Optimizing the Number of Cooperating Terminals for Energy Aware Task Computing in Wireless Networks

Anders Brødløs Olsen, Frank H.P. Fitzek, Peter Koch

Department of Communication Technology, Aalborg University
Niels Jernes Vej 12, 9220 Aalborg Øst, Denmark
Email: {abo|ff|pk}@kom.aau.dk

***Abstract*— It is generally accepted that energy consumption is a significant design constraint for mobile handheld systems, therefore motivations for methods optimizing the energy consumption making better use of the restricted battery resources are evident. A novel concept of distributed task computing is previously proposed (D^{2}VS), where the overall idea of selective distribution of tasks among terminals is made. In this paper the optimal number of terminals for cooperative task computing in a wireless network will be investigated. The paper presents an energy model for the proposed scheme. Energy consumption of the terminals with respect to their workload and the overhead of distributing tasks among terminals are taken into account. The paper shows, that the number of cooperating terminals is in general limited to a few, though alternating with respect to the various system parameters.**

Keywords: Cooperative Networks, Energy Optimization, Multi-processors, Scheduling, Dynamic Voltage Scaling

1. INTRODUCTION

Standing on the doorstep to the next generation of mobile wireless systems, referred to as fourth generation or just 4G, new envisioned multiple services will cause an increasing algorithmic complexity. The complexity increase the price of the terminal, as more and faster hardware is needed, which eventually also increases the energy consumption, and thereby reduces the available operation time of the device. Both, the price and operation time are the most important criteria for the customers' decision to buy a terminal.

Therefore, cooperative concepts were recently introduced, with the overall purpose to join forces in order to reach a common goal. Existing wireless networks assume that each individual mobile terminal (likely owned by users who may not necessarily want to share their resources with other users) will follow pre-described protocols. However, users may modify behavior or protocols for self- or group–interests. Most already proposed cooperative methods are introduced to combat the inherent nature of wireless networks, where channel errors are a significant issue, using basic principles of sharing resources for diversity [1], [2].

This work will investigate the cooperative concept for achieving energy aware computing in wireless networks. In Figure 1 a principle cooperative scenario for a cellular environment is illustrated. It is assumed that terminals receive a given service from the base station using the centralized wireless link. The service can be abstracted into a task-set, where tasks are defined by timing and workload specifications. The terminals then use a short range wireless technology for cooperating on a joint enhancement of the overall energy consumption. The concept is explained in detail in [3] and [4], showing that energy savings of more than 40% is obtainable for a two terminals case, compared to a solution where the service is executed on a single energy aware terminal.

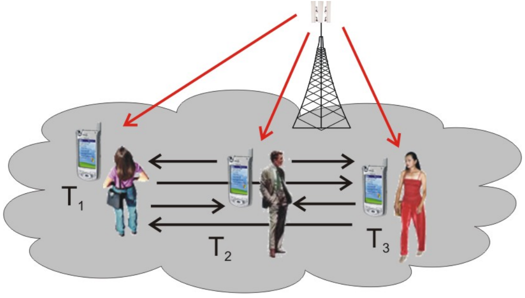

Figure 1: Principle cooperation between terminals within a cellular network

The overall concept of operation is: *Mobile terminals that cooperate on energy aware task execution must individually gain from the cooperation, and are therefore willing to exchange task-sets. If not fulfilled, terminals act selfishly and execute their tasks independently.* **Nevertheless, once the terminals have decided to cooperate the question of how many terminals could optimally be involved in the cooperation process arises.**

The principle approach is to; 1) employ various scheduling methods and optimizing cost functions for assigning tasks to multiple terminals, 2) identify the optimal number of wireless terminals to cooperate, 3) distribute the tasks using the short range wireless network, and 4) reduce energy consumption on the individual terminals using the well known methods of Dynamic Voltage Scaling (DVS).

To the best of the authors knowledge, only one attempt on energy aware cooperative computing is made for wireless sensor networks [5]. In this work, an ILP-based scheduling approach is proposed, together with a three phase optimizing heuristic. This method is made for epoch-based real-time applications, with the result that system life times are improved by a factor of 10 in the best case. Conceptually, our work is similar, although it deviates in a number of ways: 1) In [5], tasks are described using Directed Acyclic Graphs (DAG), where we use the definition of real-time task-sets. Using DAG and multiple heuristic seems as a static approach of allocating tasks among nodes, where we eventually are targeting a more dynamic environment and scheduling mechanism. 2) From the DVS literature the issue of task slack time reclaiming is shown to be important. By static approaches this is not possible, where dynamic DVS methods are able to utilize runtime accumulation of task slack times for further energy reduction. 3) Finally, a sound conclusion is that task allocation on other terminals is not always a benefit, which thereby calls for dynamic decision mechanisms for task distribution.

