Weekly Report

We focus on the total energy consumption of multi-core systems with heterogeneous voltage scaling. With leakage energy considered, we prove the total energy is a convex function of supply voltage for uniprocessor system. Such convexity can be extened to multi-core system without considering on-chip interconnection overhead. It is hard to analyze the convexity of multi-core system with bus contention in heterogeneous V_{dd} case, due to the approximity of our model for heterogeneous V_{dd} case. Therefore, I suggest we first study the voltage setup and DVS problem without bus contention, assuming a distribute memory architecture. Once we have clear understand and complete solution of the problem, it is easy to extend them by incooperating the bus overhead. Furthermore, I suggest we can borrow the methodologies from studies of process variation to our workload variation awared voltage scaling.

1 Power and Performance Model

We assume the workload contains periodic tasks and one PE needs N cycles to finish one task. To finish one task, the total energy E as function of supply voltage V is given as follows:

$$E = E_{dynamic} + Eleakage \quad (1)$$

$$E_{dynamic} = N \cdot C_{eff} \cdot V^2 \quad (2)$$

$$E_{leakage} = P_{leakage} \cdot \frac{N}{Frequency} \quad (3)$$

$$P_{leakage} = V \cdot K_3 e^{K_4 V_{dd}} e^{K_5 V_{bs}} + |V_{bs}| \cdot I_j \quad (4)$$

$$Frequency = k \cdot \frac{(V - V_{th})^2}{V} \quad (5)$$

where leakage power model is from [1], K_3, K_4, K_5 are empirical constants, V_{bs} is the substract body bias, I_j is the reverse bias junction current, V_{th} is the subthreshold voltage. All of them are constants independent of supply voltage V. Therefore, we can simply the experssion for leakage energy $E_{leakage}$ with new constants A, B, and C as follows:

$$E_{leakage} = \frac{ANV^2 e^{BV} + CNV}{k(V - V_{th})^2} \quad (6)$$

Therefore, total energy E can be presented as:

$$E = N \cdot C_{eff} \cdot V^2 + \frac{ANV^2 e^{BV} + CNV}{k(V - V_{th})^2} \quad (7)$$

2 Convexity of Total Energy for Single Core System

For single core system, it is obvious that total energy E is a convex function of V, if only dynamic energy is considered. With the consideration of leakage energy, we can prove that the total energy E as in (7) is still a convex function of V.

An important porperty of convex function is the *affine operation*, which states if two functions $f(x)$ and $g(y)$ are both convex, then the function $F(z) = f(z) + g(z)$ is convex. Based on this affine operation, we only need to prove that both $E_{dynamic}$ and $E_{leakage}$ are convex. Obviously $E_{dynamic}$ is convex. In the follows, we show the derivation to prove that $E_{leakage}$ is convex.

Convexity of $E_{leakage}$ is equivalent to $\frac{d^2E_{leakage}}{dV^2} > 0$. For $E_{leakage} = \frac{ANV^2e^{BV}+CNV}{k(V-V_{th})^2}$, we have the follows:

$$\begin{aligned}
\frac{dE_{leakage}}{dV} &= \frac{2ANVe^{BV} + ABNV^2e^{BV} + CN}{(V-V_{th})^2} - \frac{2(ANV^2e^{BV} + CNV)}{(V-V_{th})^3} \qquad (8)\\
\frac{d^2E_{leakage}}{dV^2} &= \frac{2ANe^{BV} + 4ABNVeBV + AB^2NV^2eBV}{(V-V_{th})^2}\\
&\quad -\frac{2\cdot(2ANVe^{BV} + ABNV^2e^{BV})}{(V-V_{th})^3} - \frac{4ANVe^{BV} + 2ABNV^2e^{BV}}{(V-V_{th})^3} + \frac{6ANV^2e^{BV}}{(V-V_{th})^4}\\
&= \frac{ANe^{BV}(2+4BV+B^2V^2)}{(V-V_{th})^2} - \frac{4ANVe^{BV}(2+BV)}{(V-V_{th})^3} + \frac{6ANV^2e^{BV}}{(V-V_{th})^4}
\end{aligned}$$

