Advances in Research on Scheduling and Communication Performance Optimization for Distributed Bus Network with WHRT Characteristics

Geng Liang

School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, China

Liangeng1976@163.com

Abstract

Distributed bus network with WHRT characteristics is widely used in modern process control. Theoretical research and industrial application shows that this kind of network is the most effective way to transmit HRT and WHRT information. In this paper, advances in the research on scheduling and communication performance optimization for distributed bus network with WHRT characteristics are described and analyzed. It is pointed out that currently available research results are strictly constrained by some assumptions and applying conditions, which means their feasibility and practicality is expected to be greatly improved further before their implementation in any real engineering. Scheduling and performance optimization of WHRT network under weak constraint should be studied and investigated more. In this paper, some possible solutions to these problems are also proposed in detail and more approaches to optimize the network performance are investigated.

Keywords

Weak constraint, Distributed bus network, WHRT system, Communication scheduling, Optimization.

1. Introduction

Application of real time systems is showing new characteristics with the development of medium and large scaled distributed measurement and control technologies. Data transmission over distributed bus network has its own priorities and communicatin performance must be ensured under the condition that some real time transmission requests from limited tasks can not be met, which makes new demands for the development of real time system. Obviously, traditional simple *HRT* and *SRT* can not precisely describe the real time application in such network environments.Correspondingly, optimal scheduling algorithms for real time application based on *HRT* and *SR*T find their limitations in network real time applications. The scheduling theories for such real time system are facing new chanllenges.

With the increasingly wide use of intelligent measurement and control network base on distributed bus (such as fieldbus), networked control systems (NCS) are also increasingly find their use in industrial applications. However, hard real time application can not be supported purely in practice for the indeterminisim caused by network transmission. To meet the requirements and constraints from task execution time, communication tasks can be divided into two types: hard real time and weakly hard real time. Determinism and limited queuing delay is the main characteristics for the distributed bus network with weakly hard real time performance, which is widely used in process control today. Theoretical research and practice has shown that the hard real time and weakly hard real time data transmission characterized by periodic and aperiodic data transmission is the most effective way for data transmission, among which are Foundation Fiedlbus (FF) and WorldFIP fieldbus.

2. Status quo of related researches with their pending problems

The so-called "real time" system is referred as not a class of rapid computing system, but that can quickly respond to the requests from extern events within given time, i.e. the deadline [1]. Traditional real time system can be divided into Hard Real Time (HRT) and Soft Real Time (SRT) according to requirement for the deadlines. In HRT system, serious consequence, even system breakdown, may be caused if deadline for event (also called task) can not be met. Some tasks are allowed not to meet the deadline in SRT system, which will cause performance degradation in return. The research on real time scheduling theories mainly focused on HRT system in early years. Worst-case executive time (WCET) and worst case utilization (WCU) were used to analyze the system performance in the worst case to ensure the deadline can be met. Accordingly, real time system can be designed and selected.

To the earlist age, constraint model *(m-k)-firm* was used by *Hamdoui* first to address the QoS problem in aspect that some tasks are allowed to be lost in a limited time interval and a distance-based priority (DBP) method was proposed [3,4]. Later *Bernat* defined such real time system with window constraints as weakly hard real time (WHRT) system[5], meaning task deadlines can meet some pre-specified distribution requirements or not within limited time windows.

Internationally, research on scheduling and optimization of WHRT task mainly is focused on industrial process control, among which WHRT-HRT task scheduling characterized by periodic and aperiodic traffic are the main stream. Fundamental characteristics of WHRT industrial communication network lie in that centralized medium access control scheme is adopted for periodic traffic, making it time deterministic, while aperiodic polling and inquiry mode is used for aperiodic traffic,with typical application like FF and WorldFIP bus network [6]. As for transmission of aperiodic HRT emergent data, research is mainly focused on average transmission delay and worst case transmission delay.

