Systematic Classification of Self-Adapting Algorithms for Power-Saving Operation Modes of ICT Systems

Paul J. Kühn
University of Stuttgart, Institute of Communication Networks and Computer Engineering (IKR)
Paffenwaldring 47, 70569 Stuttgart, Germany
paul.j.kuehn@ikr.uni-stuttgart.de

ABSTRACT

Reduction of power consumption is a paramount objective in the ICT area, especially in computer and communication systems. This paper provides a fundamental approach to model power-consuming system resources and self-adapting algorithms for dynamic activation and deactivation of such resources dependent on the actual load level. The algorithms are based on Finite State Machines (FSM) represented by state transition diagrams with various types of hysteresis which can be implemented easily in real systems and which need only the available state information.

Categories and Subject Descriptors

C.2.3 [**Computer Communication Networks**]: Network Operation – *Power Saving*

General Terms

Algorithms, Management, Measurement

Keywords

Power-saving algorithms, Self-adaptation, Finite state machine, Controlled queueing system

1. INTRODUCTION

Increasing traffic demand in the internet, in wireless and mobile communication systems, and in data processing result in a tremendous power consumption. Today, the whole field of ICT is responsible for about 10% of the world's energy consumption and the corresponding effects on the carbon footprint. Even with more efficient electronic and photonic technologies, the increase in power demand cannot be satisfied without massive power-saving operations. On the other hand, most communication links, routers and servers are operated at top speed, although their peak performance is only required sporadically and their average utilization amounts only to about 30%.

For these reasons, many efforts are currently being directed at power-saving operation modes as cloud computing, network virtualization, dynamic activation and deactivation of power-consuming resources, adaptive link rate aggregation and sleep modes, better known under the headlines as "Green ICT".

Methods for power saving in ICT systems have become a hot topic in the recent years, they range from the device level (on-chip power control), circuit level (low power circuit design), network level (transmission and coding, switching and routing, protocols) up to the application level (user behaviour) and to system operation management (power supply, cooling). In this paper, we will focus mainly on the network level of ICT systems, specifically on communication network aspects, c.f. [1, 2]. The derived models lead to resource management problems and can be described as stochastic service systems. The models as such have a wider application than specifically communication systems and can, thus, be applied also to, e.g., distributed computer systems or cloud computing.

Principles for power saving in communication systems by adaptive resource management can be classified into:

- Dynamic Transmission Link Control (DTLC)
- Adaptive Link Rate (ALR)

With DTLC, transmission links (including the associated port interfaces) are dynamically activated and deactivated dependent on the current load level. As activation/ deactivation mechanisms require signalling between the link terminal interfaces, the frequency of activations/deactivations has to be carefully controlled to avoid excessive overhead and delays. Overhead can be reduced by effective forecasting methods on the load development or by more sophisticated sleep modes by which power consumption and overhead can be reduced. By ALR, power saving is achieved through dynamically switching the link transmission rate to adapt to the current load. This method is based on the fact that power consumption grows linearly with the frequency.

This paper aims primarily at a fundamental approach to the organization of dynamic, load-dependent activation/ deactivation of network resources such as transmission channels of a multi-channel link or processing devices of a router with multi-core processors which will be treated as "servers". Servers are the main power-consuming elements. Buffers can also be considered: buffer devices can be operated in a low-power sleep mode most of the time and become empowered only for reading/writing operations.

The control of power-consuming devices can be considered as a resource management problem which can be modeled

e-Energy '11, May 31 - June 01 2011, New York, NY, USA

as a controlled queueing system. The control is based on a Finite State Machine (FSM) represented by a state transition diagram. Transitions are caused by arrivals of data units (frames) and by service terminations ("state-based control"). This method can be extended to take estimations about the development of the near future into consideration ("measurement-/trend-based control").

The rest of the paper is structured as follows: In Chapter 2, fundamental models will be derived for dynamic activations/deactivations of servers dependent on the actual load level (state-based control). In Chapter 3, extended models are derived including a look-ahead mechanism where the future development is estimated from the development of the recent past ("measurement-/trend-based control"). In Chapter 4, the application of these models for performance evaluation of the various power-saving methods will be discussed shortly. The paper ends with a conclusion and an outlook on further work in progress.

