop iterations to the requesting worker; (3) When the worker finishes the assigned chunk of loop iterations, it notifies the master and *requests* another chunk of loop iterations; (4) If there are still unexecuted loop iterations, the master calculates and assigns a new chunk of loop iterations to that worker; otherwise it sends a *termination* signal; (5) When a worker receives a *termination* signal, it *sends back* the results of executing the assigned chunks to the master.

To integrate the master-worker execution model into PSIA, certain changes are needed to be made to Algorithm 1. The proposed algorithm is shown in Algorithm 2 in which, the code parts in blue font color (Lines 1, 2, and 3) represent the modifications required for Algorithm 1 to employ the master-worker execution model.

Recall from II-B that the current work differs from previous work [11], [12] as following: (1) The master is dedicated to handle the worker requests; (2) The workers do not communicate with each others; (3) The input data is replicated; (4) The results are collected from the workers at the end. These

Algorithm 2: Modification to the spin-image calculation for integration with the master-worker execution model and the DLS techniques

```

1  adCalculateSpinImages (W, B, S, OP, M, spinImages,
   start, end)
   Inputs : W: image width, B: bin size, S: support angle,
           OP: list of oriented points,
           M: number of oriented points,
           spinImages: list of spin-images to be filled
2  for imageCounter = start → end do
3    P = OP[imageCounter]
4    tempSpinImage[W, W]
5    init(tempSpinImage)
6    for j = 0 → M do
7      X = OP[j]
8      npi = getNormal(P)
9      npj = getNormal(X)
10     if acos(npi · npj) ≤ S then
11       k = ⌊ (W/2 - npi · (X - P)) / B ⌋
12       l = ⌊ (√(||X - P||2 - (npi · (X - P))2) / B ⌋
13       if 0 ≤ k < W and 0 ≤ l < W then
14         tempSpinImage[k, l]++
15       end
16     end
17   end
18   add(spinImages, tempSpinImage)
19 end
  
```

distinctions are made to more closely align with the earlier PSIA implementation and to allow a meaningful comparison with APSIA.

As discussed next in Section IV, the main memory of recent computing resources satisfies the memory requirements of dense 3D objects. Therefore, replicating the information of the 3D object and storing the generated spin-images on the worker side result in lightweight messages between the master and the workers. Moreover, a dedicated master resource offers rapid responses to the workers, especially when executing on large number of workers. The usefulness of the master-worker execution model and the integration of the communication protocol from Fig. 3 in APSIA is described as two pseudo code algorithms than can be found online¹.

IV. SETUP OF EXPERIMENTS

A. Input Data Set

As discussed in Section II-A, the time complexity of SIA is $O(NM)$. It is important to consider the 3D objects of high density regarding the number of 3D points. In Table II,

¹<https://c4science.ch/diffusion/3863>

TABLE II
3D OBJECTS IN THE 3D MESH WATERMARKING DATA SET [18]

Object	Approximate number of points ($\times 10^3$)
Cow	3
Casting	5
Bunny	35
Hand	37
Dragon	50
Crank	50
Rabbit	71
Venus	101
Horse	113
Ramesses	826

the objects of the 3D mesh watermarking [18] data set are presented. The 3D mesh watermarking data set consists of ten dense 3D objects. These 3D objects vary regarding the number of points from approximately $3K$ to approximately $826K$ points.

Out of the 3D objects in Table II, the *Ramesses* object is considered as the extreme case in terms of 3D points density for the APSIA. *Ramesses* object contains the largest number of oriented points, approximately $826K$, and is considered for comparing the performance of the proposed APSIA and the earlier PSIA. Similar to [2], the present work considers the three spin-image generation parameters W , B and S to be 5, 0.1 and 2π , respectively. In addition, the present work considers the number of generated spin-images N to be 10% of the total number of oriented points M (see Table I) [2].

B. Hardware Platform Specifications

Two different types of computing resources are used in this work to assess and compare the performance of the proposed APSIA and the earlier PSIA. The first platform type, denoted Type1, represents a two-socket processor (20 cores) Intel Xeon E5-2640 with 64 GB RAM. The second platform type, denoted Type2, is a standalone Intel Xeon Phi 7210 (64 cores) with 96 GB RAM. The platform types Type1 and Type2 are part of a computing cluster that consists of 26 nodes: 22 of Type1 and 4 nodes of Type2. All nodes are interconnected in a non-blocking fat-tree topology. The network characteristics are: Intel OmniPath fabric, 100 GBit/s link bandwidth, and 100 ns (for homogeneous resources) and 300 ns (for heterogeneous resources) link latency. This computing cluster is actively used for research and educational purposes. Therefore, only eight nodes of Type1 and four nodes of Type2 were dedicated to the present work.

