A Runtime Adaptation Framework for Native C and Bytecode Applications

Rean Griffith
Columbia University
rg2023@cs.columbia.edu

Gail Kaiser
Columbia University
kaiser@cs.columbia.edu

Abstract

The need for self-healing software to respond with a reactive, proactive or preventative action as a result of changes in its environment has added the non-functional requirement of adaptation to the list of facilities expected in self-managing systems. The adaptations we are concerned with assist with problem detection, diagnosis and remediation. Many existing computing systems do not include such adaptation mechanisms, as a result these systems either need to be re-designed to include them or there needs to be a mechanism for retro-fitting these mechanisms. The purpose of the adaptation mechanisms is to ease the job of the system administrator with respect to managing software systems. This paper introduces Kheiron, a framework for facilitating adaptations in running programs in a variety of execution environments without requiring the re-design of the application. Kheiron manipulates compiled C programs running in an unmanaged execution environment as well as programs running in Microsoft's Common Language Runtime and Sun Microsystems' Java Virtual Machine. We present case-studies and experiments that demonstrate the feasibility of using Kheiron to support self-healing systems. We also describe the concepts and techniques used to retro-fit adaptations onto existing systems in the various execution environments.

1 Introduction

System adaptation has been highlighted as a necessary feature of autonomic software systems [17]. In the realm of self-healing software we are concerned primarily with adaptations that effect problem diagnosis – via consistency checks or *ghost transactions* [1] – and remediation – in the form of reconfiguration or repair. In many situations adaptations must occur while the system executes so as to maintain some degree of availability. Having a critical software system operate in a degraded mode is preferable to taking the system offline to perform scheduled (or unscheduled) reconfiguration or repair activities [33, 18].

System designers have two alternatives when it comes to realizing software systems capable of adaptation. Adaptation mechanisms can be built into the system – as done in the K42 operating system [2] – or such functionality can be retro-fitted onto them using externalized architectures like KX [11] or Rainbow [5]. While arguments can be made for either approach, the retrofit approach provides more flexibility. "Baked-in" adaptation mechanisms restrict the analysis and reuse of said mechanisms. Further, it is difficult to evolve (via updates and extensions) the adaptation mechanisms without affecting the execution and deployment of the target system [32].

With any system there is a spectrum of adaptations that can be performed. Frameworks like KX perform coarse-grained adaptations e.g. re-writing configuration files and restarting/terminating operating system processes. In this paper, we focus on fine-grained adaptations, those interacting with individual components, sub-systems or methods e.g. restarting/refreshing individual components or subsystems, or augmenting methods.

Whereas the retro-fit approach is attractive because it does not require a re-design of the system and it is possible to separately evolve the target system and the adaptation mechanisms, it is not always easy to achieve. A major challenge is that of actually **retro-fitting fine-grained adaptation mechanisms onto existing/legacy systems**. [2]

Managing the performance impact of the mechanisms used to effect fine-grained adaptations in the running system presents an additional challenge. Since we are interacting with individual methods or components we must be cognizant of the performance impact of effecting the adaptations e.g. inserting instrumentation into individual methods may slow down the system; being able to selectively add/remove instrumentation allows the performance impact to be tuned throughout the system's execution.

This paper is primarily concerned with addressing the

[1] A ghost transaction is a special form of a self-test/diagnosis targeting a specific subset of subsystems or components.

[2] For purposes of discussion we define a legacy system as any system for which the source code may not be available or for which it is undesirable to engage in substantial re-design and development.

challenges of retro-fitting fine-grained adaptation mechanisms onto existing software systems and managing the performance impacts associated with retro-fitting these adaptation mechanisms. In this paper we posit that we can leverage the the unmodified execution environment to transparently facilitate the adaptations of existing/legacy systems. We describe three systems we have developed for this purpose. **Kheiron/C** manipulates running compiled C programs on the Linux platform, **Kheiron/CLR** manipulates running .NET applications and finally **Kheiron/JVM** manipulates running Java applications.

Our contribution is the ability to transparently retro-fit new functionality (for the purpose of diagnosing problems and resolving problems where possible) onto existing software systems. The techniques used to facilitate the retro-fit exhibit negligible performance overheads on the running systems. Finally, our techniques address effecting adaptations in a variety of contemporary execution environments. New functionality, packaged in separate modules, collectively referred to as an *adaptation engine*, is loaded by Kheiron. At runtime, Kheiron can transfer control over to the adaptation engine, which effects the desired adaptations in the running application.

The remainder of the paper is organized as follows; §2 motivates retro-fitting fine-grained adaptation mechanisms onto existing systems and presents a number of specific adaptations and their potential benefits. §3.1 gives a working definition of an execution environment and describes two classes of execution environments – *managed* and *unmanaged*. §3.2 outlines some challenges associated with performing adaptations at the execution environment level. §4 describes the mechanisms and concepts used to adapt running bytecode-based applications, using our Kheiron/JVM implementation and its performance overhead. Kheiron/CLR, our first adaptation framework, targets Microsoft Intermediate Language (MSIL) bytecode applications and is discussed in [14, 13, 15]. §5 compares and contrasts runtime adaptation in an unmanaged execution environment with runtime adaptation in a managed execution environment. We also present Kheiron/C and discuss some experimental results of the performance impact imposed on target systems and §5.4 describes a special case of adaptation – dynamically adding fault detection and recovery to running compiled C programs via selectively emulating individual functions.§6 covers related work and finally §7 presents our conclusions.

2 Motivation

There are a number of specific fine-grained adaptations that can be retro-fitted onto existing systems to aid problem detection, diagnosis and in some cases remediation via performing reconfigurations or (temporary) repairs. In this paper we describe how our Kheiron implementations can be used to facilitate a number of fine-grained adaptations in running systems via leveraging facilities and properties of the execution environments hosting these systems.