The organization of the paper is as follows: In Section 2 a brief explanation of the methodology proposed for energy aware computing in wireless terminals is made. Section 3 formulates a parameterized energy model of the scheme. Section 4 show the effects of the model parameters and finally in Section 5 the concluding remarks are made.

2. METHODOLOGY

The method is an extension of our privious work [3], [4], and for further detail we refer to these. This methodology is referred to as "distributed DVS" or D^{2}VS for short. Figure 2 illustrates our overall abstraction of the energy aware cooperative terminals, which is an abstraction of a multi-processor environment containing a number of processing units interconnected by a certain network topology. By this, mobile terminals are modeled as processing units, and the interconnecting network is a short range wireless communication technology. Efficient energy reduction methods are able to minimize the energy consumption on the individual terminals, as well for the wireless network.

Overall, an application being initiated on a given terminal, modeled as a task-set, and eventual introducing workload on this terminal. By careful selection, tasks are distributed among terminals, balancing the workload onto the terminals. Efficient energy management is responsible for energy conservation on the terminals, and Dynamic Voltage Scaling (DVS) is introduced for this purpose. DVS is opted for, as it is proven to produce near optimal energy schedules for single processors [6], and also being able to harvest the energy advantages expected for multi-processors environments [7].

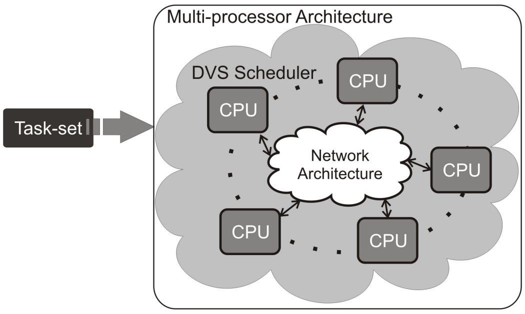

Figure 2: Abstraction of a multiple processors environment, connected by abstract network architecture. A DVS method schedules the task-set onto the multi-processor architecture and performs the energy conservation.

3. MODEL

The energy model applied in this work is based on a number of realistic assumptions:

1. Terminals or network notes are composed of a processing unit (PU) and a network device, capable of distributing tasks between terminals.
2. Task-sets are fully distributable, i.e. having sufficient degree of parallelism. Therefore, task-set can be randomly distributed among the n cooperating terminals.
3. A master/slave network structure is assumed, such that the master only communicates with a number of slaves and the slaves are having no communication with each other. At the time of scheduling the slaves are idle, meaning that no local workload is present on the slave PU's.
4. Homogeneous PU's are assumed, having identical energy consumption both for the PU and network device.
5. The network device is independently able to handle task transmissions, meaning that the PU's can execute tasks while the network transceiver is active. Further more ideal energy control is assumed, implying that energy only is spend while the network device is active.

Next, based on the system description from the previous sections an energy model for task distribution is formulated. Energy consumption is assumed to emerge from two primarily sources, 1) the energy consumption of the PU's and 2) the energy on the interconnecting network devices. This is expressed in Equation 1;

$$E_{Tot} = E_{PU} + E_{Net}, \tag{1}$$

where E_{PU} is the total combined energy on the PU's, and E_{Net} is the energy cost for distributing tasks among the slave PU's in the cooperative group.

In an environment containing n PU's connected by $(n-1)$ communication links, the total energy is expressed in Equation 2;

$$E_{Tot} = nE_{PU} + (n-1)E_{Net} \tag{2}$$

The reason for $n-1$ communication links is due to the overall system assumption 3, overall giving a star-like network structure. Following, energy models for the PU's and network devices will be further elaborated.

