

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 10-04-2023	2. REPORT TYPE Final Report	3. DATES COVERED (From - To) 15-Jun-2020 - 14-Dec-2022
---	--------------------------------	---

4. TITLE AND SUBTITLE Final Report: A Comprehensive Framework for Soft Real-Time Multicore Applications Motivated by Safety-Critical DoD Systems	5a. CONTRACT NUMBER W911NF-20-1-0237
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102

6. AUTHORS	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of North Carolina - Chapel Hill 104 Airport Drive, CB 1350 Suite 2200 Chapel Hill, NC 27599 -1350	8. PERFORMING ORGANIZATION REPORT NUMBER
---	--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSOR/MONITOR'S ACRONYM(S) ARO
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 76214-NC.23

12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.
--

13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON James Anderson
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 919-590-6057

RPPR Final Report

as of 10-Apr-2023

Agency Code: 21XD

Proposal Number: 76214NC

Agreement Number: W911NF-20-1-0237

INVESTIGATOR(S):

Name: Ph.D. James Anderson

Email: anderson@cs.unc.edu

Phone Number: 9195906057

Principal: Y

Organization: **University of North Carolina - Chapel Hill**

Address: 104 Airport Drive, CB 1350, Chapel Hill, NC 275991350

Country: USA

DUNS Number: 608195277

EIN: 566001393

Report Date: 14-Mar-2023

Date Received: 10-Apr-2023

Final Report for Period Beginning 15-Jun-2020 and Ending 14-Dec-2022

Title: A Comprehensive Framework for Soft Real-Time Multicore Applications Motivated by Safety-Critical DoD Systems

Begin Performance Period: 15-Jun-2020

End Performance Period: 14-Dec-2022

Report Term: 0-Other

Submitted By: Ph.D. James Anderson

Email: anderson@cs.unc.edu

Phone: (919) 590-6057

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 0

STEM Participants:

Major Goals: Most safety-critical DoD applications have significant soft real-time (SRT) components, i.e., components for which deadline misses are sometimes acceptable. Furthermore, the development of methods for supporting SRT workloads has been identified as a pressing need for the safe use of multicore platforms in safety-critical DoD application domains. This project is directed at addressing this need.

In work on real-time systems, "soft real-time" can be defined in various ways. Over the last 20 years, the PI's research group has been the driving force behind work on a notion of SRT correctness that focuses on the metric of deadline tardiness. In particular, an SRT task system is deemed to be correct if every task in the system provably has bounded tardiness, i.e., if any deadline is missed, then the extent of miss is bounded. The PI's research on this notion of SRT correctness has been very impactful. For example, as discussed in depth below, this research led to an SRT scheduler called SCHED_DEADLINE being accepted into the Linux kernel.

The overarching objective of this project is to extend prior tardiness-centric work on SRT multiprocessor systems to enable the use of additional SRT metrics and to broaden the implementation contexts where all of these metrics can be applied. The work towards this objective has focused on four goals.

Goal 1: Devise tardiness analysis for multiprocessor systems with asymmetries. Such asymmetries are becoming increasingly common in safety-critical application domains, due to both hardware heterogeneity and the manner in which hardware (heterogeneous or not) is shared by disparate applications. Regardless of the source, asymmetries can cause real-time multicore resource-allocation problems to be significantly harder to solve.

Goal 2: Devise techniques for supporting weakly hard constraints. Such constraints require a given task to meet a specified fraction of its deadlines. We are especially interested in enabling such constraints in contexts where task execution times are determined on a non-worst-case basis, in which case probabilistic guarantees may arise. Worst-case task provisionings can cause an unacceptable amount of processing capacity to be wasted, especially for workloads that are SRT.

Goal 3: Enable tardiness reductions through intra-task parallelism. In many uses cases, it is permissible for successive invocations of the same task to execute in parallel. For example, in many computer-vision applications (as used in autonomous vehicles), successive frames of video can be processed in parallel, with any out-of-order processing corrected via buffering. Allowing greater parallelism in this manner can greatly reduce tardiness bounds.

