
多核实时系统时间分析的文献综述

吴伟

摘要：本文综述了时间分析技术在多核实时系统和车载 TSN 网络的应用现状与问题。在多核实时系统中，穷举最坏情况下的任务释放模式并不适用于复杂实时系统，因此计算最坏响应时间上界的近似值成为替代方法，但仍存在精度不够的问题。而时间分析技术应用在车载 TSN 时，主要考虑对 AVB 流量的最坏端到端延迟进行可调度性判断，但现有研究未充分考虑 TT 流量对 AVB 的影响，还需要进一步优化延迟计算。并且现有的多核 ECU 的时间分析仅针对抽象的多核系统进行分析，缺乏针对性，比如针对车载 TSN 的特有分析技术。

关键词：多核处理器；最坏响应时间；车载 TSN；多核 ECU

0 引言

随着信息技术的发展，人们对现代汽车的需求不仅包括驾驶功能的精准化、智能化，还包括对信息娱乐、驾驶员辅助这些多功能应用的追求，以获得更好的舒适性。一般的驾驶功能包括实时路况检测、车身控制、三维导航等，同时利用安装在汽车上的传感器随时感应周围的环境，进行系统运算与分析，预先让驾驶者察觉可能发生的危险，高级驾驶辅助系统（Advanced Driving Assistant System, ADAS）也应运而生。这些应用不仅对车载网络的带宽有很高要求，而且还有严苛的延迟限制。另外，为了支持这些多功能应用，在一些高级汽车中，需要连接到网络的电子控制单元（Electric Control Unit, ECU）有些已经超过 100 个，极大的增加了汽车网络的复杂性。由于诸如成本，能量效率和热性能的各种原因，车载系统中使用多核 ECU 的趋势正在增长，那么这必然会带来多核 ECU 的任务调度问题。在调度问题中，我们在确定了任务的优先级后，需要分析任务的最坏响应时间，以确保其能在相对截止期内完成，达到安全设计标准，从汽车开发的设计阶段就减小安全隐患发生的可能性。

1 时间分析的基本原理

多核实时系统是需要对外部产生的输入在特定的时间内做出响应的信息处理系统。故这些系统的正确性不但取决于其逻辑结果，还取决于产生结果的时间，即系统必须在规定的时间范围内对外部物理过程的变化做出正确响应。如果系统的时间行为不能满足要求，则可能导致灾难性后果。例如，当人们在行驶汽车的过程中，因遇到紧急状况需要停车时，刹车系统若不能根据驾驶员的操作在规定时间内做出对应的响应，那么可能会导致安全事故的发生甚至威胁

到行车人员的人身安全。所以，对系统进行时间分析与预测是必要的，并且涉及到两方面的问题：(1) 系统自身是如何运行的；(2) 如何对系统的运行情况进行预测与分析[1]。

针对第一个问题系统自身是如何运行的，即如何安排系统中各个任务在何时执行。系统需要决定在何种情况下运行哪个任务，以满足每个任务的时间特性要求，这通常与调度算法(Scheduling Algorithm)有关，并且调度算法的性能好坏取决于其能够多大程度地满足这一要求。

针对第二个问题如何对系统的运行情况进行预测与分析，即判定系统在该调度算法下是否满足所有任务的时间特性要求，这种分析也成为可调度性分析(Schedulability Analysis)。判定系统可调度性的方法主要有以下两类：(1)基于响应时间分析(Response Time Analysis,RTA)的判定；(2)基于资源利用率界限的判定。基于响应时间分析的判定，是通过计算一个任务的最坏响应时间(Worst-Case Response Time,WCRT)并与其相对截止期相比较从而判断该任务是否可调度。WCRT 是指在该任务释放的所有实例中，响应时间的最大值，而一个实例的响应时间是指从任务实例被释放到执行完成的时间间隔。将此过程应用于任务集中的每一个任务，便可以判断整个系统的可调度性。

2 多核处理器任务的 WCRT 分析

RTA 技术也是分析多处理器平台上实时任务集可调度性的重要基础工具。由于多核处理器的并行性，我们很难确定全局调度中造成 WCRT 的具体时刻，即无法精确计算当前分析任务何时受到最大干扰。为了得到精确的 WCRT 进行预测与分析，许多研究者提出蛮力逼近的方法，即通过列举表示任务释放模式组合的可能状态来详尽地分析全局调度。文献[2]通过对非常大的状态空间进行详尽搜索来进行可调度性的判定，但考虑到该算法的时间和空间复杂性，仅适用于整数周期非常小的任务集。文献[3]定义并证明了[2]中提出的有限自动机中的模拟关系的正确性，同时改进得到适用于实际大小的系统算法，提高了可调度性测试的性能。为了进一步缓解状态空间爆炸的问题，文献[4]提出了关于模型符号状态的偏序关系，将其他人模拟的状态安全地从状态空间中删除，而不会影响可调度性分析。尽管这种穷举办法通过在有限状态机中列举所有可能的到达时刻和执行间隔找到了全局调度下相对精确的 WCRT，但是这种穷举方法并不适用于真实的复杂实时系统。

