

NumaPerf: Predictive NUMA Profiling

Xin Zhao

University of Massachusetts Amherst
zhao@umass.edu

Wei Wang

University of Texas at San Antonio
wei.wang@utsa.edu

Jin Zhou

University of Massachusetts Amherst
jinzhou@umass.edu

Xu Liu

North Carolina State University
xliu88@ncsu.edu

Hui Guan

University of Massachusetts Amherst
huiguan@cs.umass.edu

Tongping Liu

University of Massachusetts Amherst
tongping@umass.edu

Abstract

It is extremely challenging to achieve optimal performance of parallel applications on a NUMA architecture, which necessitates the assistance of profiling tools. However, existing NUMA-profiling tools share some similar shortcomings, such as portability, effectiveness, and helpfulness issues. This paper proposes a novel profiling tool—NumaPerf—that overcomes these issues. NumaPerf aims to identify potential performance issues for any NUMA architecture, instead of only on the current hardware. To achieve this, NumaPerf focuses on memory sharing patterns between threads, instead of real remote accesses. NumaPerf further detects potential thread migrations and load imbalance issues that could significantly affect the performance but are omitted by existing profilers. NumaPerf also identifies cache coherence issues separately that may require different fix strategies. Based on our extensive evaluation, NumaPerf is able to identify more performance issues than any existing tool, while fixing these bugs leads to significant performance speedup.

ACM Reference Format:

Xin Zhao, Jin Zhou, Hui Guan, Wei Wang, Xu Liu, and Tongping Liu. 2021. NumaPerf: Predictive NUMA Profiling. In *2021 International Conference on Supercomputing (ICS '21), June 14–17, 2021, Virtual Event, USA*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3447818.3460361>

1 Introduction

The Non-Uniform Memory Access (NUMA) is the de facto design to address the scalability issue with an increased number of hardware cores. Compared to the Uniform Memory Access (UMA) architecture, the NUMA architecture avoids the bottleneck of one memory controller by allowing each

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICS '21, June 14–17, 2021, Virtual Event, USA

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8335-6/21/06...\$15.00

<https://doi.org/10.1145/3447818.3460361>

node/processor to concurrently access its own memory controller. However, the NUMA architecture imposes multiple system challenges for writing efficient parallel applications, such as remote accesses, interconnect congestion, and node imbalance [5]. User programs could easily suffer from significant performance degradation, necessitating the development of profiling tools to identify NUMA-related performance issues.

General-purpose profilers, such as `gprof` [12], `perf` [11], or `Coz` [9], are not suitable for identifying NUMA-related performance issues [24, 30] because they are agnostic to the architecture difference. To detect NUMA-related issues, one type of tools simulates cache activities and page affinity based on the collected memory traces [29, 33]. However, they may introduce significant performance slowdown, preventing their uses even in development phases. In addition to this, another type of profiler employs coarse-grained sampling to identify performance issues in the deployment environment [14, 18, 24, 26, 32, 35], while the third type builds on fine-grained instrumentation that could detect more performance issues but with a higher overhead [10, 30].

However, the latter two types of tools share the following **common issues**. *First, they mainly focus on one type of performance issues (i.e., remote accesses), while omitting other types of issues that may have a larger performance impact.* *Second, they have limited portability that can only identify remote accesses on the current NUMA hardware.* The major reason is that they rely on the physical node information to detect remote accesses, where the physical page a thread accesses is located in a node that is different from the node of the current thread. However, the relationship between threads/pages with physical nodes can be varied when an application is running on different hardware with different topology, or even on the same hardware at another time. That is, existing tools may miss some remote accesses caused by specific binding. *Third, existing tools could not provide sufficient guidelines to fix these issues.* Users have to spend significant effort to figure out the corresponding fix strategy by themselves.

This paper proposes a novel tool—NumaPerf—that overcomes these issues. NumaPerf is designed as an automatic tool that does not require human annotation or changing the code. It also does not require new hardware, or the change

of the underlying operating system. NumaPerf aims to detect NUMA-related issues in development phases, when applications are exercised with representative inputs. In this way, there is no need to pay additional and unnecessary runtime overhead in deployment phases. We further describe NumaPerf's distinctive goals and designs as follows.

First, NumaPerf aims to detect some additional types of NUMA performance issues, while existing NUMA profilers could only detect remote access. The first type is load imbalance among threads, which may lead to memory controller congestion and interconnect congestion. The second type is cross-node migration, which turns all previous local accesses into remote accesses. Based on our evaluation, cross-node migration may lead to $4.2\times$ performance degradation for *fluidanimate*. However, some applications may not have such issues, which require the assistance of profiling tools.

Second, it proposes a set of architecture-independent and scheduling-independent mechanisms that could predictively detect the above-mentioned issues on any NUMA architecture, even without running on a NUMA machine. NumaPerf's detection of remote accesses is based on a **key observation**: memory sharing pattern of threads is an invariant determined by the program logic, but the relationship between threads/pages and physical nodes is architecture and scheduling dependent. Therefore, *NumaPerf focuses on identifying memory sharing pattern between threads, instead of the specific node relationship of threads and pages*, since a thread/page can be scheduled/allocated to/from a different node in a different execution. This mechanism not only simplifies the detection problem (without the need to track the node information), but also generalizes to different architectures and executions (scheduling). *NumaPerf also proposes an architecture-independent mechanism to measure load imbalance based on the total number of memory accesses from threads*: when different types of threads have a different number of total memory accesses, then this application has a load imbalance issue. *NumaPerf further proposes a method to predict the probability of thread migrations*. NumaPerf computes a migration score based on the contending number of synchronizations, and the number of condition and barrier waits. Overall, NumaPerf predicts a set of NUMA performance issues without the requirement of testing on a NUMA machine, where its basic ideas are further discussed in Section 2.2.

Last but not least, NumaPerf aims to provide more helpful information to fix NUMA performance issues. First, it proposes a set of metrics to measure the seriousness of different performance issues, preventing programmers from spending unnecessary efforts on insignificant issues. Second, its report could guide users for a better fix. For load imbalance issues, NumaPerf suggests a thread assignment that could achieve much better performance than existing work [1]. For remote accesses, there exist multiple fix strategies with different levels of improvement. Currently, programmers have to figure

out a good strategy by themselves. In contrast, NumaPerf supplies more information to assist fixes. It separates cache false sharing issues from true sharing and page sharing so that users can use the padding to achieve better performance. It further reports whether the data can be duplicated or not by confirming the temporal relationship of memory reads/writes. It also reports threads accessing each page, which helps confirm whether a block-wise interleave with the thread binding will have a better performance improvement.

We performed extensive experiments to verify the effectiveness of NumaPerf with widely-used parallel applications (i.e., PARSEC [4]) and HPC applications (e.g., AMG2006 [17], lulesh [15], and UMT2013 [16]). Based on our evaluation, NumaPerf detects many more performance issues than the combination of all existing NUMA profilers, including both fine-grained and coarse-grained tools. After fixing such issues, these applications could achieve up to $5.94\times$ performance improvement. NumaPerf's helpfulness on fixing performance issues is also exemplified by multiple case studies. Overall, NumaPerf imposes less than $6\times$ performance overhead, which is orders of magnitude faster than the previous state-of-the-art in the fine-grained analysis. The experiments also confirm that NumaPerf's detection is architecture-independent, which is able to identify most performance issues when running on a non-NUMA machine.