3.1. PU energy model

Energy consumption on a PU is eventually determined by the task-set specification. The workload (or utilization) on a PU is determined by task-set timing constraints and the task workload. Therefore, in this model the task-set utilization will be used as a measure for the combined task-set workload, and therefore by model assumption 2 the energy for the PU energy consumption can be expressed by Equation 3;

$$E_{PU} = E\left(\frac{U_{Task-set}}{n}\right), \tag{3}$$

where $U_{Task-set}$ is the utilization of the task-set, and $E(\cdot)$ is the energy operator for executing the task-set using an energy optimal speed. Using model assumption 2, the task-set utilization is divided onto the n PU's. At this point it is important to note that $E(\cdot)$ is a function having a quadratic nature due to the assumed CMOS technology of the PU's. Equation 4 contain two models for energy consumption:

$$E\left(\frac{U_{Task-set}}{n}\right) = U_{PU}^2 \vee f(S|U_{PU}), \tag{4}$$

where U_{PU}^2 is a model where the utilization on the individual PU is squared. This is a model commonly used in DVS literature and can be shown to comply with realistic hardware. Secondly, the model $f(S|U_{PU})$ expresses energy as a function of the speed required for a given task-set utilization. The speed value refers to the optimal pair of clock frequency and supply voltage for a given hardware architecture.

To exemplify the PU energy model, measurements on a real hardware platform, specific an Analog Devices BF535 EZ-KIT Lite, is conducted, see Figure 3. Also in this figure the squared utilization model is illustrated, both models normalized to 1 for the highest speed value. As seen, the two models are having correlating shapes and are essentially overlapping.

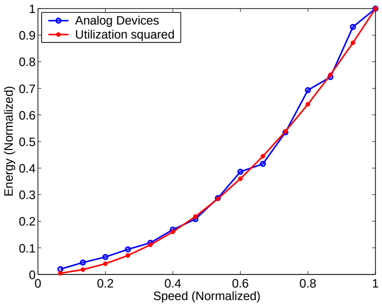

Figure 3: The utilization squared energy model, compared to measurements on an Analog Devices Blackfin 535 EZ-Kit Lite evaluation board. Normalized for speed and energy.

The two energy models in Equation 4 will be referred to as 1) utilization squared and 2) Analog Devices energy models, respectively.

3.2. Network energy model

The model assumptions state that the network must be capable of 1) handle task transmission when the PU executes tasks and 2) an ideal energy control only coursing energy consumption when the network is active. Previously, it is mentioned that a short range wireless technology is employed. From an energy view point, it is essential that the energy consumption on the network device is reduced as much as possible. Therefore, short range technologies are important, and from a physical consideration it is evident that distance between communicating notes is a factor. From known technologies like e.g. Bluetooth we safely assume support of adjustable power emission on the air-interface according to distance. On the other hand, technologies like Wireless LAN typically do not have any energy awareness, and therefore likely a poor candidate as an interconnecting network technology.

This model utilizes a high abstraction not targeting any particular technology, but includes parameters that overall will affect the energy consumption. Taking an outset in a simple communication relation, it is well known that communication is a relation of the network capacity, the amount of data that has to be transmitted, and the condition of the wireless link. By the assumption that a network is available for transmission at any time instance, issues like packet collision and other network interference will be neglected, as it is assumed that the network devices are able to handle the traffic on the network. Therefore, the most momentous parameters are 1) the time it take to transmit a given task information, 2) the power consumption of the active network device, and 3) the number of tasks that have to be transmitted between the master and slave terminals. The energy consumption of transmitting information can be expressed as in Equation 5.

$$E_{Net} = f(t_{\tau_i} \cdot P_{Net}) \cdot \underset{i=1:m}{n_{\tau_i}}\;, \tag{5}$$

where t_{τ_i} is the time it takes to transmit the τ_i'th task, P_{Net} the power consumption on the network devices, and n_{τ_i} is the number of tasks transmitted on the network (a maximum of m tasks is assumed).

The time expression of Equation 5 is a composition of the time it takes to establish the connection and the time it takes to transmit the information, as expressed in Equation 6;

$$t_{\tau_i} = t_{setup} + t_{bit} \cdot m_{\tau_i}, \tag{6}$$

where t_{setup} is the time for setting up the communication link, t_{bit} the time it takes to transmit a bit, and m_{τ_i} the number of bit for the data of the τ_i'th task. In this work, τ_i is used for describing the time load of the uniform task parts distributed among terminals.

