

An Apparatus for Thermal Management of Multiple Core Microprocessors**FIELD OF THE INVENTION**

[0001] The present invention pertains to the field of integrated circuit design. More particularly, the present invention relates to a method and apparatus for the management for thermal and power management of multiple processor cores on a single die.

BACKGROUND OF THE INVENTION

[0002] An integrated circuit (IC) is a device consisting of a number of connected circuit elements, such as transistors and resistors, fabricated on a single chip of silicon crystal or other semiconductor material. During operation, an IC consumes power causing the temperature of the IC to increase. An overheated IC can potentially result in reduced performance and even operation failure.

[0003] A microprocessor is an example of an IC. Because of higher operating frequencies, the trend in microprocessors is toward increased power consumption and dissipation with every new micro-architecture. In particular, server class processors having multiple processor cores are typically power limited by increasing processor density. A processor core typically includes an instruction register, an input/output bus, a floating point unit, an integer execution unit, a L0 cache, and a L1 cache.

[0004] To help reduce power dissipation, thermal and power management of multiple processor cores on a single IC is desired. The goal is to achieve

maximum compute throughput while keeping the junction temperature below the reliability limit for each processor core.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] The embodiments of the present invention are illustrated by way of example and not in the figures of the accompanying drawings, in which like references indicate similar elements and in which:

FIG. 1A is one embodiment of a multiple core processor floor plan having multiple on die thermal sensors;

FIG. 1B is one embodiment of a processor core having a plurality of thermal sensors;

FIG. 2 is one embodiment of a thermal sensor circuit; and

FIG. 3 is one embodiment of a thermal management unit circuit that computes the frequency of each processor core.

DETAILED DESCRIPTION

[0006] In the following detailed description, numerous specific details are set forth in order to provide a thorough understanding of the invention. However, it will be understood by those skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known methods, procedures, components and circuits have not been described in detail so as not to obscure the present invention.

[0007] Figure 1A depicts a multiple core processor 100 having multiple on-die thermal sensors. For this embodiment of the invention, the multiple core processor 100 comprises eight processor cores 110. Each of the processor

cores 110 is an isolated processor or processing unit. The multiple core processor is not limited to having eight processor cores. A processor core may comprise a floating point unit, an integer execution unit, a L0 cache, and a L1 cache. In addition, a processor core 110 may comprise a plurality of thermal sensors 115. A processor core comprising a plurality of thermal sensors 115 is depicted in figure 1B.

[0008] For this embodiment of the invention, each of the processor cores 110 is coupled to an L2 cache 120. The processor cores 110, however, are not limited to having an individual L2 cache 120. For example, the processor cores 110 may be coupled to a single cache.

[0009] The processor cores 110 and L2 caches 120 are coupled to an input/output (I/O) 130 and a thermal management unit (TMU) 140. The I/O 130 serves as a hardware interface between the multiple core processor and external devices. The TMU 140 is located in a central location of the die and receives the outputs of the thermal sensors 115. The thermal sensors 115 may be placed near hot spots of a processor core 110, such as the floating point unit or the integer execution unit. A processor core 110 may have a number of hot spots. Thus, a plurality of thermal sensors 115 may be used for each processor core 110.

[0010] The TMU 140 monitors the temperature of the thermal sensors 115 and ensures that the processor 100 delivers the maximum throughput without any hot spot exceeding the maximum allowed junction temperature. The TMU 140 may be programmed by software to optimize the highest overall throughput or to give priority to a few application threads running on the processor at the expense of others. For example, the TMU 140 may be programmed using the operating system.

[0011] The TMU 140 controls the operating frequency and operating voltage of each processor core 110. For this embodiment of the invention, each of the processor cores 110 has a different operating frequency, while the entire processor 100 only has two voltages. The top processor cores 110 of the processor 100 operate at a first voltage and the bottom processor cores 110 operate at a second voltage. The invention, however, is not limited to a processor 100 having only two operating voltages. The frequencies of the processor cores 110 are coordinated to match the maximum frequencies the cores 110 can run at each operating voltage. Thus, the frequencies of the top processor cores 110 are limited by the first voltage and the bottom processor cores are limited by the second voltage.

[0012] The thermal sensors 115 may be implemented using an adjustable trip point. When the temperature of a given hot spot exceeds this preset trip point, the TMU 140 is notified. The TMU may then adjust the frequency of that processor core 110 or the voltage of that group of cores to reduce the thermal dissipation of the hot spots to below the trip point. Note that power is defined by the expression

$$P = V^2 * C * f \quad (\text{equation 1})$$

where P represents power, V represents voltage, C represents capacitance, and f is the frequency. From equation 1, it can be seen that reducing the frequency also linearly reduces the power dissipation. Alternatively, power of a processing core 110 may be reduced by lowering the voltage value. However, when an operating voltage of a group of processor cores 110 is adjusted, the frequencies of those cores need to be adjusted correspondingly because the maximum frequency of each processor core 110 is limited by the operating voltage.