Overall, NumaPerf makes the following contributions.

- NumaPerf proposes a set of architecture-independent and scheduling-independent methods that could predictively detect NUMA-related performance issues, even without evaluating on a specific NUMA architecture.
- NumaPerf is able to detect a comprehensive set of NUMA-related performance issues, where some are omitted by existing tools.
- NumaPerf designs a set of metrics to measure the seriousness of performance issues, and provides helpful information to assist their fixes.
- We have performed extensive evaluations to confirm NumaPerf's effectiveness and overhead.

Outline

The remainder of this paper is organized as follows. Section 2 introduces the background of NUMA architecture and the basic ideas of NumaPerf. Then Section 3 presents the detailed implementation and Section 4 shows experimental results. After that, Section 5 explains the limitation and Section 6 discusses related work in this field. In the end, Section 7 concludes this paper.

2 Background and Overview

This section starts with the introduction of the NUMA architecture and potential performance issues. Then it briefly discusses the basic idea of NumaPerf to identify such issues.

2.1 NUMA Architecture

Traditional computers use the Uniform Memory Access (UMA) model. In this model, all CPU cores share a single memory controller such that any core can access the memory with the same latency (uniformly). However, the UMA architecture cannot accommodate the increasing number of cores because these cores may compete for the same memory controller. The memory controller becomes the performance bottleneck in many-core machines since a task cannot proceed without getting its necessary data from the memory.

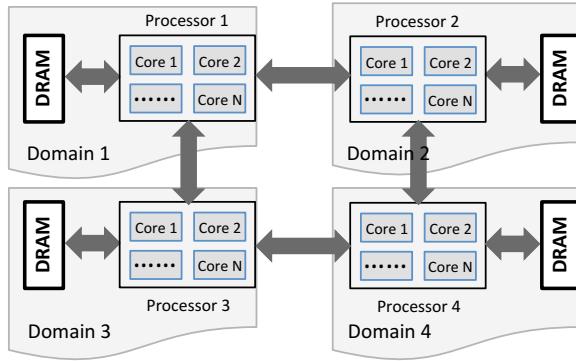


Figure 1. A NUMA architecture with four nodes/domains

The Non-Uniform Memory Access (NUMA) architecture is proposed to solve this scalability issue, as further shown in Figure 1. It has a decentralized nature. Instead of making all cores waiting for the same memory controller, the NUMA architecture is typically equipped with multiple memory controllers, where each controller serves a group of CPU cores (called a “node” or “processor” interchangeably). Incorporating multiple memory controllers largely reduces the contention for memory controllers and therefore improves the scalability correspondingly. However, the NUMA architecture also introduce multiple sources of performance degradations [5], including *Cache Contention*, *Node Imbalance*, *Interconnect Congestion*, and *Remote Accesses*.

Cache Contention: the NUMA architecture is prone to cache contention, including false and true sharing. False sharing occurs when multiple tasks may access distinct words in the same cache line [3], while different tasks may access the same words in true sharing. For both cases, multiple tasks may compete for the shared cache. Cache contention will cause more serious performance degradation, if data has to be loaded from a remote node.

Node Imbalance: When some memory controllers have much more memory accesses than others, it may cause the node imbalance issue. Therefore, some tasks may wait more time for memory access, thwarting the whole progress of a multithreaded application.

Interconnect Congestion: Interconnect congestion occurs if some tasks are placed in remote nodes that may use the inter-node interconnection to access their memory.

Remote Accesses: In a NUMA architecture, local nodes can be accessed with less latency than remote accesses. Therefore, it is important to reduce remote access to improve performance.

2.2 Basic Idea

Existing NUMA profilers mainly focus on detecting remote accesses, while omitting other performance issues. In contrast, NumaPerf has different design goals as follows. First, it aims to identify different sources of NUMA performance issues, not just limited to remote accesses. Second, NumaPerf aims to design architecture- and scheduling-independent approaches that could report performance issues in any NUMA hardware. Third, it aims to provide sufficient information to guide bug fixes.

For the first goal, NumaPerf detects NUMA issues caused by cache contention, node imbalance, interconnect congestion, and remote accesses, where existing work only considers remote accesses. *Cache contention* can be either caused by false or true sharing, which may impose a larger performance impact and require a different fix strategy. Existing work never separates them from normal remote accesses. In contrast, NumaPerf designs a separate mechanism to detect such issues by tracking possible cache invalidations caused by cache contention. It is infeasible to measure all *node imbalance* and *interconnect congestion* without knowing the actual memory and thread binding. Instead, NumaPerf focuses on one specific type of issues, which is workload imbalance between different types of threads. Existing work omits one type of remote access caused by thread migration, where thread migration will make all local accesses remotely. NumaPerf identifies whether an application has a higher chance of thread migrations, in addition to normal remote accesses. Overall, NumaPerf detects more NUMA performance issues than existing NUMA profilers. However, the challenge is to design architecture- and scheduling-independent methods.

The second goal of NumaPerf is to design architecture- and scheduling approaches that do not bind to specific hardware. Detecting remote accesses is based on the key observation of Section 1: if a thread accesses a physical page that was initially accessed by a different thread, then this access will be counted as remote access. This method is not bound to specific hardware, since memory sharing patterns between threads are typically invariant across multiple executions. NumaPerf tracks every memory access in order to identify the first thread working on each page. Due to this reason, NumaPerf employs fine-grained instrumentation, since coarse-grained sampling may miss the access from the first thread. Based on memory accesses, NumaPerf also tracks the number of cache invalidations caused by false or true sharing with the following rule: a write on a cache line with multiple copies will invalidate other copies. Since the number of cache invalidations is closely related to the number of concurrent threads, NumaPerf divides the score with the number of threads to

achieve a similar result with a different number of concurrent threads, as further described in Section 3.2.3. Load imbalance will be evaluated by the total number of memory accesses of different types of threads. It is important to track all memory accesses including libraries for this purpose. To evaluate the possibility of thread migration, NumaPerf proposes to track the number of lock contentions and the number of condition and barrier waits. Similar to false sharing, NumaPerf eliminates the effect caused by concurrent threads by dividing with the number of threads. The details of these implementations can be seen in Section 3.

For the third goal, NumaPerf utilizes the data-centric analysis as in existing work [24]. That is, it could report the callsite of heap objects that may have NUMA performance issues. In addition, NumaPerf aims to provide useful information that helps bug fixes, which could be easily achieved when all memory accesses are tracked. NumaPerf provides word-based access information for cache contentions, helping programmers to differentiate false or true sharing. It provides threads information on page sharing (help determining whether to use block-wise interleave), and reports whether an object can be duplicated or not by tracking the temporal read/write pattern. NumaPerf also predicts a good thread assignment to achieve better performance for load imbalance issues. In summary, many of these features require fine-grained instrumentation in order to avoid false alarms.