C. Implementation and Execution Details

The Intel message passing interface library (Intel-MPI, version 17.0.1) was used to compile and execute the implementation of the proposed APSIA. The Intel-MPI library has the advantage of default pinning of operating system level processes to hardware cores (i.e., process pinning). Pinning a particular MPI process to a hardware core eliminates the undesired process migration that may be performed by the operating system during execution. Moreover, to examine the

performance of the DLS in one of the worst cases, all master-worker control and data messages exchanged (cf. Fig. 3) are implemented using MPI point-to-point communication primitives. The O3 compilation flag was used to compile the code for execution on Type1 nodes. In addition, the xCommon-AVX512 flag was used to compile the code for execution on Type2 nodes.

A user-specified machine file is used to map the MPI ranks to the computing resources (cores of nodes of Type1 and Type2). All computing resources are listed in the machine file in a certain order. This order indicates the MPI rank assigned to each computing resource during the execution of the application. Executing on homogeneous resources of Type1 or Type2 where all computing resources are similar, this order has no influence on performance. However, when executing on heterogeneous resources of Type1 and Type2, all computing resources of Type2 are listed in the machine file before computing resources of Type1. The rationale behind this listing is to enable the nodes with the largest number of cores (Type2) take the first MPI ranks. In the next section, the influence of this listing is presented and discussed. The master (MPI rank = 0) is always mapped to a dedicated computing resource. This computing resource is a core of a dedicated node of Type1. This dedicated computing resource is always written at the beginning of the machine file.

Each experiment has been executed fifteen times to obtain certain descriptive measurements, such as maximum, minimum, average, median, first, and third quartiles.

D. Reproducibility Information

To enable reproduction of this work, apart from the information in Sections IV-A, IV-B, and IV-C, the source code of the proposed APSIA is available upon request from the authors under the lesser general public license (LGPL). Moreover, the raw results are already available online². The code was compiled and executed using the Intel MPI version 17.0.1. All computing nodes use CentOS Linux release 7.2.1511 as operating system.

V. EXPERIMENTAL RESULTS AND EVALUATION

A. Performance of APSIA vs. PSIA on Homogeneous Computing Resources

In this section, the performance of the PSIA is compared to the performance of the proposed APSIA for two profiles: weak and strong scalability. As discussed in Section II-C, the PSIA *statically* divides and assigns the spin-image calculations to the available computing resources. In all experiments, the PSIA is referred to as PSIA-STATIC. APSIA-SS, APSIA-GSS and APSIA-FAC denote the proposed APSIA code parallelized with the three DLS techniques: SS, GSS, and FAC, respectively.

²<https://c4science.ch/diffusion/3863>

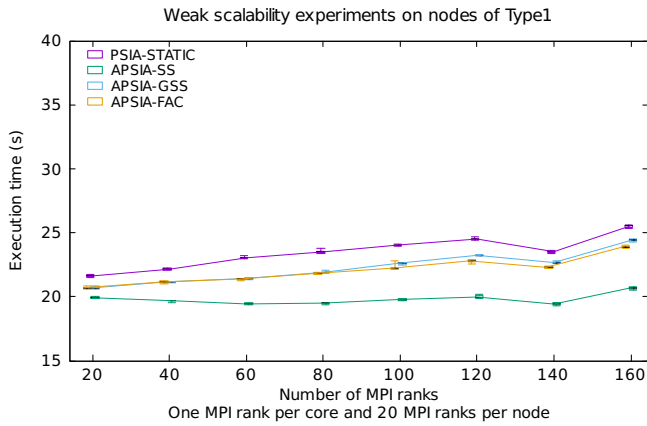


Fig. 4. Performance of the proposed APSIA and the earlier PSIA on homogeneous computing resources of Type1. The number of generated spin-images per computing node is $8K$.

