Monitored Software Components
- A Novel Software Engineering Approach -

Daniel Sundmark†, Anders Möller †*, Mikael Nolin†
†MRTC, Mälardalen University, Box 883, SE-721 23 Västerås, Sweden
*CC Systems, www.cc-systems.com
{Daniel.Sundmark, Anders.Moller, Mikael.Nolin}@mdh.se

Abstract

We propose monitoring of software components, and use of monitored software components, as a general approach for engineering of embedded computer systems. In our approach, a component’s execution is continuously monitored and experience regarding component behaviour is accumulated. As more and more experience is collected the confidence in the component grows; with the goal to eventually allow certification of the component. Continuous monitoring is also the base for contract checking, and provides means for post-mortem crash analysis; an important prerequisite for many companies to start use 3rd party component in their dependable systems.

In this paper we show how four software engineering goals can be reached by monitoring four component properties.

1. Introduction

In this paper we propose monitoring of software components and use of monitored software components as a general approach for engineering of embedded computer systems. Industrial developers of distributed, heterogeneous, reliable, resource constrained, embedded, real-time control systems (in this paper denoted embedded systems) are facing increased challenges with respect to demands on increased profitability, functionality and reliability, while at the same time having to decrease development times, project costs and time-to-market. Since development costs only constitute a fraction of the total project cost for software projects (about 20% [19]), a general approach for engineering embedded systems must consider not only the development phase; also the debugging, testing and maintenance phases need to be addressed. Furthermore, since most systems are developed incrementally, where new versions are based on previous versions, and product-line architectures [5] are becoming increasingly important, a general approach for engineering embedded systems needs to consider reuse of components between product versions and product variants. Another emerging key-issue in engineering of embedded systems is safe and predictable integration of third-party functions, and the associated legal matters regarding contract fulfilment and liability issues.

The main contributions of this paper are three-fold:

- We present a novel approach to engineering component-based systems, using monitored software components.

- Our approach takes a life-cycle perspective on the engineering process, and we identify four key-areas where monitored components will have significant impact.

- We present four measurable properties that can be used to impact these key-areas.

The outline of the rest of this paper is as follows: In Section 2 we present a life-cycle approach to engineering component-based systems, and Section 3 describes properties of embedded systems. In Section 4, we present a survey of related work in built-in monitoring support for component-based systems and existing monitoring practices in commercial component technologies. In Section 5, the impacts of monitorable components on predictable assemblies are discussed. Section 6 illustrates how to make use of the monitored information, and finally, in Section 7, we summarise and present our ideas on future work.

2. A Life-Cycle Approach to Component-Based Systems

This continuous increase of requirements for embedded-system developers can be mitigated by deploying suitable software engineering methods. It is our view that the whole system and component life-cycles need to be considered by an engineering method. Below we present four key-areas of engineering component-based systems where significant
gains can be made by using our proposed concept of monitored components.

- Certifiable components. By monitoring component-based software, information about the component properties can be extracted. This information can be used to fully (or partially) describe the components by their externally visible properties. Such a description can be used as a basis for certifying components.\textsuperscript{1} By reusing certified components, predictable component assemblies are facilitated.

- System-level testing and debugging. By monitoring individual components and component interactions, errors can be found and traced. Monitoring can also be used to support replay debugging [22], where erroneous system-executions are recreated in a lab environment to allow tracing of bugs.

- Run-time contract checking. This will allow surveillance of third party components. Both functional (e.g. range of output values) and non-functional (e.g. memory usage) properties can be monitored. During acceptance testing, the contract checking is used to validate that a component does not violate its specification. In systems that fail after system deployment, logs from the contract checking can be used in post-mortem analysis to identify failing or contract-breaking components.

- Observability. Computer systems in general, and embedded systems in particular, are infamous for the difficulty of observing their internal behaviour. This has drawbacks throughout the whole debugging, testing and maintenance phases. Systems whose behaviour is unobservable become very difficult to analyse and validate. Also after deployment, observability is an important feature, allowing inspection and performance tuning of running systems.