Since A, N, e^{BV}, and $(V-V_{th})^2$ are all greater than zero, $\frac{d^2E_{leakage}}{dV^2} > 0$ is equivalent to

$$\begin{aligned}
& 2+4BV+B^2V^2 - \frac{4V(2+BV)}{V-V_{th}} + \frac{6V^2}{(V-V_{th})^2} > 0 \qquad (9)\\
<=> \quad & (2+4BV+B^2V^2)(V-V_{th}) - 4V(2+BV) + \frac{6V^2}{(V-V_{th})^2} > 0\\
<=> \quad & B^2V^3 - B^2V^2V_{th} - 4BVV_{th} - 2V_{th} - 6V + \frac{6V^2}{(V-V_{th})^2} > 0\\
<=> \quad & (B^2V^3 - B^2V^2V_{th} - 4BVV_{th} - 2V_{th} - 6V)(V-V_{th}) + 6V^2 > 0\\
<=> \quad & B^2V^4 - 2B^2V^3V_{th} - 4BV^2V_{th} + B^2V^2V_{th}^2 + 4BVV_{th}^2 + 4V_{th}^2 + 4VV_{th} > 0\\
<=> \quad & B^2V^2(V-V_{th})^2 - 4BVV_{th}(V-V_{th}) + 4V_{th}^2 + 4VV_{th} > 0\\
<=> \quad & (BV(V-V_{th}-V_{th}))^2 + 4VV_{th} > 0
\end{aligned}$$

where the last inequation is obvious and the convexity of $E_{leakage}$ is proved. Based on the affine operation, the total energy for single core system is a convex function of supply voltage V.

3 Convexity of a Dual-Core System

First we assume the bus contention is not considered. In this case the total energy is simply the sum of energy for two single-core systems. As for each system, the energy is a convex function of their Vs, clearly the sum of them is also a convex function of the supply voltage (V_1, V_2), where V_1 and V_2 are V_{dd} for each core.

	Process variation	Workload variation
Entity	Buffer	Task
Variation	Buffer delay	Task execution time
Method	Adjust buffer size	Adjust V_{dd} on the task
Final value	Total delay	Total task execution time

Table 1: Comparison between process variation aware buffer insertion and workload variation aware voltage scaling.

After that we can consider the bus contention. The change with this new assumption is that the total execution cycle is no long a constant, but a function of both V_1 and V_2. In addition, it depends on the ratio between V_1 and V_2, ratio between benchmark memory accesses and total execution cycle with ideal memory, and the ratio between cycle time and main memory latency. More importantly, our existing bus performance model use an approximation method to convert the heterogeneous V_{dd} cases to homogeneous V_{dd} case (the homogeneous case has precise analytical model.) by deciding the *equivalent* PE number and memory access rate. Such method can hardly be applied to the threotical analysis as above.

On one hand, we can further improve our bus performance model for analysis. On the other hand, we can assume distributed memory architecture for our network processor and avoid the performance penalty by shared bus for the time being. Such architecture is not unreasonable for network packet processing as each packet is independent of others and can choose its memory location freely. As long as the whole payload of a packet resides in the same memory module, there will not be memory access to other memory modules during the process of this packet. The drawback of the distributed memory architecture is the complexity of control logic which decides the memory module for each packet. In our voltage setup and DVS problems, I suggest we first assume the distributed architecture and solve the problems.

4 Workload Variation Aware Voltage Scaling

There is large similarity between process variation and workload variation. Table 1 shows the comparison between process variation aware buffer insertion and workload variation aware voltage scaling. Clearly we can leverage the methodologies in process variation aware buffer insertion for our voltage scaling study. I am still thinking of details and looking at related papers in the recent DAC.

References

[1] S. Martin, K. Flautner, T. Mudge, and D. Blaauw, "Combined dynamic voltage scaling and adaptive body biasing for low power microprocessors under dynamic workloads," in *ICCAD*, Nov 2002.