These adaptations include (but are not limited to): **Inserting or removing system instrumentation** [28] to discover performance bottlenecks in the application or detect (and where possible repair) data-structure corruption. The ability to remove instrumentation can decrease the performance impact on the system associated with collection information. **Periodic refreshing** of data-structures, components and subsystems. One example of this is a *micro-reboot* [3], which could be performed at a fine granularity e.g., restarting individual components or sub-systems, or at a coarse granularity e.g., restarting entire processes periodically. **Replacing** failed, unavailable or suspect components and subsystems (where possible) [15]. **Input filtering/audit** to detect misused APIs. **Initiating ghost transactions** against select components or subsystems and collecting the results to obtain more details about a problem. **Selective emulation of functions** – effectively running portions of computation in an emulator, rather than on the raw hardware to detect errors and prevent them from crashing the application.

3 Background

3.1 Execution Environments

At a bare minimum, an execution environment is responsible for the preparation of distinguished entities – *executables* – such that they can be run. Preparation, in this context involves the loading and laying out in memory of an executable. The level of sophistication, in terms of services provided by the execution environment beyond loading, depends largely on the *type* of executable.

We distinguish between two types of executables, *managed* and *unmanaged* executables, each of which require or make use of different services provided by the execution environment. A managed executable, e.g. a Java bytecode program, runs in a *managed execution environment* such as Sun Microsystems' JVM whereas an unmanaged executable, e.g. a compiled C program, runs in an *unmanaged execution environment* which consists of the operating system and the underlying processor. Both types of executables consist of metadata and code. However the main differences are the amount and specificity of the metadata present and the representation of the instructions to be executed.

Managed executables/applications are represented in an abstract intermediate form expected by the managed execution environment. This abstract intermediate form consists of two main elements, *metadata* and *managed code*. Metadata describes the structural aspects of the application

including classes, their members and attributes, and their relationships with other classes [21]. Managed code represents the functionality of the application's methods encoded in an abstract binary format known as *bytecode*.

The metadata in unmanaged executables is not as rich as the metadata found in managed executables. Compiled C/C++ programs may contain symbol information, however there is neither a guarantee nor requirement that it be present. Finally, unmanaged executables contain instructions that can be directly executed on the underlying processor unlike the bytecode found in managed executables, which must be interpreted or Just-In-Time (JIT) compiled into native processor instructions.

Managed execution environments differ substantially from unmanaged execution environments [3]. The major differentiation points are the metadata available in each execution context and the facilities exposed by the execution environment for tracking program execution, receiving notifications about important execution events including; thread creation, type definition loading and garbage collection. In managed execution environments built-in facilities also exist for augmenting program entities such as type definitions, method bodies and inter-module references whereas in unmanaged execution environments such facilities are not as well-defined.

3.2 Challenges of Runtime Adaptation via the Execution Environment

There are a number of properties of execution environments that make them attractive for effecting adaptations on running systems. They represent the lowest level (short of the hardware) at which changes could be made to a running program. Some may expose (reasonably standardized) facilities (e.g. profiling APIs [24, 26]) that allow the state of the program to be queried and manipulated. Further, other facilities (e.g. metadata APIs [23]) may support the discovery, inspection and manipulation of program elements e.g. type definitions and structures. Finally, there may be mechanisms which can be employed to alter to the execution of the running system.

However, the low-level nature of execution environments also makes effecting adaptations a risky (and potentially arduous) exercise. Injecting and executing adaptations must not corrupt the execution environment nor the system being adapted. The execution environment's rules for what constitutes a "valid" program must be respected while guaranteeing consistency-preserving adaptations in the target software system. Causing a crash in the execution environment typically has the undesirable side-effect of crashing the target application and any other applications being hosted.

[3]The JVM and CLR also differ considerably even though they are both managed execution environments.

At the level of the execution environment the programming-model used to specify adaptations may be quite different from the one used to implement the original system. For example, to effect changes via an execution environment, those changes may have to be specified using assembly instructions (moves and jump statements), or bytecode instructions where applicable, rather than higher level language constructs. This disconnect may limit the kinds of adaptations which can be performed and/or impact the mechanism used to inject adaptations.

4 Adapting Managed Applications

Kheiron/JVM leverages facilities exposed by Sun Microsystems' v1.5 implementation of the JVM, the Java HotspotVM, to dynamically attach/detach an engine capable of performing adaptations. Examples of adaptations include: adding instrumentation and performing consistency checks to improve problem detection and diagnosis, performing reconfigurations such as component replacements or component swaps, and performing repairs (where possible) to a target Java application while it executes.

4.1 Java HotspotVM Execution Model

The unit of execution (sometimes referred to as a module) in the JVM is the *classfile*. Classfiles contain both the metadata and bytecode of a Java application. Two major components of the Java HotspotVM interact with the metadata and bytecode contained in the classfile during execution, the *classloader* and the *global native-code optimizer*.

The classloader reads the classfile metadata and creates an in-memory representation and layout of the various classes, members and methods on demand as each class is referenced. The global native-code optimizer uses the results of the classloader and compiles the bytecode for a method into native assembly for the target platform.

The Java HotspotVM first runs the program using an interpreter, while analyzing the code to detect the critical hot spots in the program. Based on the statistics it gathers, it then focuses the attention of the global native-code optimizer on the hotspots to perform optimizations including JIT-compilation and method inlining [25]. Compiled methods remain cached in memory, and subsequent method calls jump directly into the native (compiled) version of the method.