The ultimate goal of component monitoring is to be able to compose predictable assemblies by reusing information gathered from well-tested software components. The proposal of this paper is that this can be achieved by the iterative process of refinement described in Figure 1. When a new component (or a 3rd party component) is included in an assembly, its run-time properties (such as execution time or memory consumption) are estimated by well-founded guesses. During testing, these guesses are validated and refined. As the tested component is deployed in a target-system assembly, its behaviour is continuously monitored, allowing for further refinement of the component run-time behaviour description. This refinement process will eventually lead to a certified component, which can be used to compose predictable assemblies.

Monitoring of components will allow information about the dynamic behaviour of the component to be recorded. This information allows static and dynamic properties of newly (or partly) constructed systems to be predicted. Interesting aspects to monitor (on component level) and predict (on system level) include timing properties, such as end-to-end response times, and resource utilisation, such as memory consumption.

3. Embedded Systems

This paper addresses software engineering aspects of resource-constrained, embedded, distributed real-time control systems. To make clear in what context the provided monitoring approach is supposed to work, we also provide a brief example of a typical embedded system and an introduction to component monitoring. In this section we also discuss prerequisites, placed on the component technology and on the hardware, to be able to monitor the embedded system software.

3.1. CBSE for Embedded Systems

In Component-Based Software Engineering (CBSE), software applications are built by composing software components into component assemblies. CBSE is gaining more and more acceptance in the business segment of office/Internet applications [2][3]. Unfortunately, the market segment of embedded real-time systems is, to a large extent, left behind this positive development. Reusing components, i.e. one of the main drivers for introducing CBSE, is both complex and expensive for embedded real-time systems [6].

However, by building embedded-system software out of well-tested components, we could gain an increase in the predictability of the behaviour of the software; provided that experience from component behaviour has been collected.

\textsuperscript{1} For some high integrity systems, monitored properties need to be combined with static analysis to obtain safe bounds on properties.
In the area of embedded real-time systems, predictable run-time behaviour is crucial. A component assembly is predictable if its run-time behaviour can be predicted from the properties of its components and their patterns of interactions [4]. Predictability is achieved by analysis, and analysis techniques require information about the system. When analysing a system built from well-tested and functionally correct components, the main issues are associated with composability. The composition process should be able to guarantee the fulfilment of non-functional requirements of the system, such as communication, synchronisation, memory, and timing [6]. However, research projects tend to focus on how to design and analyse component technologies, leaving predictable assemblies using run-time information gathered from well-tested and trusted components unexplored [9].

3.2. Embedded System Example

In order to exemplify the typical settings, in which the software components are considered, we have studied some characteristic vehicular electronic systems [17]. An electronic vehicular control-system can be characterised as a resource constrained, safety-critical, distributed real-time system. The computer nodes, called Electronic Control Units (ECUs), are distributed to reduce cabling and to allow for division into subsystems. Vehicular systems are usually heterogeneous, meaning that nodes of different architecture and computational power cooperate in controlling the vehicle. The ECUs vary from extremely light-weighted nodes, like intelligent sensors (i.e. processor-equipped, bus-enabled sensors), to PC-like hardware for non-control applications, such as telematics, and information systems.

Figure 2 gives an overview of the hardware resources of a typical ECU, with requirements on sensing and actuating, and with a relatively high computational capacity.

<table>
<thead>
<tr>
<th>Example Power train ECU in a Vehicular Control-System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor: 25 MHz 16-bit processor</td>
</tr>
<tr>
<td>Memory devices:</td>
</tr>
<tr>
<td>✔ Flash: 1 MB used for application code</td>
</tr>
<tr>
<td>✔ RAM: 128 kB used for the run-time memory usage</td>
</tr>
<tr>
<td>✔ EEPROM: 64 kB used for system parameters</td>
</tr>
<tr>
<td>Serial interfaces: RS232 or RS485, used for service purpose</td>
</tr>
<tr>
<td>Communications: Controller Area Network (CAN) (one or more interfaces)</td>
</tr>
<tr>
<td>I/O: A number of digital and analogue in and out ports</td>
</tr>
</tbody>
</table>

Figure 2. Specification of a typical power train ECU.