Studying the worst case transmission delay of aperiodic HRT data is significantly helpful in improving network parameters and methods and avoiding the worst cases. As for the WHRT bus network adopting aperiodic inquiry mode, F.Vasque used worst-case response time (WCRT) to evaluate its real time communication performance and investigated the generation and computation of WCRT [7]. However, F.Vasque considered that the aperiodic traffic is transferred only after all the periodic traffic in a macrocycle, which leads to very pessimistic results, which makes its practical application infeasible. P.Pedro studied the WCRT of HRT aperiodic data in FIP network [8]. He considered equal lengths for periodic windows since they do not discuss the BAT construction and thus they cannot evaluate the length of the periodic window in each microcycle. Thus, the overall results are also pessimistic, which also greatly constrain its practically applicable condition. Meanwhile, F.Vasque and P.Pedro did not consider the "atom" property of communication in the aperiodic windows of each microcycle, which can cause "time scrapes" in microcycles and change the distribution of aperiodic data within a macrocycle. E.Tovar and Almeida improved the research by F.Vasque and P.Pedro. They investigated the issues dealing with aperiodic data transmission under the condition that periodic traffic is taken into consideration based on a specified bus arbitrator table (BAT) [9]. In their studies, the "atom" property of communication in the aperiodic windows of each microcycle is fully considered and WCRT of aperiodic traffic was given as follows.

$$MIT\left(Va^{k}\right) \geq Ra^{k} = \sigma^{k} + len\_abi$$

where

$$\sigma^{k} = Tp_{j} + J_{Vp_{j}} + Cp_{j},\ \text{with}\ Vp_{j} : Tp_{j} = \min_{Vp_{i}\ \text{produced in}\ k} \{Tp_{i}\}$$

and

$$len\_abi = \left(N' - 1\right) \times \mu Cy +$$
$$+ \sum_{i=1}^{np} \left(bat[i, N] \times Cp_{i}\right) + \left(2 \times na - \sum_{l=1}^{N'-1} nap(l)\right) \times Ca^{*}$$

Nevertheless, the research results of E.Tovar was based on some strict assumptions and constraints: 1) Assume that the node related to periodic variables generate only one aperiodic variable once 2) Assume that the time consumption on identification of aperiodic variables (V^+) and the transmission of the indicated variables (V^-) is identical. Meanwhile, the case that V^+ and V^- were accumulated after dead interval were not considered in Tovar's research although the extreme condition that the end of the dead interval happens to be the start point of a new macrocycle was considered.Tovar only assumed that there were n_a complete combination of { V^+, V^-} accumulated after dead interval. The fact is that the accumulated V^+ or V^- does not make complete combinations, as shown in Fig.1.

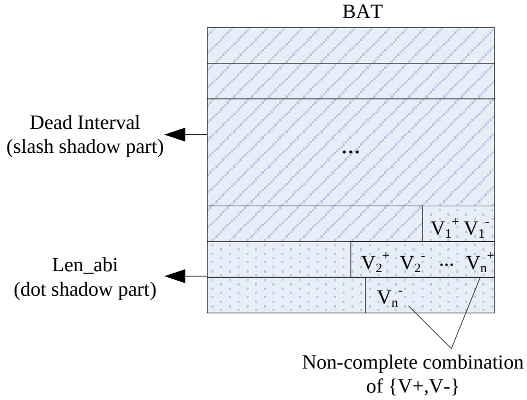

Fig.1 Non-complete combination of {V+,V-} in BAT

Z.Wang studied the WCRT of aperiodic data in WorldFIP network and proposed algorithms to evaluate WCRT of emergent and normal aperiodic data by constructing a non-pre-emptive queue of multiple priority model for transfer of aperiodic messages. The deficiency of WorldFIP in guaranteeing a timely response of important aperiodic message was found and a feasible strategy for implementation of the improved queue model was also proposed in his research [10], as shown in Fig.2.

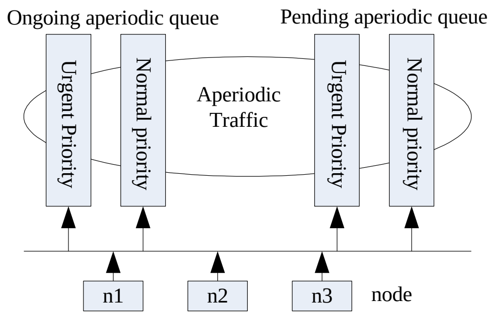

Fig.2 Improved queue model for WorldFIP

However, the assumption of the research is the same with that in Tovar's research. Z.Wang also studied WCRT of aperiodic data in polling network and proposed the evaluation method for the upper bound of WCRT . Effect of periodic data transmission on the WCRT of aperiodic data transmission and "time scrape" in microcycle are not considered in the research. L.Chen presented a bandwidth allocation algorithm based on priority mechanism and two bandwidth allocation strategies were proposed and compared in performance.But the author did not analyze the delay of aperiodic data under different token priorities and TTRT values.