2. MODELS FOR SERVER ACTIVATION/ DEACTIVATION

2.1 Resource-Sharing Queueing Model

We consider a finite buffer queueing model with a total number of n servers and s buffer places for data units ("Frames"), c.f. Figure 1. Arriving frames are directly served and occupy an available idle server; if all activated servers are occupied, the frame will be buffered and has to wait to become scheduled for service. If all buffer places are occupied, the arrived frame will be dropped and has no further effect on the state of the queueing model.

Arrivals may occur according to any type of arrival process (single or batch arrivals) considered as a stochastic point process). Service times at the servers may also be generally distributed. Buffered frames will be organized in the buffer strictly on the basis of First-In, First-Out (FIFO) queueing strategy. As soon as a server becomes idle, it will be either immediately re-occupied by the frame waiting at the head of the queue or it will be deactivated. The number of activated servers X depends on the current state (X, Z) of the system where Z denotes the actual number of buffered frames.

Activations and deactivations of servers as well as scheduling of buffered frames for service are controlled by a State Machine (Control). The Control receives values of the actual state variables X and Z and decides about the activation of an idle server, the deactivation of a server, or the re-occupation of a server when it becomes idle by assignment of the frame waiting at the head of the queue at the instant of a service termination. In summary, the generic model for dynamic activation/deactivation of servers acc. to Figure 1 is a generalized queueing model ("controlled queueing model").

2.2 Single Hysteresis Model

The basic idea of power saving is to activate only as many servers for a particular load level such that the service quality (performance) is still acceptable given a certain load level $\rho = \lambda/\mu$ where λ is the arrival rate of frames per second and μ the service rate of a service; $1/\lambda$ is the average interarrival time, and $1/\mu$ denotes the average service time.

In the basic queueing system without hysteresis, the number of occupied servers follows from the standard finite buffer model $G_A/G_S/n/s$ and is described by the server state probability distribution $P(x), x = 0, 1, \cdots, n$. The distribution $P(x)$ varies dependent on ρ and on the interarrival and service time characteristic G_A and G_S, respectively (G: general, or arbitrary type of stochastic process). If at any frame arrival/service termination in state $x < n$ a server would be activated/deactivated, the rate of activations/deactivations would be quite high. To avoid too frequent events of activation/deactivation, arriving frames will be buffered at states $x < n$ shortly so that a new server activation is not necessary for occasional arrival peaks. If, however, the number of buffered frames would exceed a threshold w, the next idle server will be activated to avoid overloads or higher delays, respectively. Similarly, a server will only be deactivated when the queue of waiting frames drops down to a lower level.

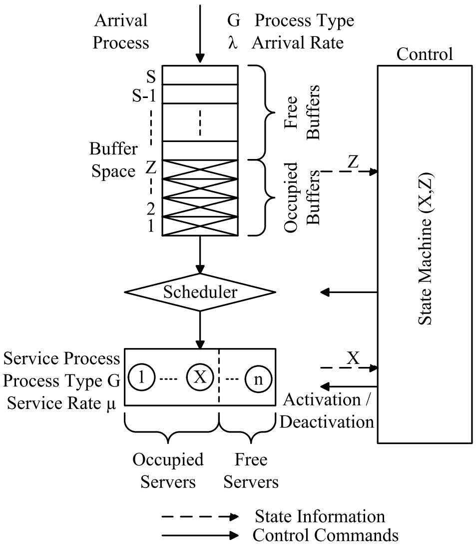

Figure 1: Generic Model for Dynamic Activation/Deactivation of Servers

The described effects are perfectly covered by a single hysteresis in the state transition diagram for the system state $S = (X, Z)$, c.f. Figure 2 (w denotes the "width" of the hysteresis). As the total number of buffers amounts to s, $s > w$ must hold. The server activations and deactivations are highlighted for the corresponding state transitions. The hysteresis changes the shape of the server state distribution as the probability mass concentrates around the range $x < \rho < x + 1$ (Note: in a pure delay system, the average number of occupied servers $Y = E[X] = \rho$). As a consequence, only little probability mass is left over for the probabilities $p(i)$ for $i < x$ and $i > x+1$. The frequency of server activations/ deactivations will be reduced accordingly.