An example of a typical vehicular system communication solution is shown in Figure 3, where two buses are separated by a gateway. The gateway is an architectural pattern that is used for several reasons, e.g., separation of criticality and real-timeliness, increased available bus bandwidth, increased fault tolerance, or compatibility with standards [1]. Communicating functions may require support for global synchronisation or fault tolerance mechanisms.

Figure 3. Example sketch of a vehicle network.

Looking at the software part of the system, there are some aspects that need to be considered when building the assembly out of monitorable components. A source of uncertainty is the frequency of interrupts in the assembly. Typically, a vehicular system is heavily loaded with interrupts. When interrupts hit the assembly, these will pre-empt the execution of the running component, thereby possibly perturbing its monitoring.

Dynamic memory allocation (and the garbage collection that this brings) is usually not allowed in control applications, since it compromises the determinism and predictability of the application behaviour. The only type of memory that is allowed to dynamically shrink and grow in the system is the stack space (albeit within a statically allocated stack memory area).

3.3. Prerequisites for Monitoring Component-Based Embedded Systems

Monitoring component-based software requires support in the component technology, and the framework used during run-time. Usually, when looking at today’s component technologies suitable for embedded systems with resource constrained ECUs, considerable code optimisations are done during compile time. This is mainly done to minimise the size of the application source code. This code optimisation might lead to a loss of the design-time component concept, meaning that clearly identifiable components with
specified in- and out-ports are reduced to regular source code functions, subjected to, e.g., function in-lining and redundant instruction-sequence coalescing.

Thus, to be able to monitor the components in the form described during design-time and to be able to reuse the information gathered during run-time in the next generation of applications, information about the design-time components have to be included in the source code. This should however not be a problem, if the component technology satisfies the requirements described in [17], i.e. a straight forward port-based object approach, illustrated in Figure 4, using a pipes-and-filters model of computation.

Component monitoring also poses some requirements on the system hardware. A fraction of the system memory needs to be allocated for monitor recording purposes. In addition, the target system should support some suitable means of communication through which monitor recordings can be uploaded to a host computer. Furthermore, some of the debugging techniques discussed in this paper will benefit significantly from the support of an instruction counter.

4. Related Work

To summarise the available techniques that can be used to monitor software components, we have studied some commercial component technologies that include support for component monitoring and the state-of-the-art methods for component monitoring. The methods and the different technologies are described in Section 4.1 and Section 4.2.

4.1. Monitoring Techniques for Component-Based Systems

Currently, only a few component technologies provide support for run-time monitoring of component behaviour. However, there is numerous ways of performing this monitoring and there is a multitude of run-time aspects to monitor.

Gao et al. identify three different methods for component tracking and monitoring [9]: (A) framework-based code insertion, where monitoring code (e.g. from a class library) can be inserted by component engineers, (B) automatic code insertion, where monitoring code is inserted into the program by a specialised monitoring tool, and (C) automatic component wrapping, where monitoring code is automatically added to the external interface of components.

According to Gao et al., each of these methods has its own pros and cons. As for framework-based code insertion, it is highly flexible and can be used for all types of monitoring. However, the method requires access to the component source code, and the programming overhead is high. Automatic code insertion also requires access to the source code, and is much more complex and inflexible compared to the framework-based code insertion. However, the programming overhead is low, since the tracking code is automatically inserted. Automatic component wrapping, on the other hand, has no need for component source code in order to insert tracking code. Therefore, not only in-house components, but also Commercial-Off-The-Shelf (COTS) components can be monitored. On the downside, automatic component wrapping is not suitable for monitoring anything within components, since the monitoring is performed exclusively outside the component.

