Techniques Supporting \texttt{threadprivate} in OpenMP

Xavier Martorell, Marc Gonzàlez, Alex Duran, Jairo Balart, Roger Ferrer, Eduard Ayguadé and Jesús Labarta
Barcelona Supercomputing Center - Computer Architecture Dept. - Technical Univ. of Catalunya
Campus Nord, Modul C6, c/Jordi Girona 1-3 08034 Barcelona Spain
\{xavim, marc, aduran, jbalart, rferrer, eduard, jesus\}@ac.upc.edu

Abstract

This paper presents the alternatives available to support \texttt{threadprivate} data in OpenMP and evaluates them. We show how current compilation systems rely on custom techniques for implementing thread-local data. But in fact the ELF binary specification currently supports data sections that become threadprivate by default. ELF naming for such areas is Thread-Local Storage (TLS). Our experiments demonstrate that implementing threadprivate based on the TLS support is very easy, and more efficient. This proposal goes in the same line as the future implementation of OpenMP on the GNU compiler collection.

In addition, our experience with the use of \texttt{threadprivate} in OpenMP applications shows that usually it is better to avoid it. This is because threadprivate variables reside in common blocks and they impede the compiler to fully optimize the code. So it is better to keep \texttt{threadprivate} as a temporary technique only to ease porting MPI codes to OpenMP.

1. Introduction and motivation

OpenMP [16] is a mature programming model for shared memory architectures, providing incremental parallelization over sequential codes. An OpenMP programmer usually starts from an initial sequential code, in C/C++ or Fortran and adds pragmas/directives to express the parallelism. A number of parallel constructs are offered, namely, parallel regions, loops, sections and combinations of them. Usually, it supports the exploitation of multiple levels of parallelism.

Nearly during the past decade, from the very beginning of OpenMP, we have been developing a threading library and a compilation infrastructure supporting this shared-memory programming model. NthLib [13] [14] [1] was designed to efficiently support multiple levels of parallelism, and as much fine grained as possible. The compilation infrastructure (Nanos Compiler [6]) was initially based on Parafrase-2 [17], and it is still used in the experiments for this paper. We are currently migrating to a new compilation infrastructure named Nanos Mercurium [2].

During all this time, there has been one OpenMP construction that has consistently resisted to be implemented in our environment. That was the \texttt{threadprivate} directive. Every attempt was giving too complex and inefficient implementations, or not general enough. We have also tried to make the work on the compiler side simple and let the run-time system to deal with most of the features. Also, initially, \texttt{threadprivate} constructions were rarely found in OpenMP applications. But this is not currently the case, specially since more and more applications written initially in MPI are being recoded using OpenMP. For this reason it was important to find a way to implement \texttt{threadprivate} in a general and efficient way.

Our current proposal is simple, both from the compiler and the runtime system perspective, mainly because it is based on a new feature provided by the Executable and Linking Format (ELF) specification, which allows to implement thread-local storage (TLS) efficiently.

The rest of the paper is organized as follows: Section 2 outlines the current support for the \texttt{threadprivate} directive found in several commercial compilers supporting OpenMP. This section also presents the related work. Section 3 presents our approach to implement \texttt{threadprivate} based on the TLS extension to ELF files. Section 4 shows the evaluation of our proposal with respect to not using the \texttt{threadprivate} directive at all. Finally, Section 5 concludes and presents future work.

2. Current support for \texttt{threadprivate}

Currently, OpenMP compilers support \texttt{threadprivate} variables by letting the runtime system to allocate a memory region for each thread. This is usually done during the initialization of the runtime system. The original variable is declared as a \texttt{common} area in the binary file. It is used as the place where the sequential code and the master thread access to find the values. The amount of
The program code is then changed, so that every access to the threadprivate variable is done by getting a pointer to the thread area first and then applying an offset to access to actual value. The code transformation used in the IBM XL compilers is shown in Figure 1. This transformation must be done once for every different threadprivate common block accessed in a subroutine.

This way of implementing the threadprivate access has a large overhead compared to a usual variable access, residing either globally or in the stack. A simple microbenchmark consisting of the subroutine presented in Figure 2 has been developed to reveal this overhead. The subroutine (tps) does a single variable access for each invocation. This way, its execution time is heavily dependent on the way such an access is implemented, either as a regular variable or as a threadprivate variable. Accessing the variable as a threadprivate value has, in this particular situation, a cost which easily doubles the cost of accessing a variable in a regular common block. The effects of this overhead in the execution time are shown in Figure 3, using different operating systems, OpenMP compilers, and architectures. All executions are done with a single thread, so that only the overhead introduced by the different code generation is shown. In these tests, the subroutine shown in Figure 2 is executed 5 million times and the execution time is measured.

