Dynamic Power Management based on variations in Workloads and Hardware

Roadmap

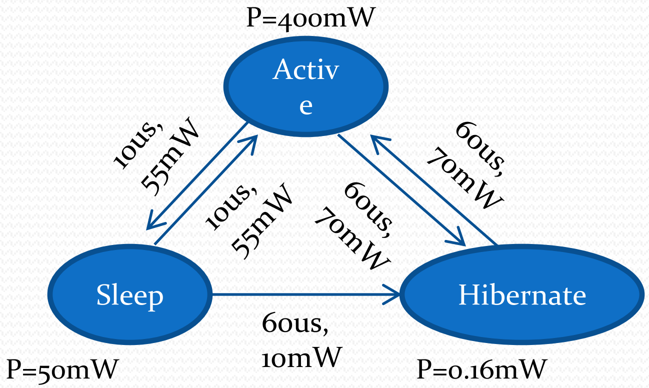
- Motivation
- System-level DPM
- 2-State PSM
- Oracle power policy
  - Break even time (policy parameter)
- Power Management Policies
  - TimeOut
  - Predictive
  - Stochastic
- Adapting to process variations

Motivation

- Longer operation of battery operated and performance constrained portable devices.
- Reduce operation cost of stationary systems.
- Minimize the impact on environment – heat dissipation and cooling-induced noise.
- Avoid performance deterioration due to uncontrolled rise in temperature.

System Level DPM

- Power Manageable Components (PMC)
  - Multiple modes of operations that span power/performance trade-off. Modeled as Power State Machine (PSM)
  - Transition overheads



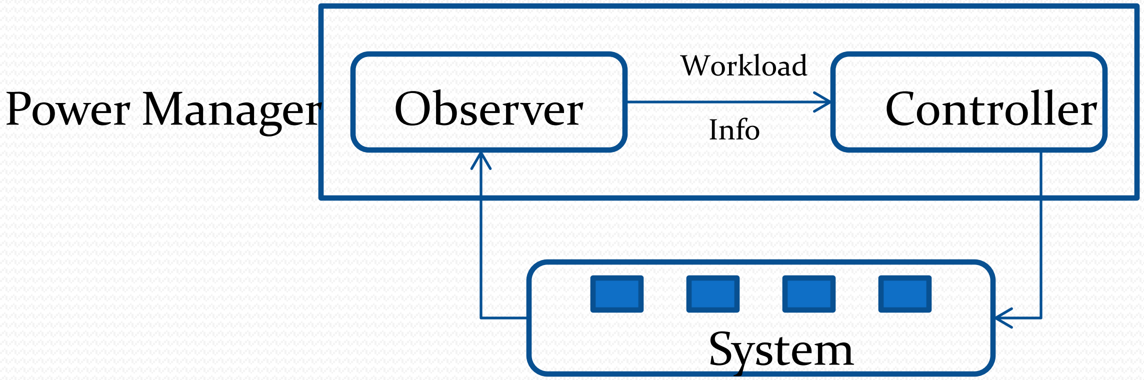
PSM

- Power Manageable System
  - Set of interacting PMC's and non-PMC's
- Power Manageable Network
  - Actively communicating systems
  - Predict communication patterns
  - Decentralized power management

