Policy Objective (abstract)

Resource management in many embedded systems application domains (including consumer electronics, robotics, automotive and telecommunications) faces the challenges of highly variable resource demand, highly variable resource availability, distributed architecture, and strong requirements for resource.

Hence, the objective of the research is to exploit the excellence of the different teams for advancing the state of the art in the theory and practice of techniques leading to adding the required flexibility to the resource management in these systems. The challenge is to develop an efficient resource manager that can be adopted in next generation kernels to perform adaptive QoS control of time sensitive applications with dynamic characteristics.

The former activities “Flexible Scheduling Technologies” and “Adaptive Resource Management for Consumer Electronics” have been merged into this activity.
Table of Contents

1. Overview of the Activity ........................................................................................................3
   1.1 ARTIST Participants and Roles .........................................................................................3
   1.2 Affiliated Participants and Roles .......................................................................................3
   1.3 Starting Date, and Expected Ending Date ........................................................................3
   1.4 Baseline ............................................................................................................................4
   1.5 Problem Tackled in Year 3 ...............................................................................................4
   1.6 Comments From Year 2 Review
      1.6.1 Reviewers’ Comments ...............................................................................................5
      1.6.2 How These Have Been Addressed .............................................................................6

2. Summary of Activity Progress .............................................................................................7
   2.1 Previous Work in Year 1 .................................................................................................7
   2.2 Previous Work in Year 2 ...............................................................................................7
   2.3 Current Results
      2.3.1 Technical Achievements .........................................................................................11
      2.3.2 Individual Publications Resulting from these Achievements ................................15
      2.3.3 Interaction and Building Excellence between Partners ........................................19
      2.3.4 Joint Publications Resulting from these Achievements ........................................19
      2.3.5 Keynotes, Workshops, Tutorials ..............................................................................21

3. Future Work and Evolution ...............................................................................................23
   3.1 Problem to be Tackled over the next 12 months (Sept 2007 – Aug 2008) .................23
   3.2 Current and Future Milestones ......................................................................................24
   3.3 Indicators for Integration ...............................................................................................24
   3.4 Main Funding ..................................................................................................................24

4. Internal Reviewers for this Deliverable ............................................................................26
1. Overview of the Activity

1.1 ARTIST Participants and Roles
Activity Leader: Gerhard Fohler - Technische Universität Kaiserslautern
   Role: resource management framework, video streaming, wireless networks
Team Leader Giorgio Buttazzo - Scuola Superiore S. Anna
   Role: scheduling mechanisms for temporal isolation
Team Leader Michael Gonzalez Harbour - University of Cantabria
   Role: end-to-end timing analysis of distributed systems
Team Leader Luis Almeida - University of Aveiro
   Role: Bandwidth adaptation in networks
Team Leader Eduardo Tovar - Polytechnic Institute of Porto
   Role: communication protocols for wireless networks
Team Leader Alejandro Alonso - UP Madrid
   Role: adaptive techniques for QoS
Team Leader Alan Burns – University of York (UK)
   Role: fixed priority schemes and real-time languages

1.2 Affiliated Participants and Roles
Team Leader Marisol García-Valls - U. Carlos III, Madrid
   Role: QoS management
Team Leader Pau Martí - Universitat Politècnica de Catalunya
   Role: real-time and control
Team Leader Paolo Gai - Evidence, SME
   Role: tools for resource management
Team Leader Liesbeth Steffens - NXP
   Role: consumer electronics, video streaming, multi resource management
Team Leader Ivo De Lotto – University of Pavia (Italy)
   Role: dynamic priority schemes
Team Leader Lucia Lo Bello – Univ. of Catania (Italy)
   Role: communication protocols and stochastic scheduling

1.3 Starting Date, and Expected Ending Date
The activity started in Sept 2006 as the merger of the previous Cluster Integration Activities “Flexible Scheduling Technologies”, and “Adaptive Resource Management for Consumer Electronics”, and will be complete when research topics of industrial relevance and example solutions have been identified and developed and the cluster is able to define a scheduling framework that is capable of managing quality of service requirements for multiple resources. Expected ending date: August 2008
1.4 Baseline

In some application domains, such as multimedia, applications are very expensive in terms of resource consumption. In other applications domains, such as automotive, mobile telephony or even building automation, the resources are scarce and there is a growing pressure to integrate resources even further and optimize their use. In both cases, timeliness directly relates to user perceived quality, e.g., smoothness of the video stream. Furthermore, efficient resource usage is key issue not only for cost considerations, but also for competition on a feature bases: better resource usage – more features.

Both resource demands, e.g., MPEG-2 video streams, and resource availability, e.g., available bandwidth on wireless links, fluctuate rapidly and unpredictably; worst case assumptions will lead to extreme over provisioning. Consequently, methods for adaptive resource management are required.

Trading resource usage (processing, communication and memory/storage, inter-device and intra-device) against offered output is known as QoS (Quality of Service). The different resources cannot be considered separately, interferences and inter-resource tradeoffs have to be taken into account because they affect the application output. The tradeoffs have to be made at different time scales, in order to match the time scales of the system dynamics.

Theory for independent scheduling algorithms is well defined in the areas of event triggered and time triggered systems, but few theoretical results have been achieved in trying to integrate these approaches. Some partial results exist for simplified architectures, but it is necessary to enhance them by taking into account all of the requirements of modern real-time systems including distributed ones. In addition to the development of theory, a framework needs to be built in order to allow a flexible way to handle different scheduling algorithms for different kinds of resources, and evaluate their applicability to real application domains.