Considering the use of these methods with respect to the restrictions posted by component-based embedded systems, it should be noted that automatic component wrapping can not be used in order to extract any component information other than that available at the component ports. This makes the method unsuitable for monitoring other component properties than those available from outside the component. Automatic code insertion, on the other hand, could be used for all types of monitoring, but would introduce a trade-off between the complexity of the instrumentation tool and the amount of data needed to record. Ideally, especially in resource-constrained systems, the amount of data to record should be minimised. However, this calls for an elaborate analysis of the internal workings of the component, requiring an inflexible (with respect to portability) and highly advanced instrumentation tool. Using framework-based code insertion, no instrumentation tool is required, allowing ad-hoc optimisations in the monitoring code. In a resource-constrained environment, this might be useful, but it must be kept in mind that such optimisations might lead to unpredictable probe-effects in system ordering and timing [8].

Jhumka et al. [14] propose the use of executable assertions in order to monitor component behaviour. The assertions are included in component wrappers, enabling them to test the validity of the input and output values of the component. By using these wrapper assertions, the pre- and post-condition sanity checks transforms a regular compo-
C. Monitoring Support in Commercial Component Technologies

There is a handful of available component technologies suitable for distributed embedded real-time systems. Some of these technologies include various supports for monitoring the software. The reason for choosing these is that they are deployed in industry today, and that they well satisfy the industrial requirements stated by the embedded-system domain [17].

The Rubus Component Model (CM) [16] and the Rubus Operating System (OS) have support for some of the described monitoring aspects. Rubus CM and OS are developed by Arcticus Systems2 and are used for developing heavy vehicle software systems by, e.g., Volvo Construction Equipment3 (VCE). When using the Rubus CM and OS, all resource allocation of the application and the operating system is done at compile-time.

The temporal properties needed to obtain static timing analysis and schedule generation, Best-Case Execution Time (BCET) and Worst-Case Execution Time (WCET), are monitored by the Rubus OS during run-time. Apart from the temporal aspects of the software, maximum stack usage for each thread and the peak usage of, e.g., queues can be monitored using Rubus. The OS also gives support for monitoring the CPU utilisation.

In multi-threaded embedded software, various types of relations, such as precedence and exclusion relations, exist. To be able to guarantee the behaviour of the system with respect to these issues, the Rubus CM includes support for monitoring event traces of the program execution, i.e., the execution order and the release times of the components. This information is dumped on an external interface (e.g., CAN or a serial interface like RS485) during run-time. Since events are related only via time-stamps, this service requires a high-resolution hardware timer. There will be a significant amount of data associated with this monitoring, and the accuracy of the log reflects the size of the buffer used to store it.

PECOS4 (PErvasive COmponent Systems) [10] is a collaborative project between ABB Corporate Research Centre5 and academia. The goal for the PECOS project is to enable component-based technology for embedded systems, especially for field devices, i.e., embedded reactive systems. The project tries to consider non-functional properties, such as memory consumption and timeliness, very thoroughly in order to enable assessment of the properties during construction time.

---

2 Arcticus Systems, www.arcticus.se
3 Volvo CE, www.volvo.com
4 The PECOS Project, www.pecos-project.org
5 ABB Corporate Research, www.abb.com
Non-functional properties cannot only be attached to components, but also to ports and connectors, e.g., examining the min and max values for an out port, (i.e., a built in sanity check). Since PECOS is developed to support resource constrained embedded real-time systems, scheduling information and memory consumption are crucial properties to monitor. Hence, PECOS enables support for instrumenting components during run-time. Every component is instrumented to extract information about the WCET and its cycle time. The components are also instrumented with respect to their code size and data (i.e., information on the heap).

5. Monitoring Software Components

Although a multitude of component properties are of interest when building reliable and reusable software components, there are some aspects that would significantly help increasing reusability and lower the time spent on integration testing. We have identified four main aspects to monitor, in order to support the key areas defined in Section 1.