The experiments shown in this Figure are executed in a Linux platform based on Pentium Xeon [11] and the Linux Intel Fortran compiler 8.1 [10], a Linux platform based on POWER5 [8] and the IBM XL compiler 10.1 compiler [9], and a Solaris platform based on a ULTRA SPARC processor [22] and the SUNWpro 10 Studio [20]. All these environments use a similar technique for implementing threadprivate variables, as explained above in this section.

As it can be observed, the results consistently show that the implementation of threadprivate variables in these three different architectures adds much more overhead with respect to an access to a common block. The execution time of the microbenchmark in each one of the architectures depends only on the processor speed. The benchmark is very small, it fits in the cache memory of the processors, so there are no memory effects. This is the reason why the Pentium architecture running at 2.8 Ghz. is the fastest executing this microbenchmark.

There exist some research OpenMP environments that use also similar techniques to implement the threadprivate directive. Among them, the Balder runtime system [12] is based on the OdinMP compiler to transform all accesses to threadprivate symbols towards an area previously allocated. This code transformation adds a call to the runtime system to determine a thread pointer to be used for such access. The Omni [19] compiler and runtime system also generates a code based on a runtime system call to get access to the threadprivate data.
From the perspective of a Fortran compiler, *threadprivate* values always reside in a common block. This is true even if the *threadprivate* directive is applied to individual variables, as it is allowed in OpenMP. This is to ensure that the values are kept across parallel regions. In C, they reside in a global variable. Fortran common blocks and C global variables (scalars, arrays, and structures) are the same from the perspective of the linker. This is the way C programs can access values in Fortran commons, and viceversa, allowing interoperability between these languages. So it is not a matter of the *threadprivate* directive, but a matter of optimizing access to common blocks and global variables. With variables residing in common blocks, the compilers are tied to the common block behaviour and they can not apply certain code optimizations: It is not possible to place a *threadprivate* scalar value in a processor register unless the compiler does interprocedural analysis: The code inside a function can cache the value in a register while it does not call another function. When a new call is reached, the value must be stored in its place in memory, so that the new function (possibly not available to the compiler, but to the linker) will get the initial value of the variable from the common block. *Threadprivate* simply makes more complex loading and storing the value from memory, increasing the overhead, as the previous experiment shows.

In the next section, we present an alternative implementation that reduces this overhead to a small fraction.

### 3. Implementing *threadprivate* using the ELF TLS specification

Starting in January 2002, the ELF binary file specification [3][23][18] includes the possibility to declare and use thread-local storage. Thread-local storage contains thread-local variables. A thread-local variable is implemented in such a way that each thread in the program can assign a different value to it. Each copy of the variable has a different address. It is the responsibility of the runtime system to allocate and manage the memory area to support the local storage of each thread. Programmers can easily define thread-local variables using the C language (non-standard) __thread attribute when declaring global or static variables [3]. Variables declared with this attribute are allocated by the compiler in two new object file sections: a .tdata section is used to place initialized TLS variables, and a .tbss section holds non-initialized variables. Any scalar, array or structure variable in C can be attributed as __thread.

While thinking in a way to implement the *threadprivate* directive for our OpenMP runtime system (Nanos), we thought that using this feature should provide several benefits:

- Ease the development of our prototype source-to-source compilers (NanosCompiler [6] and Mercurium [2]).
- Simplify code generation with respect to other techniques.
- Achieve better performance than current implementations.
- Achieve automatic interoperability between C/C++ and Fortran (assuming the compilers of these languages all use the same conventions for *threadprivate*).

In the next subsections, we describe the modifications introduced in the compiler and inside the run-time system.

#### 3.1. Modifications to the compiler

The approach to achieve the previous benefits from the compiler was simple: every variable in a common block annotated as *threadprivate* should be attributed as __thread in the code generated by the compilers. This is what NanosCompiler and Mercurium do to translate the *threadprivate* directive. Any variable reference is kept the same. __thread variables will be allocated in the .tbss and .tdata sections and the ELF specification will take care of the rest.

This solution is fine in the case of C/C++: we can use native compilers supporting the __thread attribute in the target system to generate machine code (gcc supports it).

The problem appears in Fortran: even the Fortran specification 2003 [4] does not include any support for threading, so our compiler cannot generate anything to express thread-local storage. The Fortran 2005 specification, which is currently under discussion (and planned for 2008) comes with the solution. It includes a subset of Co-Array Fortran [15], in particular the possibility of expressing per-image variables, where each image can be a different thread running on the same address space, or a process running in a different node.

But meanwhile, we thought useful to look inside the gcc compiler [5] to see whether it could be easy to include some (partial) *threadprivate* support. At that time gcc 4.0 was already stable enough to be used with our benchmarks, so we selected it for examination.