1.5 Problem Tackled in Year 3

The objectives of the activity in the second year crystallised around the following issues:

- **Application requirements.** The characteristics of adaptive applications diverge significantly from the classic, static models: the amount of resources needed varies largely and differently for resource. In a joint effort with the FRESCOR project, application domains covered in the activity were considered.

- **Adaptive Management of Multiple Resources.** The resources typically used in end-to-end delivery of streams often exhibit fluctuating availability and interdependencies. Wireless networks, for example, are influenced by interference, mobility, or physical structures, which cannot be known before system deployment.

  Even on single devices, a number of resources will be interdependent making the issue of multi resource management important. The focus of efforts has been put towards integrating CPU scheduling and cache management for efficient cache use and predictability.

- **QoS Middleware.** In addition to these effects on single devices, the fluctuating resource availability has to be dealt with on a system wide level as well, allowing the system to make QoS adaptation decisions in response to fluctuations in the nodes' service load, under the
control of the user. Since complex dynamic real-time scenarios may prevent the possibility of computing optimal service configurations before execution, iterative refinement approaches with the ability to trade off deliberation time for the quality of the solution will be more practical. Another issue is the adaptation of distributed QoS negotiation policies and the integration of this middleware with communication protocols.

- **Flexible scheduling Technologies.** Handling multiple concurrent activities with different criticality and timing constraints (e.g., periodic, aperiodic, time driven, event driven tasks) in the same system requires a development of an integrated flexible scheduling framework. Work in the first two years of this activity had focused on the integration of diverse scheduling schemes, such as fixed priority based, dynamic priority based, and offline construction into a single coherent set, on the development of new theory for the case of energy-aware scheduling, and on the requirements for the flexible scheduling framework. Work in the current year has focused on the integration of different resources into the same flexible scheduling framework, by extending and further refining the requirements for the framework, and by starting its implementation. Work in the following year will focus on finishing the implementation and performing the evaluation of the technologies developed.

- **Overload management techniques.** A number of overload management techniques have been tested to evaluate their impact on real-time systems. In particular, *Resource Reservations* and *Elastic Scheduling* techniques have been evaluated to cope with transient and permanent overload conditions. *Resource Reservations* techniques basically isolate the temporal behavior of a task (or subset of tasks) protecting the rest of the systems from potential overruns. On the other hand, *Elastic Scheduling* provides an effective solution to cope with permanent overload conditions. According to this method, task utilizations are treated as flexible springs that can be compressed (by enlarging periods) to reduce the load up to a desired value. Such novel techniques (not yet available in commercial operating systems) have been implemented into Shark as basic scheduling modules to be tested and evaluated in actual control applications.

### 1.6 Comments From Year 2 Review

#### 1.6.1 Reviewers’ Comments

Comments in relation with Deliverable D10-ART-Y2 Adaptive Resource Management for Consumer Electronics:

**ACCEPTED**

“This task concentrated this year in defining common requirements for resource management especially in the domain of multimedia application, wireless networks and middleware QoS. The deliverable is concise, very clear, well structured. Collaboration and funding are identified. Future work is very factual and clear.”

Comments in relation with Deliverable D9-ART-Y2 Flexible Scheduling Technologies:

**ACCEPTED**

This task concentrated this year on the integration of different scheduling policies to cope with different application requirements.
The document is very concise (which is nice) but the current results on section 2.2 do not provide enough information.

- The requirements for integrated-resource scheduling framework are not included in the document.
- The baseline for integrated-resource scheduling framework lack details or at least some references.
- The new theoretical development is much clearer but need reference to articles. Same comments for the following section.

The document is accepted. However it should be completed with references and more detailed results. The timetable needs to be updated.

1.6.2 How These Have Been Addressed

A revised version of deliverable “D9-ART-Y2. Flexible Scheduling Technologies” was produced with the following enhancements:

- The requirements for integrated-resource scheduling framework were included in section 2.2.1.
- The description of the baseline for integrated-resource scheduling framework, in section 2.2.1, was extended and missing references were added pointing to the bibliography section (2.2.2).
- The description of the new theoretical developments (in section 2.2.1) had references added.
- The description of flexible architectures and communication protocols for networks used in distributed embedded real-time systems (in section 2.2.1) had references added.
2. **Summary of Activity Progress**

2.1 **Previous Work in Year 1**

In the first period, the technical results were achieved in the following areas: video stream demand analysis, identification of scheduling algorithms and kernel mechanisms for stream adaptations based on integrated, flexible scheduling; adaptive resource management for network bandwidth management, multi resource management, in particular with respect to cache aware scheduling; middleware support for QoS management. Furthermore, the ART cluster has been in active contacts with relevant industry to gather understanding of realistic requirements and to identify research topics and baselines relevant for industrial and academic research. Partners has been giving presentations at the Philips Software Conference – Real-time Workshop and had meetings with Nokia, Ericsson mobile platforms and Visual tools from Spain. The goal has been to go as far as possible towards the actual engineers for better understanding and prepare for a specific industry – academia workshop with selected participants.

New scheduling mechanisms for integrating overload management techniques with energy-aware strategies were investigated in the context of real-time systems. The new scheduling mechanisms were analysed to guarantee timing constraints while minimizing energy consumption, and a kernel infrastructure was developed into the Shark operating system in order to facilitate their implementation. Moreover, the assessment of the (m,k)-firm model was done for its implementation over the FTT-Ethernet protocol.

