This version contains an evaluation that is slightly updated and corrected from the version appearing at ISCA 2004.

iWatcher: Efficient Architectural Support for Software Debugging *

Pin Zhou, **Feng Qin**, **Wei Liu**, **Yuanyuan Zhou**, and **Josep Torrellas**
Department of Computer Science
University of Illinois at Urbana-Champaign
{pinzhou,fengqin,liuwei,yyzhou,torrellas}@cs.uiuc.edu

Abstract

Recent impressive performance improvements in computer architecture have not led to significant gains in ease of debugging. Software debugging often relies on inserting run-time software checks. In many cases, however, it is hard to find the root cause of a bug. Moreover, program execution typically slows down significantly, often by 10-100 times.

To address this problem, this paper introduces the *Intelligent Watcher (iWatcher)*, novel architectural support to monitor dynamic execution with minimal overhead, automatically, and flexibly. iWatcher associates program-specified monitoring functions with memory locations. When any such location is accessed, the monitoring function is automatically triggered with low overhead. To further reduce overhead and support rollback, iWatcher can leverage Thread-Level Speculation (TLS). To test iWatcher, we use applications with various bugs. Our results show that iWatcher detects many more software bugs than Valgrind, a well-known open-source bug detector. Moreover, iWatcher only induces a 4-80% execution overhead, which is orders of magnitude less than Valgrind. Even with 20% of the dynamic loads monitored in a program, iWatcher adds only 66-174% overhead. Finally, TLS is effective at reducing overheads for programs with substantial monitoring.

1. Introduction

1.1. Motivation

Despite costly efforts to improve software-development methodologies, software bugs in deployed codes continue to thrive, often accounting for as much as 40% of computer system failures [24]. Software bugs can crash systems, making services unavailable or, in the form of "silent" bugs, corrupt information or generate wrong outputs. According to NIST [26], software bugs cost the U.S. economy an estimated $59.5 billion annually, or 0.6% of the GDP!

There are several approaches to debug codes. One approach is to perform checks statically. Examples of this approach include explicit model checking [25, 42] and program analysis [3, 8, 12]. Most static tools require significant involvement of the programmer to write specifications or annotate programs. In addition, most static tools are limited by aliasing problems and other compile-time limitations. This is especially the case for programs written in unsafe languages such as C or C++, the predominant programming languages in industry. As a result, many bugs often remain in programs even after aggressive static checking.

Another approach is to monitor execution dynamically, with instrumentation inserted in the code that monitors invariants and reports violations as errors. The strength of this approach is that the analysis is based on actual execution paths and accurate values of variables and aliasing information. Examples of dynamic monitors include Purify [14], Valgrind [36], Intel thread checker [17], DIDUCE [13], Eraser [33], CCured [5, 27], and other tools [1, 6, 23, 30, 31].

Unfortunately, most dynamic checkers suffer from two limitations. First, they are often computationally expensive. One major reason is their large instrumentation cost. Another reason is that dynamic checkers may instrument more places than necessary due to lack of accurate information at instrumentation time. As a result, some dynamic checkers slow down a program by 6-30 times [13, 33], which makes such tools undesirable for production runs. Moreover, some timing-sensitive bugs may never occur with these slowdowns.

Second, most dynamic checkers rely on compilers or preprocessing tools to insert instrumentation and, therefore, are limited by imperfect variable disambiguation. Consequently, some accesses to a monitored location may be missed by the instrumentation tool. Because of this reason, some bugs are caught much later than when they actually occur, which makes it hard to find the root cause of the bug. The following C code gives a simple example.

```
int x, *p;
           /* assume invariant: x = 1 */
...
p = foo(); /* a bug: p points to x incorrectly */
*p = 5;    /* line A:  corruption of x */
...
InvariantCheck(x == 1);  /* line B */
z = Array[x];
...
```

While x is corrupted in line A, the bug is not detected until the invariant check at line B. Due to the difficulty of performing perfect pointer disambiguation, it may be hard for a dynamic checker to know that it needs to insert an invariant check after line A.

To assist software debugging, several processor architectures such as Intel x86 and Sun SPARC provide support for watchpoints to monitor several programmer-specified memory locations [15, 18, 39, 45]. When a watched memory location is accessed, the hardware triggers an exception that is handled by the debugger. It is then up to the programmer to manually check the program state. While watchpoints are a good starting point, they have several limitations. First, they do not support *low-overhead* checks on variable

*This work was supported in part by NSF under grants CCR-0325603, EIA-0072102, and CHE-0121357; by DARPA under grant F30602-01-C-0078; by an IBM SUR grant; and by additional gifts from IBM and Intel.

values *automatically*. Since exceptions are expensive, it would be very inefficient to use them for dynamic bug detection during production runs. Second, most architectures only support a handful of watchpoints (four in Intel x86). Therefore, it is hard to use watchpoints for dynamic monitoring in production runs, which requires efficiency and watching many memory locations.

1.2. Our Approach

This paper introduces the *Intelligent Watcher (iWatcher)*, novel architectural support to monitor dynamic execution with *minimal overhead*, *automatically*, and *flexibly*. iWatcher associates program-specified monitoring functions with memory locations. When any such location is accessed, the monitoring function is automatically triggered with low overhead. To further reduce overhead and support rollback, iWatcher can leverage Thread-Level Speculation (TLS). The main advantages of iWatcher are:

- It monitors *all* accesses to the watched memory locations. Consequently, it catches hard-to-find bugs such as updates through aliased pointers and stack-smashing attacks commonly exploited by viruses.
- It has low overhead because it (i) only monitors memory instructions that *truly* access the watched memory locations, (ii) uses minimal-overhead hardware-supported triggering of monitoring functions, and (iii) can leverage TLS to execute monitoring functions in parallel with the program.
- It is flexible in that it can support a wide range of checks, including program-specific checks. Moreover, iWatcher is language independent, cross-module and cross-developer.

We evaluate iWatcher using buggy applications with memory corruption, memory leaks, buffer overflow, value invariant violations, outbound pointers, and smashed stacks. iWatcher detects all the bugs evaluated in our experiments with only a 4-80% execution overhead. In contrast, a well-known open-source bug detector called Valgrind induces orders of magnitude more overhead, and can only detect a subset of the bugs. Moreover, even with 20% of the dynamic loads monitored in a program, iWatcher only adds 66-174% overhead. Finally, TLS is effective at reducing overheads for programs with substantial monitoring.

This paper is organized as follows. Section 2 briefly describes some background. Sections 3, 4, and 5 describe iWatcher's functionality, architectural design, and advantages. Sections 6 and 7 present the evaluation methodology and experimental results. Section 8 discusses related work and Section 9 concludes.

2. Background

2.1. Dynamic Execution Monitoring

Many methods have been proposed for dynamic code monitoring. The most commonly used ones are assertions, dynamic checkers, and watchpoints.

Assertions. Assertions are inserted by programmers to perform sanity checks at certain places. If the condition specified in an assertion is false, the program aborts. Assertions are one of the most commonly used methods for debugging. However, they can add significant overhead to program execution. Moreover, it is often hard to identify all the places where assertions should be placed.

Dynamic Checkers. Dynamic checkers are automated tools that detect common bugs at run time. For example, DIDUCE [13] automatically infers likely program invariants, and uses them to detect program bugs. Purify [14] and Valgrind [36] monitor memory accesses to detect memory leaks and some simple instances of memory corruption, such as freeing a buffer twice or reading an uninitialized memory location. StackGuard [6] can detect some buffer overflow bugs, which have been a major cause of security attacks. Eraser [33] can detect data races by dynamically tracking the set of locks held during program execution. These tools usually use compilers or code-rewriting tools such as ATOM [40], EEL [20] and Dyninst [2] to instrument programs with checks.

While this approach is promising, dynamic checkers often suffer from the following limitations: (1) aliasing problems, especially in C/C++ programs, (2) high run-time overhead, (3) hard-coded bug detection functionality, (4) language specificity, and (5) difficulty to work with low-level code.

Hardware-Assisted Watchpoints. Hardware-assisted watchpoints [15, 18, 39] use simple hardware support to watch a user-selected memory location. When a watched location is accessed by the program, an exception is handled by an interactive debugger such as gdb. Then, the state of the process can be examined by programmers using the debugger. The hardware support is provided through a few special debug registers. Watchpoints are designed to be used in an interactive debugger. For non-interactive execution monitoring, they are both inflexible and inefficient. They do not provide a way to associate an automatic check to the access of a watched location. Moreover, they require an expensive exception when a watched location is accessed. Finally, most architectures only support a few watchpoints (four in Intel's x86).