It happens that gcc 4.0 includes gfortran, with good support for Fortran 95. As gfortran and gcc share the same backend for generating assembly code in each architecture, it seemed reasonable to modify it, including our proposed extensions.

The changes to gfortran were simple:

- Recognize a new token to express that a variable is thread-local.
- Mark the thread-local symbol with the same flag used in C.
- Allow the generation of common thread-local data. This was already supported by the compiler,
but it was forbidden by a previous conditional statement.

- Add code generation support for `.tdata` and `.tbss` sections.

With these changes, gfortran generates common blocks into the `.tbss` and `.tdata` sections and the appropriate code to access them correctly. In addition, they are accessible from object files generated by GCC using variables with the same name and the __thread attribute. We know that the development of OpenMP in gfortran is going in the same direction [7].

With our current approach, the subroutine shown in Figure 2 is transformed in the intermediate and assembly code presented in Figure 4. The assembly code is presented in its Intel Pentium version.

In the Figure, we can observe that the code generated is much more simple compared to the code presented in Figure 1. In particular, as the thread has a dedicated register (gs) pointing to the TLS area, there is no need to call a function to obtain this pointer (_xlGetThStorageBlock in Figure 1).

### 3.2. Runtime system support

As our OpenMP runtime system (NthLib [1]) runs directly on top of the kernel-level abstraction offering the parallelism, it has been modified to create the kernel level entities with the TLS characteristics.

The usual case is the kernel-level thread, which can be a thr in Solaris, or a clone in Linux. In Solaris, just linking against the librt system library is enough to have available a memory region for each thread to act as the thread-local storage. The mechanism used is described in [21].

In the Linux implementations on Pentium and Power processors, we have modified NthLib to allocate a region of thread-local storage for each processor requested for OpenMP. This has been an easy task, as the address space of the parallel applications was already well organized. Figure 5 shows the resulting organization of the address space. We try to leave as much space as possible on the lower addresses (left part of the Figure) to the application code and data. We let the operating system map the dynamic libraries on its standard address. At the higher addresses (right hand side of the figure), we map the thread stacks and the thread-local storage (TLS) areas.

This is different from what we have seen in the Pthreads library in Linux, which maps the TLS areas inside the thread stacks. This has a strange effect for the programmer: the TLS size depends on the size of the pthread stacks, and not on the size of the actual `.tbss` and `.tdata` sections. This means that the pthread library attribute pthread_stack_size is in fact controlling the size of both the stack and the TLS areas for each pthread.

### 4. Evaluation

We have evaluated the TLS proposal for threadprivate, on one hand with the same microbenchmark used for the motivation, and on the other hand, using the NAS benchmarks written in Fortran which use the threadprivate directive (BT, SP, LU and CG). We have used Classes S and A of the NAS benchmarks to show how the technique influence the performance of applications with different working-set sizes.
The results of the *threadprivate* microbenchmark generated using our modified gfortran compiler are presented in Figure 6. As the benchmark has been now compiled with gfortran, the results labeled *common* are a little bit different than the previous results presented, labeled *common* in Figure 3. The important observation here is that the performance to access *threadprivate* variables is now very similar to when accessing a regular common block.

### 4.2. Evaluation of the NAS benchmarks

We have taken the NAS benchmarks version 3.2, where four of the benchmarks written in Fortran use the *threadprivate* directive. The directive is used in different ways, so that different amount of data is privatized in a per-thread manner, depending on the benchmark. Table 1 shows how BT and SP use *threadprivate* for two common structures (*work_1d* and *work_lhs*) containing around 10 different arrays. In addition, BT has 3 double precision scalar values on the *work_lhs* common structure. LU and CG use *threadprivate* to make some scalar values private. Those values are the amount of threads participating on the parallel region and the thread self identifier (obtained through the primitives `omp_get_num_threads()` and `omp_get_thread_num()`, respectively). In addition, CG has the *threadprivate* structure `urando` containing two double precision numbers to compute random numbers in a per-thread basis.

The evaluation of the benchmarks has been carried on a POWER5-based machine with 2 chips, giving a total of 8 hardware threads, running the Linux 2.6 operating system.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Scalars</th>
<th>Arrays</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT</td>
<td><code>tmp1, tmp2, tmp3</code></td>
<td><code>work_1d, work_lhs</code></td>
</tr>
<tr>
<td>SP</td>
<td>-</td>
<td><code>work_1d, work_lhs</code></td>
</tr>
<tr>
<td>LU</td>
<td><code>threadinfo</code></td>
<td>-</td>
</tr>
<tr>
<td>CG</td>
<td><code>info, urando</code></td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1: Usage of the *threadprivate* directive in the Fortran NAS3.2 benchmarks

**Evaluation of the CLASS S version.** To evaluate the proposal described in this paper, we decided to compare the performance obtained with the new implementation of *threadprivate* with the same application coded without the *threadprivate* directive. It is in fact easy to remove such directives from the source code in a way that the application is still running correctly and providing successful results: The arrays in BT and SP can be made just private, as there is no reuse of their values across parallel sections, so they must not be maintained.