System Level DPM

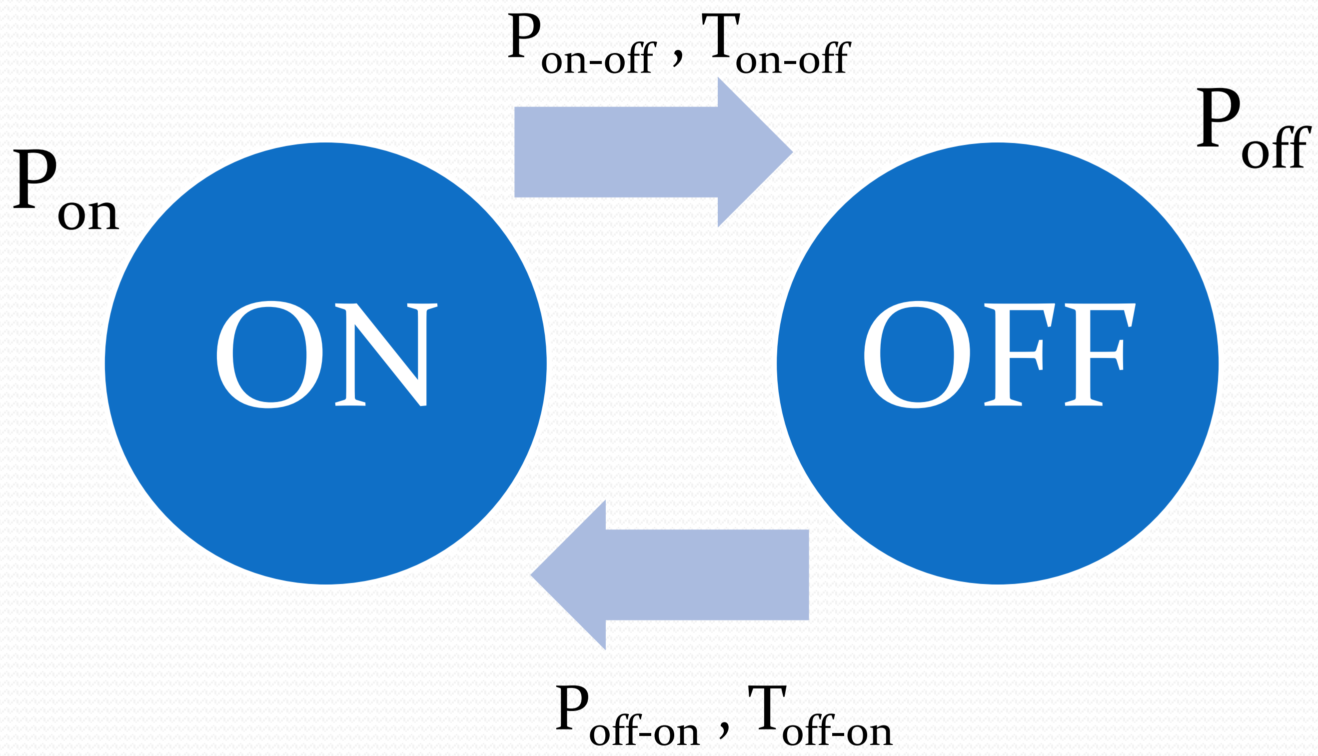
- Task of system level power manager
  - If a PMC is idle and it saves power and thus, cost, to be in a low-power state, then switch the state.

This requires, for every PCM
  - Knowledge of the idle periods i.e. workload model
  - Power/performance characteristics of each state i.e. power state machine



2-State PSM

- We shall be using 2 state PSM



Oracle Power Policy

- Assumptions
  - Workload is known *a priori*
  - Power characteristics of states are known and fixed
- Break Even Time (T_{be})
  - T_{idle} that is long enough to compensate for transition power overheads. So, in case, transition happens at the beginning of T_{idle} and $T_{idle}=T_{be}$ then, net power consumed in transitions and low power state (OFF) would be equal to the power consumed in high power state (ON), if transition won't have happened.
  - Policy parameter
- Policy says if $T_{idle} > T_{be}$, then transition will happen.
- Ideal policy

Power Management Policies

- Classified based on the prediction method whether $T_{idle} > T_{be}$.
  - Time Out
  - Predictive
  - Stochastic

  Each one can be adaptive or non-adaptive

Time Out

- If, $T_{idle} > \tau \;\Rightarrow\; T_{idle} > \tau + T_{be}$, where τ is the time out value.
- Disadvantage: Power wasted during the time out period

Types of time out policies:

- Fixed Time Out (non-adaptive)
- Adaptive Time Out (ATO)
  - τ_{new} is updated depending upon ratio of τ_{old} and previous idle period.
  - Ratio too small, increase τ, else if too large, decrease τ.
- Device Dependent Time Out (DDT)
  - τ is set to T_{be}
  - Atmost twice the energy of 'oracle' policy is consumed

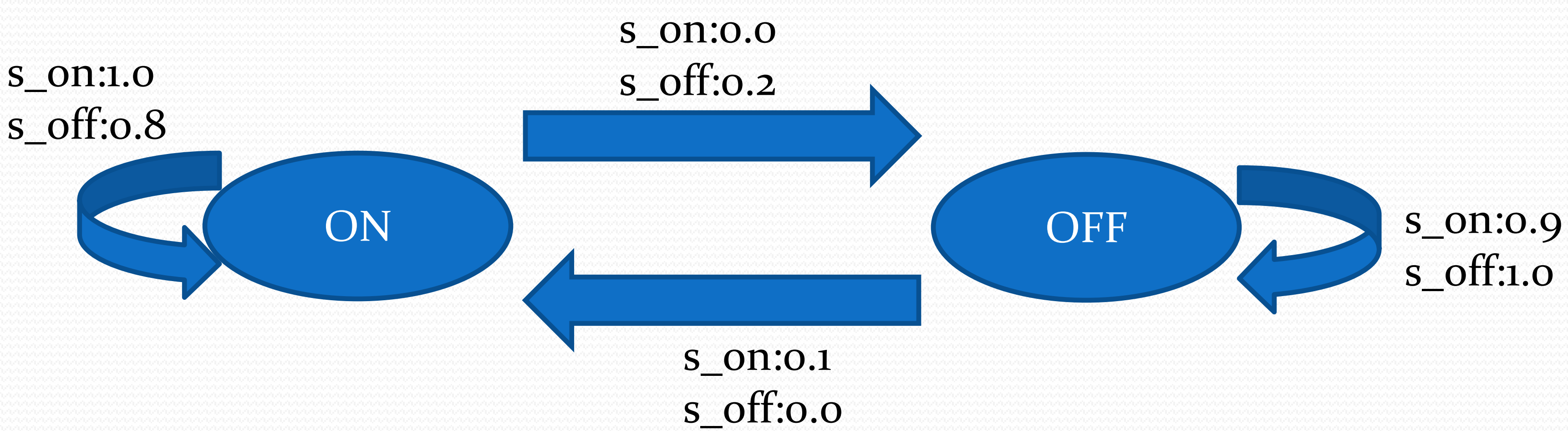
Predictive

- Predict the length of idle period before it begins, thus eliminating the time out period.
- Predictions based on previous busy, predicted idle periods and/or actual idle period(s)
- Adaptive learning tree
  - Similar to multi-bit branch prediction in µP
  - Based on history of the node and confidence level
- Exponential Average policy

$$\rho[n+1] = a * I[n] + (1-a) * \rho[n], a < 1$$

Stochastic

- Models arrival of requests and device power state changes as controlled Markov processes.



- Optimal probability for the device to sleep is obtained from solving stochastic optimization problem

Stochastic

- Static Approach
  - *A priori* knowledge of Markov model for Service Provider (system states) and Service Requester (workload)
- Adaptive Approach
  - Accounts for non-stationary nature of the workloads. Uses *policy pre-characterization, parameter learning and policy interpolation*.

Stochastic

- Can account for
  - Multiple low power states
  - Uncertainty in system power consumption and transition times
  - Compute globally optimum solution
  - Explore trade-offs between power and performance in a controlled fashion
- Disadvantage: Performance and power obtained by this policy are expected values

Hardware Variations Aware Methodology

- Mean value based policy is sub-optimal
- Policy parameter value, T_{be} needs to derived to optimize power consumption distribution instead of deterministic measure of power.
- Two approaches to
  - Design specific: A set of parameter values are derived to be used across all instances, in order to optimize a specific metric of power consumption distribution.
  - Chip specific: Specific to an instance of a chip, set of optimal parameter values is derived depending on its power characteristics