Rules of average delay changed with some important network parameter, such as TTRT, in aperiodic HRT emergent data transmission is quite important and significant for improving communication efficiency and real time performance. Related researches are mainly focused on aperiodic polling network as yet. S.H.Hong and B.D.Jang used probability distribution and probability density function (PDF) based mathmatical model to analyze average delay in time-critical data transmission statistically, and presented the PDF and rules of average delay for time-critical data transmission changed with TTRT [11]. The research results of Hong and Jang were based on the following assumption: 1) periodic traffic is not considered at all 2) transimssion time for the time-critical and time-available data are fixed 3) the traffic load of the same priority calss is identical at all nodes 4) the FF is assumed to operate on a single-service discipline 5) actual token rotation time (ATRT) is an independent and identically distributed (i.i.d.) random variable. Y.H.Lee improved Hong's research by proposing the ATRT with token priority 0 and 1 are not i.i.d. based on the assumption 2),3),and 4), and presented the evaluation algorithm for the PDF of ATRT with full consideration of periodic traffic as follows.

$$T_0(t) = \sum_{i=0}^{N_0}\left(1 - \frac{R + iL_0}{p(Q_{(R+iL_0)/p} + 1)}\right)\Phi_{N_0}^{i}\,\delta(t - (R + iL_0 + cQ_{(R+iL_0)/p})) + \sum_{i=0}^{N_0}\left(\frac{R + iL_0}{p(Q_{(R+iL_0)/p} + 1)}\right)\Phi_{N_0}^{i}\,\delta(t - (R + iL_0 + c(Q_{(R+iL_0)/p} + 1)))$$

where

$$\Phi_N^i = \binom{N}{i}\rho_{00}^i(1-\rho_{00})^{N-i}$$

and

$$T_1(t) = \sum_{i=0}^{N_0}\sum_{j=0}^{N_1}\left(1 - \frac{R + iL_0 + jL_1}{p(Q_{(R+iL_0+jL_1)/p} + 1)}\right)\Phi_{N_0N_1}^{ij} \times \delta(t - (R + iL_0 + jL_1 + cQ_{(R+iL_0+jL_1)/p})) + \sum_{i=0}^{N_0}\sum_{j=0}^{N_1}\left(\frac{R + iL_0 + jL_1}{p(Q_{(R+iL_0+jL_1)/p} + 1)}\right)\Phi_{N_0N_1}^{ij} \times \delta(t - (R + iL_0 + jL_1 + c(Q_{(R+iL_0+jL_1)/p} + 1)))$$

where

$$\Phi_{NM}^{ij} = \binom{N}{i}\rho_{01}^i(1-\rho_{01})^{N-i}\binom{M}{j}\rho_{11}^j(1-\rho_{11})^{M-j}$$

Meanwhile, Lee extended the analysis of transmission delay from time-critical data to time-available data and validated the proposed ATRT mathmatical model with numerical simulations [11]. In Lee's later research, the analysis was extended to urgent, normal and time-available data. Notably, the above-mentioned assumption 2) and 3) are quite strict and assumption 4) also constrained data access mode of network nodes.Therefore, it can be seen that the mathmatical models proposed by Hong and Lee have limitations in application. As for the aperiodic polling network, the extra process "periodic indicating-aperiodic inquiry" in aperiodic traffic makes the HRT and WHRT data transmission more complex. To our best knowledge, there are not related research on average delay of aperiodic traffic in aperiodic inquiry network as yet.