The hysteresis of Figure 2 is a special case of a dual-threshold policy which is a common control policy to reduce oscillations and has been applied, e.g., in overload control schemes for switching systems and is also applied in power-saving algorithms by the ALR strategy, c.f. [1]. Multiple hysteresis models have already been studied earlier in the computer performance literature by complex methods as Green's function or stochastic complementation, see [3, 6]. In [4, 5] we develop simple recursive algorithms for such problems which are much easier to exploit; the analyses in [5] provide even results on the delay distribution functions

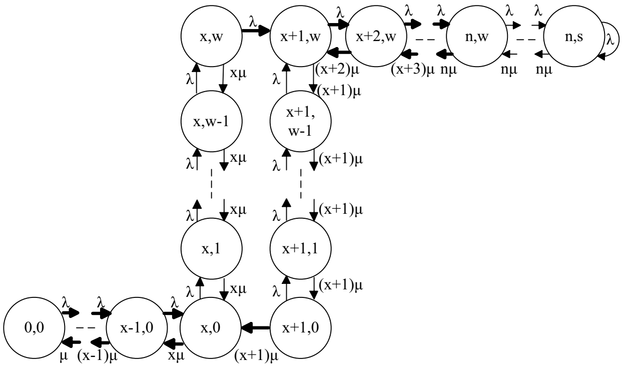

Figure 2: State Transition Diagram for the Single Hysteresis Model

which are important to bound the delay jitter caused by the power-saving operations. This paper addresses specifically a generalization of such models with particular focus on the application to power-saving operations.

2.3 Multiple Serial Hysteresis Model

The single hysteresis model of Figure 2 is only useful for a known load level ρ for $x < \rho < x+1$. For a self-adapting algorithm, the state transition diagram has to be generalized that it holds for any load level ρ automatically. This can be achieved by a proper design of the state transition diagram. In Figure 3, the single hysteresis principle is applied to any state $x = 1, 2, \cdots, n-1$ by serial repetition of the single hysteresis principle at states $(i, w^{(i-1)})$, $i = 1, 2, \cdots, n-1$, where $w^{(j)} = w_1 + w_2 + \cdots + w_j, j = 1, \cdots, n-1, w^{(0)} = 0$; w_j describes the "width" of the hysteresis in state $x = j$. For the total number of buffers s, the relation $s \geq w^{(n-1)} = w_1 + w_2 + \cdots + w_{n-1}$ must hold. Note, that in the serial hysteresis model all hysteresis widths w_i have to be chosen according to $w_i > 1, i = 1, 2, \cdots, n-1$.

In Figure 3, the server activations and deactivations are highlighted again for the corresponding state transitions. As the probability mass for the load level region $x < \rho < x+1$ will be concentrated on the states of the hysteresis between x and $x+1$, the resulting total rate of server activations/deactivations will be reduced drastically compared to the model without any hysteresis. For an automatic self-adaptation of the server activation/deactivation algorithm with respect to the load level ρ, the serial hysteresis model is not the only possibility. Alternative models of parallel hysteresis and combined serial/parallel hysteresis models can be constructed.

2.4 Multiple Parallel Hysteresis Model

In Figure 4, a state transition diagram is shown where hysteresis is applied in parallel. It is derived from the model of Figure 3 where the deactivation of servers takes place only at instants where a server becomes idle and when the buffer is empty. By this strategy, the service of buffered frames is enforced which results in a reduction of frame delays.

From the multiple parallel hysteresis model a combined serial/parallel hysteresis model can be derived if server deactivations are initiated at the server termination instants of an arbitrary state which extends the parameter space for control. A specific multiple parallel model results from a specific choice of the width parameters w_i choosing $w_1 > 0$ and $w_2 = w_3 = \cdots = w_{n-1} = 0$.