2.2 **Previous Work in Year 2**

**Temporal Constraints for Video streaming**

Philips and TUKL have studied temporal constraints of video streaming. As sources for the constraints we looked into semantic stream dependencies from MPEG decoding, as well as the temporal impact of devices and their resources in the end-to-end delivery chain of a stream. The work was carried out with industrial partners in the area. The results will feed into other activities in the cluster, in particular w.r.t. to scheduling and networking. [http://rts.eit.uni-kl.de/research/mediaprocessing](http://rts.eit.uni-kl.de/research/mediaprocessing)

**Integrated real-time scheduling and cache management**

Philips and TUKL have continued work on integrating real-time scheduling and cache management on multiprocessor platforms. To this end, we have carried out experiments to study cache behaviour on the actual platform and formulated a number of scenarios with increasing complexity. A joint PhD student is carrying out the work.

**Adaptive service configuration for Quality-of-Service aware middleware (Porto)**

In order to support dynamic services with adaptive QoS requirements, we have proposed a dynamic scheduler which is able to react to load variations. Isolation between different services is still achieved by guaranteeing a minimal service quality to accepted services and by an efficient overload control that considers the challenges and opportunities of dynamic distributed embedded systems. This scheduler has also been extended taking into consideration simple dependencies between services’ QoS attributes.
Server Based Flexible Scheduling
Schedulability analysis techniques were developed for server-based systems that can be used to schedule different kinds of flexible timing requirements, such as those needed to integrate control systems with multimedia activities. In particular, this work focused on hierarchical scheduling analysis and design techniques. A further issue was the dimensioning of the parameters of a server for minimizing the average response time of the served activities. A statistical approach was addressed in order to compute the probability of missing a given deadline. Partners were SSSA, Cantabria, TUKL of the cluster and the partners of the FIRST and FRESCOR EU STREP consortia. www.frescor.org http://rts.eit.uni-kl.de/research/adaptive-rts

Adaptive resource management for networks
Work concerned the analysis of the achievable QoS guarantees in wireless networks. After defining a proper model for the main resources (i.e., CPUs, disk and network), a number of existing scheduling algorithms for the three types of resources were analysed under fluctuating workload to evaluate their behaviour in terms of service guarantee. Then, the achievable end-to-end QoS guarantees were investigated as a function of the guarantees provided by the underlying resource schedulers. Further activities dealt with network protocols to efficiently support dynamic bandwidth management with strict QoS guarantees in Ethernet-based networks, which is still an important networking technology in the field of distributed multimedia systems. A wireless time-token communication protocol that allows providing real-time guarantees for real-time messages and tune the allocated bandwidth according to the required QoS was developed. Aveiro, Porto, SSSA, and TUKL carried out work. http://www.hurray.isep.ipp.pt/activities/art-wise/
http://www.rts.eit.uni-kl.de/research/mediaprocessing

Adaptive service configuration for Quality-of-Service aware middleware
Complex dynamic real-time scenarios may prevent the possibility of computing optimal service configurations before execution; an iterative refinement approach with the ability to trade off deliberation time for the quality of the solution was specified. The approach is to quickly find good initial solution and to propose heuristic evaluation functions that optimize the rate at which the quality of the current solution improves as the algorithms have more time to run. The work also addressed the problem of dynamically changing system conditions, allowing the system to make QoS adaptation decisions in response to fluctuations in the nodes service load, under the control of the user. Monitoring the stability period and resource load variation of Service Level Agreements for different types of services was used to dynamically adapt future stability periods, according to a feedback control scheme. Work was done by Madrid and Porto (www.hurray.isep.ipp.pt/activities/qos)

Middleware
System adaptation requires full knowledge of the system state, therefore work was also carried out in a framework to gather actual resource usage information, and interact with the operating system, extending the traditional POSIX trace model with a partial reflective model for operating system monitoring. The work was done by Porto. HOLA-QoS is a framework for managing QoS and resources and it has been used in media processing which UPM and UC3m have developed jointly. It is implemented as a layered architecture, so that layers can be replaced, as far as the API is kept. The higher layers are meant to deal with quality, while lower layers are mainly related with resource management. In particular, the lowest layer is intended to manage budgets or resource shares assigned to
applications. This layer has provided accounting and enforcing facilities to ensure that budgets are guaranteed. Sometimes this functionality is provided by what is called resource kernels. Cluster partners have developed kernels that provide these facilities and, hence, could be suitable to act as the lower layer of a HOLA-QoS based system. Work that was under development was to port HOLA-QoS on top MARTE (Cantabria) and SHARK (Pisa) kernels. One result of this work is the possibility of experimenting with the adaptation techniques that these advanced resource kernels provide. Some publications on HOLA-QoS can be found at [http://www.dit.upm.es/str](http://www.dit.upm.es/str).

**Resource availability prediction**

The resources typically used in-home entertainment applications (e.g., video/audio streaming) exhibit fluctuating availability. It is desirable to have mechanisms for indicating the available bandwidth during system runtime.

A comparative analysis of bandwidth estimation techniques for WiFi links has been carried out. In particular, the analyzed estimation techniques include several statistical and control-based algorithms. The analysis has identified the best suitable techniques taking into account the specific behavior of WiFi links. Work was carried out by UPC and TUKL. Analysis available at [http://www.upcnet.es/~pmc16/nde_06.pdf](http://www.upcnet.es/~pmc16/nde_06.pdf).

**Requirements for integrated-resource scheduling framework**

A workshop on “Requirements for Flexible Scheduling in Complex Embedded Systems” was held in Massy (Paris) in June 2006, with the objective of developing a set of requirements for building a flexible scheduling framework for applications demanding various types of tasks, constraints, and scheduling paradigms within the same system, and paying attention to the integration of multiple resources. The workshop was very successful and brought together 20 participants.