Summary. We classify the dynamic monitoring methods into two categories:

- *Code-Controlled Monitoring (CCM).* Monitoring is performed only at special points in the program. Assertions and most dynamic checkers belong to CCM because they only check at assertions or instrumentation points.
- *Location-Controlled Monitoring (LCM).* Monitoring is associated directly with memory locations and therefore all accesses to such memory locations are monitored. Hardware-assisted watchpoints and iWatcher belong to this category.

LCM has two advantages over CCM: (1) LCM monitors *all* accesses to a watched memory location using all possible variable names or pointers, whereas CCM may miss some accesses because of pointer aliasing; (2) LCM monitors only those memory instructions that *truly* access a watched memory location, whereas CCM may need to instrument at many unnecessary points due to the lack of accurate information at instrumentation time. Therefore, LCM can be used to detect both invariant violations and illegal accesses to a memory location, whereas it may be difficult and too expensive for CCM to check for illegal accesses.

2.2. Thread-Level Speculation (TLS)

TLS is an architectural technique for speculative parallelization of sequential programs [4, 38, 41, 44]. TLS support can be built

on a multithreaded architecture, such as simultaneous multithreading (SMT) or chip multiprocessor (CMP) machines. With TLS, the execution of a sequential program is divided into a sequence of *microthreads* (also called tasks, slices, or epochs). These microthreads are then executed speculatively in parallel, while special hardware detects violations of the program's sequential semantics. Any violation results in squashing the incorrectly executed microthreads and re-executing them. To enable squash and re-execution, the memory state of each speculative microthread is typically buffered in caches or special buffers. When a microthread finishes its execution and becomes safe, it can commit. Committing a microthread merges its state with the safe memory. To guarantee sequential semantics, microthreads commit in order.

iWatcher can leverage TLS to reduce monitoring overhead and to support rollback and re-execution of a buggy code region [32]. For our design, we assume an SMT machine, and that the speculative memory state is buffered in caches. However, we believe our iWatcher design can be easily ported to other TLS architectures.

In our design, each cache line is tagged with the ID of the microthread to which the line belongs. Moreover, for each speculative microthread, the processor contains a copy of the initial state of the architectural registers. This copy is generated when the speculative microthread is spawned and is freed when the microthread commits. It is used in case the microthread needs to be rolled back.

The TLS mechanisms for in-cache state buffering and rollback can be reused to support incremental rollback and re-execution of the buggy code [32]. To do this, basic TLS is modified slightly by postponing the commit time of a successful microthread. In basic TLS, a microthread can commit when it completes and all its predecessors have committed. We say that such a microthread is *ready*. To support the rollback of buggy code, a ready microthread commits only in one of two cases: when we need space in the cache and when the number of uncommitted microthreads exceeds a certain threshold. With this support, a ready but uncommitted microthread can still be asked to rollback. This feature can be used to support one of the iWatcher modes (Section 4.5).

3. iWatcher Functionality

iWatcher provides high-flexibility and low-overhead dynamic execution monitoring. It associates program-specified monitoring functions with memory locations. When any such location is accessed, the monitoring function associated with it is automatically triggered and executed.

iWatcher provides two system calls to turn on and off monitoring on a memory location, namely *iWatcherOn* and *iWatcherOff*. These calls can be inserted in programs either automatically by an instrumentation tool or manually by programmers. The following is the *iWatcherOn* interface:

```
iWatcherOn(MemAddr, Length, WatchFlag, ReactMode,
   MonitorFunc, Param1, Param2, ... ParamN)
/* MemAddr: starting address of the memory region*/
/* Length: length of the memory region */
/* WatchFlag: types of accesses to be monitored */
/* ReactMode: reaction mode */
/* MonitorFunc: monitoring function */
/* Param1...ParamN: parameters of MonitorFunc */
```

If a program makes such a call, iWatcher associates monitoring function $MonitorFunc()$ with a memory region of $Length$ bytes starting at $MemAddr$. The $WatchFlag$ specifies what types of accesses to this memory region should be monitored. Its value can be "READONLY", "WRITEONLY", or "READWRITE", in which case the monitoring function is triggered on a read access, write access or both, respectively.

At a *triggering access* (an access to a monitored memory location), the hardware automatically initiates the monitoring function associated with this memory location. The architecture passes the values of $Param1$ through $ParamN$ to the monitoring function. In addition, it also passes information about the triggering access, including the program counter, the type of access (load or store; word, half-word, or byte access), reaction mode, and the memory location being accessed. It is the monitoring function's responsibility to perform the check.

A monitoring function can have side effects and can read and write variables without any restrictions. To avoid recursive triggering of monitoring functions, no memory access performed inside a monitoring function can trigger another monitoring function.

From the programmers' point of view, the execution of a monitoring function follows sequential semantics, just like a very lightweight exception handler (Section 4 describes why monitoring in iWatcher is very lightweight). The semantic order is: the triggering access, the monitoring function, and the rest of the program after the triggering access.

Upon successful completion of a monitoring function, the program continues normally. If the monitoring function fails (returns FALSE), different actions are taken depending on the $ReactMode$ parameter specified in iWatcherOn(). iWatcher supports three modes: $ReportMode$, $BreakMode$ and $RollbackMode$:

- $ReportMode$: The monitoring function reports the outcome of the check and lets the program continue. This mode can be used for profiling and error reporting without interfering with the execution of the program.

- $BreakMode$: The program pauses at the state right after the triggering access and control is passed to an exception handler. Users can potentially attach an interactive debugger, which can be used to find more information.

- $RollbackMode$: The program rolls back to the most recent checkpoint, typically much before the triggering access. This mode can be used to support deterministic replay of a code section to analyze an occurring bug [32], or to support transaction-based programming [29].

A program can associate multiple monitoring functions with the same location. In this case, upon an access to the watched location, all monitoring functions are executed following sequential semantics according to their setup order.

When a program is no longer interested in monitoring a memory region, it turns off the monitoring using

```
iWatcherOff(MemAddr, Length, WatchFlag, MonitorFunc)
/* MemAddr: starting address of the watched region*/
/* Length: length of the watched region */
/* WatchFlag: types of accesses to be unmonitored */
/* MonitorFunc: the monitoring function */
```

After this operation, the $MonitorFunc$ associated with this memory region of $Length$ bytes starting at $MemAddr$ and

$WatchFlag$ is deleted from the system. Other monitoring functions associated with this region are still in effect.

Besides using the iWatcherOff() call to turn off monitoring for a specified memory region, the program can also use a $MonitorFlag$ global switch that enables or disables monitoring on all watched locations. This switch is useful when monitoring overhead is a concern. When the switch is disabled, no location is watched and the overhead imposed is negligible.

Note that iWatcher only provides a very flexible mechanism for dynamic execution monitoring. It is not iWatcher's responsibility to ensure that a monitoring function is written correctly, just like an *assert(condition)* call cannot guarantee that the condition in the code makes sense. Programmers can use invariant-inferring tools such as DIDUCE [13] and DAIKON [9] to automatically insert iWatcherOn() and iWatcherOff() calls into programs.

With this support, we can rewrite the example of Section 1 using iWatcherOn()/iWatcherOff() operations. There is no need to insert the invariant check. iWatcherOn() is inserted at the very beginning of the program so that the system can continuously check x's value whenever and however the memory location is accessed. This way, the bug is caught at line A.

```
int x, *p;
           /* assume invariant: x = 1 */
iWatcherOn(&x, sizeof(int), READWRITE,
           BreakMode, &MonitorX, &x, 1);
...
p = foo(); /* a bug: p points to x incorrectly */
*p = 5;    /* line A: a triggering access */
...
z = Array[x]; /* line B: a triggering access */
...
iWatcherOff(&x, sizeof(int), READWRITE, &MonitorX);

bool MonitorX(int *x, int value){
  return  (*x == value);
}
```

4. Architectural Design of iWatcher

To implement the functionality described above, there are at least four challenges: (1) How to monitor a location? (2) How to detect a triggering access? (3) How to trigger a monitoring function? (4) How to support the three reaction modes? In this section, we first give an overview of the implementation and then show how it addresses these challenges.

4.1. Overview of the Implementation

iWatcher is implemented using a combination of hardware and software. Logically, it has four main parts. First, to detect triggering accesses on small monitored memory regions, we tag cache lines in both L1 and L2 caches with WatchFlags; to detect triggering accesses on large monitored memory regions, we use a small *Range Watch Table (RWT)*. Second, the hardware triggers monitoring functions on the fly and provides a special *Main_check_function* register to store the common entry point for all monitoring functions. Third, we leverage TLS to reduce overheads and support the three reaction modes. Finally, we use software to manage the associations between watched locations and monitoring functions.