Efficiency of distributed measurement and control network with WHRT characteristics has close relationship not only with the scheduling method used but also with multiple factors such as network communication mode, scale of system, number of nodes and real time requirements for node tasks. For instance, real time data can be transmitted not only by variable exchange but also by peer-to-peer message transmission, and communication efficiency relate to multiple factors such as selection of communication chips and amount of data to be transmittted. As for the research on optimization of network efficiency, present studies only focused on scheduling with some important parameters such as network communication mode, scale of system and data amount etc. taken as pre-set conditions and no deeper research has been done to the effects of these parameters on communication efficiency. Q.LI investigated the delays associated with the use of Foundation Fieldbus (FF) H1 networks within control loops and present some valid means to reduce transmission delay in FF-H1 and Profibus-PA network. We studied the usage optimization of WHRT network based on multi-element information fusion in aspect of measurement and control system and investigated the quantative relationship between network usage and several important network parameters. The existence of optimal value in network usage was proved and boundaries condition was given in our work[13].

Researches on optimization of scheduling algorithm are focused on Link Activity Scheduler (LAS). T. Wang proposed a method to set the priorities of tasks based on RM approach. J. Chen proposed an improved method to compute ATRT, which change delay distribution from cycle-based mode to node-based mode[14], as shown in Fig.3.

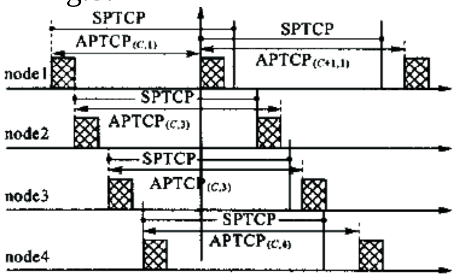

Fig.3 Relationships between TTRT and ATRT in J.Chen's improvement

However, only space element, which was represented by the priority of passing-token, was considered in Chen's research. Time element, i.e. PDF of data arrival, is not taken into account.In fact, variation in time element, in most of time, can cancel the improvement in space element. That demonstrates the limitation of optimizing network performance purely by focusing on LAS. More effective ways to solve such problem must are expected to be found.

As for research on the scheduling of HRT network traffic, methods such as LCM-HCF,RM,EDF etc. are widely used in design of schedule table. S.H.Hong presented a bandwidth allocation scheme for FF bus, in which traditional grouped data transmission scheduling mode was renovated into evenly-distributed data transmission scheduling mode instead based on a proposed schedule table design, as shown in Fig.4.

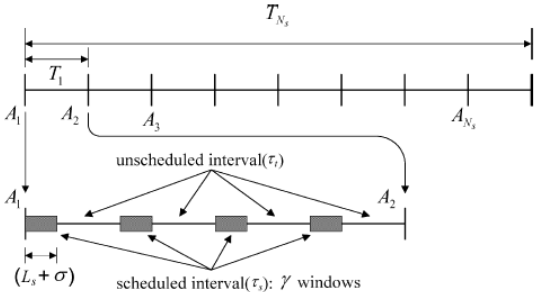

Fig.4 Bandwidth allocation scheme proposed by Hong for FF bus

Hong's scheme can restrict the delay of periodic traffic to a pre-specified time range. The update cycle of periodic data in each node is designed according to exponential rule, which may lead to prominent increment of scheduled periodic data when update cycle of periodic data increases exponentially, as shown in Fig.5.

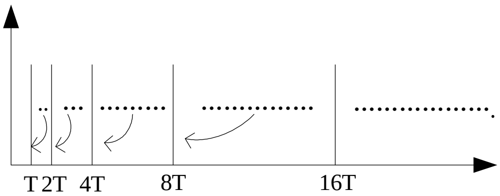

Fig.5 The update cycle of periodic data in each node designed according to exponential rule

3. Some possible solutions and approaches for the pending problems

From our above analysis, it can be seen that the current main problem existing in the present researches lies in its strong constraints for their results, which restrict their application to great extent. Therefore, studying of optimization of network communication under weak constraints is quite practically meaningful. Following are some pending problems dealing with the above mentionded research and possible approaches to solve them.