In the case of the variables representing the number of threads and the thread identifier, in LU and CG, they are simply replaced by private variables that take the value of `omp_get_num_threads()` and `omp_get_thread_num()`, and that are assigned each time a parallel region using the variable starts. There are three parallel regions using this variables in LU and CG.

Table 1 presents the evaluation of these benchmarks. It shows the megaflops per second obtained by the benchmark, depending on the number of hardware threads used and the technique used for privatization.

Initially, we started comparing the *threadprivate* common version with the *private-only* version. The three benchmarks SP, LU and CG had a similar performance, independently of the kind of privatization used. Observe that this means that accessing arrays in *threadprivate* with our technique offers similar performance to accessing private arrays located in the stack. On the other hand, the access to *threadprivate* scalar values is similar to getting their value using simple library calls like `omp_get_num_thread()`.

The BT benchmark was the exception: The performance of the *threadprivate* version was clearly lower, even when executed in a single processor. Observe in that the performance of *threadprivate* in BT.S with 1 thread is lower than 700 Mflops/s, while using private variables is 835 Mflops/s. After some thoughts and tests, we found that the problem was that mixing scalar and array variables in *threadprivate* commons was not a good idea. In particular, any variable residing in a common in Fortran has a common association, that may forbid some optimizations to the compiler. For this reason, we also present the intermediate version (labeled *threadprivate variables*), where the *threadprivate* directive is applied to the individual variables, instead of to the full common. In this case, the performance is closer to the *private-only* solution. The difference comes from the fact that a common for each *threadprivate* variable is still introduced in our code.
transformation for such variables, while in the private-only solution they may be just allocated into a processor register.

**Evaluation of NAS CLASS A.** To check that the performance obtained in the CLASS S is maintained when the working-set of the benchmarks is increased, we present in Figure 8 the results for the CLASS A of the benchmarks.

In this Figure, we can observe that the behaviour of the benchmarks is very similar. BT reflects the same problem related to common association when making the common threadprivate. Still, forcing the disassociation of the variables in the intermediate approach (labeled threadprivate variables) is not so good as making them really private.

It is interesting to comment that the performance of all the benchmarks is reduced when executed in a small number of threads, compared to the class S version. The cause of this is that, as the working-set increases, the execution suffers from an increment in the number of cache misses. This effect is nevertheless amortized through the parallelization in BT, SP, and LU. With 4 to 8 threads, these class A applications perform better than the class S ones. It is not the same with CG, in which in class A, the amount of sharing increases with the increase in the working-set and the performance is comparatively degraded, with respect to the class S.

5. Conclusions and Future Work

In this paper, we have studied the current techniques used by various OpenMP compilers to implement threadprivate variables. A first conclusion is that threadprivate must be used with care. Do not consider that a threadprivate variable can be as efficient as a private variable. Threadprivate implies that variables must be placed in common blocks to be kept save across parallel regions. This may impede some optimizations, like keeping values in registers and the code will be slower. So use threadprivate only when the variables involved need to keep their values across functions and parallel regions.

In addition, we have proposed an implementation of threadprivate based on thread-local storage (TLS), a support introduced relatively recently in the ELF specification. We have implemented such proposal in the gfortran compiler and our OpenMP runtime system (Nanos). The evaluation of this proposal on a microbenchmark shows that the fine-grain overhead is highly reduced and the measurements are close to those obtained when accessing private variables. Also, it is shown that the performance obtained from the NAS benchmarks (Classes S and A) are comparable to the ones using strictly private variables. Nevertheless, it is also important to note that using threadprivate variables, and depending on the way they are finally placed on common blocks, there are applications (e.g., BT in our experiments) that suffer from the lower level of optimization achieved by the compiler.

Our current work is the porting of this feature to our new line of compilers, Nanos Mercurium, supporting both C and Fortran95, and achieve full interoperability between both languages. We will also maintain the modifications to gfortran across different versions of the compiler, while possible, with the goal to continue being able to work in a source-to-source way. In this line, adopting a syntax in accordance with Co-Array Fortran for our intermediate language seems the appropriate direction to follow to allow Fortran applications to be conscious about threads.

We would also like to continue with the evaluation of the proposal using other kinds of applications (e.g., integer, multimedia applications) to have a better understanding.
with respect the differences of keeping the data private or making it threadprivate.

Acknowledgements

This work has been supported by the Ministry of Education of Spain under contract CICYT-TIN2004-07739-C02-01.

References