Figure 1 gives an overview of the iWatcher hardware. Each L1 and L2 cache line is augmented with WatchFlags. They identify

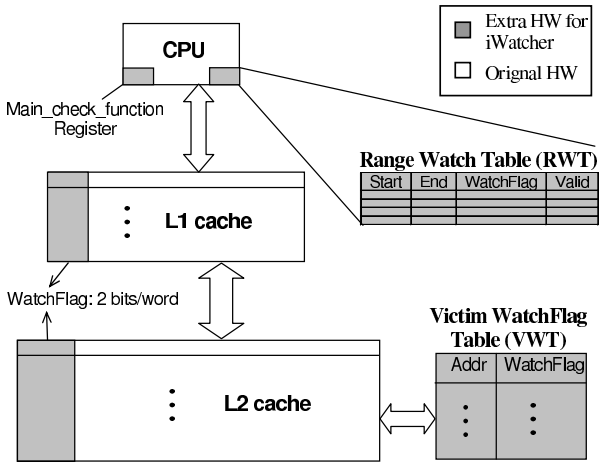


Figure 1. iWatcher hardware architecture.

words belonging to small monitored memory regions. There are two WatchFlag bits per word in the line: a read-monitoring one and a write-monitoring one. If the read (write)-monitoring bit is set for a word, all loads (stores) to this word automatically trigger the corresponding monitoring function. The processor also has a Main_check_function register that holds the address of the *Main_check_function()*, which is the common entry point to all program-specified monitoring functions. In addition, iWatcher also has a *Victim WatchFlag Table (VWT)*, which stores the WatchFlags for watched lines of small regions that have at some point been displaced from L2.

To detect accesses to large (multiple pages) monitored memory regions, iWatcher uses a set of registers organized in the RWT. Each RWT entry stores the virtual start and end addresses of a large region being monitored, plus two bits of WatchFlags and one valid bit. We will see that the RWT is used to prevent large monitored regions overflowing the L2 cache and the VWT. The WatchFlags of these lines do not need to be set in the L1 or L2 cache unless the lines are also included in a small monitored regions. When the RWT is full, additional large monitored regions are treated the same way as small regions.

To reduce monitoring overhead, iWatcher can use TLS to speculatively execute the main program in parallel with monitoring functions. Moreover, iWatcher can also leverage TLS to roll back the buggy code with low overhead, for subsequent replay.

While TLS was also used by Oplinger and Lam to hide overheads [29], iWatcher uses a different TLS spawning mechanism. Specifically, iWatcher uses dynamic hardware spawning, which requires no code instrumentation. Oplinger and Lam, instead, insert thread-spawning instructions in the program statically. In general, their approach is less efficient and hurts some conventional compiler optimizations. Many of the new issues that appear with dynamic hardware spawning are discussed in Sections 4.3 and 4.4.

The software component of iWatcher includes the iWatcherOn/Off() system calls, which set or remove associations of memory locations with monitoring functions. iWatcher uses a software table called *Check Table* to store detailed monitoring information for each watched memory location. The information stored includes MemAddr, Length, WatchFlag, ReactMode, MonitorFunc, and Parameters. Using software

simplifies the hardware and enables the system to use sophisticated data structures. An iWatcherOn/Off() call adds or removes the corresponding entry to or from the check table.

The iWatcher software also implements the Main_check_function() library call, whose starting address is stored in the Main_check_function register. When a triggering access occurs, the hardware sets the program counter to the address in this register. The Main_check_function() is responsible to call the program-specified monitoring function(s) associated with the accessed location. To do this, it needs to search the check table and find the corresponding function(s).

4.2. Watching a Range of Addresses

When a program calls iWatcherOn() for a memory region equal or larger than *LargeRegion*, iWatcher tries to allocate an RWT entry for this region. If there is already an entry for this region in the RWT, iWatcherOn() sets the entry's WatchFlags to the logical OR of its old value and the WatchFlag argument of the call. If, instead, the region to be monitored is smaller than *LargeRegion*, iWatcher loads the watched memory lines into the L2 cache (if they are not already in L2). We do not explicitly load the lines into L1 to avoid unnecessarily polluting L1. As a line is loaded from memory, iWatcher accesses the VWT to read-in the old WatchFlags, if they exist there. Then, it sets the WatchFlag bits in the L2 line to be the logical OR of the WatchFlag argument of the call and the old WatchFlags. If the line is already present in L2 and possibly L1, iWatcher simply sets the WatchFlag bits in the line to the logical OR of the WatchFlag argument and the current WatchFlag. In all cases, iWatcherOn() also adds the monitoring function to the check table.

When a program calls iWatcherOff(), iWatcher removes the corresponding monitoring function entry from the check table. Moreover, if the monitored region is large and there is a corresponding RWT entry, iWatcherOff() updates this RWT entry's WatchFlags. The new value of the WatchFlags is computed from the remaining monitoring functions associated with this memory region, according to the information in the check table. If there is no remaining monitoring function for this range, the RWT entry is invalidated. If, instead, the memory region is small, iWatcher finds all the lines of the region that are currently cached and updates their WatchFlags based on the remaining monitoring functions. iWatcher also updates (and, if appropriately removes) any corresponding VWT entries.

Caches and VWT are addressed by the physical addresses of watched memory regions. If there is no paging by the OS, the mapping between physical and virtual addresses is fixed for the whole program execution. In our prototype implementation, we assume that watched memory locations are pinned by the OS, so that the page mappings of a watched region do not change until the monitoring for this region is disabled using iWatcherOff().

Note that the purpose of using RWT for large regions is to reduce L2 pollution and VWT space consumption: lines from this region will only be cached when referenced (not during iWatcherOn()) and, since they will never set their WatchFlags in the cache, they will not use space in the VWT on cache eviction.

It is possible that iWatcherOn()/iWatcherOff() access some memory locations sometimes as part of a large region and sometimes as a small region. In this case, the iWatcherOn()/iWatcherOff() software handlers, as they add or remove entries to or from the check table, are responsible for ensuring the consistency between RWT entries and L2/VWT WatchFlags.

4.3. Detecting Triggering Accesses

iWatcher needs to identify those loads and stores that should trigger monitoring functions. A load or store is a triggering access if the accessed location is inside any large monitored regions recorded in the RWT, or the WatchFlags of the accessed line in L1/L2 are set.

In practice, the process of detecting a triggering access is complicated by the fact that modern out-of-order processors introduce access reordering and pipelining. To help in this process, iWatcher augments each reorder buffer (ROB) entry with a *Trigger* bit, and each load-store queue entry with 2 bits that store WatchFlag information.

To keep the hardware reasonably simple, the execution of a monitoring function should only occur when a triggering load or store reaches the head of the ROB. At that point, the values of the architectural registers that need to be passed to the monitoring function are readily available. In addition, the memory system is consistent, as it contains the effect of all preceding stores. Moreover, there is no danger of mispredicted branches or exceptions, which could require the cancellation of an early-triggered monitoring function.

For a load or store, when the TLB is looked up early in the pipeline, the hardware also checks the RWT for a match. This introduces negligible visible delay. If there is a match, the access is a triggering one. If there is no match, the WatchFlags in the caches will be examined to determine if it is a triggering access.

A load typically accesses the memory system before reaching the head of the ROB. It is at that time that a triggering load will detect the set WatchFlags in the cache. Consequently, in our design, as a load reads the data from the cache into the load queue, it also reads the WatchFlag bits into the special storage provided in the load queue entry. In addition, if the RWT or the WatchFlag bits indicate that the load is a triggering one, the Trigger bit associated with the load's ROB entry is set. When the load (or any instruction) finally reaches the head of the ROB and is about to retire, the hardware checks the Trigger bit. If it is set, the hardware triggers the corresponding monitoring function.