RPPR Final Report as of 10-Apr-2023

Goal 4: Conduct experimental evaluations. We will integrate the research under the prior goals to obtain a comprehensive SRT scheduling and analysis framework that allows different SRT metrics to be applied. This framework will expose various tradeoffs among metrics that we will evaluate by conducting overhead-aware schedulability studies and by experimenting with synthetic and case-study workloads. (In a schedulability study, the timing constraints of many million randomly generated task systems is checked, with overheads that impact scheduling considered. These studies can elucidate important resource-allocation tradeoffs.)

Accomplishments: With the dismantling of ARO's Computational Architectures and Visualization Program, our funding we unexpectedly reduced by a significant amount. In fact, the funding on this project was exhausted as of October 2021. As a result, this project effectively ceased at that time. Therefore, we have no new accomplishments to report beyond our last report.

Training Opportunities: As noted above, this project effectively ceased a year and a half ago.

Results Dissemination: As noted above, this project effectively ceased a year and a half ago.

Honors and Awards: PI Anderson was elected as a Fellow of the Asia-Pacific Artificial Intelligence Association (AAIA).

Protocol Activity Status:

Technology Transfer: As noted above, this project effectively ceased a year and a half ago.

ARTICLES:

Publication Type: Journal Article

Peer Reviewed: Y

Publication Status: 1-Published

Journal: Real-Time Systems

Publication Identifier Type: DOI

Publication Identifier: <https://doi.org/10.1007/s11241-022-09381-y>

Volume: 58

Issue:

First Page #: 105

Date Submitted: 7/11/22 12:00AM

Date Published: 3/23/22 7:57PM

Publication Location:

Article Title: Exploring AMD GPU Scheduling Details by Experimenting With 'Worst Practices'

Authors: Nathan Otterness, James H. Anderson

Keywords: real-time systems, graphics processing units, parallel computing

Abstract: We provide an end-to-end description of how compute work is scheduled on AMD GPUs. In doing so, we start with a concrete demonstration of how incorrect management triggers extreme worst-case behavior in shared AMD GPUs. Subsequently, we explain the internal scheduling rules for AMD GPUs, how they led to the 'worst practices,' and how to correctly manage some of the most performance-critical factors in AMD GPU sharing.

Distribution Statement: 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support: Y

RPPR Final Report

as of 10-Apr-2023

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 1-Published
Journal: Journal of Systems Architecture
Publication Identifier Type: DOI **Publication Identifier:** <https://doi.org/10.1016/j.sysarc.2021.102292>
Volume: 120 **Issue:** **First Page #:** 102292
Date Submitted: 3/7/23 12:00AM **Date Published:** 2/6/22 2:10AM
Publication Location:
Article Title: The Price of Schedulability in Cyclic Workloads: The History-vs.-Response-Time-vs.-Accuracy Trade-Off
Authors: Tanya Amert, Ming Yang, Sergey Voronov, Saujas Nandi, Thanh Vu, James H. Anderson, F. Donelson
Keywords: autonomous driving, computer vision, cyberphysical systems, multi-object tracking, real-time systems
Abstract: Real-time processing graphs can have cycles due to back edges that provide history information. If immediate back history is required, then such a cycle must execute sequentially. Due to this requirement, any graph that contains a cycle with utilization exceeding 1.0 is unschedulable, i.e., bounded graph response times cannot be guaranteed. This dilemma can be obviated by allowing older back history, resulting in a history-vs.-response-time-vs.-accuracy trade-off. This paper explores this trade-off.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: Y

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 41st IEEE Real-Time Systems Symposium
Date Received: 04-Jun-2021 **Conference Date:** 01-Dec-2020 **Date Published:** 01-Dec-2020
Conference Location: Online due to COVID.
Paper Title: Towards Practical Multiprocessor EDF with Affinities
Authors: Stephen Tang, James H. Anderson
Acknowledged Federal Support: Y