In practical network implementations the cost of transmitting and receiving comes with different cost, as expressed in Equation 7;

$$P_{Net} = P_T + P_R, \tag{7}$$

where P_T is the power for transmitting information and P_R is the power of receiving information on the network device. In this work, a superposition of the two is used. As joint energy overhead of the cooperative network is the objective, it is therefore not distinguished what the individual energy consumption of the terminals are.

Finally, a constraint on the network active time is needed, leading to the inequality in Equation 8;

$$\sum_{i=1}^{n-1} t_{\tau_i} \leq 1, \tag{8}$$

where the summation of task time overheads is less or equal to 1, which denotes full activity on the network and has to be or less than fully utilized.

4. RESULTS

For the experimental setup, in this paper, the cooperative computing model expressed in the previous section is used. Using Equation 2 as base, energy models for both the utilization squared and Analog Devices model of Equation 4 is used. For the network overhead Equation 5 is primary used, with the assumption that only one task is distributed to the individual terminals. Further more, simulations is only carried out so Equation 8 is satisfied.

4.1. Ideal and cost free communication

These simulations will show the effect of distributing tasks on n PU's when assuming an ideal and cost free communication. Both the energy models of Equation 4 is used, where the latest uses the actual measured values of the Analog Devices processor showed in Figure 3. In the simulations we will limit the PU count to 10. For the Analog Devices model it should be noted that not all speed settings are physically available, and therefore the values as specified in Table 1 will be used. All normalized to full speed equal 1 and also full speed energy to 1.

Table 1: Speed values and energy consumptions for Analog Devices model, showing quantizations due to hardware limitations

Speed/# PU	1	1/2	1/3	1/4	1/5
Actual speed	1	.533	.333	.266	.2
Energy	1	.287	.118	.094	.065

Speed/# PU	1/6	1/7	1/8	1/9	1/10
Actual speed	.2	.2	.133	.133	.133
Energy	.065	.065	.044	.044	.044

Using the model of Equation 2 without contributions form the network, the results can be seen in Figure 4. As expected for the squared utilization model, we see a decreasing energy consumption with respect to the number of PU's. As the number of PU's grows large, the energy consumption will limit to an asymptotic value. For the Analog Devices model, the limitation is in the physical possible values of the speed settings. In Figure 4 it is seen that up to three PU's the two models correlate, whereas for higher numbers of PU's the gains of distributing tasks disappears as the hardware cannot support the speeds required. E.g. using four units will increase the joint energy consumption as related to using only three, where also five PU's is slightly better than three PU's.

Comparing the Analog Devices model's minimum to a single PU solution, the energy gain is obvious, and the ratio for the five PU's case is providing approximately 3 times better energy consumption.

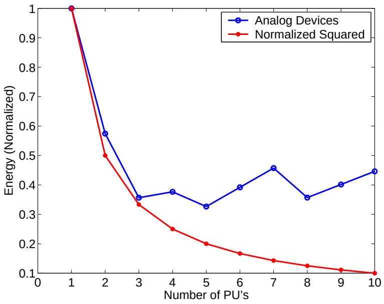

Figure 4: Energy consumption as a function of PU's, both illustrated for the utilization squared and the Analog Devices energy models. Showing the effect of hardware quantization.

4.2. Effects of network model parameters

In these experiments the effect of the network parameters is investigated. The task-set is having a utility factor of one, meaning when executed on a single PU it will be fully utilized. In Figures 5 and 6 test using the utilization squared and Analog Devices energy model is used, respectively. In both figures, network power of $\{2, 1, 0.75, 0.5\}$ is used, represented in the top left sub-plot succeeded to the bottom right sub-plot. In each sub-plot network activity time of $\{0.5, 0.4, 0.3, 0.25, 0.2, 0.15, 0.1, 0.05\}$ is used, see in each sub-plot in sequence from top to bottom. Especially it should be noted that curves only are shown such that Equation 8 is satisfied, e.g. an activity factor of 0.5 only supports transmission onto two network links. In each plot the vertical line indicates the energy consumption of the task-set if executed on a single PU.

From Figures 5 and 6 it is obvious that activity time and power levels on the network link affect the joint energy consumption of distribution task substantially. Taking the top left sub-plot, where the power level of the network device is a factor two of the PU's running at full speed. Even in such an setup, the proposed method will contributed with energy reduction of a ratio of 1.88 at 3 PU's for activity time factor of 0.05 and a ratio of 1.43 at 2 PU's for activity time factor of 0.1. Contrary, the ratios are 1.79 and 1.32 for the Analog Devices model, but both at 3 PU's.