5.1. Temporal Behaviour

Having knowledge of the temporal behaviour of an execution is particularly important for real-time systems. If the worst-case and best-case execution times of a set of reusable components are known, the possibility of successfully predicting the temporal behaviour of the component assembly will radically increase. Also, other execution time metrics, such as average execution time, standard deviation, execution time histogram or other types of statistical representations of component execution time behaviour can be helpful to estimate statistical temporal properties of component assemblies [18].

When considering timeliness for embedded real-time systems, it is important to be able to verify (1) that each component meets its timing requirements, (2) that each node (which is built up from several components) meets its deadlines, and (3) to be able to analyse the end-to-end timing behaviour of functions in a distributed system. In order to make sure that all deadlines are met, temporal analysis is needed.

This type of analysis is performed using schedulability analysis techniques, and requires information about the component’s execution time. Ideally, the bounds for worst-case and best-case execution times should be statically computed by an analysis tool; this is the only way to be sure that the execution-time bounds are safe (i.e. guaranteed not to be violated at run-time), see e.g. [7]. Unfortunately, tools for execution-time analysis are immature and few commercial tools exist. Hence, the industrial practice is to rely on measurements of execution-times. However, structured measurement of execution-times is a tedious, error-prone and expensive process, which has to be re-done after each modification to a component. Using monitored components, the correctness of the execution time values can be improved gradually, i.e., the more execution hours, the better the accuracy [18]; this is achieved without any extra effort for execution-time measurement.

In general, execution behaviour information is used for schedulability analysis and scheduling. In hard real-time systems, where it is mandatory that deadlines are met, deterministic schedulability analysis and scheduling (using worst-case assumptions for execution times) is preferable. However, in practice, many systems would settle for high probabilities instead of absolute deadline guarantees. Therefore, stochastic schedulability analysis and scheduling can be used. Depending on the type of analysis intended, either worst-case or statistical timing metrics should be collected during monitoring.

5.2. Memory Usage

Since we are targeting resource-constrained systems, it is important to be able to analyse the memory consumption and to check the sufficiency of the system memory, as well as the ROM memory. This check should be done pre-runtime to avoid failures during run-time. Memory is allocated in a static (pre-run-time or during run-time initialisation) or a dynamic (run-time) fashion. As mentioned in Section 3.1, dynamic memory allocation is usually not allowed when developing embedded real-time systems. In order to improve the possibility of achieving predictable assemblies, information of static memory allocation (e.g., component binary size) is necessary, but since this information can be provided by means of compiler output, this property is typically not necessary to monitor.

The stack memory, however, is statically allocated, but used in a dynamic fashion. In order not to end up in a stack overflow situation, stack size is often pessimistically over-dimensioned during system configuration. In resource-constrained environments, this might lead to a situation where the high percentage of unused memory leads to increased requirements on the system hardware. Therefore, monitoring the stack usage per component is most important, since this information can be used to predict the stack usage behaviour of future assemblies. Due to the high criticality of stack overflow, we are not interested in anything but worst-case usage during the execution of the component. However, in a system allowing dynamic allocation, also heap size monitoring would be important.
5.3. Event Ordering

When testing and debugging software, it is often helpful to be aware of the occurrence and ordering of system events, such as mutex- and semaphore operations, message receipts and interrupt occurrences. Using the information provided by an event log, system designers are able to detect improperly synchronised accesses to shared data or illegal pre-emption of non-reentrant code. In addition, by including event monitoring in the component model, we ensure that all components conforming to the model will produce event-trace logs of similar formats. This will reduce the problem of ad-hoc tracing code inserted by system developers.

Using event traces, we can gain substantial insight regarding the internal workings of current assemblies. This information can be used in order to guarantee precedence relations, mutual exclusion and to enhance the efficiency of shared resource usage (e.g., field bus or third-level storage usage) in future assemblies.