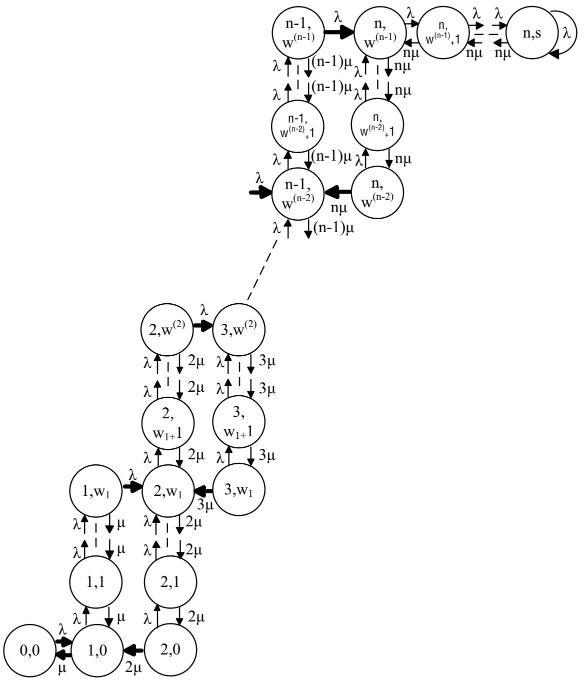

Figure 3: State Transition Diagram for the Multiple Serial Hysteresis Model

3. MODELS FOR ADVANCED SERVER ACTIVATION

State-based activation/deactivation of servers may be too slow for systems with extremely high dynamic state changes. Internet traffic patterns are characterized by burst effects, i.e., packet flows can quickly change from a low (or zero) packet rate to a high packet rate. As activation/deactivation control of transmission links requires overhead resulting from a signaling procedure for synchronized actions between both link termination interfaces, buffering peaks may result from activation delays. To overcome such activation delays, a forecast mechanism with respect to the resource requirements within the near future would be of interest in order to speed up server activations.

Forecasting these resource requirements for the near future can be based on continuous monitoring of the arriving frame rate within short intervals of time by a window mechanism with moving average or with a weighted moving average over several past intervals from which the trend can be estimated ("measurement- or trend-based control"). In fact, this trend estimation can be performed by the FSM control on the basis of the frame arrival information, c.f. Figure 1.

The most simple case of an advanced server activation can be based on the state transition diagram for the multiple parallel hysteresis model of Figure 4: If at an arbitrary state (x, z) an indication of a burst is monitored, any number of servers between 1 and z ($z < n-x$) or between 1 and $(n-x)$ ($z > n-x$) may be activated simultaneously. The advanced multiple server activation can be represented in the state transition diagram by a transition between the state (x, z) and $(y, 0), y = x+1, x+2, \cdots, n$. The number of simultaneously activated servers is a parameter which may

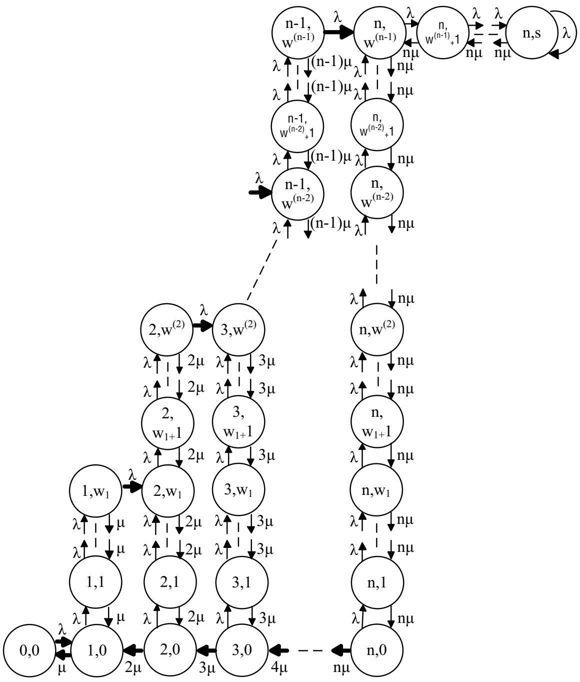

Figure 4: State Transition Diagram for the Multiple Parallel Hysteresis Model

be estimated from the expected size of the remaining burst length. With the simultaneous activation of i servers, the buffer length is reduced by i. For the rest of the burst period, frames are served with a higher (or even the maximum possible) service rate.