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2021 Design, Automation Test in Europe Conference Exhibition (DATE)
Date Received: 04-Jun-2021 **Conference Date:** 01-Feb-2021 **Date Published:** 01-Feb-2021
Conference Location: Online due to COVID.
Paper Title: Timing-Predictable Vision Processing for Autonomous Systems
Authors: Tanya Amert, Michael Balszun, Martin Geier, F. Donelson Smitm, James H. Anderson, Samarjit Chakral
Acknowledged Federal Support: Y

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2021 Design, Automation Test in Europe Conference Exhibition (DATE)
Date Received: 04-Jun-2021 **Conference Date:** 01-Feb-2021 **Date Published:** 01-Feb-2021
Conference Location: Online due to COVID.
Paper Title: Perception Computing-Aware Controller Synthesis for Autonomous Systems
Authors: Debayan Roy, Clara Hobbs, James H. Anderson, Marco Caccamo, Samarjit Chakraborty
Acknowledged Federal Support: Y

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 2021 Design, Automation Test in Europe Conference Exhibition (DATE)
Date Received: 04-Jun-2021 **Conference Date:** 01-Feb-2021 **Date Published:** 01-Feb-2021
Conference Location: Online due to COVID.
Paper Title: Timing Debugging for Cyber-Physical Systems
Authors: Debayan Roy, Clara Hobbs, James H. Anderson, Marco Caccamo, Samarjit Chakraborty
Acknowledged Federal Support: Y

RPPR Final Report

as of 10-Apr-2023

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 29th International Conference on Real-Time Networks and Systems
Date Received: 04-Jun-2021 Conference Date: 07-Apr-2021 Date Published: 07-Apr-2021
Conference Location: Online due to COVID.
Paper Title: Exploring AMD GPU Scheduling Details by Experimenting With 'Worst Practices'
Authors: Nathan Otterness, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 29th International Conference on Real-Time Networks and Systems
Date Received: 04-Jun-2021 Conference Date: 07-Apr-2021 Date Published: 07-Apr-2021
Conference Location: Online due to COVID.
Paper Title: On the Defectiveness of SCHED_DEADLINE w.r.t. Tardiness and Affinities, and a Partial Fix
Authors: Stephen Tang, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 27th IEEE Real-Time and Embedded Technology and Applications Symposium
Date Received: 04-Jun-2021 Conference Date: 18-May-2021 Date Published: 18-May-2021
Conference Location: Online due to COVID.
Paper Title: Simultaneous Multithreading in Mixed-Criticality Real-Time Systems
Authors: Joshua Bakita, Shareef Ahmed, Sims Hill Osborne, Stephen Tang, Jingyuan Chen, F. Donelson Smith,
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 24th IEEE International Symposium on Real-Time Distributed Computing
Date Received: 04-Jun-2021 Conference Date: 01-Jun-2021 Date Published: 01-Jun-2021
Conference Location: Online due to COVID.
Paper Title: UPiD^{ART}: Detecting Improper GPU Usage in Real-Time Applications
Authors: Tanya Amert, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 33rd Euromicro Conference on Real-Time Systems
Date Received: 30-Aug-2021 Conference Date: 05-Jul-2021 Date Published: 05-Jul-2021
Conference Location: Online due to COVID.
Paper Title: Tight Tardiness Bounds for Pseudo-Harmonic Tasks under Global-EDF-Like Schedulers
Authors: Shareef Ahmed, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 33rd Euromicro Conference on Real-Time Systems
Date Received: 04-Jun-2021 Conference Date: 05-Jul-2021 Date Published: 05-Jul-2021
Conference Location: Online due to COVID.
Paper Title: Light Reading: Optimizing Reader/Writer Locking for Read-Dominant Real-Time Workloads
Authors: Catherine Nemitz, Shai Caspin, James H. Anderson, Bryan Ward
Acknowledged Federal Support: **Y**