3.1 WCRT and average delay in WHRT network

As for aperiodic inquiry-typed WHRT network, it is important to obtain the general computation model for WCRT and distribution of average delay of aperiodic traffic with time under weak constraints. The following constraints used in former research must be cancelled: 1)forced constraint on number of aperiodic data associated with periodic data 2)assumption that time of aperiodic data identifying is identical with that of aperiodic data transimtting. Meanwhile, the extreme case that maximum number of aperiodic data is generated in the nodes which is associated with periodic data in each microcycle should be considered. The related problem can be re-evaluated according to following process:

construct multiple variable update cycle set φ with different character and element number;

constructing periodic scheduling table with RM,EFD etc and get the maximum number of variable Nvpmax which is inlcuded in the longest update cycle Tvpmax by comparison;

computing dead interval based on step 2;

constructing the generation model of aperiodic data in aperiodic window, fully considering the "atom" property of aperiodic traffic;

studying aperiodic data accumulation in Tvpmax based on step 4.

Issues dealing with WCRT of aperiodic urgent data in periodic polling network under the condition that "time scrape" in microcycle is fully considered can be addressed according to the following process:

constructing an analytical model for transmission delay of aperiodic data in FF bus MFF_AP, as shown in Fig.5.

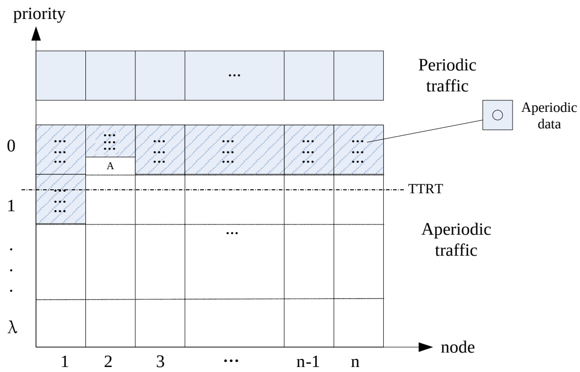

Fig.5 an analytical model for transmission delay of aperiodic data in FF bus

Analyzing the WCRT in aperiodic traffic WCRTap based on MFF_AP; investigating the distribution of WCRTap with periodic traffic included in analysis under the condition that microcycle "time scrape" is not considered;

Considering microcycle "time scrape", computing the maximum number of traffic "atom" in each aperiodic window, and then solve the following optimizing problem:

Assume that WCRTap is distributed within N' consecutive cycle in schedule table and N' satisfying:

$$N' = \max\{N \mid N_{ap} = \sum_{i=a}^{a+N} (NOA_i)\}$$

where N_{ap} is the number of aperiodic data in the case of $WCRT_{ap}$;

Evaluate the PDF of transmission delay of aperiodic urgent and normal queue with time:

$$D_0(t) = \sum_{i=1}^{N} P_i \delta(t - d_i^u), \quad D_1(t) = \sum_{i=1}^{N} P_i \delta(t - d_i^n)$$

analyzing the dependancy of $D_0(t)$ and $D_1(t)$ and correct their expressions according to the dependency.

3.2 optimization of WHRT network usage based on multi-information fusion

Efficiency of WHRT network is the effect of multi-parameters in network. Taking WorldFIP bus as a typical case for study, they are:

Network communication pattern m;

Network speedβ;

Turn-around time Tr;

Network load d;

Number of nodes n;

Analytical process for such problems can be as follows:

study the characteristics and limitations of different network communication pattern m in application;

Establish mono-numerical relationships between m, β, Tr, d and n with network usage λ;

Finding the main element affectingλby regressive analytical method; deciding appropriate combinatin of m based on d and n to achieve the optimal network usage.

3.3 Scheduling optimization cooperation from network nodes instead of LAS

This problem analysis includes the following steps:

analyze the effect of the following parameters on reducing transmission delay of aperiodic traffic:

current token priority TPRI;

current token position in rotation TPOS;

accumulation time of the current token TRT;

target token rotation time TTRT;

deadlines for HRT urgent data in basic nodes;

designing a dynamic discrete event simulation model and sampling the above mentioned network parameters from the model

contructing a BP neural network with n inputs, m hidden layers and 1 output. Trainning the NN with samples obtained in step 1, as shown in Fig.6. inputs are the above mentioned network parameters and output is transmission delay of aperiodic data Map_i in a basic node; supervisor is the transmission delay measured from the simulation model.