Due to the reasons mentioned above, NumaPerf utilizes fine-grained memory accesses to improve the effectiveness and provide better information for bug fixes. NumaPerf employs compiler-based instrumentation in order to collect memory accesses to address performance and flexibility concern. An alternative approach is to employ binary-based dynamic instrumentation [7, 25, 27], which may introduce more performance overhead but without an additional compilation step. NumaPerf inserts an explicit function call for each read/write access on global variables and heap objects, while accesses on stack variables are omitted since they typically do not introduce performance issues. To track thread migration, NumaPerf also intercepts synchronizations. To support data-centric analysis, NumaPerf further intercepts memory allocations to collect their callsites.

Figure 2 summarizes NumaPerf’s basic idea. NumaPerf includes two components, NumaPerf-Static and NumaPerf-Dynamic. NumaPerf-Static is a static compile-time based tool that inserts a function call before every memory access on heap and global variables, which compiles a program into an instrumented executable file. Then this executable file will be linked to NumaPerf-Dynamic so that NumaPerf could collect memory accesses, synchronizations, and information of memory allocations. NumaPerf then performs detection on NUMA-related performance issues, and reports to users in the end. More specific implementations are discussed in Section 3.

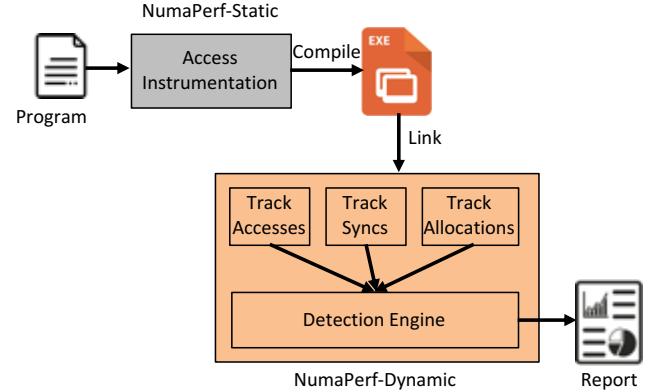


Figure 2. Overview of NumaPerf

3 Design and Implementation

This section elaborates NumaPerf-Static and NumaPerf-Dynamic. NumaPerf leverages compiler-based instrumentation (NumaPerf-Static) to insert a function call before memory access, which allows NumaPerf-Dynamic to collect memory accesses. NumaPerf utilizes a pre-load mechanism to intercept synchronizations and memory allocations, without the need of changing programs explicitly. Detailed design and implementation are discussed as follows.

3.1 NumaPerf-Static

NumaPerf’s static component (NumaPerf-Static) performs the instrumentation on memory accesses. In particular, it utilizes static analysis to identify memory accesses on heap and global variables, while omitting memory accesses on static variables. Based on our understanding, static variables will never cause performance issues, if a thread is not migrated. NumaPerf-Static inserts a function call upon these memory accesses, where this function is implemented in NumaPerf-Dynamic library. In particular, this function provides detailed information on the access, including the address, the type (i.e., read or write), and the number of bytes.

NumaPerf employs the LLVM compiler to perform the instrumentation [20]. It chooses the intermediate representation (IR) level for the instrumentation due to the flexibility, since LLVM provides lots of APIs and tools to manipulate the IR. The instrumentation pass is placed at the end of the LLVM optimization passes, where only memory accesses surviving all previous optimization passes will be instrumented. NumaPerf-Static traverses functions one by one, and instruments memory accesses on global and heap variables. The instrumentation is adapted from AddressSanitizer [31].

3.2 NumaPerf-Dynamic

This subsection starts with tracking application information, such as memory accesses, synchronizations, and memory allocations. Then it discusses the detection of each particular performance issue. In the following, NumaPerf is used to represent NumaPerf-Dynamic unless noted otherwise.

3.2.1 Tracking Accesses, Synchronizations, and Memory Allocations. NumaPerf-Dynamic implements the inserted functions before memory accesses, allowing it to track memory accesses. Once a memory access is intercepted, NumaPerf performs the detection as discussed below.

NumaPerf utilizes a preloading mechanism to intercept synchronizations and memory allocations before invoking correspond functions. NumaPerf intercepts synchronizations in order to detect possible thread migrations, which will be explained later. NumaPerf also intercepts memory allocations, so that we could attribute performance issues to different call-sites, assisting data-centric analysis [24]. For each memory allocation, NumaPerf records the allocation callsite and its address range. NumaPerf also intercepts thread creations in order to set up per-thread data structure. In particular, it assigns each thread a thread index.

3.2.2 Detecting Normal Remote Accesses. *NumaPerf detects a remote access when an access's thread is different from the corresponding page's initial accessor*, as discussed in Section 2. This is based on the assumption that the OS typically allocates a physical page from the node of the first accessor due to the default first-touch policy [19]. Similar to existing work, NumaPerf may over-estimate the number of remote accesses, since an access is not a remote one if the corresponding cache is not evicted. However, this shortcoming can be overcome easily by only reporting issues larger than a specified threshold, as exemplified in our evaluation (Section 4).

NumaPerf is carefully designed to reduce its performance and memory overhead. NumaPerf tracks a page's initial accessor to determine a remote access. A naive design is to employ hash table for tracking such information. Due to a large number of memory accesses, a naive design (e.g. hash table) could easily introduce significant performance overhead. Instead, NumaPerf maps a continuous range of memory with the shadow memory technique [34], which only requires a simple computation to locate the data. NumaPerf also maintains the number of accesses for each page in the same map. We observed that a page without a large number of memory accesses will not cause significant performance issues. Based on this, NumaPerf only tracks the detailed accesses for a page, when its number of accesses is larger than a pre-defined (configurable) threshold. Since the recording uses the same data structures, NumaPerf uses an internal pool to maintain such data structures with the exact size, without resorting to the default allocator.

For pages with excessive accesses, NumaPerf tracks the following information. *First*, it tracks the threads accessing these pages, which helps to determine whether to use block-wise allocations for fixes. *Second*, NumaPerf further divides each page into multiple blocks (e.g., 64 blocks), and tracks the number of accesses on each block. This enables us to compute the number of remote accesses of each object more

accurately. *Third*, NumaPerf further checks whether an object is exclusively read after the first write or not, which could be determined whether duplication is possible or not. *Last not least*, NumaPerf maintains word-level information for cache lines with excessive cache invalidations, as further described in Section 3.2.3.

Remote (Access) Score: NumaPerf proposes a performance metric – remote score – to evaluate the seriousness of remote accesses. An object's remote score is defined as the number of remote accesses within a specific interval, which is currently set as one millisecond. Typically, a higher score indicates a more serious performance issue, as shown in Table 1. For pages with both remote accesses and cache invalidations, we will check whether cache invalidation is dominant or not. If the number of cache invalidations is larger than 50% of remote accesses, then the major performance issue of this page is caused by cache invalidations. We will omit remote accesses instead.