**Baseline for integrated-resource scheduling framework**

The FIRST (Flexible Integrated Real-Time Scheduling Technologies) IST project that finished in 2005 produced as its main result a contract-based scheduling framework, called FSF that was capable of scheduling multiple application components with various kinds of requirements for CPUs and, to a limited extent, for networks in distributed systems. FSF defines the basis of a new flexible scheduling strategy, and opens a new line of future research in design methodologies that help in translating application requirements into the best possible set of contracts to achieve the highest level of resource utilization and quality. This framework was selected as the baseline for the more ambitious framework that is being developed in this activity and that will take into account the integrated scheduling of multiple resources. The new framework will also include enhanced on-line schedulability analysis and derivation methods for server parameters, as well as the development of high-level transaction management middleware. It will be fully implemented on top of several networks and network protocols, to make it usable for distributed applications.

The FTT framework, in which a master manages the synchronous activities in a distributed system or cluster, was also extended to micro-segmented switched Ethernet-based distributed systems, having revealed potential to provide efficient support to the contract model, to dynamic QoS management and to integrated resource scheduling in distributed environments. The VTPE protocol (Virtual Token-Passing Ethernet), which supports event-triggered communication with real-time guarantees and high bandwidth utilization, was also extended with appropriate mechanisms to support isochronous traffic, more adequate for some applications, e.g. multimedia transmission. Moreover, the impact of flexible scheduling on dependability for distributed safety-critical applications was also assessed using FTT-CAN and
appropriate mechanisms were developed, namely master replication, fail-silence enforcement and a protocol to minimize inconsistencies during updates of the masters table.

Work was also carried out in a framework that facilitates the distribution of resource intensive services across a community of nodes, forming temporary coalitions for a cooperative execution. In this framework, the system must able to react to load variations, degrading its performance in a controlled fashion if needed. Server-based scheduling approaches were developed to provide isolation between different services and an efficient overload control, considering the resource requirements of the supported services.

**New theoretical developments**

The contract-based scheduling framework needs to be implemented using a specific scheduling strategy, and the most effective approach for this case is the server-based hierarchical scheduling in which an application or application component is scheduled over a protected bandwidth-preserving server (such as a periodic server, a sporadic server, or a constant bandwidth server) and individual threads in that component are scheduled by a higher-level scheduler that uses the bandwidth provided by the server. Theory was developed towards being able to analyze such scheduling schemes.

Work was done by the University of York together with Technische Universiteit Eindhoven (TU/e) on the analysis underpinning the use of CAN in real-time systems.

SSSA developed the following theoretical results: energy-aware scheduling algorithms for processors with dynamic voltage scaling and discrete frequency levels; a method for minimizing the deadline of periodic tasks with the objective of reducing delay and jitter; a general methodology for performing sensitivity analysis of fixed priority periodic systems with configurable periods and computation times, allowing the system designer to derive the feasibility region of a task set and compute the maximum parameter variations that keep the system feasible.

Work carried out at the University of Aveiro also exposed a couple of anomalies related to the definition of critical instant in hierarchical scheduling scopes found in communication systems that led to optimistic worst-case response time analysis in the past. Adequate methods were devised to cope with such anomalies.

The Polytechnic Institute of Porto provided new theoretical developments on: a new multiprocessor scheduling approach with a higher utilization bound and with few preemptions, able to trade the utilization bound for preemptions; new flexible admission control algorithms for IEEE 802.15.4 networks improving the bandwidth utilization compared to the explicit allocation used in the IEEE 802.15.4 protocol; a new server-based scheduling approach for handling isolation and overload control on distributed cooperative systems;

**Flexible architectures and communication protocols for networks used in distributed embedded real-time systems**

This research was partly done in collaboration between the following ARTIST 2 partners: University of Pavia, University of Catania (affiliated partner) and Malardalen University, Sweden. It consists of the following activities:

- Integration of networked subsystems in a resource constrained environment.
- Facilitating subsystem integration by decoupling priority and identifier in CAN messages.
- Interconnection of real-time networks in factory automation and in the automotive domain.
- Design issues and transmission protocols for wireless networks used in factory communication.
• Bluetooth to support real-time traffic in factory communication.

• Modelling of wireless real-time communications for land monitoring systems.

In addition, a joint work was done by Pavia, Pisa and Aveiro to support connectivity tracking in mobile ad-hoc wireless networks subject to real-time constraints.

2.3 Current Results

2.3.1 Technical Achievements

Application requirements (all)
We have expanded the collection of application requirements beyond the video streaming domain. Activities were carried out in particular together with the FRESCOR project, including a joint requirements workshop and a meeting with the industrial advisory board of FRESCOR on the topic. A previously endangered workshop between scientists and engineer in the domain of media processing has been revived and will be held Nov 15, 16 2007.

http://www.frescor.org/index.php?mact=News,cntnt01,detail.0&cntnt01articleid=8&cntnt01origid=51&cntnt01returnid=51

http://soller.eit.uni-kl.de/mediawiki/index.php/Requirements__Workshop

Integrated CPU scheduling and cache management (TUKL, NXP (formerly Philips))
NXP and TUKL have continued work on integrating real-time scheduling and cache management on multiprocessor platforms. We have carried out thorough experiments to study the impact of cache usage and scheduling on predictability on the actual platform. We developed a first algorithm with exhibits encouraging behavior.

A joint PhD student is carrying out the work. Publications are in preparation and have been submitted.

Adaptive resource management for networks (UC2m, Aveiro)
An architecture for flexible functional composition on networked real-time applications was developed by UC3M group and in collaboration with Aveiro. A joint PhD thesis has been produced between UC3M and Aveiro. Also, UC3M is developing an improved version of HOLA-QoS architecture is being ported to real-time Java platforms to test its suitability for medium side embedded platforms.