The v1.5 implementation of the Java HotspotVM introduces a new API for inspecting and controlling the execution of Java applications – the Java Virtual Machine Tool Interface (JVMTI) [26]. JVMTI replaces both the Java Virtual Machine Profiler Interface (JVMPI) and the Java Virtual Machine Debug Interface (JVMDI) available in older releases. The JVMTI is a two-way interface: clients of the

JVMTI, often called *agents*, can receive notifications of execution events in addition to being able to query and control the application via functions either in response to events or independent of events. JVMTI notification events include (but are not limited to): classfile loading, class loading, method entry/exit.

The Java HotspotVM does not have a built in API for manipulating type definitions. As a result, to perform operations such as reading class and method attributes, parsing method descriptors, defining new methods for types, emitting/rewriting the bytecode for method implementations, creating new type references and defining new strings we were required to roll our own APIs based on information provided in the Java Virtual Machine Specification [22].

4.2 Kheiron/JVM Operation

Kheiron/JVM is implemented as a single dynamic linked library (DLL), which includes a JVMTI agent. It consists of 2658 lines of C++ code and is divided into four main components. The **Execution Monitor** receives classfile load, class load and class prepare events from the JVM. The **Metadata Helper** wraps our metadata import and interface, which is used to parse and read the classfile format. **Internal book-keeping structures** store the results of metadata resolutions. The **Bytecode and Metadata Transformer** wraps our metadata emit interface to write new metadata, e.g., adding new methods to a type, adding references to other classes and methods. It also generates, inserts and replaces bytecode in existing methods as directed by the Execution Monitor. Bytecode changes are committed using the RedefineClasses function exposed by the JVMTI. Active method invocations continue to use the old implementation of their method body while new invocations use the latest version.

Kheiron/JVM performs operations on type definitions, object instances and methods at various stages in the execution cycle to make them capable of interacting with an adaptation engine. In particular, to enable an adaptation engine to interact with a class instance, Kheiron/JVM augments the type definition to add the necessary "hooks". Augmenting the type definition is a two-step operation.

Step 1 occurs at classfile load time, signaled by the **ClassFileLoadHook** JVMTI callback that precedes it. At this point the VM has obtained the classfile data but has not yet constructed the in-memory representation of the class. Kheiron/JVM adds what we call *shadow methods* for each of the original public and/or private methods. A shadow method shares most of the properties – including a subset of the attributes and the method descriptor – of the corresponding original method. However, a shadow method gets a unique name. Figure 1, transition A to B, shows an example of adding a shadow method **_SampleMethod** for the original method **SampleMethod**.

Extending the metadata of a type by adding new methods must be done before the type definition is installed in the JVM. Once a type definition is installed, the JVM will reject the addition or removal of methods. Attempts to call RedefineClasses will fail if new methods or fields are added. Similarly, changing method signatures, method modifiers or inheritance relationships is also not allowed.

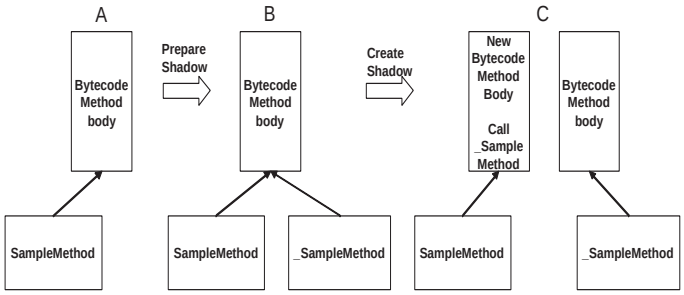

Figure 1. *Preparing and Creating a Shadow Method*

Step 2 of type augmentation occurs immediately after the shadow method has been added, while still in the ClassFileLoadHook JVMTI callback. Kheiron/JVM uses bytecode-rewriting techniques to convert the implementation of the original method into a thin *wrapper* that calls the shadow method, as shown in Figure 1, transition B to C.

```
SampleMethod( args )
   <room for prolog>
   push args
   call _SampleMethod( args )
   <room for epilog>
   return value/void
```

Figure 2. *Conceptual Diagram of a Wrapper*

Kheiron/JVM's wrappers and shadow methods facilitate the adaptation of class instances. In particular, the regular structure and single return statement of the wrapper method, see Figure 2, enables Kheiron/JVM to easily inject adaptation instructions into the wrapper as prologues and/or epilogues to shadow method calls.

To add a prologue to a method new bytecode instructions must prefix the existing bytecode instructions. The level of difficulty is the same whether we perform the insertion in the wrapper or the original method. Adding epilogues, however, presents more challenges. Intuitively, we want to insert instructions before control leaves a method. In the simple case, a method has a single return statement and the epilogue can be inserted right before that point. However, for methods with multiple return statements or exception handling routines, finding every possible return point can be an arduous task [27]. Using wrappers thus delivers a

cleaner approach since we can ignore all of the complexity in the original method.

To initiate an adaptation, Kheiron/JVM augments the wrapper to insert a jump into an adaptation engine at the *control point(s)* before and/or after a shadow method call. Effecting the jump into the adaptation engine is a three-step process. **Step 1**: Extend the metadata of the classfile currently executing in the JVM such that a reference to the classfile containing the adaptation engine is added using our **IMetaDataEmit::DefineTypeRef** and **IMetaDataEmit::DefineNameAndTypeRef** methods. **Step 2**: Add references to the subset of the adaptation engine's methods that we wish to insert calls to, using **IMetaDataEmit::DefineMethodRef**. **Step 3**: Augment the bytecode and metadata of the wrapper function to insert bytecode instructions to transfer control to the adaptation engine before and/or after the existing bytecode that calls the shadow method. The adaptation engine can then perform any number of operations, such as inserting and removing instrumentation, caching class instances, performing consistency checks over class instances and components, or reconfigurations and diagnostics of components. To persist the bytecode changes made to the method bodies of the wrappers, the Execution Monitor uses the RedefineClasses method of the JVMTI.