1) *Weak Scalability*: For conducting weak scalability experiments, the number of the generated spin-images and the number of the computing resources are increased such that their ratio is kept constant at $8K$ spin-images per computing node. The number of the generated spin-images in this ratio represents approximately 1% of the total spin-images that can be generated from the *Ramesses* object. This work percentage is selected to result in a suitable, yet representative, execution time per experiment, given that each experiment has been executed fifteen times.

A comparison between the *parallel execution time* of the proposed APSIA and the PSIA achieved by executing them on different node counts of the two platform types is presented in Fig. 4 and 5. The execution time of the PSIA-STATIC is significantly higher than that of APSIA-SS, APSIA-GSS, and APSIA-FAC. For PSIA-STATIC on Type1 nodes, increasing the number of the generated spin-images from $8K$ to $64K$ (i.e., by a factor of 8) and increasing the number of the computing resources from 20 to 160 (i.e., by a factor of 8) result in an undesired performance degradation. Specifically, the execution time increased from 21 to 25 seconds, an almost 20% increase. The APSIA-SS did not exhibit such performance degradation. Specifically, the execution time increased from 20 to 20.5 seconds, an almost 1% increase. Similarly to the performance on Type1 nodes, for PSIA-STATIC on Type2 nodes, increasing the number of the generated spin images from $8K$ to $32K$ (i.e., by a factor of 4) and increasing the number of the computing resources from 64 to 256 (i.e., by a factor of 4) result in an undesired performance degradation. In particular, the execution time increased from 30 to 35 seconds, approximately a 17% increase. Executing the APSIA-SS on Type2 nodes resulted in poor performance compared with the execution on Type1 nodes. In particular, the execution time increased from 27.5 to 30 seconds, approximately a 9% increase. However, APSIA-SS still outperforms all other versions of the proposed APSIA.

Such a difference in the performance of different APSIA

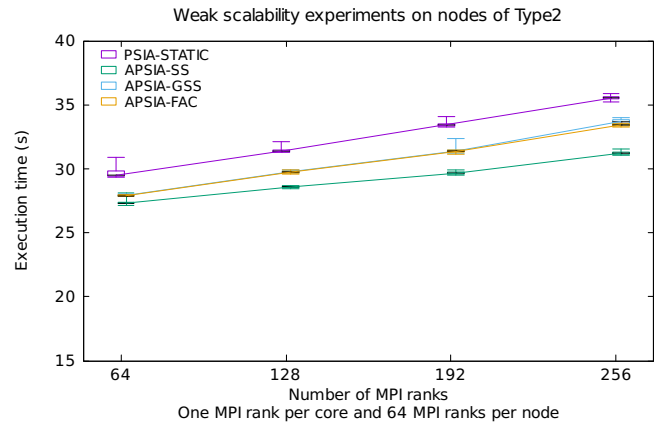


Fig. 5. Performance of the proposed APSIA and the earlier PSIA on homogeneous computing resources of Type2. The number of generated spin-images per computing node is $8K$.

versions and the PSIA-STATIC can be explained due to the load imbalance by the static division and the static assignment of the generation of the spin-images in PSIA-STATIC. According to Algorithm 2, not all computing resources will execute Line 14, because they depend on the condition mentioned on Line 13. The operations on Line 14 represent a memory read and a memory write. The proportion between the memory operations and the computations performed by each resource affects and determines its performance.

In general, the SS algorithm incurs high communication overhead caused by the large volume³ and/or number of messages⁴ between the master and the worker. In this work, however, the input data is *replicated* and the master exchanges only lightweight messages (a few bytes per message) with the workers to indicate the chunk sizes they need to execute. The number of such lightweight messages corresponds to the total number of chunks of tasks calculated by the dynamic loop scheduling algorithm and is different across DLS techniques. The superiority of the APSIA-SS over the other two APSIA versions can be explained by its fine-grain self-scheduled task assignment design as well as by the high speed of the network infrastructure used in the experiments. Both these aspects result in a more balanced execution time among the computing resources, hence, a shorter parallel execution time, using APSIA-SS.

2) *Strong Scalability*: To perform strong scalability experiments, the number of generated spin-images is kept constant while the number of the computing resources is increased. The number of generated spin-images is set at $80K$, which represents approximately 10% of the total spin-images that can be generated from the *Ramesses* object.