Architectural model of the flexible scheduling framework (All)
An architectural model of a flexible scheduling framework has been developed. The framework is capable of handling multiple concurrent activities with different criticality and timing in the same system, integrating the management of different kinds of resources such as processors, networks, memory, energy, and shared objects with time protection. This framework provides the ability to compose several applications or components into the system, and to flexibly schedule the available resources while guaranteeing hard real-time requirements. The framework is independent of the underlying implementation, and can run on different underlying scheduling strategies. It is based on establishing service contracts that represent the complex and flexible requirements of the applications, and which are managed by the underlying system to provide the required level of service. The framework provides

11/26
mechanisms to distribute the share of resources among those application parts that are capable of using them. The distribution is made using QoS parameters specified by the application. The framework also separates the management of resources from the actual scheduled elements, and this independence eases the dynamic allocation of resources, which is specially useful for integrating component-based design methods into the framework. An interface between this framework and the applications that run on top of it has been designed.

The work has been carried out by the members of the FRESCOR EU project, together with other members of the ART cluster, in particular UPM and Aveiro. Of course it is necessary to distinguish those resources that we can manage for scheduling purpose from those that we can just model, or that do not require specific scheduling.

Implementation of the framework (UC, UPVLC, CTU, SSSA, Aveiro)

The flexible scheduling framework described in the previous clause has been implemented for different execution platforms and networks. UPVLC and Cantabria have implemented the framework on RT-Linux and MaRTE OS. SSSA has implemented it on top of Linux and the AQUOSA resource management middleware. An adaptation layer has been created to ease the porting to different platforms. Aveiro has implemented the framework on the FTT-SE protocol. Cantabria and UPVLC have implemented it on top of Ethernet networks, with two different kinds of protocols: RT-EP and a time-division multiplexing protocol. CTU is implementing the protocol for wireless networks. York has implemented an on-line schedulability analysis module for the framework, together with an API that makes it possible to plug in other schedulability analysis techniques. The separate implementation on processors and networks requires integration that will be addressed in the following year. An adaptation layer for network protocols has been defined for this purpose.

New theoretical developments (York, SSSA)

A number of theoretical developments have occurred during this last year.

- For hierarchical scheduling, a resource sharing model has been defined and the scheduling equations appropriate for this model have been derived and verified. This is applicable to fixed priority scheduling and builds upon one of last year’s developments.
- For systems that allow online/dynamic admissions, efficient scheduling algorithms for fixed priority scheduling have been developed and verified during experimentation. These algorithms allow spare capacity to be allocated to applications in an effective and efficient way, thereby enhancing the adaptability of such applications.
- For hierarchical systems that are implemented on an EDF scheduler, necessary and sufficient analysis has been developed that mirrors the fixed priority approach, and thereby allow mixed EDF and fixed priority systems to be deployed.
- For the CAN communication protocol, a previously unknown error in the standard way industry verifies CAN-based systems has been identified and revisited analysis has been developed and published.
- Jitter-aware scheduling algorithms have been proposed and compared at SSSA. The main contribution in this area was to develop two algorithms, with different performance and complexity, to reduce finishing time jitter in control tasks. The proposed techniques exploit deadline reduction methods to limit the active execution window of jitter sensitive tasks. Schedulability analysis has been derived and performance has been evaluated with respect to other existing techniques. A more theoretical work has been carried out at SSSA to identify the space of feasible deadlines in a set of periodic real-time tasks.
Feedback control of server-based real-time systems (UPC)

Feedback control of computing systems is an emerging area that aims at applying control theory to computing systems in order to meet performance specifications regardless of varying applications demands or unpredictable resource availability. This approach has been applied to server-based real-time systems, in order to automatically control the partial utilization of each server.

Design of a quality of service manager (SSSA, TUKL, UPM, Cantabria)

HOLA-QoS is a framework for managing QoS and resources and it has been used in media processing which UPM and UC3m have developed jointly. It is implemented as a layered architecture, so that layers can be replaced, as far as the API is kept. The work that has been finished is to replace the lower levels of HOLA-QoS with two kernels with resource management facilities (also called resource kernels): MARTE (Cantabria) and SHARK (Pisa). The integration has been validated with some test applications. The final part of this work is the integration of the higher levels of HOLA-QoS that is currently being redesigned. One result of this work is the possibility of experimenting with the adaptation techniques that these advanced resource kernels provide. Some publications on HOLA-QoS can be found at http://www.dit.upm.es/str.


Architecture for dynamic service composition (UC3M, Aveiro)

The group at UC3M (affiliated to UPM) explored using the service-oriented paradigm to develop distributed real-time applications and, together with Aveiro, devised an architecture to support dynamic service composition. This kind of architecture can also be an alternative to provide some level of flexible resource management by supporting different profiles of each service, with different QoS and resource usage, and allowing a dynamic recomposition of the services/profiles involved in an application according to a predefined objective, e.g., maximize QoS of an application given the currently available resources, minimize the resources needed by a set of applications, etc. The architecture is based on a global entity, called the composer, which, together with a QoS manager, decides when and how to recompose a given application. The real-time coordination of the composition changes in the distributed system is carried out with the FTT-SE protocol. (see joint publications below).