3.2.3 Detecting False and True Sharing Issues. Based on our observation, cache coherence has a higher performance impact than normal remote accesses. Further, false sharing requires a different fixing strategy, and is typically addressed by padding. NumaPerf detects false and true sharing separately, which is different from all NUMA profilers.

NumaPerf detects false/true sharing with a similar mechanism as Predator [23], but adapting it for the NUMA architecture. Predator tracks cache invalidations as follows: if a thread writes a cache line that has been loaded by multiple threads, this write operation introduces a cache invalidation. But this mechanism under-estimates the number of cache invalidations. Whenever multiple cores hold the same cache line, this write operation may introduce multiple cache invalidations and therefore multiple remote accesses. Therefore, NumaPerf tracks the number of threads loaded the same cache line, and increases cache invalidation by the number of threads that has loaded this cache line. Further, NumaPerf differentiates remote cache invalidations from local ones since they may have different performance impacts: if a thread should load the memory from a remote node, then it is a remote cache invalidation. Otherwise, it is a local cache invalidation. NumaPerf only focuses on remote cache invalidations that have a higher impact on the NUMA machine.

False/True Sharing Score: NumaPerf further proposes false/true sharing scores for each corresponding object, which is lacked in Predator [23]. The scores are computed by dividing the number of cache invalidations with the product of time (milliseconds) and the number of threads. The number of threads is employed to reduce the impact of parallelization degree, with the architecture-independent method. NumaPerf differentiates false sharing from true sharing by recording word-level accesses. Note that NumaPerf only records word-level accesses for cache lines with the number of writes larger than a pre-defined threshold, due to the performance concern.

3.2.4 Detecting Issues Caused by Thread Migration. As discussed in Section 1, NumaPerf identifies applications with excessive thread migrations, which are omitted by all existing NUMA profilers. Thread migration may introduce excessive remote accesses. After the migration, a thread is forced to reload all data from the original node, and access its stack remotely afterwards. Further, all deallocations from this thread may be returned to freelists of remote nodes, causing more remote accesses afterwards.

Thread Migration Score: NumaPerf evaluates the seriousness of thread migrations with thread migration scores. This score is computed as the following formula:

$$S = p \sum_{t \in T} m_t / (rt \cdot |T|)$$

[[why would the total number of threads looks like the absolute value of the number of threads? : where S is the thread migration score, p is the parallel phase percentage of the program, T is threads in the program, $|T|$ is the number of total threads, m_t is the possible migration times for thread t , and rt is total running seconds of the program.]]

NumaPerf utilizes the total number of lock contentions, condition waits, and barrier waits as the possible migration times, instead of identifying the real migration times. Therefore, NumaPerf predicts the potential performance impact caused by thread migration. NumaPerf utilizes the preloading mechanism to intercept these synchronizations, and detects potential lock contention by checking the status of lock. The parallel phase percentage indicates the necessity of performing the optimization. For instance, if the parallel phase percentage is only 1%, then we could at most improve the performance by 1%. In order to reduce the effect of parallelization, the score is further divided by the number of threads. Based on our evaluation, this parameter makes two platforms with different number of threads have very similar results.

When an application has a large number of thread migrations, NumaPerf suggests users to utilize thread binding to reduce remote accesses. As shown in Table 1, thread migration may degrade the performance of an application (i.e., fluidanimate) by up to 418%. This shows the importance to eliminate thread migration for such applications. However, some applications in PARSEC (as not shown in Table 1) have very marginal performance improvement with thread binding.

3.2.5 Detecting Load Imbalance. Load imbalance is another factor that could significantly affect the performance on the NUMA architecture, which could cause node imbalance and interconnect congestion. NumaPerf detects load imbalance among different types of threads, which is omitted by existing NUMA-profilers.

The detection is based on an assumption: *every type of threads should have a similar number of memory accesses in*

a balanced environment. NumaPerf proposes to utilize the number of memory accesses to predict the workload of each types of threads. In particular, NumaPerf monitors memory accesses on heap objects and globals, and then utilizes the sum of such memory accesses to check the imbalance.

NumaPerf further predicts an optimal thread assignment with the number of memory accesses. A balance assignment is to balance memory accesses from each type of threads. For instance, if the number of memory accesses on two type of threads has a one-to-two portion, then NumaPerf will suggest to assign threads in one-to-two portion. Section 4.2 further evaluates NumaPerf's suggested assignment, where NumaPerf significantly outperforms another work [1].

4 Experimental Evaluation

This section aims to answer the following research questions:

- **Effectiveness:** Whether NumaPerf could detect more performance issues than existing NUMA-profilers? (Section 4.1) How helpful of NumaPerf's detection report? (Section 4.2)
- **Performance:** How much performance overhead is imposed by NumaPerf's detection, comparing to the state-of-the-art tool? (Section 4.3)
- **Memory Overhead:** What is the memory overhead of NumaPerf? (Section 4.4)
- **Architecture In-dependence:** Whether NumaPerf could detect similar issues when running on a non-NUMA architecture? (Section 4.5)

Experimental Platform: NumaPerf was evaluated on a Intel(R) Xeon(R) 8153 machine with 8 nodes and 128 physical cores in total, except in Section 4.5. This machine is installed with 512GB memory. Any two nodes in this machine are less than or equal to 3 hops, where the latency of two hops and three hops is 2.1 and 3.1 respectively, while the local latency is 1.0. The OS for this machine is Linux Debian 10 and the compiler is Clang-LLVM 10.0.0. The hyperthreading was turned off for the evaluation.

4.1 Effectiveness

We evaluated NumaPerf on multiple HPC applications (e.g., AMG2006 [17], lulesh [15], and UMT2013 [16]) and a widely-used multithreaded application benchmark suite — PARSEC [4]. Applications with NUMA performance issues are listed in Table 1. The performance improvement after fixing all issues is listed in “Improve” column, with the average of 10 runs. Here, we divide the original runtime by the runtime after optimization to get the improvement rate. For each issue, the table listed the type of issue and the corresponding score, the allocation site, and the fix strategy. Note that the table only shows cases with page sharing score larger than 1500 (if without cache false/true sharing), false/true sharing score larger than 1, and thread migration score larger than 150. Further, the performance improvement of each specific