4.3 Preliminary Results

We are able to show, that like our other framework for facilitating adaptations in a managed execution environment, Kheiron/CLR, Kheiron/JVM imposes only a modest performance impact on a target system when no adaptations, repairs or reconfigurations are active. We have evaluated the performance of our prototype by quantifying the overheads on program execution using two separate benchmarks.

The experiments were run on a single Pentium III Mobile Processor, 1.2 GHz with 1 GB RAM. The platform was Windows XP SP2 running the Java HotspotVM v1.5 update 4. In our evaluation we used the Java benchmarks SciMark v2.0 [4] and Linpack [5].

SciMark is a benchmark for scientific and numerical computing. It includes five computation kernels: Fast Fourier Transform (FFT), Jacobi Successive Over-relaxation (SOR), Monte Carlo integration (Monte Carlo), Sparse matrix multiply (Sparse MatMult) and dense LU matrix factorization (LU). **Linpack** is a benchmark that uses routines for solving common problems in numerical linear algebra including linear systems of equations, eigenvalues and eigenvectors, linear least squares and singular value decomposition. In our tests we used a problem size of 1000.

[4] http://math.nist.gov/scimark2/

[5] http://www.shudo.net/jit/perf/Linpack.java

Running an application under the JVMTI profiler imposes some overhead on the application. Also, the use of shadow methods and wrappers converts one method call into two. Figure 3 shows the runtime overhead for running the benchmarks with and without profiling enabled. We performed five test runs for SciMark and Linpack each with and without profiling enabled. Our Kheiron/JVM DLL profiler implementation was compiled as an optimized release build. For each benchmark, the bar on the left shows the performance normalized to one, of the benchmark running without profiling enabled. The bar on the right shows the normalized performance with our profiler enabled.

Our measurements show that our profiler contributes $\sim$2% runtime overhead when no adaptations are active, which we consider negligible. Note that we do not ask the Java HotspotVM to notify us on method entry/exit events since this can result in a slow down in some cases in excess of 5X. If adaptations were actually being performed then we expect the overheads measured to depend on the specific adaptations being performed.

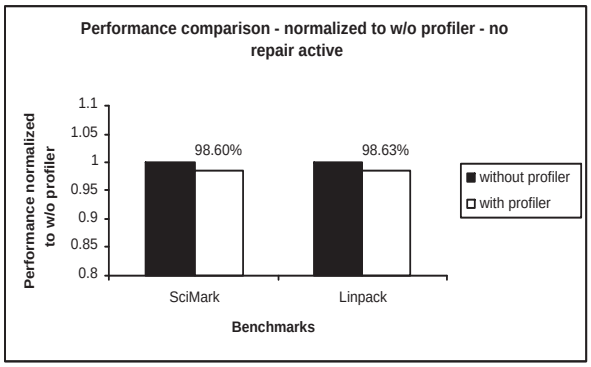

Figure 3. *Overheads when no repair active*

By implementing Kheiron/JVM we are able to show that our conceptual approach of leveraging facilities exposed by the execution environment, specifically profiling and execution control services, and combining these facilities with metadata APIs that respect the verification rules for types, their metadata and their method implementations (bytecode) is a sufficiently low-overhead approach for adapting running programs in contemporary managed execution environments.

5 Adapting Unmanaged Applications

Effecting adaptations in unmanaged applications is markedly different from effecting adaptations in their managed counterparts, since they lack many of the characteristics and facilities that make runtime adaptation qualitatively easier, in comparison, in managed execution environments. Unmanaged execution environments store/have ac-

cess to limited metadata, no built-in facilities for execution tracing, and less structured rules on well-formed programs.

In this section we focus on using Kheiron/C to facilitate adaptations in running compiled C programs, built using standard compiler toolkits like *gcc* and *g++*, packaged as Executable and Linking Format (ELF) [37] object files, on the Linux platform.

5.1 Native Execution Model

One unit of execution in the Linux operating system is the ELF executable. ELF is the specification of an *object file format*. Object files are binary representations of programs intended to execute directly on a processor as opposed to being run in an implementation of an abstract machine such as the JVM or CLR. The ELF format provides parallel views of a file's contents that reflects the differing needs of program linking and program execution.

Program loading is the procedure by which the operating system creates or augments a process image. A process image has segments that hold its text (instructions for the processor), data and stack. On the Linux platform the loader/linker maps ELF sections into memory as segments, resolves symbolic references, runs some initialization code (found in the *.init* section) and then transfers control to the *main* routine in the *.text* segment.

One approach to execution monitoring in an unmanaged execution environment is to build binaries in such a way that they emit profiler data. Special flags ,e.g. -pg, are passed to the gcc compiler used to generate the binary. The executable, when run, will also write out a file containing the times spent in each function executed. Since a compile-time/link-time flag is used to create an executable that has logic built in to write out profiling information, it is not possible to augment the data collected without rebuilding the application. Further, selectively profiling portions of the binary is not supported.

To gain control of a running unmanaged application on the Linux operating system, tools use built-in facilities such as *ptrace* and the */proc* file system. ptrace is a system call that allows one process to attach to a running program to monitor or control its execution and examine and modify its address space. Several monitored events can be associated with a traced program including; the end of execution of a single assembly language instruction, entering/exiting a system call, and receiving a signal. ptrace is primarily used to implement breakpoint debuggers. Traced processes behave normally until a signal is caught – at which point the traced process is suspended and the tracing process notified [6]. The /proc filesystem is a virtual filesystem created by the kernel in memory that contains information about the system and the current processes in their various stages of execution.