因此，有研究提出了基于近似计算对任务集进行时间分析的技术。文献[5]首先利用问题窗口分析方法，提出该技术的通用计算方法，即首先计算每个任务对当前任务的干扰上界，然后对其求和得到整个任务集对当前任务的总干扰上界。文献[6]提出由高优先级任务的响应时间来约束当前任务的响应时间，并

通过迭代计算得到每个任务的 WCRT 上界。然而高优先级任务的响应时间还包括其释放到真正执行的时间，这部分时间并没有干扰到当前任务的执行的，因此当前任务受到的干扰还能细化为高优先级任务的执行时间，我们把任务在一段时间内的执行时间之称为负载。因此要在忙碌期内分析当前任务的 WCRT 上界，也就是使当前任务受到的干扰最大，即更高优先级任务的负载最大。文献[1]证明了当前任务的忙碌期起始时刻为该任务某实例的释放时刻时求得的 WCRT 上界是安全的，也就是真正的 WCRT 小于等于该上界值。并且还指出在忙碌期 $[t_0, t_1]$ 内，将对当前任务造成干扰的高优先级任务最大负载分为前部负载(carry-in)、中部负载(body)及后部负载(carry-out)。其中，前部负载表示实例的释放在 t_0 之前，执行完成在忙碌期内；中部负载表示实例的释放和执行都在忙碌期内；后部负载表示实例的释放在忙碌期内，但完成执行在 t_1 之后。根据文献[7]的结论，一个具有前部任务实例的任务当满足下列所有条件，其在忙碌期内的负载最大：（1）该任务的每个实例都执行其最坏情况执行时间；（2）该任务的所有实例以最小释放间隔释放；（3）该任务的前部任务实例所有的工作尽可能晚地执行；（4）该任务的后部任务实例在释放后立即不受干扰地执行至完成；（5）该任务的后部任务实例的执行结束时间与忙碌期的结束时间对齐。文献[8]将拥有前部任务实例的高优先级任务个数限定为最多 $m-1$ 个（ m 表示处理器个数），从而进一步减小前部负载带来的干扰，提高方法的精确度。文献[9]和文献[10]将文献[8]对前部任务的数量限定分别应用到了全局固定优先级(Global Fixed Priority, G-FP)调度和全局截止期优先(Global Earliest First, G-EDF)调度算法下的具体干扰计算中，并结合了文献[11]提到的负载函数来量化高优先级任务的请求数量。文献[9]和文献[10]的方法都极大的减小了计算干扰当前任务的负载的悲观性。文献[12]考虑前部负载的实际执行情况，对前部负载估算进行了进一步的优化，并应用到了文献[9]和文献[10]中提到的两种调度算法中。然而，以上这些研究针对拥有前部任务实例的高优先级任务数量、前部负载的估算以及负载的最坏情况进行了讨论与改进，并没有考虑到后部负载的优化。由文献[7]的结论，最坏情况下的后部任务实例是释放了就立即执行的，并且忙碌期的结束时间与后部任务实例的执行结束时间对齐，那么在这种假设下，所有后部任务实例中必然存在部分执行时间会与当前分析的任务实例并行执行，这部分后部负载就不能算入对当前分析的任务实例造成干扰的负载中。总的来说，现有多核处理器的时间分析技术还存在需要改进的地方，与精确的 WCRT 值还存在差距。而且，这些工作普遍采用 EDF 以及 FP 这些单处理器调度的最优算法作为调度算法进行时间分析，相比于实际情形中的复杂应用场景还需要基于更复杂的调度算法进行时间分析。