Dual-rate switching control (UPC, Aveiro)

The group at UPC (affiliated to TUKL) and Aveiro developed an online adaptation strategy for feedback control loops that minimizes the respective resources used, both CPU and network. The strategy is based on using a pair of controller for each feedback loop, a nominal one and an adapted low bandwidth controller for periods of stationarity. The savings in resource usage are highly dependent on the dynamics of the application but can be explored by a flexible resource management framework (see joint publications below). Further joint work in a similar direction is being developed to carry out flexible resource management in the scope of feedback control applications themselves, with rate adaptation, to distribute bandwidth dynamically among several feedback control loops running on a single CPU (a joint publication has been submitted to a journal).
Feedback control of server-based real-time systems (UPC)

Feedback control of computing systems is an emerging area that aims at applying control theory to computing systems in order to meet performance specifications regardless of varying applications demands or unpredictable resource availability. This approach has been applied to server-based real-time systems, in order to automatically control the partial utilization of each server.

Overrun handling approaches for soft real-time periodic systems (U. of Catania)

In soft real-time environments featuring highly variable workload and resource requirements, tasks typically require probabilistic guarantees on meeting deadlines. When building the schedule for soft real-time applications it is therefore desirable avoiding worst-case assumptions on task execution times, in order to fully exploit the processor and increase system utilization as much as possible. To this end, “typical” task execution times (obtained either through previous measurements or from the knowledge of the execution time distributions and averages) can be used in the schedulability analysis instead of worst-case ones, which are very unlikely to occur. However, according to this approach, tasks can overrun, that is, the actual execution time required by some task instances can exceed the one assumed in the schedule. Due to overruns, it may happen that the system becomes overloaded, thus causing an unbounded number of deadline misses. As a result, a suitable overrun handling mechanism is needed in order not to jeopardize the system performance.

Traditional overrun handling approaches for real-time systems enforce some isolation property at the job or task level. This activity aimed at showing that by “relaxing” task isolation, it is possible to efficiently deal with overruns in soft real-time systems with highly variable task execution times, and proposed Randomized Dropping (RD), a novel overrun handling mechanism. RD is able to bound task overruns in a probabilistic manner, thus providing “soft” task isolation. It has been shown how to combine RD with priority-driven and rate-based scheduling algorithms, and how to analyze the resulting system [CT1]. Performance evaluation and comparison between simulation and analytical results were discussed.

Stochastic response time analysis of hybrid task sets in priority-driven soft real-time systems (U. of Catania)

Starting from the approach in [CT2], this activity led to the development of a more generic task model than the traditional one, which assumes each task being characterized by a fixed activation period and a Worst-Case Execution Time. According to the new model, a task is characterised by an Arrival Profile (AP) and an Execution Time Profile (ETP), both given by random variables with known distributions. To cope with the unbounded interference introduced by aperiodic tasks in the system, this activity targeted the encapsulation of sporadic and aperiodic tasks using server-based techniques. The calculus for obtaining the ETP of servers used to handle hybrid task sets was derived, and the results obtained are then applied to a previous stochastic analysis framework, making it applicable to a more general domain [CT3]. Then, the calculation of stochastic Response Time Profiles (RTPs) of tasks hierarchically scheduled using server-based techniques in a stochastic analysis framework was addressed. Results are given in [CT4].

Server-based approaches for flexible scheduling in distributed embedded real-time systems (U. of Catania)

Previous work from the University of Catania proposing the adoption of the Total Bandwidth Server on the traditional Master/Slave (M/S) Bluetooth operating mode to support flexible scheduling of hybrid (periodic+aperiodic) messages on Bluetooth (BT) networks was extended
towards a more flexible and efficient approach. Although successful in providing support to soft real-time traffic, the original approach based on TBS suffered from some limitations, in terms of scan cycle duration and waste of bandwidth, due to the M/S mode. This year research at the University of Catania therefore focused on a novel approach, called M/S\(_{\text{EDF+CBS}}\)+S/S\(_{\text{EDF}}\), which combines deadline-aware scheduling with the Slave/Slave operating mode, thus overcoming the limitations introduced by the M/S mode. The approach handles both periodic and aperiodic traffic in a flexible and integrated way combining Earliest Deadline First (EDF) with the Constant Bandwidth Server (CBS) [CT6]. The new approach proved to be capable of increasing the amount of bandwidth available for both periodic and aperiodic transmissions, while keeping periodic traffic isolated from aperiodic traffic. The use of real-time scheduling techniques, i.e. EDF and CBS, allows periodic traffic deadlines to be met in the various scenarios examined, while in the same conditions the standard BT approach experiences significant periodic packet loss and deadline miss values [CT7].

The activity on the problem of integrating networked subsystems in resource constrained environments, carried out in collaboration with Malardalen University, Sweden, continued investigating in a joint publication the choice of the Server-CAN scheduler to use, examining the difference of the temporal guarantees provided by the usage of different Server-CAN schedulers [CT5]. Moreover, it was addressed how to configure the N-Server parameters for a subsystem and the integration process of subsystems using the Server-CAN concept was illustrated using a simple example of how to configure the Server-CAN framework for a “real” subsystem. The cost of using the Server-CAN framework, in terms of temporal degradation and overhead, was also assessed.

**A middleware to combine energy-awareness and real-time support for distributed embedded systems used in monitoring applications (U. of Catania)**

The main focus of this activity was a middleware, called an Aggregation layer, to be introduced between the MAC and Routing layers. Such a middleware, in combination with suitable routing techniques, is able to support soft real-time traffic while reducing the energy consumption of the wireless nodes, which have to work for long periods without the possibility of replacing their batteries. This layer mainly deals with reducing the amount of energy dissipated, while the Routing layer is entrusted with achieving the desired delivery speed, to support the transmission of soft real-time traffic. The proposed middleware has been implemented in the RTPAW framework [CT8].