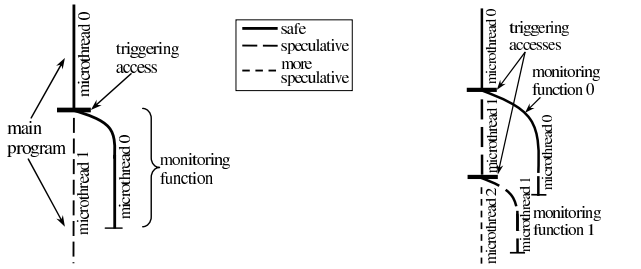
Stores present a special difficulty. A store is not sent to the memory system until it reaches the head of the ROB. At that point, it is retired immediately, but it may still cause a cache miss, in which case it may take a long time to actually complete. In iWatcher, this would mean that, for stores that do not hit in the RWT, the processor may have to wait a long time to know whether it is a triggering access, especially for stores that do not hit in the RWT. During that time, no subsequent instruction could be retired, as the processor may have to trigger a monitoring function. To reduce this delay as much as possible, we change the micro-architecture so that, as soon as a store address is resolved early in the ROB, a prefetch is issued to the memory system. Such prefetch brings the data into the cache, and the WatchFlag bits are read into the special storage in the store queue entry. If the RWT or the WatchFlag bits indicate that the store is a triggering one, the Trigger bit in the ROB entry is also set. With this support, the processor is much less likely to have to wait when the store reaches the head of the ROB. While issuing this prefetch may have implications for the memory consistency model

supported in a multiprocessor environment, we consider the topic to be beyond the scope of this paper.

Note that bringing the WatchFlag information into the load-store queue entries enables correct operation for loads that get their data directly from the load-store queue. For example, if a store in the load-store queue has the read-monitoring WatchFlag bit set, then a load that reads from it will correctly set its own Trigger bit.

4.4. Executing Monitoring Functions

When a triggering load or store is retired, its associated monitoring function has to be automatically initiated. Using TLS mechanisms, the iWatcher hardware automatically spawns a new microthread (denoted as microthread 1 in Figure 2(a)) to speculatively execute the rest of the program after the triggering access, while the current microthread (denoted as microthread 0 in Figure 2(a)) executes the monitoring function non-speculatively. To provide sequential semantics (the remainder of the program is semantically after the monitoring function), data dependencies are tracked by TLS and any violation of sequential semantics results in the squash of the speculative microthread (microthread 1).



(a) *Executing a monitoring function.* (b) *Triggering a monitoring function from a speculative microthread.*

Figure 2. Examples of monitoring function execution.

Microthread 0 executes the monitoring function by starting from the address stored in the Main_check_function register. It is the responsibility of the Main_check_function() to find the monitoring functions associated with the triggering access and call all such functions one after another. Note that, although semantically, a monitoring function appears to programmers like a user-specified exception handler, the overhead of triggering a monitoring function is tiny with our hardware support. Indeed, while triggering an exception handler typically needs OS involvement, triggering a monitoring function in iWatcher is done completely in hardware: the hardware automatically fetches the first instruction from the Main_check_function(). iWatcher can skip the OS because monitoring functions are not related to any resource management in the system and, in addition, do not need to be executed in privileged mode. Moreover, the Main_check_function() and the check table are in the same address space as the monitored program. Therefore, a "bad" program cannot use iWatcher to mess up other programs.

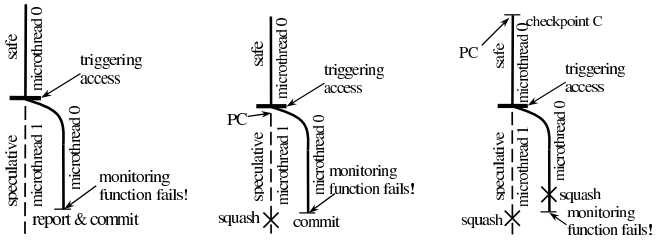
Microthread 1 speculatively executes the continuation of the monitoring function, i.e., the remainder of the program after the triggering access. To avoid the overhead of flushing the pipeline, iWatcher dynamically changes the microthread ID of all the instructions currently in the pipeline from 0 to 1. Unfortunately, it is possible that some un-retired load instructions after the triggering access may have already accessed the data in the cache and, as per TLS, already updated the microthread ID in the cache line to be 0. Since the microthread ID on these cache lines should now be 1, the hardware re-touches the cache lines that were read by these un-retired loads, correctly setting their microthread IDs to 1. There is no such problem for stores because they only update the microthread IDs in the cache at retirement.

It is possible that a speculative microthread issues a triggering access, as shown on Figure 2(b). In this case, a more speculative microthread (microthread 2) is spawned to execute the rest of the program, while the speculative microthread (microthread 1) enters the Main_check_function. Since microthread 2 is semantically after microthread 1, a violation of sequential semantics will result in the squash of microthread 2. In addition, if microthread 1 is squashed, microthread 2 is squashed as well. Finally, if microthread 1 completes while speculative, iWatcher does not commit it; it can only commit after microthread 1 becomes safe.

Note that, in a CMP-based iWatcher, microthreads should be allocated for cache affinity. In our Figure 2(a) example, speculative microthread 1 should be kept on the same CPU as the original program, while microthread 0 should be moved to a different CPU. This is because microthread 1 continues to execute the program and is likely to reuse cache state.

4.5. Different Reaction Modes

If a monitoring function fails, iWatcher takes different actions depending on the function's ReactMode. Figure 3 illustrates the three supported reaction modes. $ReportMode$ is the simplest one. iWatcher treats it the same way as if the monitoring function had succeeded: microthread 0 commits and microthread 1 becomes safe. If the reaction mode is $BreakMode$, iWatcher commits microthread 0 but squashes microthread 1. The program state and the program counter (PC) of microthread 1 are restored to the state it had immediately after the triggering access (Figure 3(b)). The cache updates of microthread 1 are discarded. At this point, programmers can use an interactive debugger to analyze the bug.



(a) *ReportMode.* (b) *BreakMode.* (c) *RollbackMode.*

Figure 3. Different reaction modes supported by iWatcher.

If the reaction mode is $RollbackMode$, iWatcher squashes microthread 1 and also rolls back microthread 0 to the most recent checkpoint (the checkpoint at PC in Figure 3(c)). iWatcher can use support similar to ReEnact [32] to provide this reaction mode.

4.6. Other Issues

Displacements and Cache Misses. When a watched line of small regions is about to be displaced from the L2 cache, its WatchFlags are saved in the VWT. The VWT is a small set-associative buffer.

Feature	Assertions	Hardware Watchpoints	DIDUCE	iWatcher
Hardware Support?	None	Simple support	TLS support	TLS and memory watch support
Type of checks	Code-controlled	Location-controlled	Code-controlled	Location-controlled
Reaction modes	Abort	Interrupt	Break or transaction abort	Report, break or rollback
Programmer's effort	High	High	Low	Moderate or low with automatic instrumentation
Language dependent?	No	No	Yes (Java)	No
Flexibility	Very flexible, program specific	Inflexible, only supports a few watchpoints, relies on programmers or debuggers for checks	Moderately flexible, currently only supports simple invariance checks	Very flexible, program specific
Cross-module and cross-developer	No	Yes	No	Yes
Completeness	Hard to make sure all possible places are checked	Detects all accesses	May miss some accesses due to aliasing problems	Detects all accesses

Table 1. Comparison of iWatcher to three other approaches. Completeness refers to whether an approach monitors all accesses to a watched memory location by construction.

If the VWT needs to take an entry while full, it selects a victim entry to be evicted, and delivers an exception. The OS then turns on page protections for the pages that correspond to the WatchFlags to be evicted from the VWT. Future accesses to these pages will trigger page protection faults, which will enable the OS to insert their WatchFlags back into the VWT. However, in our experiments, we find that a 1024-entry VWT is never full. The reason is that the VWT only keeps the WatchFlags for watched lines of small regions that have at some point been displaced from L2.

On an L2 cache miss, as the line is read from memory, the VWT is checked for a address match. If there is a match, the WatchFlags for the line are copied to the destination location in the cache. We do not remove the WatchFlags from the VWT because the memory access may be speculative and be eventually undone. If there is no match, the WatchFlags for the loaded line are set to the default "un-watched" value. Note that this VWT lookup is performed in parallel with the memory read and, therefore, introduces negligible visible delay.

Aside from these issues, caches work as in conventional TLS systems. In particular, speculative lines cannot be displaced from the L2. If space is needed in a cache set that only holds speculative lines, a speculative microthread is squashed to make room. More details can be found in [32].

Check Table Implementation. The check table is a software table. Our current implementation uses one entry for each watched region. The entries are sorted by start address. To speed-up check table lookup, we exploit memory access locality to reduce the number of accessed table entries during one search. A table entry contains all arguments of the iWatcherOn() call. If there are multiple monitoring functions associated with the same location, they are linked together. Since the check table is a pure software data structure, it is easy to change its implementation. For example, another implementation could be to organize it as a hash table. It can be hashed with the virtual address of the watched location.