With respect to metadata, ELF binaries support various processors with 8-bit bytes and 32-bit architectures. Complex structures, etc. are represented as compositions of 32-bit, 16-bit and 8-bit "types". The binary format also uses special sections to hold descriptive information about the program. Two important sections are the *.debug* and *.symtab* sections, where information used for symbolic debugging and the symbol table, respectively, are kept.

The symbol table contains the information needed to locate and relocate symbolic references and definitions. The fields of interest in a symbol table entry (Figure 4) are *st_name*, which holds an index into the object file's symbol string table where the symbol name is stored, *st_size*, which contains the data object's size in bytes and *st_info*, which specifies the symbol's type and binding attributes.

```
typedef struct {
          Elf32_Word    st_name;
          Elf32_Addr    st_value;
          Elf32_Word    st_size;
          unsigned char st_info;
          unsigned char st_other;
          Elf32_Half    st_shndx;
} Elf32_Sym;
```

Figure 4. *ELF Symbol Table Entry [37]*

Type information for symbols can be one of: STT_NOTYPE, when the symbol's type is not defined, STT_OBJECT, when the symbol's type is associated with a data object such as variable or array, STT_FUNC, for a function or other executable code, and STT_SECTION, for symbols associated with a section. As we can see, the metadata available in ELF object files is not as detailed or as expressive as the metadata found in managed executables.

5.2 Kheiron/C Operation

Our current implementation of Kheiron/C relies on the Dyninst API [1] (v4.2.1) to interact with target applications while they execute. Dyninst presents an API for inserting new code into a running program. The program being modified is able to continue execution and does not need to be recompiled or relinked. Uses for Dyninst include, but are not limited to, runtime code-patching and performance steering in large/long running applications.

Dyninst employs a number of abstractions to shield clients from the details of the runtime assembly language insertion that takes place behind the scenes. The main abstractions are *points* and *snippets*. A point is a location in a program where instrumentation can be inserted, whereas a snippet is a representation of the executable code to be inserted. Examples of snippets include **BPatch_funcCallExpr**, which represents a function call, and **BPatch_variableExpr**, which represents a variable or area of memory in a thread's address space.

To use the Dyninst terminology, Kheiron/C is implemented as a *mutator* (Figure 5), which uses the Dyninst API to attach to, and modify a running program. On the Linux platform, where we conducted our experiments, Dyninst relies on ptrace and the /proc filesystem facilities of the operating system to interact with running programs.

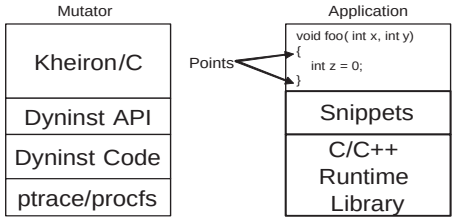

Figure 5. *Kheiron/C*

Kheiron/C uses the Dyninst API to search for global or local variables/data structures (in the scope of the insertion point) in the target program's address space, read and write values to existing variables, create new variables, load new shared libraries into the address space of the target program, and inject function calls to routines in loaded shared libraries as prologues/epilogues (at the points shown in Figure 5) for existing function calls in the target application. As an example, Kheiron/C could search for globally visible data structures e.g. the head of a linked list of abstract data types, and insert periodic checks of the list's consistency by injecting new function calls passing the linked-list head variable as a parameter.

To initiate an adaptation Kheiron/C attaches to a running application (or spawns a new application given the command line to use). The process of attaching causes the thread of the target application to be suspended. It then uses the Dyninst API to find the existing functions to instrument (each function abstraction has an associated call-before instrumentation point and a call-after instrumentation point). The target application needs to be built with symbol information for locating functions and variables to work. If necessary, Kheiron/C locates any "interesting" global structures or local variables in the scope of the intended instrumentation points. It then loads any external library/libraries that contain the desired adaptation logic and uses the Dyninst API to find the functions in the adaptation libraries for which calls will be injected in the target application. Next, Kheiron/C constructs function call expressions (including passing any variables) and inserts them at the instrumentation points. Finally, Kheiron/C allows the target application to continue its execution.

5.3 Preliminary Results

We carry out a simple experiment to measure the performance impact of Kheiron/C on a target system. Using the C version of the SciMark v2.0 benchmark we compare the time taken to execute the un-instrumented program, to the time taken to execute the instrumented program – we instrumented the SOR_execute and SOR_num_flops functions such that a call to a function (AdaptMe) in a custom shared library is inserted. The AdaptMe function is passed an integer indicating the instrumented function that was called. Our experiment was run on a single Pentium 4 Processor, 2.4 GHz with 1 GB RAM. The platform was Suse Linux 9.2 running a 2.6.8-24.18 kernel and using Dyninst v4.2.1. All source files used in the experiment (including the Dyninst v4.2.1 source tree) were compiled using gcc v3.3.4 and glibc v2.3.3.

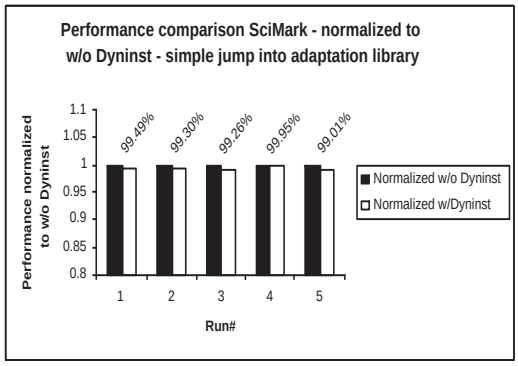

Figure 6. *Overheads Simple Instrumentation*

As shown in Figure 6 the overhead of the inserted function call is negligible, $\sim$1%. This is expected since the x86 generated behind the scenes effects a simple jump into the adaptation library followed by a return before executing the bodies of SOR_execute and SOR_num_flops. We expect that the overhead on overall program execution would depend largely on the operations performed while inside the adaptation library. Further, the time the SciMark process spends suspended while Kheiron/C performs the instrumentation is sub-second, $\sim$684 msecs $\pm$ 7.0686.