**Title of the Technical Achievement (partners)**

Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms. Describe the Technical Achievement here, in clear and concise terms.

http://www.weblink.eu/weblink/weblink/weblink/webpage/

**2.3.2 Individual Publications Resulting from these Achievements**

**TUKL**


C. Otero-Perez, J v. EIndhoven, G. Fohler: "Execution-aware cache reservations for media applications", submitted

Radu Dobrin, Gerhard Fohler: Handling Non-periodic Events Together with Complex Constrained Tasks in Distributed Real-Time Systems, 16th International Conference on Control Systems and Computer Science, Bucharest, Romania, May, 2007

Universidad de Cantabria


Porto


Universidade de Aveiro


UC3M


UPC


York


**University of Catania**


2.3.3 Interaction and Building Excellence between Partners

TUKL and NXP (formerly as partner Philips) work together on integrated resource management via a joint PhD student. Resulting publications have been submitted.

In the scope of the collaboration between Aveiro and Cantabria, this year a PhD student from Aveiro (Ricardo Marau) spent two weeks in Cantabria working on the integration of the FRESCOR framework and the FTT-SE protocol. Later, two PhD students from Cantabria (Daniel Sangorrín and Hector Perez) attended a meeting in Aveiro to discuss further details of this work.

The collaboration between Aveiro and UC3M for dynamic service composition is carried out in the scope of the co-supervision of one PhD thesis (Iria Estevez). During this year there were a few short meetings, in Aveiro and in Madrid, to coordinate specific aspect of the work.

The collaboration between Aveiro and UPC for rate adaptation in feedback control loops was, during this year, mainly remote, following several mutual visits in the year before.

In the scope of another collaboration, between Aveiro and UPVLC (group from Alcoy), a professor from UPVLC (Javier Silvestre) visited Aveiro for one month, to carry out joint work in dynamic bandwidth management in video transmission within industrial environments.

2.3.4 Joint Publications Resulting from these Achievements

C. Otero-Perez, J v. Elindhoven, G. Fohler: “Execution-aware cache reservations for media applications”, submitted


2.3.5 Keynotes, Workshops, Tutorials

Keynote: Predictable response times in event-driven real-time systems.
Automotive 2006 - Security and Reliability in Automotive Systems
Stuttgart, October, 2006
M. González Harbour

Keynote: QoS-Based Resource Management
University of Thessaloniki
Thessaloniki, Greece– May 4th, 2007

The keynote talk was given in the context of a periodic departamental workshop open to researchers and university students.

The topic of the talk was resource management in embedded applications that are tolerant to some degree of flexibility and, therefore, admit QoS-based operation. The talk introduced the problems of these systems and the solutions ranging from the older contract-based approach, to centralised resource management, to distributed resource management, and higher level protocols for managing application execution.
Marisol Garcia Vals

Workshop NeRES 2007 – Networks for Reconfigurable Embedded Systems
Aveiro, Portugal – April 2007

This workshop was organized in the scope of the activity Dynamic and Pervasive Networks (appears in more detail in that report) and it targeted discussing the network requirements to support reconfigurability in distributed embedded systems, as well as the adequacy of current protocols and middlewares for that purpose. Several presentations were delivered focusing on middleware layers to support flexible resource management, which was one kind of reconfiguration that was addressed.
URL: http://www.artist-embedded.org/artist/Motivation-and-Goal.html

First Italian Workshop on Real-Time Embedded Systems
RETIS Lab, Scuola Superiore Sant'Anna, Pisa, July 2, 2007

Organizers: Giorgio Buttazzo, Giuseppe Lipari, Lucia Lo Bello
Objectives: Build an Italian community on real-time embedded systems and favor interactions between Italian industry and academic researchers.
Topics: real-time scheduling, operating systems, sensor networks, design methodologies
Results: The workshop attracted 16 universities and 12 industries working in the field, 28 short presentations were given and participants had time to meet, know each other and exchange information on their research interests. A second workshop is planned for next year.
URL: http://feanor.sssup.it/wirtes07/

WCET 2007: Worst Case Execution Time Analysis
RETIS Lab, Scuola Superiore Sant'Anna, Pisa, July 3, 2007

Organizers: Christine Rochange, TRACES group, IRIT, Toulouse, France
Objectives: Bring together people from academia, tool vendors and users in industry that are interested in all aspects of timing analysis for real-time systems.
Topics: Timing analysis, calculation methods for WCET, testing methods for WCET analysis, tools for timing analysis, compiler optimizations for worst-case paths.
Results: The workshop attracted 32 participants from different European countries and
technical papers have been published in proceedings.
URL: http://www.irit.fr/wcet2007/

RTN 2007: Real-Time Networks
RETIS Lab, Scuola Superiore Sant’Anna, Pisa, July 3, 2007
Organizers: Ye-Qiong Song, LORIA, Nancy, France
Objectives: RTN focuses on the current technological challenges of developing communication infrastructures that are real-time, reliable, pervasive and interoperable.
Topics: Distributed systems, communication protocols, wireless sensor networks, mobile ad-hoc networks.
Results: The workshop attracted 30 participants from different European countries and technical papers have been published in proceedings.
URL: http://rtn2007.loria.fr/