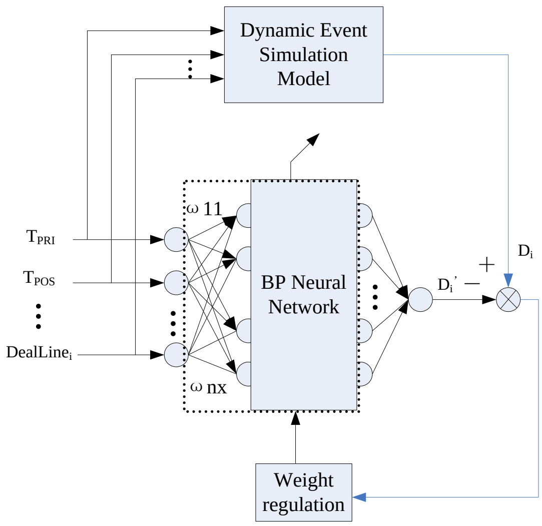

Fig.6 BP neural network for establishing the effect of network parameters on transmission delay

4. Conclusion

From above description and analysis of present-day research on scheduling and communication performance optimization for distributed network, it can be seen that the currently available research results are strictly constrained by some assumptions and applying conditions, which means their feasibility and practicability are expected to be greatly improved further before their implementation in any real engineering. Scheduling and performance optimization of WHRT network under weak constraint should be studied and investigated more. In this paper, we also propose some possible solutions to these problems and investigate more approaches to optimize the network performance.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities (2014MS23), Beijing Natural Science Foundation (4122074).

References

[1] Buttazzo Giorgio. Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications. Kluwer Academic Publishers, 1997.

[2] Furth. Real-Time UNIX System Design and Application Guide.Kluwer Academic Publisher , 1991.

[3] Moncef Hamdaoui and Pammeswaran Ramartathan.A service policy for real-time customers with(m,k)-firm deadlines. Proceedings of24th International Symposium on Fault Tolerant Computing, p196-205, 1994.

[4] Moncef Hamdaoui, Parameswaran Rarraanathan. A dynamic priority assignment technique for streams with(m,k)-firm deadlines.IEEE Transaction On Computers, 44(4), p1443-1451, 1995.

[5] Bemat Guillem,Alan Bums,Albert Llarnosi.Weakly hard real-time systems. IEEE Transaction on Computers, 50(4), p308-321,2006.

[6] Yeqiong Song. Evaluation of FIP and Application Dimensioning. PhD thesis,Institut National Polytechnique de Lorraine,1991.

[7] F.Vasques, G.Juanole. Pre-run-time Schedulability Analysis in Fieldbus Networks.Proceedings of the 20th Annual Conference of the IEEE Industrial Electronics Society (IECON'94), p1200-1204, 1994.

[8] Pedro, P. and A. Burns. Worst Case Response Time Analysis of Hard Real-time Sporadic Traffic in FIP Networks. Proceedings of the 9th Euromicro Workshop on Real-Time Systems, p3-10, June 1997.

[9] E. Tovar and F. Vasques, Contribution for the worst-case response time analysis of real-time sporadic traffic in WorldFIP networks, Euromicro Conference on Real-Time Systems RTS'99 (Work-in-Progress Session), p1-4 , 1999.

[10] Z.Wang, Y.Song, H.Yu, Y.Sun. Worst-case response time of aperiodc message in WorldFIP and its improvement in real-time capability. ISA Transactions, v 43, n 4, p 623-637, 2004.

[11] S.H.Hong,B.D.Jang.Time-critical data transmission in the foundation fieldbus,Proceedings of ISIE2001, p555-559.

[12] Y.H.Lee, S.H.Hong. Dependency on prioritized data in the delay analysis of foundation fieldbus,Control Engineering Practice,18(2010),p845-851.

[13] Geng Liang, Hong Wang, Wen Li, Dazhong Li. Communication Performance Analysis and Comparison of 2 Patterns for Data Exchange between Nodes in WorldFIP Fieldbus Network.. ISA Transactions, vol.49,no.4,2010,p567-576.

[14] Z.Wang, X.Shen, J.Chen Real-time performance evaluation of urgent aperiodic messages in FF communication and its improvement,Computer Standards & Interfaces 27 (2005) , p105–115.