A comparison between the *parallel cost* of executing the proposed APSIA and the earlier PSIA on Type1 and Type2 nodes is presented in Fig. 6 and 7, respectively. The parallel

³Depending on the input data distribution strategy, which can be either centralized, partitioned, or replicated.

⁴At least equal to the total number of parallel tasks within the application.

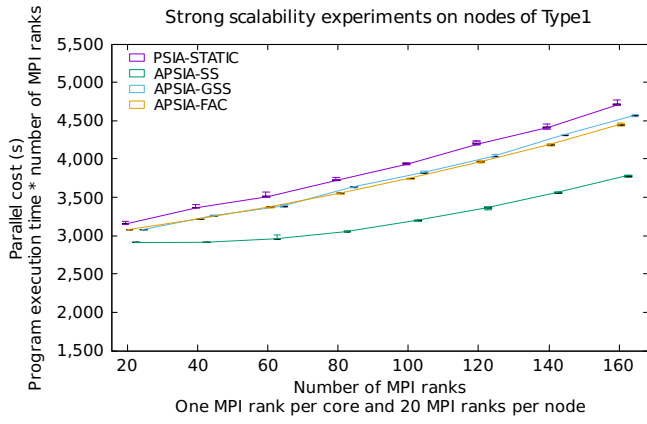


Fig. 6. Performance of the proposed APSIA and the earlier PSIA on homogeneous computing resources of Type1. The number of generated spin-images is 80K.

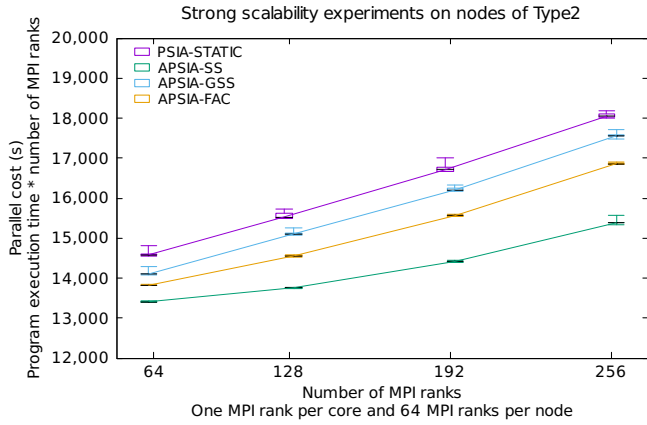


Fig. 7. Performance of the proposed APSIA and the earlier PSIA on homogeneous computing resources of Type2. The number of generated spin-images is 80K.

cost is calculated as the number of the computing resources used to execute a parallel application multiplied by the total parallel execution time of that application. The selection of parallel cost as a performance metric (over the parallel execution time) is due to the fact that it reflects the benefits of using additional computing resources versus the time needed to execute the parallel algorithm. A good strong scalability profile of a program corresponds to an almost constant parallel cost for any number of computing resources. It can be seen in Fig. 6 and 7, that PSIA-STATIC does not exhibit a strong scalability profile for both computing resource types. Similar to the weak scalability results in Section V-A1, the three versions of the proposed APSIA outperform PSIA-STATIC.

The performance advantage of APSIA-SS over APSIA-GSS and APSIA-FAC is attributed to the small message sizes exchanged between the master and the workers and to the high speed of the network infrastructure used in the experiments. The performance gap between the APSIA-SS, and APSIA-GSS and APSIA-FAC can be explained similarly to

the performance gap between the same algorithms in the weak scalability experiments in Section V-A1. The performance gap may, however, be reduced in certain other cases where the network infrastructure has a lower performance than the one used in this work.

In both weak (Section V-A1) and strong (Section V-A2) scalability experiments, the APSIA-SS achieves a speed up of approximately 1.26 on the largest number of computing resources, compared to the performance of PSIA-STATIC on Type1 nodes. On Type2 nodes, APSIA-SS achieves a speed up of approximately 1.16 compared against PSIA-STATIC.

B. Performance of APSIA vs. PSIA on Heterogeneous Computing Resources

The performance of the weak scalability and the strong scalability experiments executed on heterogeneous computing resources is shown in Fig. 8 and 9, respectively. These performance results are very similar to the results obtained on homogeneous computing resources.