OSPERT 2007: Operating Systems Platforms for Embedded Real-Time Applications
RETIS Lab, Scuola Superiore Sant’Anna, Pisa, July 3, 2007
Organizers: Scott A. Brandt, University of California, Santa Cruz, CA, USA and Kevin Elphinstone, University of New South Wales, Kensington, NSW, Australia.
Objectives: This workshop is intended as a forum for researchers and practitioners of RTOS to discuss the recent advances in RTOS technology and the challenges that lie ahead.
Topics: Support for component based development; Scalability, from very small scale embedded systems to full-fledged OSes; Real-Time on Linux; Interaction with reconfigurable hardware; Support for embedded multi-processor architectures; Security and fault tolerance for embedded real-time systems; Power-aware operating systems..
Results: The workshop attracted 18 participants from different European countries and technical papers have been published in proceedings.
URL: http://www.cs.ucsc.edu/~sbrandt/OSPERT.html

RETIS Lab, Scuola Superiore Sant’Anna, Pisa
July 3, 2007
URL: http://feanor.sssup.it/ecrts07/tutorial.shtml
3. Future Work and Evolution

3.1 Problem to be Tackled over the next 12 months (Sept 2007 – Aug 2008)

In the next 12 months we will conclude the analysis of application requirements of adaptive applications in alignment with current and future EU IST projects, in which partners are involved.

We will develop algorithms for the integration of CPU scheduling and cache management and analyse their effectiveness on actual boards. We will also continue tackling the problem of dynamic QoS configuration, considering multiple dependencies between services’ resource needs. These dependencies will be considered both between services in the same node, and between distributed services. The problem will be tackled at two levels: (i) anytime algorithms for QoS adaptation considering dependencies; and (ii) scheduling of services with multiple dependencies in QoS attributes.

After the work in the current year in which an architectural model of a flexible scheduling framework has been developed and its first implementation modules have been developed, work in the following year concerning the implementation of the framework will concentrate on the finalization of the implementation for additional resources such as memory and energy, the integration of the developed modules, and the evaluation of the approach.

A second goal in this period is to complete and evaluate the adaptive QoS manager that understands about the quality concepts of the application and translates them to the scheduling domain information that the underlying system can understand and implement, generating the contracts in an automatic or semiautomatic way.

A third goal of this period is the development of better schedulability analysis and spare capacity distribution techniques that can be integrated into the framework and take into account an integrated view of all the resources. In particular, there are strong interactions between energy and the performance of other resources such as processors or networks, and these interactions need to be correctly integrated into the analysis techniques.

There is a strong relation between the flexible scheduling activity in ARTIST2 and the FRESCOR EU project that started in June 2006. Many of the partners in the activity are also partners of the project, and the goals are somehow shared. Indeed ARTIST2 has contributed to creating the necessary critical mass to set up a project like FRESCOR. Given the shared goals it is necessary to define clearly the role of each project in the overall activity. We see the following role of the flexible scheduling activity in ARTIST2:

- Bring together a wide body of expertise from the whole ARTIST2 community to come out with a broad and ambitious set of requirements for the flexible scheduling framework
- Use that expertise to establish and later evaluate the usefulness and applicability of a contract-based specification for the framework, that allows applications to express their requirements, and the scheduler to satisfy them in the most optimal way.
- Use the ARTIST2 community as one of the ways of disseminating the results of the FRESCOR project

With these roles in mind we think that this is an ideal situation in which the FRESCOR project can benefit from the ARTIST2 NoE expertise, and the network can benefit from being able to influence the project, and from being able to exploit its results.
3.2 Current and Future Milestones

1. achieved Year1: Identify case studies, perform preliminary assessment. Preliminary work on the integration of diverse scheduling schemes. Achieved for the case of CPU and energy, and CPU and networks.

2. achieved Year2: Define a set of case studies and from them deduce the QoS requirements and their mapping into operational parameters of the computing and communication infrastructures.
Temporal requirements for video streaming from stream demands and the temporal impact of devices have been identified. Adaptive methods for resource management and QoS middleware have been developed. Relations have been identified, e.g., on wireless networks.
Demonstrate the combination of specific scheduling schemes applied both to CPU as well as to the network, to suit diverse application requirements in the same system. This has been achieved both in theoretical developments for server-based hierarchical scheduling, and in practice through the FIRST scheduling framework.

3. achieved Year3: Expend the requirements into a meaningful set of requirements of adaptive application domains that allow the creation of global mechanisms for resource management. Define a set of requirements and an architectural model for the framework for flexible scheduling that integrates multiple resources, including CPUs and networks, multiple processors, shared resources with time protection, memory protection, and energy/power-aware scheduling.

4. Year4: Integration of the application adaptation processes into a general QoS resource management structure. Provide a framework that allows the seamless integration of flexible scheduling schemes for integrated resources, allowing the choice of appropriate scheduling methods for individual activities in the different resources.

3.3 Indicators for Integration

Expected results and visible impact are: influence on integrating the operating system design, network, and middleware, joint results on scheduling algorithms, analysis tools, and kernel and network support for the integration and coexistence of diverse system-wide scheduling schemes.

3.4 Main Funding

The FRESCOR EU IST project, in which the following ARTIST2 partners are involved:
University of Cantabria, University of York, Scuola Superiore Santa Anna, Technical University of Kaiserslautern, Technical University of Valencia and the Czech Technical University in Prague
– www.frescor.org

The ACTORS EU IST project, starting early 2008, in which the following ARTIST2 partners are involved: Scuola Superiore Santa Anna, Technical University of Kaiserslautern, Lund University
THREAD Spanish project, in which the following ARTIST2 partners are involved: Technical University of Madrid, University of Cantabria, Technical University of Valencia.

The CooperatES Project, started July 2007, a Portuguese funded FCT project involving the Polytechnic Institute of Porto.
4. Internal Reviewers for this Deliverable

Giorgio Buttazzo, SSSA

Michael Gonzalez, UC