4. PERFORMANCE ANALYSIS AND OPTIMIZATION

The analysis of the controlled queueing model of Figure 1 can generally be performed by the discrete event-by-event simulation. Under the special assumption of memoryless arrival and service processes (i.e., negative-exponentially distributed interarrival times of frames and negative-exponentially distributed frame transmission times), all models can be analyzed exactly by recursive algorithms. In [4, 5], the models of Figures 2-4 have been analyzed mathematically by effective recursive algorithms. A numerical example is given in Figure 5 which demonstrates the effectiveness of the self-adaptive algorithms. The example shows the mean waiting time of delayed frames versus the load factor ρ for different widths w of the hysteresis for a 10-server model. The power requirement is directly proportional to the load factor ρ, but has to be paid by an increased delay (case $w = 0$); however, the hysteresis effect produces quite flat curves for the delay and allows an optimum for power saving under prescribed performance limits.

Besides the power saving by dynamic activations/ deactivations of servers and buffers, it is shown quantitatively that power saving has to be paid by a price as enforced buffering increases the frame delays. However, with the parameter space of the hysteresis widths w_i and the various deactivation and advanced activation principles, the parameters can be optimized to meet specific restrictions in the activation/deactivation rate, the mean frame delay or even percentiles of the frame delays.

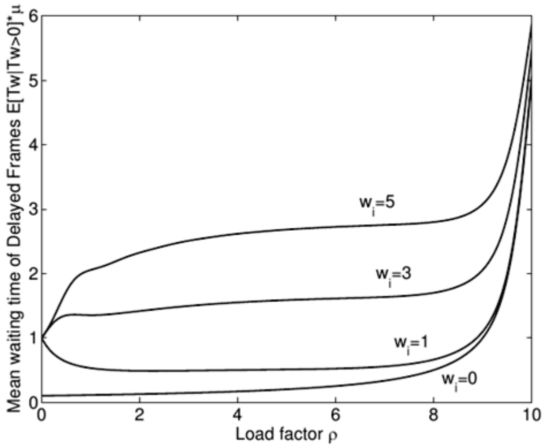

Figure 5: Example Results for the Multiple Parallel Hysteresis Model Mean Waiting Time of Delayed Frames $E[T_w \mid T_w > 0] \times \mu$ versus Load Factor ρ

5. CONCLUSION

In this discussion paper, a variety of algorithms has been discussed systematically for self-adapting activation/ deactivation of servers by means of a queueing system controlled by a Finite State Machine. The basic method of "state-based control" is based on various configurations of serial, parallel and combined serial/parallel hysteresis. To enforce server activations faster in case of burst frame arrivals, a "measurement-/trend-based control" has been proposed which can be modelled by an extension of the state transition diagrams. The FSM-controlled queueing models can be analyzed mathematically by recursive algorithms (see [4]). Further analyses are currently under study and will be reported in a forthcoming paper [5]. An effective and fast mathematical analysis is important for the optimization of system parameters to meet performance bounds.

6. REFERENCES

[1] C. Callegari et al. A new Markov model for evaluating the ALR dual-threshold policy. In *IEEE GLOBECOM Workshops*, 2009.

[2] Y. Fukuda et al. Performance evaluation of power saving scheme with dynamic transmission capacity control. In *IEEE GLOBECOM Workshops*, 2009.

[3] O. C. Ibe and J. Keilson. Multi-server threshold queues with hysteresis. *Performance Evaluation*, 21(3):185 – 213, 1995.

[4] P. J. Kühn and M. Mashaly. Modeling and performance evaluation of self-adapting algorithms for the optimization of power-saving operation modes. In *Proceedings of the 1st European Teletraffic Seminar*, 2011.

[5] P. J. Kühn and M. Mashaly. Performance modeling of power-saving resource management strategies. (submitted), 2011.

[6] J. C. Lui and L. Golubchik. Stochastic complement analysis of multi-server threshold queues with hysteresis. *Performance Evaluation*, 35:19–48, March 1999.