5.4 Injecting Selective Emulation

To enable applications to detect low-level faults and recover at the function level or, to enable portions of an application to be run in a computational sandbox, we describe an approach that allows portions of an executable to be run under the STEM x86 emulator. We use Kheiron/C to dynamically load the emulator into the target process' address space and emulate individual functions. STEM (Selective Transactional EMulation) is an instruction-level emulator – developed by Locasto et al. [34] – that can be selectively invoked for arbitrary segments of code. The emulator can be used to monitor applications for specific types of failure prior to executing an instruction, to undo any mem-

ory changes made by the function inside which the fault occurred (by having the emulator track memory modifications) and, simulate an error return from the function (error virtualization)[34].

The original implementation of STEM works at the source-code level i.e. a programmer must inject the necessary STEM "statements" into the portions of the application's source code expected to run under the emulator. To inject STEM into a running compiled C application, we need to make a few changes to STEM and Dyninst. The change to STEM is purely cosmetic: The original version of STEM is deployed as an GNU AR archive of the necessary object files; however, since the final binary does not contain an ELF header – needed for executables and shared object (dynamically loadable) files – we need to change the way STEM is built by using gcc with the -shared switch at the final link step. Once the STEM emulator is built as a true shared object, it can then be dynamically loaded into the address space of a target program using the Dyninst API.

The changes to Dyninst to support STEM are a bit more involved. Using the Dyninst API, Kheiron/C can inject calls to the **emulate_begin** and **emulate_end** function calls, which must surround the code to be run in the emulator. However, **emulate_init** (which must precede the call to emulate_begin) and **emulate_term** (which must come immediately after the call to emulate_end) are macros, which save the CPU state (registers) and restore or commit the emulator registers to the CPU registers respectively. These macros cannot be injected by Dyninst since they are intended to be expanded inline by the C/C++ preprocessor before compilation begins. In the original STEM implementation, these macros expand into inline assembly, which moves the CPU (x86) registers (eax, ebx, ecx, edx, esi, edi, ebp, esp, eflags) and segment registers (cs,ds,es,fs,gs,ss) into/from STEM data structures. STEM needs to store the contents of these registers to record the previous state of the machine and to calculate the address of the next instruction to fetch once it is loaded.

To load STEM dynamically we need to support its register save/restore process. To insert STEM Kheiron/C creates an area of memory in the target program's address space, *storage area*, large enough to hold the 9 general purpose (GP) x86 registers and the 6 segment registers. It uses the Dyninst API to find the functions to be run under the emulator and passes the address of storage area into Dyninst via a field added to the BPatch_point class, which is the concrete implementation of Dyninst's point abstraction. When Dyninst sets up the *trampoline* [6] used to inject x86 assembly into running programs we cause it to also save the GP registers and segment registers at offsets to the storage area. Kheiron/C loads the STEM emulator shared library and a custom library (dynamically linked to the STEM shared library) that has a function (RegisterSave). RegisterSave is passed the address of the storage area and copies data over from the storage area into STEM registers – so that a subsequent call to emulate_begin will work. Next it injects the call to the RegisterSave function passing the address of the storage area. Kheiron/C then uses Dyninst API to find and inject calls to the emulate_begin and emulate_end functions. Finally, Kheiron/C injects a custom snippet – a class inheriting from the BPatch_snippet class, which is the concrete implementation of Dyninst's snippet abstraction – to facilitate the restore/commit of STEM registers to the CPU registers and then allows the target program to continue.

At the end of this process, the instrumented function, when invoked, will cause the STEM emulator to be loaded and initialized with CPU and segment register values. After the initialization, the injected call to emulate_begin will cause STEM to being its instruction fetch-decode-execute loop thus running the function under the emulator.

[6] A trampoline is a small piece of code constructed on the fly on the stack.

6 Related Work

Our Kheiron prototypes are concerned with facilitating very fine-grained adaptations in existing/legacy systems, whereas systems such as KX [11] and Rainbow [5] are concerned with coarser-grained adaptations. However, the Kheiron prototypes could be used as low-level mechanisms orchestrated/directed by these larger frameworks.

One popular approach to performing fine-grained adaptations in managed applications is to use Aspect Oriented Programming (AOP). AOP is an approach to designing software that allows developers to modularize cross-cutting concerns [12] that manifest themselves as non-functional system requirements. In the context of self-managing systems AOP is an approach to designing the system such that the non-functional requirement of having adaptation mechanisms available is cleanly separated from the logic that meets the system's functional requirements. An AOP engine is still necessary to realize the final system. Unlike Kheiron, which can facilitate adaptations in existing systems at the execution environment-level, the AOP approach is a design-time approach, mainly relevant for new systems.

AOP engines *weave* together the code that meets the functional requirements of the system with the aspects that encapsulate the non-functional system requirements. There are three kinds of AOP engines: those that perform weaving at compile time (static weaving) e.g. AspectJ [10], Aspect C# [16], those that perform weaving after compile time but before load time, e.g. Weave .NET [7], which pre-processes managed executables, operating directly on bytecode and metadata and those that perform weaving at runtime (dynamic weaving) using facilities of the execution environment, e.g. A dynamic AOP-Engine for .NET [9] and CLAW

[19]. Kheiron/JVM is similar to the dynamic weaving AOP engines only in its use of the facilities of execution environment to effect adaptations in managed applications while they run.