5. Advantages of iWatcher

Based on the previous discussion, we can list the advantages of iWatcher. One of them is that it provides location-controlled monitoring. Therefore, *all* accesses to a watched memory location are monitored, including "disguised" accesses due to dangling pointers or wrong pointer manipulations.

Another advantage of iWatcher is its low overhead. iWatcher only monitors memory operations that truly access a watched memory location. Moreover, iWatcher uses hardware to trigger monitoring functions with minimal overhead. Finally, iWatcher uses TLS to execute monitoring functions in parallel with the rest of the program, effectively hiding most of the monitoring overhead.

iWatcher is flexible and extensible. Programmers or automatic instrumentation tools can add monitoring functions. iWatcher is convenient even for manual instrumentation because programmers do not need to instrument every possible access to a watched memory location. Instead, they only need to insert an iWatcherOn() for a location when they are interested in monitoring this location and an iWatcherOff() when the monitoring is no longer needed. In between, all possible accesses to this location are automatically monitored. In addition, iWatcher supports three reaction modes, giving flexibility to the system.

iWatcher is cross-module and cross-developer. A watched location inserted by one module or one developer is automatically honored by all modules and all developers whenever the watched location is accessed.

iWatcher is language independent since it is supported directly in hardware. Programs written in any language, including C/C++, Java or other languages can use iWatcher. For the same reason, iWatcher can also support dynamic monitoring of low-level system software, including the operating system.

iWatcher can be used to detect illegal accesses to a memory location. For example, it can be used for security checks to prevent illegal accesses to some secured memory locations. In our experiments, we have used iWatcher to protect the return address in a program stack to detect stack-smashing attacks [6, 11, 28, 47].

Table 1 summarizes the differences between iWatcher and the three related approaches discussed in Section 2. In the table, we select DIDUCE [13] as a state-of-the-art representative of dynamic checkers. DIDUCE is a tool for debugging Java programs. It can dynamically infer some simple program invariants and then dynamically perform invariant checks to detect bugs. Like other dynamic checkers, DIDUCE requires instrumentation. The instrumentation overhead slows down execution by 6-20 times. To reduce the overhead, the authors use TLS to execute the checks in parallel with the main program [29].

From the table, we see that iWatcher differs from DIDUCE in several major ways. Most importantly, while DIDUCE performs code-controlled monitoring, iWatcher performs location-controlled monitoring, therefore enjoying the benefits described in Section 2.1. In particular, if we applied the DIDUCE mechanism to C/C++ codes, due to aliasing problems, we would need to insert potentially many more checks than necessary, as we lack accurate information at instrumentation time. In addition, we may miss some accesses that need to be instrumented.

Another difference between iWatcher and DIDUCE is that the former is language independent and can work for any programs, even operating systems. DIDUCE, instead, currently works only for Java programs, and some of its techniques are not applicable to C/C++ programs. Finally, iWatcher is cross-module and cross-developer.

We note that iWatcher is complementary to DIDUCE. DIDUCE could provide iWatcher with automatic invariant inferences, while iWatcher could provide DIDUCE with an efficient location-based monitoring capability.

6. Evaluation Methodology

6.1. Simulated Architecture

To evaluate iWatcher, we have built an execution-driven simulator that models a workstation with a 4-context SMT processor augmented with TLS support and iWatcher functionality. The parameters of the architecture are shown in Table 2. As seen in the table, each microthread is allocated 32 load-store queue entries. We model the overhead of spawning a monitoring-function microthread as 5 cycles of processor stall visible to the main-program thread. The reaction mode used in all experiments is ReportMode, so that all programs can run to completion.

CPU frequency	2.4GHz	ROB size	360
Fetch width	16	I-window size	160
Issue width	8	Int FUs	6
Retire width	12	Mem FUs	4
Ld/st queue entries	32/thread	FP FUs	4
Spawn overhead	5 cycles	Reaction mode	ReportMode
L1 cache	32KB, 4-way, 32B/line, 3 cycles latency		
L2 cache	1MB, 8-way, 32B/line, 10 cycles latency		
VWT	1024 entries, 8-way, 2B/entry		
LargeRegion	64Kbytes		
RWT	4 entries, 32bits for the start and end address		
Memory	200 cycles latency		

Table 2. Parameters of the simulated architecture. Latencies are given as unloaded round-trips from the processor.

To isolate the benefits of TLS, we also evaluate the same architecture without TLS support. On a triggering access, the processor first executes the monitoring function, and then proceeds to execute the rest of the program. Finally, we simulate the same architecture with no iWatcher or TLS support. For the evaluation without TLS support, the single microthread running is given a 64-entry load-store queue.

6.2. Valgrind

In our evaluation, we compare the functionality and overhead of iWatcher to Valgrind [36], an open-source memory debugger for x86 programs. Valgrind is a binary-code dynamic checker to detect general memory-related bugs such as memory leaks, memory corruption and buffer overflow. It simulates every single instruction of a program. Because of this, it finds errors not only in a program but also in all supporting dynamically-linked libraries. Valgrind takes control of a program before it starts. The program is then run on a synthetic x86 CPU, and its every memory access is checked. All detected errors are reported.

Valgrind provides an option to enable or disable memory leak detection. We also enhanced Valgrind to enable or disable variable uninitialization checks and invalid memory access checks (checks for buffer overflow and invalid accesses to freed memory locations).

In our experiments, we run Valgrind on a real machine with a 2.6 GHz Pentium 4 processor, 32-Kbyte L1 cache, 2-Mbyte L2 cache, and 1-Gbyte main memory. Since iWatcher runs on a simulator, we cannot compare the absolute execution time of iWatcher with that of Valgrind. Instead, we compare their relative execution overheads over runs without monitoring.

6.3. Tested Applications

We have conducted two sets of experiments. The first one uses applications with bugs to evaluate the functionality and overheads of iWatcher for software debugging. The second one systematically evaluates the overheads of iWatcher to monitor applications without bugs.

The applications used in our first set of experiments contain various bugs, including memory leaks, memory corruption, buffer overflow, stack-smashing attacks, value invariant violations and outbound pointers. These applications are: bc-1.03 (an arbitrary precision calculator language), cachelib (a cache management library developed at UIUC), and gzip (a SPECINT 2000 application running the Test input data set). Of these codes, bc-1.03 and cachelib already had bugs, while we injected some common bugs into gzip.

Table 3 shows the details of the bugs and monitoring functions. For gzip, we evaluate the case of single bugs: stack-smashing, memory corruption, buffer overflow (dynamic buffer overflow and static array overflow), memory leak, or value invariant violation. We also evaluate the case of a combination of bugs (memory leak, memory corruption, and dynamic buffer overflow). Table 3 shows the names given to each buggy application.

For fair comparison between Valgrind and iWatcher, in Valgrind we enable only the type of checks that are necessary to detect the bug(s) in the corresponding application. For example, for gzip-ML, we enable only the memory leak checks. Similarly, for gzip-MC and gzip-BO1, we enable only the invalid memory access checks. In all our experiments, variable uninitialization checks are always disabled.

To detect bugs such as stack smashing, memory corruption, dynamic buffer overflow, memory leak, or static array overflow, our iWatcher monitoring functions are very general. They monitor all possible relevant locations without using program-specific semantic information. In addition, all iWatcherOn/Off() calls can be inserted by an automated tool without any semantic program information. We enforce these rules to have a fair comparison with Valgrind, which does not have any semantic program information. Therefore, we feel that the comparison is fair.

To detect other bugs, such as value invariant violations and outbound pointers, we need program-specific information. Valgrind cannot detect these bugs, whereas iWatcher can.

For gzip with memory leak, iWatcher not only detects all dynamic memory buffers that are not freed; it also ranks buffers based on their access recency. Buffers that have not been accessed for a long time are more likely to be memory leaks than the recently-accessed ones.

Finally, our second set of experiments evaluates iWatcher overheads by monitoring memory accesses in two unmodified SPECINT 2000 applications running the Test input data set, namely gzip and parser. We measure the overhead as we vary the percentage of *dynamic* loads monitored by iWatcher and the length of the monitoring function.