Application	Improve	Specific Issues						
		#	Issue	Score	Allocation Site	Fix Strategy	Improve	New
lulesh	594%	1	remote access	7390	par_rap.c:1385	block interleave	160%	
		2	thread migration	6		thread binding	132%	✓
		3	remote access	1840	lulesh.cc:543-545	block interleave	429%	
		4	remote access	1504	lulesh.cc:1029-1034	block interleave	504%	
		5	remote access	4496	lulesh.cc:2251-2264	block interleave	406%	418%
		6	false sharing	26		padding	103%	
		7	remote access	1229	lulesh.cc:2089	block interleave	392%	407%
		8	false sharing	12		padding	104%	
		9	thread migration	3328		thread binding	382%	✓
UMT2013	131%	10	thread migration	18		thread binding	131%	✓
		11	remote access	N/A	ZoneData_mod.F90:95	N/A	N/A	N/A
bodytrack	109%	12	remote access	10800	FlexImageStore.h:146	page interleave	106%	
		13	false sharing	24				✓
		14	thread migration	297		thread binding	105%	✓
dedup	116%	15	thread imbalance			adjust threads	116%	✓
facesim	105%	16	thread migration	607		thread binding	105%	✓
ferret	206%	17	thread imbalance			adjust threads	206%	✓
fluidanimate	429%	18	remote access	90534	pthreads.cpp:292	page interleave	340%	
		19	true sharing	2941				✓
		20	remote access	180	pthreads.cpp:294	page interleave	112%	160%
		21	false sharing	20		padding	158%	
		22	thread migration	73		thread binding	418%	✓
streamcluster	167%	23	remote access	427	streamcluster.cpp:984	page interleave	100%	103%
		24	false sharing	31		padding	102%	
		25	remote access	7169	streamcluster.cpp:1845	duplicate	158%	
		26	thread migration	229		thread binding	132%	✓

Table 1. Detected NUMA performance issues when running on an 8-node NUMA machine. NumaPerf detects 15 more performance bugs that cannot be detected by existing NUMA profilers (with a check mark in the last column). Note that NumaPerf missed one issue (item 11) that can be detected by existing tools.

issue is listed as well. We also present multiple case studies that show how helpful is NumaPerf’s report in Section 4.2.

Overall, we have the following observations. First, it reports no false positives by only reporting scores larger than a threshold. Second, NumaPerf detects more performance issues than the combination of all existing NUMA profilers [10, 14, 18, 24, 26, 30, 32, 35]. The performance issues that cannot be detected by existing NUMA profilers are highlighted with a check mark in the last column of the table, although some can be detected by specific tools, such as cache false/true sharing issues [8, 13, 21–23]. This comparison with existing NUMA profilers is based on the methodology. Existing NUMA profilers cannot separate false or true sharing with normal remote accesses, and cannot detect thread migration and load imbalance issues.

When comparing to a specific profiler, NumaPerf also has better results even on detecting remote accesses. For lulesh, HPCToolkit detects issues of # 4 [24], while NumaPerf detects three more issues (# 3, 5, 7). Fixing these issues improves the performance by up to 504% (with the threads

binding). Multiple reasons may contribute to this big difference. First, NumaPerf’s predictive method detects some issues that are not occurred in the current scheduling and the current hardware, while HPCToolkit has no such capabilities. Second, HPCToolkit requires to bind threads to nodes, which may miss remote accesses caused by its specific binding. Third, NumaPerf’s fine-grained profiling provides a better effectiveness than a coarse-grained profiler like HPCToolkit. NumaPerf may have false negatives caused by its instrumentation. NumaPerf cannot detect an issue (# 11 in Table 1) of UMT2013 reported by HPCToolkit [24]. The basic reason is that NumaPerf cannot instrument Fortran code. NumaPerf’s limitations are further discussed in Section 4.2.

4.2 Case Studies

In this section, multiple case studies are shown how programmers could fix performance issues based on the report.

4.2.1 Remote Accesses. For remote accesses, NumaPerf not only reports remote access scores, indicating the seriousness of the corresponding issue, but also provides additional information to reduce remote accesses. Remote accesses can be reduced with different strategies, such as padding (false sharing), block-wise interleaving, duplication, and page interleaving.

```
Allocation Site: lulesh.cc:2251
Remote score: 4496
False sharing score: 26
True Sharing score: 0.00
Pages accessed by threads:
  0--8, 8--16, 16--23, 23--31 .....
```

Listing 1. Remote access issue of lulesh

NumaPerf provides a data-centric analysis, as existing work [24]. That is, it always attributes performance issues to its allocation callsite. NumaPerf also shows the seriousness with its remote access score.

NumaPerf further reports more specific information to guide the fix. As shown in Listing 1, NumaPerf further reports each page that is accessed by which threads. Based on this information, block-wise interleave is a better strategy for the fix, which achieves a better performance result. However, for Issue # 18 or # 20 of `fluidanimate`, there is no such access pattern. Therefore, these issues can be fixed with the normal page interleave method.

```
Allocation site:streamcluster.cpp:1845
Remote score: 7169
False sharing score: 0.00
True Sharing score: 0.00
Continuous reads after the last write: 2443582804
```

Listing 2. Remote access issue of streamcluster

Listing 2 shows another example of remote accesses. For this issue (# 25), a huge number of continuous reads (2330M) were detected after the last write. Based on such a report, the object can be duplicated to different physical nodes, which improves the performance by 158%, which achieves significantly better performance than page interleave.

For cache coherency issues, NumaPerf differentiates them from normal remote accesses, and further differentiates false sharing from true sharing. Given the report, programmers could utilize the padding to eliminate false sharing issues. As shown in Table 4, many issues have false sharing issues (e.g., #6, #8, #13, #21, #24). Fixing them with the padding could easily boost the performance. However, we may simply utilize the page interleave to solve true sharing issues.

4.2.2 Thread Migration. When an application has frequent thread migrations, it may introduce excessive thread migrations. For such issues, the fix strategy is to bind threads to nodes. Typically, there are two strategies: round robin and packed binding. Round robin is to bind continuous threads to different nodes one by one, ensuring that different nodes have

a similar number of threads. Packed binding is to bind multiple threads to the first node, typically the same as the number of hardware cores in one node, and then to another node afterwards. Based on our observation, round robin typically achieves a better performance than packed binding, which is the default binding policy for our evaluations in Table 1. Thread binding itself achieves the performance improvement by up to 418% (e.g., `fluidanimate`), which indicates the importance for some applications.

4.2.3 Load Imbalance. For load imbalance, NumaPerf not only reports the existence of such issues, but also suggests a thread assignment that could help reduce these issues. As discussed in Section 3.2.5, NumaPerf predicts the assignment based on the number of sampled memory accesses: *the number of threads for each type should have a similar proportion as the number of memory accesses from each type of threads*. That is, if two type of threads has a one-to-two proportion on memory accesses, then the number of threads should also follow this one-to-two proportion.

For `dedup`, NumaPerf reports that memory accesses of anchor, chunk, and compress threads have a proportion of 92.2:0.33:3.43, when all libraries are instrumented. Based on this, the portion of the chunk and compress threads should be around 1 to 10. Since `dedup` implements a pipe-line model that multiple stages are using the shared queues to communicate, many threads competing for the same queue may actually introduce high contention. Thus, we actually set the number of chunk threads to be 2, which helps to reduce the possible lock contention. Based on this, we further set the number of compress threads to be 18, and the number of anchor to be 76. The corresponding queues are **[[18:2:2:4 (76:18:2:??)]]**. With this setting, `dedup`'s performance is improved by 16%. We further compare its performance with the suggested assignment of another existing work—SyncPerf [1]. SyncPerf assumes that different types of threads should have the same waiting time. SyncPerf proposes the best assignment should be 24:24:48, which could only improve the performance by 5%.