RPPR Final Report

as of 10-Apr-2023

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 23rd IEEE International Symposium on Real-Time Distributed Computing
Date Received: 04-Jun-2021 Conference Date: 19-May-2020 Date Published: 19-May-2020
Conference Location: Online due to COVID.
Paper Title: The Price of Schedulability in Multi-Object Tracking: The History-vs.-Accuracy Trade-Off
Authors: Tanya Amert, Ming Yang, Saujas Nandi, Thanh Vu, James H. Anderson, F. Donelson Smith
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 26th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications
Date Received: 04-Jun-2021 Conference Date: 19-Aug-2020 Date Published: 19-Aug-2020
Conference Location: Online due to COVID.
Paper Title: Exploiting Simultaneous Multithreading in Priority-Driven Hard Real-Time Systems
Authors: Sims Osborne, Shareef Ahmed, Saujas Nandi, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 42nd IEEE Real-Time Systems Symposium
Date Received: 11-Jul-2022 Conference Date: 07-Dec-2021 Date Published: 07-Dec-2021
Conference Location: Dortmund, Germany.
Paper Title: Retry-Free Software Transactional Memory for Real-Time Systems
Authors: Claire Nord, Shai Caspin, Catherine Nemitz, Howard Shrobe, Hamed Okhravi, James Anderson, Nathar
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 42nd IEEE Real-Time Systems Symposium
Date Received: 11-Jul-2022 Conference Date: 07-Dec-2021 Date Published: 07-Dec-2021
Conference Location: Dortmund, Germany.
Paper Title: Extending EDF for Soft Real-Time Scheduling on Unrelated Multiprocessors
Authors: Stephen Tang, Sergey Voronov, James Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 42nd IEEE Real-Time Systems Symposium
Date Received: 11-Jul-2022 Conference Date: 07-Dec-2021 Date Published: 07-Dec-2021
Conference Location: Dortmund, Germany.
Paper Title: TimeWall: Enabling Time Partitioning for Real-Time Multicore+Accelerator Platforms
Authors: Tanya Amert, Zelin Tong, Sergey Voronov, Joshua Bakita, F. Donelson Smith, and James Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 42nd IEEE Real-Time Systems Symposium
Date Received: 11-Jul-2022 Conference Date: 07-Dec-2021 Date Published: 07-Dec-2021
Conference Location: Dortmund, Germany.
Paper Title: AI Meets Real-Time: Addressing Real-World Complexities in Graph Response-Time Analysis
Authors: Sergey Voronov, Stephen Tang, Tanya Amert, James Anderson
Acknowledged Federal Support: **Y**

RPPR Final Report

as of 10-Apr-2023

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 28th IEEE Real-Time and Embedded Technology and Applications Symposium
Date Received: 11-Jul-2022 Conference Date: 04-May-2022 Date Published: 04-May-2022
Conference Location: Online due to COVID.
Paper Title: Minimizing DAG Utilization by Exploiting SMT
Authors: Sims Osborne, Joshua Bakita, Jingyuan Chen, Tyler Yandrofski, and James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: 34th Euromicro Conference on Real-Time Systems
Date Received: 11-Jul-2022 Conference Date: 06-Jul-2022 Date Published: 06-Jul-2016
Conference Location: Modena, Italy.
Paper Title: Overrun-Resilient Multiprocessor Real-Time Locking
Authors: Zelin Tong, Shareef Ahmed, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 3-Accepted
Conference Name: 43rd IEEE Real-Time Systems Symposium
Date Received: 29-Aug-2022 Conference Date: 05-Dec-2022 Date Published: 05-Dec-2022
Conference Location: Houston, Texas.
Paper Title: Exact Response-Time Bounds of Periodic DAG Tasks under Server-Based Global Scheduling
Authors: Shareef Ahmed, James H. Anderson
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 0-Other
Conference Name: 43rd IEEE Real-Time Systems Symposium
Date Received: 29-Aug-2022 Conference Date: 05-Dec-2022 Date Published: 05-Dec-2022
Conference Location: Houston, Texas.
Paper Title: Enabling GPU Memory Oversubscription via Transparent Paging to an NVMe SSD
Authors: Joshua Bakita, James H. Anderson
Acknowledged Federal Support: **Y**

Partners

I certify that the information in the report is complete and accurate:

Signature: James H. Anderson

Signature Date: 4/10/23 3:37PM

The other parts of this report contain all needed information.