This type of monitoring provides a foundation to include full support for a replay debugging method [22][15][21] in the component technology. Replay debugging is a general term denoting methods for recording the execution behaviour of multi-tasking or truly parallel systems in order to use this information to reproduce system failures during debugging. Most replay methods require both event ordering information (such as interrupt, context switch and synchronisation information) and data flow information (such as task state and external input information) in order to reproduce executions. Provided that the assembly infrastructure (e.g., real-time operating system mechanisms) includes support for replay debugging, including sufficient monitoring in the components will ensure that the entire assembly can be debugged by means of execution replay.

5.4. Sanity Check

A sanity check is a way of determining the soundness of the functional operation of a component during run-time with respect to the component input and its current state. In other words, given a specific input, is the corresponding output realistic? During testing, having access to the input values of the component that produce erroneous output facilitates efficient testing.

Monitoring this during run-time will allow us to store erroneous operations of the component and (hopefully) to correct these errors in future assemblies. If we are unable to correct the faulty component, we could still able to prevent unsafe system behaviour by taking appropriate actions based on knowledge of the errors. This type of monitoring could also include properties like Mean-Time Between Failures (MTBF).

<table>
<thead>
<tr>
<th>Exec. Time</th>
<th>Memory</th>
<th>Event Ordering</th>
<th>Sanity Check</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certifiable Components</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Debug/Testing</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>R-T Contract Checking</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Observability</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 1. Mapping key-areas of interest to key component aspects.

In addition, this type of monitoring could be used to ensure that 3rd party software components provide the service they are supposed to. Typically a component is equipped with a provided interface, specifying the services provided by that component, and a required interface, specifying the resources needed by the component in order to provide the correct services. Formalising and standardising these interfaces allows for contractual-based component development, where the behaviour of 3rd party components included in the assembly can be specified by contracts. Using sanity checks of the inputs and outputs at the component interfaces allows for run-time contract checking of 3rd party components.

6. Using Monitored Information

In Section 2, four key-areas that would benefit from component-level monitor support were listed. Table 1 maps these areas to the four monitored component aspects discussed in this Section. For instance, execution time- and memory information can be used in order to automatically check whether 3rd party components do not violate their required interface (e.g., by memory leaks or deadline misses), and sanity checks can be used to check the provided interface during run-time. By using event ordering and sanity check traces, the observability and ability to easily test and debug the assembly can be considerably enhanced. Regardless of whether replay debugging methods are used or not, event traces are helpful during debugging in order to visualise the behaviour of the component assembly during run-time.

As for certifiable components, all monitoring aspects can be helpful in order to successfully predict the future behaviour of components in different types of assemblies. Including monitoring in a component technology will ensure that all components conforming to that technology will include identical monitoring support. Hence, component properties can be easily compared using standardised means of comparison.
7. Conclusion and Future Work

In this paper we have proposed monitoring of software components, and reuse of monitored components, as a general approach towards engineering of resource constrained, embedded, distributed, real-time control systems. The concept is general in the sense that it addresses not only the development phase; rather the whole product life cycle, including debugging, testing and maintenance, is considered. The concept also extends well into product-line settings, where components and architectures are reused over a set of related product and product variants. We have identified four key-areas within engineering of embedded systems: (i) certifiable components, (ii) system-level testing and debugging, (iii) run-time contract checking, and (iv) observability. We have also discussed how to meet the challenges within these areas, by identifying four main component-aspects that are of particular interest to monitor: (1) the execution time behaviour of the components, (2) the static and dynamic memory usage, (3) the event ordering of the execution and (4) a sanity check of the components output based on the input. We have provided a summary of the state-of-the-art of monitoring support for component models and presented a brief survey of the practices used in today’s component models for embedded real-time systems.

As for future work, there are a number of issues we would like to address. We intend to look further into the problems of using the same component on top of different hardware platforms, where some old monitoring information might be reused, while other information needs to be discarded on the new platform. Furthermore, the trade-off between minimisation of monitoring memory and CPU usage and the level of detail of monitor information will be investigated. In order to evaluate our ideas, we plan to use the SaveComp Component Model, described in [11], as a test platform. The component model is designed for embedded control-systems in vehicles, and much focus has been spent on solving reliability, timelines and safety issues.

References