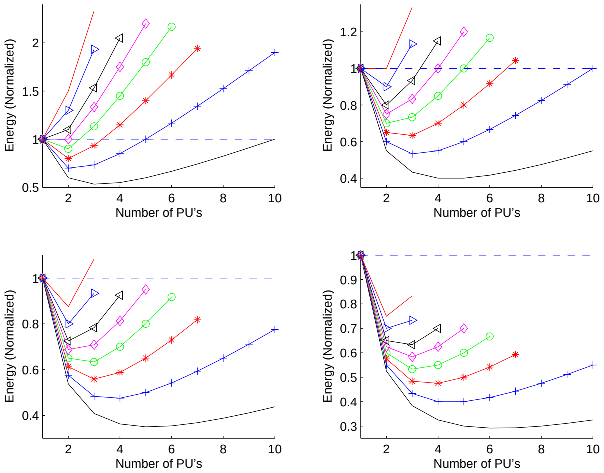

Figure 5: Network power and time parameters on the number of PU's, using the utilization squared energy function. Power shown in values $\{2, 1, 0.75, 0.5\}$ seen in sequence from top left to bottom right. Task time load in values $\{0.5, 0.4, 0.3, 0.25, 0.2, 0.15, 0.1, 0.05\}$ seen from top to bottom in each plot.

Observing the remaining sub-plots it clearly seen that activity time and network power are parabolic shaped functions of the number of PU's. Taking the lowest simulated activity factor for network power of $\{1\ 0.75, 0.5\}$ the power gains are $\{2.5, 2.86, 3.43\}$, and the number of PU's are $\{4, 5, 6\}$, respectively. Thereby, obviously, the optimal number of PU's is determined by these two parameters, but it is also clear that the number of PU's is limited to only a few PU's, specific limited to 6 PU's for the setup in the lower right sub-plot of Figure 5. Contrary, the Analog Devices model indicate that the optimal number of PU's never exceed three PU's see Figure 6.

4.3. The effect of task-set utilization

In our previous work [3] it was shown that task-set utilization factor is having an effect on the performance of the proposed method, and will therefore also have an effect on the optimal number of PU's. In Figure 7, similar plots as in Section 4.2 are shown, using identical activity values and a network power level of 0.75. Task-set utilizations of $\{1, 0.75, 0.5, 0{,}25\}$ is illustrated, with a utilization of 1 in the left top sub-plot succeeded to a utilization of 0.25 in the bottom right sub-plot. In this experiment only the "utilization squared" PU energy model is shown in the paper. In each plot the vertical line indicates the energy consumption of a single PU executing the task-set, corresponding to the utilization value.

Comparing the single PU energy at the four different utilizations levels it is clearly seen that the ratio between a single PU and the minimum energy of multiple PU's become significant less related to the utilization value. Using the minimum activity time as

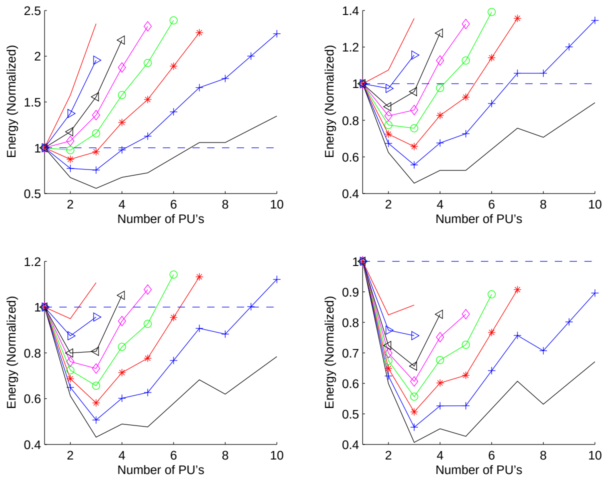

Figure 6: Network power and time parameters on the number of PU's, using the Analog Devices squared energy function. With simelar set-up as Figure 5.