3 时间分析在车载网络的应用

传统的车载网络包括比如：控制器局域网（Controller Area Network, CAN）、面向媒体的系统传输总线（Media Oriented System Transport, MOST）、局域互联网络（Local Interconnect Network, LIN）、FlexRay 等。然而考虑到未来汽车正在朝着智能化、网络化迅速发展，传统的车载网络已经很难满足新阶段的汽车应用发展需求，需要加速以太网在汽车领域应用的标准化，电气和电子工程师协会（Institute of Electrical and Electronic Engineers, IEEE）已经制定出了一套的新标准，时间敏感网络（Time-Sensitive Networking, TSN）协议。TSN 作为一种新型以太网协议标准，它在传统高带宽低时延的以太加以时间同步、流量调度与整形和通信路径的选择，预留和容错等措施，保留以太网的高带宽低时延优点的同时，增加实时性和可靠性保证，使得以太网能够适用于车载实时通信系统。TSN 根据关键度将流量分为时间触发(Time-Triggered, TT)流量、AVB 流量和最佳流量(Best-Effort, BE)三类。TT 流量支持硬实时应用程序需要非常低的延迟和抖动。TT 流量具有最高优先级，它是基于调度表(Gate-Control Lists, GCL)发送的，该表依赖于全局同步时钟。AVB 流量适用于需要有限制的端到端延迟的应用程序，但其优先级低于 TT 流量。BE 流量优先级最低，用于不需要任何定时保证的应用程序。

针对 AVB 流量的延迟分析，只有当它的最坏端到端延迟（worstcase end-to-end delay, WCD）小于它的相对截止期限时才是可调度的。文献[13][14]提出了 AVB 网络中分析 AVB 流量的延迟分析方法，但忽略了在 TSN 中 TT 流量对 AVB 流量的影响。由于 AVB 流量适用于有限制的端到端延迟的应用程序，所以在进行最坏延迟分析时得到一个安全的分析结果，考虑 TT 流对 AVB 的影响是需要的。文献[15]提出了考虑以太网中静态调度 TT 帧的速率受限(Rate-Constrained, RC)流量 WCD 的网络演算分析方法，但它并不适用于 TSN 中。因为 RC 流量与 AVB 流量不同，并且 TSN 调度 TT 流量也与以太网不一样，TSN 通过 GCL 调度 TT 流量，而以太网直接调度每个单独的 TT 流。文献[16]考虑 TT 流量的影响并扩展了 AVB 延迟数学方程，但通过对这个应用场景的分析发现这种方法计算得到的 WCD 过于悲观，小于精确的 WCD，也是不安全的。由于 AVB 使用基于信用的整形器(Credit Based Shaper,CBS)来防止低优先级流的饥饿。在标准 802.1Qav[17]中，指出权衡 AVB 与 BE 流量之间的延迟的 CBS 参数固定关系，例如发送斜率=空闲斜率-传输速率。一些研究尝试推广这个关系为 AVB 和 BE 提供更多交换机会。但这可能会导致信用溢出造成不安全的 WCD。文献[18]推导出了一种不会导致 AVB 信用溢出的 CBS 参数推广方式，提出一种在 TSN 网络中安全的 AVB 流量的时间分析方法。以上的研究指出了当前分析 AVB 流量的 WCD 还存在的主要的问题，有一些研究给出了解决方案但都在某些方面存在问题，还需要更深入的研究进行改进和完善。

对于多核 ECU 任务，早期研究集中于研究具有共享资源的多核 ECU [19], [20]。考虑到 AUTOSAR 建议使用分区调度，因此文献[19]和[20]都关注分区调度范式下固定优先级 ECU 任务的 WCRT 分析方法。文献[19]和[20]之间的区别在于，前者使用抢先式调度，而后者使用非抢先式调度。基于抢先式和分区式调度范式下固定优先级 ECU 任务，[21]研究了用于多模式应用中 ECU 任务的 WCRT，该模型可以随着时间的推移而改变其功能。文献[22]使用约束编程求解器来计算任务序列的端到端延迟的范围，并以在多核 ECU 上演示了该方法，这是第一个形式化分析方法，利用负载假设来排除不可行的数据传播路径，而无需了解最坏情况的执行时间或最坏情况的响应时间。多核 ECU 任务的时间分析的关键困难在于难以建立准确的任务模型。原因是多核 ECU 的任务是由通信消息触发的，而通信消息什么时候开始传输和结束传输通常很难准确预测。而 TSN 作为一种新型以太网协议标准，有更高的实时性保证，但也为通信消息触发的多核 ECU 任务的 WCRT 分析带来了更复杂的问题，同时针对 TSN 上多核 ECU 的 WCRT 分析目前还处于初步探索阶段，还要做进一步的研究工作。

4 总结

本文从多核处理器的时间分析技术、TSN 网络的 AVB 流量延迟分析、多核 ECU 的时间分析三个方面展开了文献的调研和综述，体现了时间分析技术在多核处理器系统和车载网路的应用，为时间分析技术在车载 TSN 多核 ECU 的应用研究寻找突破口和可借鉴的解决思路。