For another example—`ferret`, NumaPerf suggests a proportion of 3.3 : 1.9 : 47.4 : 75.3 for its four types of threads. With this suggestion, we are configuring the threads to be 4 : 2 : 47 : 75. With this assignment, `ferret`'s performance increases by 106% compared with the original version. In contrast, SyncPerf suggests an assignment of 1 : 1 : 2 : 124. However, following such an assignment actually degrades the performance by 354% instead.

Overall, NumaPerf suggested optimal assignment for load imbalance issues achieves the better performance than SyncPerf [1].

4.3 Performance Overhead

We also evaluated the performance of NumaPerf on PARSEC applications, and the performance results are shown

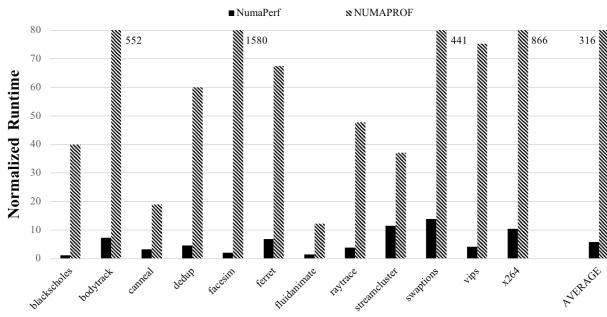


Figure 3. Performance overhead of NumaPerf and others.

in Figure 3. On average, NumaPerf’s overhead is around 585%, which is orders-of-magnitude smaller than the state-of-the-art fine-grained profiler — NUMAPROF [30]. In contrast, NUMAPROF’s overhead runs 316× slower than the original one. NumaPerf is designed carefully to avoid such high overhead, as discussed in Section 3. Also, NumaPerf’s compiler-instrumentation also helps reduce some overhead by excluding memory accesses on stack variables.

There are some exceptions. Two applications impose more than 10× overhead, including Swaption and x264. Based on our investigation, the instrumentation with an empty function imposes more than 5× overhead. The reason is that they have significantly more memory accesses compared with other applications like blackscholes. Based on our investigation, swaption has more than 250× memory accesses than blackscholes in a time unit. Applications with low overhead can be caused by not instrumenting libraries, which is typically not the source of NUMA performance issues.

4.4 Memory Overhead

Apps	Memory Usage (MB)		
	Default	NumaPerf	NUMAPROF
blackscholes	617	689	685
bodytrack	36	139	260
canneal	887	1476	2383
dedup	917	1806	2388
facesim	2638	2826	3005
ferret	160	301	445
fluidanimate	470	667	753
raytrace	1287	1610	2089
streamcluster	112	216	928
swaptions	28	67	255
vips	226	283	463
x264	2861	3039	3108
Total	10238	13120	16762

Table 2. Memory consumption of different profilers.

We further evaluated NumaPerf’s memory overhead with PARSEC applications. The results are shown in Table 2. In total, NumaPerf’s memory overhead is around 28%, which is

much smaller than the state-of-the-art fine-grained profiler — NUMAPROF [30]. NumaPerf’s memory overhead is mainly coming from the following resources. First, NumaPerf records the detailed information in page-level and cache-level, so that we could provide detailed information about these issues. Second, NumaPerf also stores allocation callsites for every object in order to attribute performance issues back to the data.

We notice that some applications have a larger percentage of memory overhead, such as *streamcluster*. For it, a large object has very serious NUMA issues. Therefore, recording page and cache level detailed information contributes to the major memory overhead. However, overall, NumaPerf’s memory overhead is totally acceptable, since it provides much more helpful information.

4.5 Architecture Sensitiveness

We further confirm whether NumaPerf is able to detect similar performance issues when running on a non-NUMA or UMA machine. We further performed the experiments on a two-processor machine, where each processor is Intel(R) Xeon(R) Gold 6230 and each processor has 20 cores. We explicitly disabled all cores in node 1 but only utilizing 16 hardware cores in node 0. This machine has 256GB of main memory, 64KB L1 cache, and 1MB of L2 cache. The experimental results are further listed in Table 3. For simplicity, we only listed the applications, the issue number, and serious scores in two different machines.

Table 3 shows that most reported scores in two machines are very similar, although with small variance. The small variance could be caused by multiple factors, such as parallelization degree (concurrency). However, this table shows that all serious issues can be detected on both machines. This indicates that NumaPerf achieves its design goal, which could even detect NUMA issues without running on a NUMA machine.

5 Limitation

NumaPerf bases on compiler-based instrumentation to capture memory accesses. Therefore, it shares the same shortcomings and strengths of all compiler-based instrumentation. On the one side, NumaPerf can perform static analysis to reduce unnecessary memory accesses, such as accesses of stack variables. NumaPerf typically achieves much better performance than binary-based instrumentation tools, such as NUMAPROF [30]. On the other side, NumaPerf requires the re-compilation (and the availability of the source code), and will miss memory accesses without the instrumentation. That is, it can not detect NUMA issues caused by non-instrumented components (e.g., libraries), suffering from false negatives. However, most issues should only occur in applications, but not libraries.

Application	Specific Issues			
	#	Type	Score (NUMA)	Score (UMA)
AMG2006	1	remote access	7390	5405
	2	thread migration	6	6
lulesh	3	remote access	1840	2443
	4	remote access	1504	2353
	5	remote access	4496	4326
	6	false sharing	26	51
	7	remote access	1229	2136
	8	false sharing	12	27
	9	thread migration	3328	5213
UMT2013	10	thread migration	18	
11	remote access	10800	8203	
bodytrack	12	false sharing	24	153
	13	thread migration	297	190
	14	thread imbalance	92:1:3	88:4:4
dedup	15	thread migration	607	274
facesim	16	thread imbalance		
fluidanimate	17	remote access	90534	15765
	18	true sharing	2941	1753
	19	remote access	180	95
	20	false sharing	20	80
	21	thread migration	73	34
streamcluster	22	remote access	427	270
	23	false sharing	31	153
	24	remote access	7169	10259
	25	thread migration	229	214

Table 3. Evaluation on architecture Sensitiveness. We evaluated NumaPerf on a non-NUMA (UMA) machine, which has very similar results as that on a NUMA machine. For ferret, NumaPerf reports a proportion of 3 : 2 : 48 : 75 on the 8-node NUMA machine, and 5 : 4 : 50 : 77 on the UMA machine.

6 Related Work

This section discusses NUMA-profiling tools at first, and then discusses other relevant tools and systems.

6.1 NUMA Profiling Tools

Simulation-Based Approaches: Bolosky et al. propose to model NUMA performance issues based on the collected trace, and then derive a better NUMA placement policy [6]. NUMAgrind employs binary instrumentation to collect memory traces, and simulates cache activities and page affinity [33]. MACPO reduces the overhead of collecting memory traces and analysis by focusing on code segments that have known performance bottlenecks [29]. That is, it typically requires programmer inputs to reduce its overhead. Simulation-based approaches could be utilized for any architecture, which are very useful. However, they are typically extremely slow, with thousands of performance slowdown, which makes them un-affordable even for development phases. Further, they still

require to evaluate the performance impact for a given architecture, which will significantly limit its usage. NumaPerf utilizes a measurement based approach, which avoids significant performance overhead of simulation-based approaches.