Application	Bug Class	Type of Monitoring	Bug Description	Monitoring Function
gzip-STACK	stack smashing	general	In function "huft_free()", the return address in the program stack is corrupted.	When entering a function, call iWatcherOn() on the location holding the return address. Turn off monitoring immediately before the function returns.
gzip-MC	memory corruption	general	In function "huft_free()", dereference a pointer after it is freed up.	Monitor all freed locations. Any access to such locations is a bug. After a free buffer is re-allocated, the monitoring for the buffer is turned off.
gzip-BO1	dynamic buffer overflow	general	In function "huft_build()", access an element past the boundary of the dynamically-allocated buffer.	Add some padding to all buffers. The padded locations are monitored by iWatcher. Any access to them is a bug.
gzip-ML	memory leak	general	In function "huft_free()", only free the first node of the linked list but not other nodes.	Monitor all accesses to heap objects. Each access to a heap object updates its time-stamp. Objects that have not been accessed for a long time are likely to be memory leaks.
gzip-COMBO	combination of bugs	general	Combination of the bugs in gzip-ML, gzip-MC, and gzip-BO1.	Combines the monitoring in gzip-ML, gzip-MC, and gzip-BO1.
gzip-BO2	static array overflow	general	In function "huft_build()", write outside of a static array.	Similar to gzip-BO1.
gzip-IV1	value invariant violation	program specific	In function "huft_build()", variable "hufts" is corrupted due to memory corruption.	Any write to this location triggers an invariant check.
gzip-IV2	value invariant violation	program specific	In function "inflate()", an unusual value is stored into the variable "hufts".	Similar to gzip-IV1.
cachelib-IV	value invariant violation	program specific	In option.c:line 90, initialize variable "conf→algos" to 0.	Similar to gzip-IV1.
bc-1.03	outbound pointer	program specific	In dc-eval.c:line 498-503, pointer "s" is outside of the array in some cases.	Use a "range_check()" function to check the value of "s" each time "s" is written.

Table 3. Bugs and monitoring functions. Type of monitoring indicates whether the monitoring function uses program-specific semantic information or, instead, monitors all possible relevant locations without using program-specific semantic information. We call the second approach "general" monitoring.

7. Experimental Results

7.1. Overall Results

Table 4 compares the effectiveness and the overhead of Valgrind and iWatcher. For each of the buggy applications considered, the table shows whether the schemes detect the bug and, if so, the overhead they add to the program's execution time. Recall from Section 6 that Valgrind's times are measured on a real machine, while iWatcher's are simulated.

Application	Valgrind		iWatcher	
	Bug Detected?	Overhead (%)	Bug Detected?	Overhead (%)
gzip-STACK	No	-	Yes	80.0
gzip-MC	Yes	1466	Yes	8.7
gzip-BO1	Yes	1514	Yes	10.4
gzip-ML	Yes	936	Yes	37.1
gzip-COMBO	Yes	1650	Yes	42.7
gzip-BO2	No	-	Yes	10.5
gzip-IV1	No	-	Yes	10.5
gzip-IV2	No	-	Yes	9.6
cachelib-IV	No	-	Yes	3.8
bc-1.03	No	-	Yes	23.2

Table 4. Comparing the effectiveness and overhead of Valgrind and iWatcher.

Consider effectiveness first. Valgrind can detect memory corruption, dynamic buffer overflow, memory leak bugs, and the combination of them. iWatcher, instead, detects all the bugs considered. iWatcher's effectiveness is largely due to its flexibility to specialize the monitoring function.

The table also shows that iWatcher has a much lower overhead than Valgrind. For bugs that can be detected by both schemes, iWatcher only adds 9-43% overhead, a factor of 25-169 smaller than Valgrind. For example, in gzip-COMBO, where both iWatcher and Valgrind monitor every access to dynamically-allocated memory, iWatcher only adds 43% overhead, which is 39 times less than Valgrind. iWatcher's low overhead is the result of triggering monitoring functions only when the watched locations are actually accessed, and of using TLS to hide monitoring overheads. The difference in overhead between Valgrind and iWatcher is larger in gzip-MC, where we are looking for a pointer that de-references a freed-up location. In this case, iWatcher only monitors freed memory buffers, and any triggering access uncovers the bug. As a result, iWatcher's overhead is 169 times smaller than Valgrind's. Finally, our results with Valgrind are consistent with the numbers (25-50 times slowdown) reported in a previous study [37].

If we consider all the applications, we see that iWatcher's overhead ranges from 4% to 80%. This overhead comes from three effects. The first one is the contention of monitoring-function microthreads and the main program for processor resources (such as functional units or fetch bandwidth) and cache space. Such contention has a high impact when there are more microthreads executing concurrently than hardware contexts in the SMT processor. In this case, the main-program microthread cannot run all the time. Instead, monitoring-function and main-program microthreads share the hardware contexts on a time-sharing basis.

Columns 2 and 3 of Table 5 show the fraction of time that there is more than one microthread running or more than four microthreads running, respectively. These figures include the main-program microthread. Note that having more than four microthreads running does not mean that the main-program microthread starves: the scheduler will attempt to share all the contexts among all microthreads fairly. From the table, we see that three applications use more than 1 microthread for more than 1% of the time. Of those, there are two that use more than 4 microthreads for a significant fraction of the time. Specifically, this fraction is 15.2% for gzip-COMBO and 16.9% for gzip-ML. Note that these applications have high iWatcher overhead in Table 4. bc-1.03 is a short program, and even a little contention has a significant impact on execution time.

A second source of overhead is the iWatcherOn/Off() calls. These calls consume processor cycles and, in addition, bring memory lines into L2, possibly polluting the cache. The overhead caused by iWatcherOn/Off() can not be hidden by TLS. In practice, their effect is small due to the small number of calls, except in gzip-

Application	% Time With > 1 Microthread	% Time With > 4 Microthreads	# Triggering Accesses per 1M Instructions	# iWatcher-On/Off() Calls	Size of iWatcherOn/Off() Call (Cycles)	Size of Monitoring Function (Cycles)	Max Monitored Memory Size at a Time (Bytes)	Total Monitored Memory Size (Bytes)
gzip-STACK	0.1	0.0	0.2	4889642	20.7	22.4	40	19558568
gzip-MC	0.1	0.0	0.4	239	1291.3	24.4	246880	246880
gzip-BO1	0.1	0.0	0.4	486	210.4	177.0	80	1944
gzip-ML	23.1	16.9	13008.9	243	582.6	47.4	6613600	6847616
gzip-COMBO	26.2	15.2	13009.6	243	1082.3	45.2	6847616	6847616
gzip-BO2	0.1	0.0	0.2	880	59.0	24.8	32	3520
gzip-IV1	0.1	0.0	0.7	132	40.5	21.7	4	528
gzip-IV2	0.1	0.0	1.1	1	83.0	23.0	4	4
cachelib-IV	0.4	0.0	91.6	10	129.0	16.5	40	40
bc-1.03	2.2	0.0	907.2	1	81.0	134.2	4	4

Table 5. Characterizing iWatcher execution.

STACK. Indeed, Columns 5 and 6 of Table 5 show the absolute number of iWatcherOn/Off() calls and the average size of an individual call. Except for gzip-STACK, the product of number of calls times the size per call is tiny compared to the hundreds of millions of cycles taken by the application to execute. For these cases, it can be shown that, even if every line brought into L2 by iWatcherOn/Off() calls causes one additional miss, the overall effect on program execution time is very small.

The exception is gzip-STACK, where the number of iWatcherOn/Off() calls is huge (4,889,642). These calls introduce a large overhead that cannot be hidden by TLS. Moreover, iWatcherOn/Off() calls partially cripple some conventional compiler optimizations such as register allocation. The result is worse code and additional overhead. Overall, while for most applications the iWatcherOn/Off() calls introduce negligible overhead, for gzip-STACK, they are responsible for most of the 80% overhead of iWatcher.

Finally, there is a third, less important source of overhead in iWatcher, namely the spawning of monitoring-function microthreads. As indicated in Section 6, each spawn takes 5 cycles. Column 4 of Table 5 shows the number of triggering accesses per million instructions. Each of these accesses spawns a microthread. From the table, we see that this parameter varies a lot across applications. However, given the small cost of each spawn, the total overhead is small.

Overall, we conclude that the overhead of iWatcher can be high (37-80%) if the application needs to execute more concurrent microthreads than contexts provided by the SMT processor, or the application calls iWatcherOn/Off() very frequently. For the other applications analyzed, the overhead is small, ranging from 4% to 23%.

Finally, the last three columns of Table 5 show other parameters of iWatcher execution: average monitoring function size, maximum monitored memory size at a time, and total monitored memory size, respectively. We can see that, while most monitoring functions take less than 25 cycles, there are a few applications where monitoring functions take 45-177 cycles. In some cases such as gzip-ML and gzip-COMBO, these relatively expensive monitoring functions occur in applications with frequent triggering accesses. When this happens, the fraction of time with more than 4 microthreads is high, which results in high iWatcher overhead (Table 4).