APSIA-GSS exhibits an interesting behavior on heterogeneous computing resources compared to that on homogeneous computing resources. In particular, its performance is almost similar to the performance of PSIA-STATIC. This is due to the order in which the available Type1 and Type2 resources request work from the master. As discussed in Section II, the GSS algorithm assigns the largest chunk of loop iterations to the first requesting worker. Recall from Section IV-C that the heterogeneous worker computing resources listed in the machine file used in this work commence with Type2 followed by Type1. Also, the master is a dedicated computing resource (core) mapped on a separate node of Type1 and it is always written in the machine file before the worker computing resources of Type1 and Type2. This listing of resources in the machine file is meant to enable the use of the computing nodes with the largest number of computing cores, i.e., 64 cores for Type2 compared to 20 cores for Type1. Changing this listing may enhance the performance of APSIA-GSS without changing the main semantic and trend of the results where PSIA-STATIC and APSIA-SS perform the worst and the best, respectively.

In both the weak and the strong scalability experiments, the APSIA-SS achieves a speed up of approximately 2 on the largest number of computing resources, compared to the APSIA-STATIC on nodes of Type1 and Type2.

VI. CONCLUSION AND FUTURE WORK

The static assignment of the spin-image generation tasks using PSIA [2] causes severe load imbalance during execution. The load imbalance worsens when executing the PSIA on heterogeneous computing resources. By employing dynamic loop scheduling and the master-worker execution model, the proposed APSIA is able to reduce the load imbalance when executing on homogeneous and on heterogeneous computing resources, as well as to deliver a high performance at increased scales. The proposed APSIA employs three different DLS techniques: SS, GSS, and FAC. For the largest problem

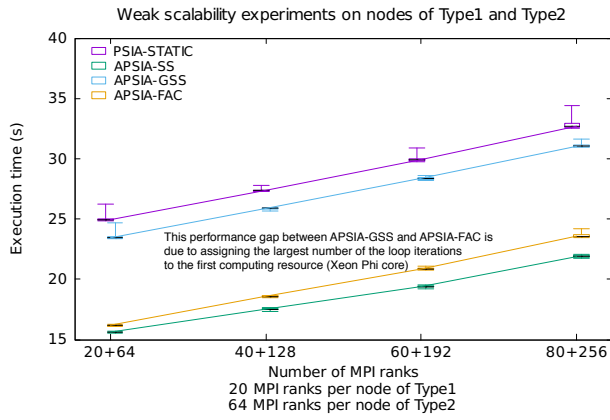


Fig. 8. Performance of the proposed APSIA and the earlier PSIA on heterogeneous computing resources of Type1 and Type2, respectively. The number of generated spin-images per computing node is $8K$.

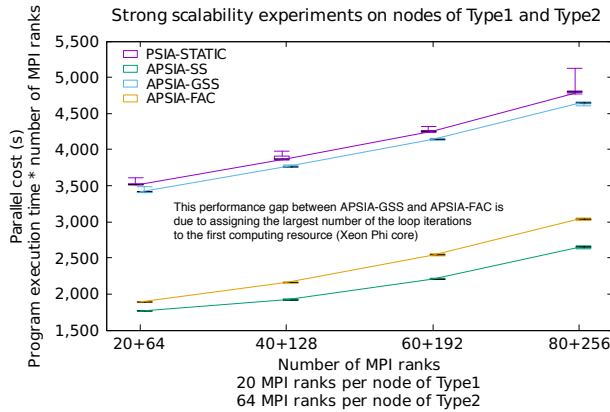


Fig. 9. Performance of the proposed APSIA and the earlier PSIA on heterogeneous computing resources of Type1 and Type2, respectively. The number of generated spin-images is $80K$.

size ($80K$ spin-images), the performance of the APSIA-SS outperforms the performance of the earlier PSIA by a factor of 1.2 and 2 on homogeneous and heterogeneous computing, respectively.

Due to the high speed network used in this work, the APSIA-SS shows the best performance. More investigation is needed and planned to assess the performance of the proposed APSIA across different hardware setups, in particular regarding the network infrastructure and the listing of resource types in the machine file. Also, additional and more complex DLS techniques will be integrated with the APSIA. As discussed in Section V-B, the performance of the APSIA-GSS is affected on heterogeneous computing resources by the type of resource requesting work in the initial chunk allocations. Further work is, therefore, needed to understand the effects of different resource listings in the machine file on the performance of APSIA.