Fine-Grained Approaches: TABARNAC focuses on the visualization of memory access behaviors of different data structures [2]. It uses PIN to collect memory accesses of every thread on the page level, and then relates with data structure information together to visualize the usage of data structures. It introduces the runtime overhead between 10× and 60×, in addition to its offline overhead. Diener et al. propose to instrument memory accesses with PIN dynamically, and then characterize distribution of accesses of different NUMA nodes [10]. The paper does not present the detailed overhead. Numaprof also uses the binary instrumentation (i.e., PIN) to collect and identify local and remote memory accesses [30]. Numaprof relies on a specific thread binding to detect remote accesses, which shares the same shortcoming as other existing work [24, 35]. Numaprof also shares the same issues with other tools, which only focuses on remote accesses while omitting other issues such as cache coherence issues and imbalance issues. In addition, Numaprof is only a code-based profiler that could only report program statements with excessive remote memory access, which requires programmers to figure out the data (object) and a specific strategy. Due to this shortcoming, it makes the comparison with Numaprof extremely difficult and time-consuming. In contrast, although NumaPerf also utilizes fine-grained measurement, it detects more issues that may cause performance issues in any NUMA architecture, and provides more useful information for bug fixes.

Coarse-Grained Approaches: Many tools employ hardware Performance Monitoring Units (PMU) to identify NUMA-related performance issues, such as VTune [14], Memphis [26], MemProf [18], Xu et al. [24], NumaMMA [32], and LaProf [35], where their difference are further described in the following. Both VTune [14] and Memphis [26] only detects NUMA-performance issues on statically-linked variables. MemProf proposes the employment of hardware Performance Monitoring Units (PMU) to identify NUMA-related performance issues [18], with the focus on remote accesses. It constructs data flow between threads and objects to help understand NUMA performance issues. One drawback of MemProf is that it requires an additional kernel module that may prevent people of using it. Similarly, Xu et al. also employ PMU to detect NUMA performance issues [24], but without the change of the kernel. It further proposes a new metric, the NUMA latency per instruction, to evaluate the seriousness of NUMA issues. This tool has a drawback that it statically binds every thread to each node, which may miss some NUMA issues due to its static binding. NumaMMA also collects traces with PMU hardware, but focuses on the visualization of memory accesses [32]. LaProf focuses on multiple issues that may

cause performances issues in NUMA architecture [35], including data sharing, shared resource contention, and remote imbalance. LaProf has the same shortcoming by binding every thread statically. Overall, these sampling-based approaches although imposes much lower overhead, making them applicable even for the production environment, they cannot detect all NUMA performance issues especially when most of them only focus on remote accesses. In contrast, NumaPerf aims to detect performance issues inside development phases, avoiding any additional runtime overhead. Also, NumaPerf focuses more aspects with a predictive approach, not just limited to remote accesses in the current hardware. Our evaluation results confirm NumaPerf’s comprehensiveness and effectiveness.

6.2 Other Related Tools

RTHMS also employs PIN to collect memory traces, and then assigns a score to each object-to-memory based on its algorithms [28]. It aims for identifying the performance issues for the hybrid DRAM-HBM architecture, but not the NUMA architecture, and has a higher overhead than NumaPerf. Some tools focus on the detection of false/true sharing issues [8, 13, 21–23], but skipping other NUMA issues.

SyncPerf also detects load imbalance and predicts the optimal thread assignment [1]. SyncPerf aims to achieve the optimal thread assignment by balancing the waiting time of each types of threads. In contrast, NumaPerf suggests the optimal thread assignment based the number of accesses of each thread, which indicates the actual workload.

7 Conclusion

Parallel applications running on NUMA machines are prone to different types of performance issues. Existing NUMA profilers may miss significant portion of optimization opportunities. Further, they are bound to a specific NUMA topology. Different from them, NumaPerf proposes an architecture-independent and scheduling-independent method that could detect NUMA issues even without running on a NUMA machine. Comparing to existing NUMA profilers, NumaPerf detects more performance issues without false alarms, and also provides more helpful information to assist bug fixes. In summary, NumaPerf will be an indispensable tool that could identify NUMA issues in development phases.

References

- [1] Mohammad Mejbah ul Alam, Tongping Liu, Guangming Zeng, and Abdullah Muzahid. Syncperf: Categorizing, detecting, and diagnosing synchronization performance bugs. In *Proceedings of the Twelfth European Conference on Computer Systems, EuroSys '17*, pages 298–313, New York, NY, USA, 2017. ACM.
- [2] David Beniamine, Matthias Diener, Guillaume Huard, and Philippe O. A. Navaux. Tabarnac: Visualizing and resolving memory access issues on numa architectures. In *Proceedings of the 2nd Workshop on Visual Performance Analysis, VPA '15*, New York, NY, USA, 2015. Association for Computing Machinery.
- [3] Emery D. Berger, Kathryn S. McKinley, Robert D. Blumofe, and Paul R. Wilson. Hoard: a scalable memory allocator for multithreaded applications. In *ASPLOS-IX: Proceedings of the ninth international conference on Architectural support for programming languages and operating systems*, pages 117–128, New York, NY, USA, 2000. ACM Press.
- [4] Christian Biemni and Kai Li. PARSEC 2.0: A new benchmark suite for chip-multiprocessors. In *Proceedings of the 5th Annual Workshop on Modeling, Benchmarking and Simulation*, June 2009.
- [5] Sergey Blagodurov, Sergey Zhuravlev, Mohammad Dashti, and Alexandra Fedorova. A case for numa-aware contention management on multicore systems. In *Proceedings of the 2011 USENIX Conference on USENIX Annual Technical Conference, USENIXATC'11*, pages 1–1, Berkeley, CA, USA, 2011. USENIX Association.
- [6] William J. Bolosky, Michael L. Scott, Robert P. Fitzgerald, Robert J. Fowler, and Alan L. Cox. Numa policies and their relation to memory architecture. In *Proceedings of the Fourth International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS IV*, pages 212–221, New York, NY, USA, 1991. ACM.
- [7] Derek Bruening, Timothy Garnett, and Saman Amarasinghe. An infrastructure for adaptive dynamic optimization. In *Proceedings of the International Symposium on Code Generation and Optimization: Feedback-Directed and Runtime Optimization, CGO '03*, page 265–275, USA, 2003. IEEE Computer Society.
- [8] Milind Chabbi, Shasha Wen, and Xu Liu. Featherlight on-the-fly false-sharing detection. In Andreas Krall and Thomas R. Gross, editors, *Proceedings of the 23rd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP 2018, Vienna, Austria, February 24–28, 2018*, pages 152–167. ACM, 2018.
- [9] Charlie Curtsinger and Emery D. Berger. Coz: Finding code that counts with causal profiling. In *Proceedings of the 25th Symposium on Operating Systems Principles, SOSP '15*, pages 184–197, New York, NY, USA, 2015. ACM.
- [10] Matthias Diener, Eduardo HM Cruz, Laércio L Pilla, Fabrice Dupros, and Philippe OA Navaux. Characterizing communication and page usage of parallel applications for thread and data mapping. *Performance Evaluation*, 88:18–36, 2015.
- [11] Stephane Eranian, Eric Gouriou, Tipp Moseley, and Willem de Bruijn. Linux kernel profiling with perf. <https://perf.wiki.kernel.org/index.php/Tutorial>, 2015.
- [12] Susan L. Graham, Peter B. Kessler, and Marshall K. McKusick. gprof: a call graph execution profiler. In *SIGPLAN Symposium on Compiler Construction*, pages 120–126, 1982.
- [13] Christian Helm and Kenjiro Taura. Perfmemplus: A tool for automatic discovery of memory performance problems. In *International Conference on High Performance Computing*, pages 209–226. Springer, 2019.
- [14] Intel Corporation. Intel VTune performance analyzer. <http://www.intel.com/software/products/vtune>.
- [15] Lawrence Livermore National Laboratory. Livermore unstructured Lagrangian explicit shock hydrodynamics (lulesh). <https://codesign.llnl.gov/lulesh.php>, Dec 2010.
- [16] Lawrence Livermore National Laboratory. Llnl coral benchmarks. <https://asc.llnl.gov/CORAL-benchmarks>, Dec 2013.
- [17] Lawrence Livermore National Laboratory. Llnl sequoia benchmarks. <https://asc.llnl.gov/sequoia/benchmarks>, Dec 2013.
- [18] Renaud Lachaize, Baptiste Lepers, and Vivien Quéma. Memprof: A memory profiler for numa multicore systems. In *Proceedings of the 2012 USENIX Conference on Annual Technical Conference, USENIX ATC'12*, pages 5–5, Berkeley, CA, USA, 2012. USENIX Association.
- [19] Christoph Lameter. An overview of non-uniform memory access. *Commun. ACM*, 56(9):59–54, September 2013.
- [20] Chris Lattner and Vikram Adve. LLVM: A compilation framework for lifelong program analysis & transformation. In *Proceedings of the 33rd ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '12, San Jose, CA, USA, June 10–15, 2012*, pages 57–68. ACM, 2012.