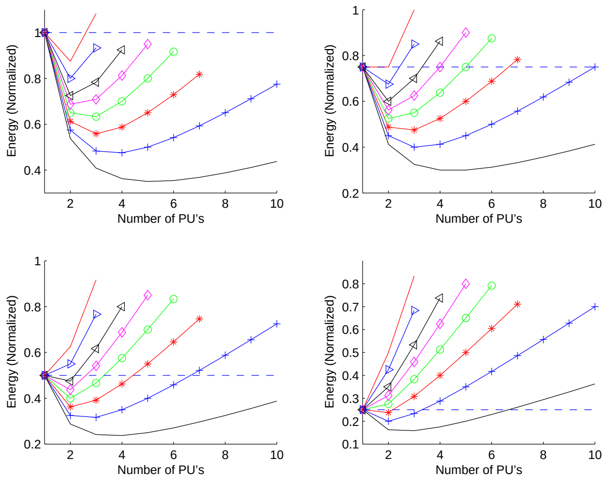

Figure 7: Task-set utilization effect on simelar task network load time as showed in Figure 5 and showed for network power 0.75.

example, it is clearly seen that the optimal number of PU's decreases related to task-set utilization, and the optimal number of PU's are $\{5, 4, 3, 2\}$ respectively of the four task-set utilizations. The energy gain also decreases from $\{2.86, 2.5, 2.07, 1.54\}$ times compared to a single PU solution, again for the minimum activity time simulated. If similar experiments for the Analog Devices PU model is made, again the optimal number of PU's never exceeds 3 and the energy ratio gain is also degraded as compared to the utilization squared model.

The reason for the decreasing performance is due to the behavior of energy consumption as seen in Figure 3. Extending the energy consumption model it will tend towards an asymptotic level, meaning that the gradient of the energy function is most steep for the first fractions of speed reductions. Therefore, the optimum number of PU's is limited to a few and also highly depended on the task-set utilization factor.

5. CONCLUSIONS

In this work we have extended our work on the concept of Distributed DVS, or D^{2}VS for short, for cooperative task computing in wireless networks. The objective of the paper is to evaluate the optimum number of terminals in a cooperative group. A model for energy consumption on the processing units with respect to their work load specification is proposed. This model uses knowledge of energy consumption in an environment using speed scaling technologies, meaning that energy consumption is having parabolic shaped energy relations to the workload on the processing unit. As tasks are distributed using wireless technologies the energy overhead is model using a general high abstraction communication model. The energy overhead of the network is abstracted into a few essential parameters.

For the experiments the proposed model parameters are explored, with the observations that the optimal number of processing units is depended on the combination of the system parameters. Two different models for the processing units are used, one being a model on the utilization factor on the processor and the other measurements of a real life processor. Introducing the various overheads parabolic shaped energy relations related to the number of processing units appears. General the network overheads displace the gain of the proposed method, and also the number of optimal units. Our results show that careful considerations must be made of the overheads, as the method can increase the energy consumption when distributing tasks among terminals. On the other hand energy saving of as much as three times can be obtained for the right combinations of overheads.

REFERENCES

[1] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," in *Communications Magazine, IEEE, Vol.42, Iss.10*, 2004, pp. 74–80.

[2] C. Politis, T. Oda, S. Dixit, A. Schieder, K.-Y. Lach, M. Smirnov, S. Uskela, and R. Tafazolli, "Cooperative networks for the future wireless world," in *Communications Magazine, IEEE, Vol.42, Iss.9*, 2004, pp. 70–79.

[3] A. B. Olsen, F. Fitzek, and P. Koch, "Evaluation of cooperative task computing for energy aware wireless networks," in *International Workshop on Wireless Ad-hoc Networks (IWWAN'05)*, May 2005.

[4] ——, "Energy aware computing in cooperative wireless networks," in *International Conference on Wireless Networks, Communications, and Mobile Computing (WirelessCOM'05)*, 2005.

[5] Y. Yu and V. Prasanna, "Energy-balanced task allocation for collaborative processing in wireless sensor networks," *Mobile Networks and Applications (MONET'05)*, vol. 10, no. 1-2, pp. 115–131, 2005.

[6] H. Aydin, R. Melhem, D. Mosse, and P. Alvarez, "Dynamic and aggressive scheduling techniques for power-aware real-time systems," in *Proceedings of Real-Time Systems Symposium (RTSS'01)*, 2001, pp. 95–105.

[7] H. Aydin and Q. Yang, "Energy-aware partitioning for multiprocessor real-time systems," in *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS03)*, 2003, pp. 113–121.