ACKNOWLEDGMENT

This work is in part supported by the Swiss National Science Foundation in the context of the Multi-level Scheduling in Large Scale High Performance Computers (MLS) grant, number 169123.

REFERENCES

- [1] R. Nishtala, P. H. Hargrove, D. O. Bonachea, and K. A. Yelick, "Scaling communication-intensive applications on BlueGene/P using one-sided communication and overlap," in *Proceedings of the 23rd International Symposium on Parallel and Distributed (IPDPS)*, Rome, Italy, May 2009, pp. 1–12.
- [2] A. Eleliemy, M. Fayze, R. Mehmood, I. Katib, and N. Aljohani, "Loadbalancing on Parallel Heterogeneous Architectures: Spin-image Algorithm on CPU and MIC," in *Proceedings of the 9th EUROSIM Congress on Modelling and Simulation*, September 2016, Oulu, Finland, pp. 623–628.
- [3] A. E. Johnson, "Spin-Images: A Representation for 3-D Surface Matching," Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.
- [4] K.-S. Choi and D.-H. Kim, "Angular-partitioned spin image descriptor for robust 3D facial landmark detection," *Electronics Letters*, vol. 49, no. 23, pp. 1454–1455, 2013.
- [5] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.
- [6] —, "Surface matching for object recognition in complex three-dimensional scenes," *Image and Vision Computing*, vol. 16, no. 9, pp. 635–651, 1998.
- [7] Y. Mei and Y. He, "A new spin-image based 3D Map registration algorithm using low-dimensional feature space," in *Proceedings of the 7th International Conference on Information and Automation (ICIA)*, Yinchuan, China, August 2013, pp. 545–551.
- [8] J. Assfalg, G. D'Amico, A. Del Bimbo, and P. Pala, "3D content-based retrieval with spin images," in *Proceedings of the International Conference on Multimedia and Expo (ICME)*, Taipei, Taiwan, June 2004, pp. 771–774.
- [9] T. L. Casavant and J. G. Kuhl, "A taxonomy of scheduling in general-purpose distributed computing systems," *IEEE Transactions on Software Engineering*, vol. 14, no. 2, pp. 141–154, Feb 1988.
- [10] O. Plata and F. F. Rivera, "Combining Static and Dynamic Scheduling on Distributed-memory Multiprocessors," in *Proceedings of the 8th International Conference on Supercomputing*, Manchester, England, 1994, pp. 186–195.
- [11] R. L. Cariño and I. Banicescu, "A load balancing tool for distributed parallel loops," in *Proceedings of the International Workshop on Challenges of Large Applications in Distributed Environments*, Washington, USA, 2003, pp. 39–46.
- [12] R. L. Cariño, I. Banicescu, T. Rauber, and G. Rünger, "Dynamic Loop Scheduling with Processor Groups," in *Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems*, California, USA, 2004, pp. 78–84.
- [13] P. Tang and P.-C. Yew, "Processor Self-Scheduling for Multiple-Nested Parallel Loops," in *Proceedings of the International Conference of Parallel Processing (ICPP)*, Urbana, USA, January 1986, pp. 528–535.
- [14] C. D. Polychronopoulos and D. J. Kuck, "Guided self-scheduling: A practical scheduling scheme for parallel supercomputers," *IEEE Transactions on Computers*, vol. 100, no. 12, pp. 1425–1439, 1987.
- [15] S. F. Hummel, E. Schonberg, and L. E. Flynn, "Factoring: A method for scheduling parallel loops," *Communications of the ACM*, vol. 35, no. 8, pp. 90–101, 1992.
- [16] D.-K. Chen and P.-C. Yew, "An Empirical Study on DOACROSS Loops," in *Proceedings of the 1991 ACM/IEEE Conference on Supercomputing*, Albuquerque, USA, 1991, pp. 620–632.
- [17] Y.-W. Fann, C.-T. Yang, S.-S. Tseng, and C.-J. Tsai, "An intelligent parallel loop scheduling for parallelizing compilers," *Journal of Information Science and Engineering*, vol. 16, no. 2, pp. 169–200, 2000.
- [18] K. Wang, G. Lavoué, F. Denis, A. Baskurt, and X. He, "A benchmark for 3D mesh watermarking," in *Proceedings of the 9th IEEE International Conference on Shape Modeling and Applications*, 2010, pp. 231–235.