The last two columns show that in some applications such as gzip-ML and gzip-COMBO, iWatcher needs to monitor many addresses. In this case, the check table will typically contain many entries. Note, however, that even in this case, the size of the monitoring function, which includes the check table lookup, is still not big. This is because our check table lookup algorithm is very efficient for the applications evaluated in our experiments.

7.2. Benefits of TLS

As indicated in Section 6, our experiments are performed using ReportMode. In this reaction mode, TLS speeds-up execution by running monitoring-function microthreads in parallel with each other and with the main program. To evaluate the effect of not having TLS, we now repeat the experiments executing both monitoring-function and main-program code sequentially, instead of spawning microthreads to execute them in parallel.

Figure 4 compares the execution overheads of iWatcher and iWatcher without TLS for all the applications. The amount of monitoring overhead that can be hidden by TLS in a program is the product of Columns 4 and 7 in Table 5. For programs with substantial monitoring, TLS reduces the overheads. For example, in gzip-COMBO, the overhead of iWatcher without TLS is 61.4%, while it is only 42.7% with TLS. This is a 30% reduction. As monitoring functions perform more sophisticated tasks such as DIDUCE's invariant inference [13], the benefits of TLS will become more pronounced.

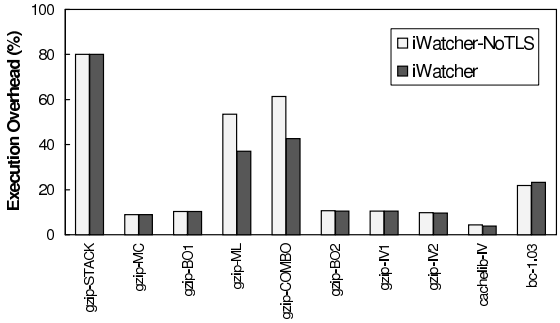


Figure 4. Comparing iWatcher and iWatcher without TLS.

For programs with little monitoring, the product of Columns 4 and 7 in Table 5 is small. For these applications, TLS does not provide benefit, because there is not much overhead that can be hidden by TLS.

Overall, we recommend supporting TLS, as it reduces the overhead of iWatcher in some applications. We also note that TLS can be instrumental in efficiently supporting RollbackMode (Section 4.5).

7.3. Sensitivity Study

To measure the sensitivity of iWatcher's overhead, we artificially vary the fraction of triggering accesses and the size of the monitoring functions. We perform the experiments on the bug-free gzip and parser applications.

In a first experiment, we trigger a monitoring function every Nth *dynamic* load in the program[1], where N varies from 2 to 10. The function walks an array, reading each value and comparing it to a constant for a total of 40 instructions. The resulting execution overhead for iWatcher and iWatcher without TLS is shown in Figure 5. The figure shows that the overhead of iWatcher with frequent triggering accesses is tolerable. Specifically, the gzip overhead is 66% for 1 trigger out of 5 dynamic loads, and 180% for 1 trigger out of 2 loads. The parser overheads are a bit higher, namely 174% for 1 trigger out of 5 loads, and 418% for 1 trigger out of 2 loads. If iWatcher does not support TLS, however, the overheads go up: 273% for gzip and 593% for parser, respectively, for 1 trigger out of 2 loads.

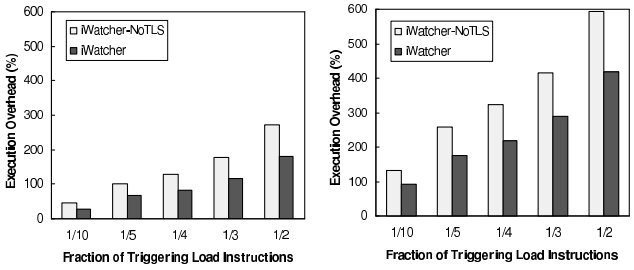


Figure 5. Varying the fraction of triggering loads.

In a second experiment, we vary the size of the monitoring function. We use the same function as before, except that we vary the number of instructions executed from 4 to 800. The function is triggered in 1 out of 10 dynamic loads. The resulting execution overhead is shown in Figure 6. The figure again shows that the iWatcher overheads are modest. For 200-instruction monitoring functions, the overhead is 65% for gzip and 159% for parser. In iWatcher without TLS, the overhead is 173% for gzip and 335% for parser. As we increase the monitoring function size, the absolute benefits of TLS increase, as TLS can hide more monitoring overhead.

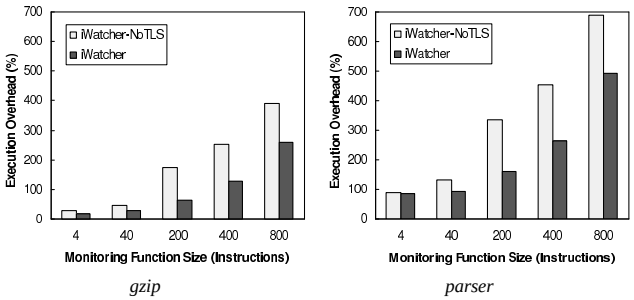


Figure 6. Varying the size of the monitoring function.

[1]For parser, we skip the program's initialization phase, which lasts about 280 million instructions, because its behavior is not representative of steady state.

8. Related Work

Our work builds upon many previous proposals to improve software robustness. Due to space limitations, we briefly describe only related works that have not been described in previous sections.

Many tools have been proposed for dynamic execution monitoring. Well-known examples include Purify [14], Intel thread checker [17], Eraser [33], StackGuard [6], DIDUCE [13], Valgrind [36], CCured [5, 27], and many others [1, 23, 30, 31]. Most of these tools rely on instrumentation to perform dynamic checks. Consequently, to check all possible accesses to a given location, they typically need to instrument more than necessary. Moreover, most dynamic checkers impose significant run-time overhead. iWatcher innovates with efficient and flexible location-controlled monitoring capability.

Schnarr and Larus have proposed using unused processor cycles to reduce overhead for *code-controlled* monitoring [34]. Our work differs from theirs in that iWatcher provides convenient, flexible architectural support to perform *location-controlled* monitoring, and uses TLS to hide monitoring overheads.

Austin *et al.* [1] have proposed detecting errors in pointer or array accesses by validating dereferences against the pointer's attributes, such as bounds. This method can prevent a pointer pointing to an array from incorrectly moving out of bounds, but cannot detect errors for general pointers such as links between structures. Moreover, it performs checks sequentially and therefore can add large execution overheads (130-540%). iWatcher can be used to reduce these overheads. Moreover, iWatcher extends dynamic monitoring from out-of-bound checks for array accesses to location-controlled checks for any general memory accesses.

To improve software debugging, several hardware supports have been proposed beyond hardware-assisted watchpoints [15, 16, 18, 39], including the proposal of Oplinger and Lam [29], ReEnact [32], and the Flight Data Recorder [48]. Oplinger and Lam use TLS to execute invariant checks to detect bugs in parallel with the main program. ReEnact uses the state buffering, rollback, and re-execute features of TLS to debug data races. The Flight Data Recorder logs coherence operations into a file that can potentially be used to support replay for debugging. Our work is different, since iWatcher's contribution is to provide efficient and flexible architectural support to monitor memory locations.

Our work is related to previous work on fine-grain access control [35, 46]. For example, Mondrian Memory Protection (MMP) [46] provides access control at word granularity using a "protection look-aside buffer" (PLB) to record protection information. MMP can potentially be used to implement location-controlled monitoring. However, like hardware-assisted watchpoints, it needs to raise an exception and, therefore can add significant overhead.

Our work is also related to some of the classic work on capability-based architectures [10, 21], protection-enhanced architectures [19], hardware support for security [11, 22, 43, 47], TLS [4, 38, 41, 44], and hardware support for instruction-level profiling [7].

9. Conclusions and Future Work

This paper has presented *iWatcher*, novel architectural support for minimal-overhead location-controlled monitoring. iWatcher de-

tects all accesses to a watched memory location, including those by astray pointer dereferences. To reduce overhead and support rollback, iWatcher leverages Thread-Level Speculation (TLS). We have evaluated iWatcher on applications with various bugs. iWatcher detects all bugs evaluated in our experiments with only a 4-80% execution overhead. In contrast, a well-known open-source bug detector called Valgrind induces orders of magnitude more overhead, and can only detect a subset of the bugs. Moreover, even with 20% of the dynamic loads monitored in a program, iWatcher only adds 66-174% overhead.