Adaptation concepts such as Micro-Reboots [3] and adaptive systems such as the K42 operating system [2] require upfront design-time effort to build in adaptation mechanisms. Our Kheiron implementations do not require special designed-in hooks, but they can take advantage of them if they exist. In the absence of designed-in hooks, our Kheiron implementations could refresh components/data structures or restart components and sub-systems, provided that the structure/architecture of the system is amenable to it, i.e., reasonably well-defined APIs exist.

Georgia Tech's 'service morphing' [29] involves compiler-based techniques and operating system kernel modifications for generating and deploying special code modules, both to perform adaptation and to be selected amongst during dynamic reconfigurations. A service that supports service morphing is actually comprised of multiple code modules, potentially spread across multiple machines. The assumption here is that the information flows and the services applied to them are well specified and known at runtime. Changes/adaptations take advantage of meta-information about typed information flows, information items, services and code modules. In contrast, Kheiron operates entirely at runtime rather than compile time. Further, Kheiron does not require a modified execution environment, it uses existing facilities and characteristics of the execution environment whereas service morphing makes changes to a component of the unmanaged execution environment – the operating system.

Trap/J [31], Trap.NET [30] produce adapt-ready programs (statically) via a two-step process. An existing program (compiled bytecode) is augmented with generic interceptors called "hooks" in its execution path, wrapper classes and meta-level classes. These are then used by a weaver to produce an adapt-ready set of bytecode modules. Kheiron/JVM, operates entirely at runtime and could use function call replacement (or delegation) to forward invocations to specially produced adapt-ready implementations via runtime bytecode re-writing.

For performing fine-grained adaptations on unmanaged applications, a number of toolkits are available, however many of them, including EEL [20] and ATOM [35], operate post-link time but before the application begins to run. As a result, they cannot interact with systems in execution and the changes they make cannot be modified without rebuilding/re-processing the object file on disk. Using Dyninst as the foundation under Kheiron/C we are able to interact with running programs – provided they have been built to include symbol information.

Our Kheiron implementations specifically focus on facilitating fine-grained adaptations in applications rather than in the operating system itself. KernInst [36] enables a user to dynamically instrument an already-running unmodified Solaris kernel in a fine-grained manner. KernInst can be seen as implementing some autonomic functionality, i.e., kernel performance measurement and consequent runtime optimization, while applications continue to run. DTrace [4] dynamically inserts instrumentation code into a running Solaris kernel by implementing a simple virtual machine in kernel space that interprets bytecode generated by a compiler for the 'D' language, a variant of C specifically for writing instrumentation code. TOSKANA [8] takes an aspect-oriented approach to deploying before, after and around advice for in-kernel functions into the NetBSD kernel. They describe some examples of self-configuration (removal of physical devices while in use), self-healing (adding new swap files when virtual memory is exhausted), self-optimization (switching free block count to occur when the free block bitmap is updated rather than read), and self-protection (dynamically adding access control semantics associated with new authentication devices).

7 Conclusions

In this paper we describe the retro-fitting of fine-grained adaptation mechanisms onto existing/legacy systems by leveraging the facilities and characteristics of unmodified execution environments. We describe two classes of execution environments – managed and unmanaged – and compare the performance overheads of adaptations and the techniques used to effect adaptations in both contexts. We demonstrate the feasibility of performing adaptations using Kheiron/C and we describe a sophisticated adaptation, injecting the selective emulation of functions into compiled C applications. Given that few legacy systems are written in managed languages (e.g. Java, C# etc.) whereas a substantial number of systems are written in C/C++, our techniques and approaches for effecting the adaptation of native systems may prove useful for retro-fitting new functionality onto these systems.

8 Acknowledgments

The Programming Systems Laboratory is funded in part by National Science Foundation grants CNS-0426623, CCR-0203876 and EIA-0202063. We would also like to thank Matthew Legendre and Drew Bernat of the Dyninst team for their assistance as we used and modified Dyninst. We thank Michael Locasto and Stelios Sidiroglou-Douskos for their assistance as we used STEM.

References

[1] B. Buck and J. K. Hollingsworth. An API for Runtime Code Patching. *The International Journal of High Performance*

Computing Applications, 14(4):317–329, Winter 2000.

[2] C. Soules et. al. System Support for Online Reconfiguration. In *USENIX Annual Technical Conference.*, 2003.

[3] G. Candea, J. Cutler, and A. Fox. Improving Availability with Recursive Micro-Reboots: A Soft-State Case Study. In *Dependable systems and networks - performance and dependability symposium (DNS-PDS)*, 2002.

[4] B. M. Cantrill, M. W. Shapiro, and A. H. Leventhal. Dynamic Instrumentation of Production Systems. In *USENIX Annual Technical Conference*, pages 15–28, 2004.

[5] S.-W. Cheng, A.-C. Huang, D. Garlan, B. R. Schmerl, and P. Steenkiste. Rainbow: Architecture-based Self-Adaptation with Reusable Infrastructure. *IEEE Computer*, 37(10):46–54, October 2004.

[6] Daniel P. Bovet and Marco Cesati. *Understanding the Linux Kernel 2nd Edition*. O'Reilly, 2002.

[7] Donal Lafferty et al. Language Independent Aspect-Oriented Programming. In *18th ACM SIGPLAN conference on Object-Oriented Programming, Systems, Languages and Applications*, October 2003.

[8] M. Engel and B. Freisleben. Supporting Autonomic Computing Functionality via Dynamic Operating System Kernel Aspects. In *4th International Conference on Aspect-Oriented Software Development*, pages 51–62, 2005.