参考文献

- [1] 关楠. 面向多核系统的实时调度算法研究[D]. 东北大学, 2012.
- [2] Baker T P, Cirinei M. Brute-force determination of multiprocessor schedulability for sets of sporadic hard-deadline tasks[C]//International Conference On Principles Of Distributed Systems. Springer, Berlin, Heidelberg, 2007: 62-75.
- [3] Geeraerts G, Goossens J, Lindström M. Multiprocessor schedulability of arbitrary-deadline sporadic tasks: complexity and antichain algorithm[J]. Real-time systems, 2013, 49(2): 171-218.
- [4] Sun Y, Lipari G. A weak simulation relation for real-time schedulability analysis of global fixed priority scheduling using linear hybrid automata[C]//Proceedings of the 22nd

-
- International Conference on Real-Time Networks and Systems. ACM, 2014: 35.
- [5] Baker T P. Multiprocessor EDF and deadline monotonic schedulability analysis[C]//RTSS 2003. 24th IEEE Real-Time Systems Symposium, 2003. IEEE, 2003: 120-129.
- [6] Bertogna M, Cirinei M. Response-time analysis for globally scheduled symmetric multiprocessor platforms[C]//28th IEEE International Real-Time Systems Symposium (RTSS 2007). IEEE, 2007: 149-160.
- [7] Bertogna M , Cirinei M , Lipari G . Improved Schedulability Analysis of EDF on Multiprocessor Platforms[J]. Proc.of Euromicro Conf.on Real Time Sys, 2005, 2005:209—218
- [8] Baruah S. Techniques for multiprocessor global schedulability analysis[C]//28th IEEE International Real-Time Systems Symposium (RTSS 2007). IEEE, 2007: 119-128.
- [9] Guan N, Stigge M, Yi W, et al. New response time bounds for fixed priority multiprocessor scheduling[C]//2009 30th IEEE Real-Time Systems Symposium. IEEE, 2009: 387-397.
- [10] Sun Y, Lipari G. Response time analysis with limited carry-in for global earliest deadline first scheduling[C]//2015 IEEE Real-Time Systems Symposium. IEEE, 2015: 130-140.
- [11] Bertogna M, Cirinei M. Response-time analysis for globally scheduled symmetric multiprocessor platforms[C]//28th IEEE International Real-Time Systems Symposium (RTSS 2007). IEEE, 2007: 149-160
- [12] Zhou Q, Li G, Li J. Improved carry-in workload estimation for global multiprocessor scheduling[J]. IEEE Transactions on Parallel and Distributed Systems, 2017, 28(9): 2527-2538.
- [13] Queck R . Analysis of Ethernet AVB for automotive networks using Network Calculus[C]// IEEE International Conference on Vehicular Electronics & Safety. IEEE, 2012.
- [14] Cao J, Cuijpers P J L , Bril R J , et al. Tight worst-case response-time analysis for ethernet AVB using eligible intervals[C]// IEEE World Conference on Factory Communication Systems. IEEE, 2016.
- [15] Boyer M , Daigmorte H , Navet N , et al. Performance impact of the interactions between time-triggered and rate-constrained transmissions in TTEthernet[J]. Die Medizinische Welt, 2016, 14(2):217-225.

-
- [16] Laursen S M , Pop P , Steiner W . Routing optimization of AVB streams in TSN networks[J]. ACM SIGBED Review, 2016, 13(4):43-48.
- [17] IEEE, "802.1Qav—Forwarding and Queuing Enhancements for TimeSensitive Streams," <http://www.ieee802.org/1/pages/802.1av.html>, 2009.
- [18] L. Zhao, P. Pop, Z. Zheng and Q. Li, "Timing Analysis of AVB Traffic in TSN Networks Using Network Calculus," 2018 IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS), Porto, 2018, pp. 25-36.
- [19] Schliecker S , Negrean M , Ernst R . Response Time Analysis on Multicore ECUs With Shared Resources[J]. IEEE Transactions on Industrial Informatics, 2009, 5(4):402-413.
- [20] Negrean M, Ernst R. Response-time analysis for non-preemptive scheduling in multi-core systems with shared resources[C]. international symposium on industrial embedded systems, 2012: 191-200.
- [21] Negrean M , Klawitter S , Ernst R . Timing analysis of multi-mode applications on AUTOSAR conform Multi-Core Systems[J]. 2013.
- [22] Friese M J , Ehlers T , Nowotka D . Estimating Latencies of Task Sequences in Multi-Core Automotive ECUs[C]// 2018 IEEE 13th International Symposium on Industrial Embedded Systems (SIES). IEEE, 2018.