Even though our experiments have only demonstrated the use of iWatcher in detecting several types of bugs, iWatcher can also be used to detect many other types of bugs such as uninitialized reads and data races. In addition, iWatcher also provides a framework for general-purpose debugging. Using performance debugging as one example, iWatcher can be used for value and address profiling, which can then guide data placement or instruction reordering to reduce cache misses. Another example is the use of iWatcher in BreakMode, which allows it to interface to interactive debuggers and efficiently support watchpoints and conditional breakpoints.

We are in the process of extending this work in several ways. First, we plan to compare iWatcher to other dynamic checkers beyond Valgrind. Moreover, we will evaluate iWatcher for multithreaded programs, which often exhibit hard-to-debug bugs such as data races and deadlocks. In addition, we plan to test more applications, especially large server programs with real bugs. In order to do this, we are in the process of upgrading our simulation infrastructure. We are also seeking help from the OS to handle the VWT overflow issue. Finally, note that, in this study, we have inserted all iWatcherOn/Off() calls manually. It is more convenient to use a compiler or an instrumentation tool to insert them. It is also interesting to combine iWatcher with an invariant-inference tool such as DIDUCE [13], which can specify watched locations and their associated monitoring functions. We are addressing these issues.

10. Acknowledgments

We thank the anonymous reviewers for useful feedback, the UIUC PROBE, I-ACOMA, and Opera groups for useful discussions, and Paul Sack and Radu Teodorescu for proofreading the paper. We also thank Dr. Konrad Lai from Intel for helpful discussions at the early stages of this project.

References

[1] T. M. Austin, S. E. Breach, and G. S. Sohi. Efficient detection of all pointer and array access errors. In *PLDI*, June 1994.
[2] B. Buck and J. K. Hollingsworth. An API for runtime code patching. *The International Journal of High Performance Computing Applications*, Winter 2000.
[3] J.-D. Choi et al. Efficient and precise datarace detection for multithreaded object-oriented programs. In *PLDI*, June 2002.
[4] M. Cintra, J. Martinez, and J. Torrellas. Architectural support for scalable speculative parallelization in shared-memory systems. In *ISCA*, June 2000.
[5] J. Condit, M. Harren, S. McPeak, G. C. Necula, and W. Weimer. CCured in the real world. In *PLDI*, June 2003.
[6] C. Cowan et al. StackGuard: Automatic adaptive detection and prevention of buffer-overflow attacks. In *USENIX Security Conference*, January 1998.
[7] J. Dean, J. E. Hicks, C. A. Waldspurger, W. E. Weihl, and G. Z. Chrysos. ProfileMe: Hardware support for instruction-level profiling on out-of-order processors. In *MICRO*, December 1997.
[8] D. Engler and K. Ashcraft. RacerX: Effective, static detection of race conditions and deadlocks. In *SOSP*, October 2003.
[9] M. D. Ernst, A. Czeisler, W. G. Griswold, and D. Notkin. Quickly detecting relevant program invariants. In *International Conference on Software Engineering*, June 2000.
[10] R. S. Fabry. Capability-based addressing. *Communications of the ACM*, July 1974.
[11] M. Frantzen and M. Shuey. StackGhost: Hardware facilitated stack protection. In *USENIX Security Symposium*, August 2001.
[12] S. Hallem, B. Chelf, Y. Xie, and D. Engler. A system and language for building system-specific, static analyses. In *PLDI*, June 2002.
[13] S. Hangal and M. S. Lam. Tracking down software bugs using automatic anomaly detection. In *International Conference on Software Engineering*, May 2002.
[14] R. Hastings and B. Joyce. Purify: Fast detection of memory leaks and access errors. In *Usenix Winter Technical Conference*, January 1992.
[15] Intel Corporation. The IA-32 Intel architecture software developer's manual, volume 2: Instruction set reference, 2001.
[16] M. S. Johnson. Some requirements for architectural support of software debugging. In *ASPLOS*, March 1982.
[17] KAI-Intel Corporation. Intel thread checker. URL: http://developer.intel.com/software/products/threading/tcwin.
[18] G. Kane and J. Heinrich. *MIPS RISC architecture*. Prentice-Hall, 1992.
[19] E. J. Koldinger, J. S. Chase, and S. J. Eggers. Architectural support for single address space operating systems. In *ASPLOS*, October 1992.
[20] J. Larus and E. Schnarr. EEL: Machine-independent executable editing. In *PLDI*, June 1995.
[21] H. M. Levy. *Capability-based computer systems*. Digital Press, 1984.
[22] D. Lie et al. Architectural support for copy and tamper resistant software. In *ASPLOS*, November 2000.
[23] A. Loginov, S. H. Yong, S. Horwitz, and T. W. Reps. Debugging via run-time type checking. In *International Conference on Fundamental Approaches to Software Engineering*, April 2001.
[24] E. Marcus and H. Stern. Blueprints for high availability. John Willey and Sons, 2000.
[25] M. Musuvathi, D. Park, A. Chou, D. R. Engler, and D. L. Dill. CMC: A pragmatic approach to model checking real code. In *OSDI*, December 2002.
[26] National Institute of Standards and Technlogy (NIST), Department of Commerce. Software errors cost U.S. economy $59.5 billion annually. NIST News Release 2002-10, June 2002.
[27] G. C. Necula, S. McPeak, and W. Weimer. CCured: Type-safe retrofitting of legacy code. In *POPL*, January 2002.
[28] A. One. Smashing the stack for fun and profit. Phrack Magazine, November 1996.
[29] J. Oplinger and M. S. Lam. Enhancing software reliability with speculative threads. In *ASPLOS*, October 2002.
[30] H. Patil and C. Fischer. Low-cost, concurrent checking of pointer and array accesses in C programs. *Software Practice & Experience*, January 1997.
[31] H. Patil and C. N. Fischer. Efficient run-time monitoring using shadow processing. In *Automated and Algorithmic Debugging*, May 1995.
[32] M. Prvulovic and J. Torrellas. ReEnact: Using thread-level speculation mechanisms to debug data races in multithreaded codes. In *ISCA*, June 2003.
[33] S. Savage, M. Burrows, G. Nelson, P. Sobalvarro, and T. Anderson. Eraser: A dynamic data race detector for multithreaded programs. *ACM Transactions on Computer Systems*, November 1997.
[34] E. Schnarr and J. R. Larus. Instruction scheduling and executable editing. In *MICRO*, December 1996.
[35] I. Schoinas et al. Fine-grain access control for distributed shared memory. In *ASPLOS*, October 1994.
[36] J. Seward. Valgrind. URL: http://valgrind.kde.org/.
[37] J. Seward. Valgrind, an open-source memory debugger for x86-GNU/Linux. URL: http://www.ukuug.org/events/linux2002/papers/html/valgrind/.
[38] G. Sohi, S. Breach, and T. Vijayakumar. Multiscalar processors. In *ISCA*, June 1995.
[39] SPARC International. *The SPARC architecture manual: Version 8*. Prentice-Hall, 1992.
[40] A. Srivastava and A. Eustace. ATOM: A system for building customized program analysis tools. In *PLDI*, June 1994.
[41] J. G. Steffan, C. B. Colohan, A. Zhai, and T. C. Mowry. A scalable approach to thread-level speculation. In *ISCA*, June 2000.
[42] U. Stern and D. L. Dill. Automatic verification of the SCI cache coherence protocol. In *Conference on Correct Hardware Design and Verification Methods*, October 1995.
[43] G. Suh, D. Clarke, B. Gassend, M. van Dijk, and S. Devadas. AEGIS: Architecture for tamper-evident and tamper-resistant processing. In *ICS*, June 2003.
[44] J.-Y. Tsai, J. Huang, C. Amlo, D. J. Lilja, and P.-C. Yew. The superthreaded processor architecture. *IEEE Transactions on Computers*, September 1999.
[45] R. Wahbe, S. Lucco, and S. L. Graham. Practical data breakpoints: Design and implementation. In *PLDI*, June 1993.
[46] E. Witchel, J. Cates, and K. Asanović. Mondrian memory protection. In *ASPLOS*, October 2002.
[47] J. Xu, Z. Kalbarczyk, S. Patel, and R. K. Iyer. Architecture support for defending against buffer overflow attacks. EASY-2 Workshop, October 2002.
[48] M. Xu, R. Bodik, and M. D. Hill. A "flight data recorder" for enabling full-system multiprocessor deterministic replay. In *ISCA*, June 2003.