the International Symposium on Code Generation and Optimization: Feedback-directed and Runtime Optimization, CGO '04, pages 75–, Washington, DC, USA, 2004. IEEE Computer Society.

[21] Tongping Liu and Emery D. Berger. Sheriff: precise detection and automatic mitigation of false sharing. In *Proceedings of the 2011 ACM international conference on Object oriented programming systems languages and applications*, OOPSLA '11, pages 3–18, New York, NY, USA, 2011. ACM.

[22] Tongping Liu and Xu Liu. Cheetah: Detecting false sharing efficiently and effectively. In *Proceedings of the 2016 International Symposium on Code Generation and Optimization*, CGO 2016, pages 1–11, New York, NY, USA, 2016. ACM.

[23] Tongping Liu, Chen Tian, Hu Ziang, and Emery D. Berger. Predator: Predictive false sharing detection. In *Proceedings of 19th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, PPOPP'14, New York, NY, USA, 2014. ACM.

[24] Xu Liu and John Mellor-Crummey. A tool to analyze the performance of multithreaded programs on numa architectures. In *Proceedings of the 19th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, PPoPP '14, pages 259–272, New York, NY, USA, 2014. ACM.

[25] Chi-Keung Luk, Robert Cohn, Robert Muth, Harish Patil, Artur Klauser, Geoff Lowney, Steven Wallace, Vijay Janapa Reddi, and Kim Hazelwood. Pin: Building customized program analysis tools with dynamic instrumentation. In *Proceedings of the 2005 ACM SIGPLAN Conference on Programming Language Design and Implementation*, PLDI '05, pages 190–200, New York, NY, USA, 2005. ACM.

[26] C. McCurdy and J. Vetter. Memphis: Finding and fixing numa-related performance problems on multi-core platforms. In *2010 IEEE International Symposium on Performance Analysis of Systems Software (ISPASS)*, pages 87–96, March 2010.

[27] Nicholas Nethercote and Julian Seward. Valgrind: A framework for heavyweight dynamic binary instrumentation. In *Proceedings of the 28th ACM SIGPLAN Conference on Programming Language Design and Implementation*, PLDI '07, page 89–100, New York, NY, USA, 2007. Association for Computing Machinery.

[28] Ivy Bo Peng, Roberto Gioiosa, Gokcen Kestor, Pietro Cicotti, Erwin Laure, and Stefano Markidis. Rthms: A tool for data placement on hybrid memory system. In *Proceedings of the 2017 ACM SIGPLAN International Symposium on Memory Management*, ISMM 2017, page 82–91, New York, NY, USA, 2017. Association for Computing Machinery.

[29] Ashay Rane and James Browne. Enhancing performance optimization of multicore chips and multichip nodes with data structure metrics. In *Proceedings of the 21st International Conference on Parallel Architectures and Compilation Techniques*, PACT '12, pages 147–156, New York, NY, USA, 2012. ACM.

[30] Othman Bouizi Sébastien Valat. Numaprof, a numa memory profiler. In *Mencagli G. et al. (eds) Euro-Par 2018: Parallel Processing Workshops. Euro-Par 2018. Lecture Notes in Computer Science*, vol 11339. Springer, Cham, pages 159–170, December 2018.

[31] Konstantin Serebryany, Derek Bruening, Alexander Potapenko, and Dmitry Vyukov. Addresssanitizer: A fast address sanity checker. In *Proceedings of the 2012 USENIX Conference on Annual Technical Conference*, USENIX ATC'12, pages 28–28, Berkeley, CA, USA, 2012. USENIX Association.

[32] François Trahay, Manuel Selva, Lionel Morel, and Kevin Marquet. Numamma: Numa memory analyzer. In *Proceedings of the 47th International Conference on Parallel Processing*, ICPP 2018, New York, NY, USA, 2018. Association for Computing Machinery.

[33] R. Yang, J. Antony, A. Rendell, D. Robson, and P. Strazdins. Profiling directed numa optimization on linux systems: A case study of the gaussian computational chemistry code. In *2011 IEEE International Parallel Distributed Processing Symposium*, pages 1046–1057, May 2011.

[34] Qin Zhao, David Koh, Syed Raza, Derek Bruening, Weng-Fai Wong, and Saman Amarasinghe. Dynamic cache contention detection in multi-threaded applications. In *The International Conference on Virtual Execution Environments*, Newport Beach, CA, Mar 2011.

[35] L. Zhu, H. Jin, and X. Liao. A tool to detect performance problems of multi-threaded programs on numa systems. In *2016 IEEE Trust-com/BiGDataSE/ISPA*, pages 1145–1152, 2016.