[9] A. Frei, P. Grawehr, and G. Alonso. A Dynamic AOP-Engine for .NET. Tech Rep. 445, Dept. of Comp Sci. ETH Zurich, 2004.

[10] G. Kiczales et al. An Overview of AspectJ. In *European Conference on Object-Object Programming*, June 2001.

[11] Gail Kaiser et. al. Kinesthetics eXtreme: An External Infrastructure for Monitoring Distributed Legacy Systems. In *The Autonomic Computing Workshop 5th Workshop on Active Middleware Services (AMS)*, June 2003.

[12] Gregor Kiczales et. al. Aspect-Oriented Programming. In *Proceedings European Conference on Object-Oriented Programming*, volume LNCS 1241. Springer-Verlag, 1997.

[13] R. Griffith and G. Kaiser. Adding Self-healing Capabilities to the Common Language Runtime. Technical Report CUCS-005-05, Columbia University, 2005.

[14] R. Griffith and G. Kaiser. Manipulating Managed Execution Runtimes to Support Self-Healing Systems. In *Workshop on Design and Evolution of Autonomic Application Software*, May 2005.

[15] R. Griffith, G. Valetto, and G. Kaiser. Effecting Runtime Reconfiguration in Managed Execution Environments. In M. Parishar and S. Hariri, editors, *Autonomic Computing: Concepts, Infrastructure, and Applications,*. CRC, 2006.

[16] Howard Kim. AspectC#: An AOSD implementation for C#. Technical Report TCD-CS-2002-55, Department of Computer Science Trinity College, 2002.

[17] J. O. Kephart and D. M. Chess. The Vision of Autonomic Computing . *Computer magazine*, January 2003.

[18] P. Koopman. Elements of the Self-Healing Problem Space. In *ICSE Workshop on Architecting Dependable Systems*, 2003.

[19] J. Lam. CLAW: Cross-Language Load-Time Aspect Weaving on Microsoft's CLR. Demonstration at AOSD 2002.

[20] J. R. Larus and E. Schnarr. EEL: machine-independent executable editing. In *ACM SIGPLAN 1995 conference on Programming language design and implementation*, pages 291–300, 1995.

[21] S. Lidin. *Inside Microsoft .NET IL Assembler*. Microsoft Press, 2002.

[22] T. Lindholm and F. Yellin. The Java Virtual Machine Specification Second Edition. http://java.sun.com/docs/books/vmspec/2nd-edition/html/VMSpecTOC.doc.html, 1999.

[23] Microsoft. Common Language Runtime Metadata Unmanaged API, 2002.

[24] Microsoft. Common Language Runtime Profiling, 2002.

[25] S. Microsystems. The Java Hotspot Virtual Machine v1.4.1. http://java.sun.com/products/hotspot/docs/whitepaper/Java_Hotspot_v1.4.1/Java_HSpot_WP_v1.4.1_1002_4.html, 2002.

[26] S. Microsystems. The JVM Tool Interface Version 1.0. http://java.sun.com/j2se/1.5.0/docs/guide/jvmti/jvmti.html, 2004.

[27] A. Mikunov. Rewrite MSIL Code on the Fly with the .NET Framework Profiling API. http://msdn.microsoft.com/msdnmag/ issues/03/09/NETProfilingAPI/, 2003.

[28] A. V. Mirgorodskiy and B. P. Miller. Autonomous Analysis of Interactive Systems with Self-Propelled Instrumentation. In *12th Multimedia Computing and Networking*, January 2005.

[29] C. Poellabauer, K. Schwan, S. Agarwala, A. Gavrilovska, G. Eisenhauer, S. Pande, C. Pu, and M. Wolf. Service Morphing: Integrated System- and Application-Level Service Adaptation in Autonomic Systems. In *Autonomic Computing Workshop, Fifth Annual International Workshop on Active Middleware Services*, June 2003.

[30] S. M. Sadjadi and P. K. McKinley. Using Transparent Shaping and Web Services to Support Self-Management of Composite Systems. In *Second IEEE International Conference on Autonomic Computing (ICAC)*, June 2005.

[31] S. M. Sadjadi, P. K. McKinley, B. H. C. Cheng, and R. E. K. Stirewalt. TRAP/J: Transparent Generation of Adaptable Java Programs. In *International Symposium on Distributed Objects and Applications*, October 2004.

[32] B. Schmerl and D. Garlan. Exploiting Architectural Design Knowledge to Support Self-Repairing Systems. In *14th International Conference of Software Engineering and Knowledge Engineering*, 2002.

[33] C. Shelton and P. Koopman. Using Architectural Properties to Model and Measure System-wide Graceful Degradation. In *Workshop on Architecting Dependable Systems*, 2002.

[34] S. Sidiroglou, M. E. Locasto, S. W. Boyd, and A. D. Keromytis. Building a Reactive Immune System for Software Services. In *USENIX Annual Technical Conference*, pages 149–161, April 2005.

[35] A. Srivastava and A. Eustace. ATOM: a system for building customized program analysis tools. In *ACM SIGPLAN 1994 conference on Programming language design and implementation*, pages 196–205, 1994.

[36] A. Tamches and B. P. Miller. Fine-Grained Dynamic Instrumentation of Commodity Operating System Kernels. In *3rd Symposium on Operating Systems Design and Implementation (OSDI)*, pages 117–130, 1999.

[37] Tool Interface Standards (TIS) Committee. Tool Interface Standard (TIS) Executable and Linking Format (ELF) Specification Version 1.2. http://www.x86.org/ftp/manuals/tools/elf.pdf, 1995.