[0013] An embodiment of a thermal sensor circuit is shown in figure 2. The thermal sensor circuit comprises a temperature-to-voltage converter 210, a level shifter 220, a ring oscillator 230, and a buffer 240. The temperature-to-voltage converter 210 may comprise a reference bias 212 and a thermal diode 214. The temperature-to-voltage converter 210 produces an output voltage having a limited range. For example, the output voltage may be in the range of 0.3 volts to 0.7 volts. The level shifter 220 takes the output of the temperature-to-voltage converter and generates a biasing voltage. This generated voltage is then input to the ring oscillator 230.

[0014] For this embodiment of the invention, the ring oscillator comprises a plurality of inverters 232. The number of CMOS inverters 232 is an odd integer greater than one in order to generate a toggling signal having a frequency. The number of inverters 232 and the delay of each inverter 232 in the oscillator 230 help to determine the generated frequency. Each of the plurality of inverters 232 is coupled to a pull down transistor 234. The strength and frequency of the signal generated by the oscillator 230 also depend, in part, on the pull down transistors 234 and the biasing voltage that controls the pull down transistors 234. As a result, the frequency of the signal generated by the oscillator 230 is a function of the temperature sensed by the thermal diode 214. An increase in temperature causes the frequency of the signal generated by the oscillator 230 to decrease, while a decrease in temperature causes the frequency of the signal generated by the oscillator 230 to increase. The buffer 240 ensures the signal generated by the oscillator 230 has enough drive strength to reach the TMU.

[0015] For another embodiment of the invention, the ring oscillator 230 comprises a plurality of differential amplifiers. Similar to the inverter design

described above, the number of differential amplifiers is an integer greater than one. Each of the differential amplifiers may be biased by the biasing voltage generated by the level shifter 220.

[0016] An embodiment of a TMU circuit in a processor is depicted in figure 3. This TMU circuit comprises a plurality of counters 310. Each of the counters 310 is coupled to a digital controller 340. Moreover, a first counter 310 may be coupled to a reference thermal sensor 320. All other counters 310 may be coupled to thermal sensors 330 located in remote areas of the die such as processor cores. The remote thermal sensors 330 are monitored by the digital controller 340. Specifically, the digital controller 340 may monitor the frequency encoded temperature information from each remote thermal sensor 330. The digital controller 340 compares the remote temperatures against the local temperature reference as sensed by the reference thermal sensor 320.

[0017] For one embodiment of the invention, the digital controller 340 may compare the time in which a first counter 310, clocked by a first signal generated by a reference thermal sensor 320, takes to count to a predetermined value against the time it takes for a second counter 310, clocked by a second signal generated by a remote thermal sensor 330 to reach the same predetermined value. The digital controller 340 controls when the counters 310 start and stop counting. If the counter 310 clocked by the reference thermal sensor 320 reaches the target value before the counter 310 clocked by the remote thermal sensor 330, the digital controller decreases an operating frequency or an operating voltage of a processing core. It follows that if the counter 310 clocked by the remote thermal sensor 330 reaches the target value before the counter 310 clocked by the reference thermal sensor 320, the digital controller increases an operating

frequency or an operating voltage of a processing core. Therefore, adjustments to each processor core voltage and frequency are determined by the digital controller 340 in order to maximize the overall compute throughput of the processor. The operating frequency of each processing core may be adjusted by adjusting the phase locked loops (PLL) that provide clocks to each core.

[0018] A look-up table within the digital controller 340 may determine the adjustment values to the frequency and the voltage. Thus, if the difference in temperature between the reference thermal sensor 320 and a remote thermal sensor 330 is large, the adjustment will be greater than if the temperature difference is small. For example, to adjust the frequency of a processing core, the digital controller 340 may look-up a multiplying ratio value of the PLL. The value of the selected multiplying ratio is a function of the difference in temperature between the reference thermal sensor 320 and a remote thermal sensor 330 in that processing core. The frequency of the clock generated by the PLL is dependent upon the selected multiplying ratio.

[0019] Alternatively, the adjustment values may be a single step value. For example, the digital controller 340 may be coupled to an oscillating structure of a PLL that generates a processor clock. Under this implementation, as long as the difference in sensed temperature between a remote thermal sensor 330 and a reference thermal sensor 320 is greater than a specified limit, an adjustment in frequency or voltage will be made by the digital controller 340 at the step value. While this single step value implementation may not be as quick in reducing power dissipation on a processor, it reduces the design complexity and takes up less die area than a look-up table implementation.

[0020] For another embodiment of the invention, the digital controller 340 compares the number of signal transitions of a signal generated by a reference thermal sensor 320 against the number of signal transitions of a signal generated by a remote thermal sensor 330 over a given period of time. The counters 310 are used to count the signal transitions from the reference thermal sensor 320 and the remote thermal sensor 330 signals. The larger the count value over the given time period, the cooler the sensed area. Thus, if the area of the reference thermal sensor 320 is determined to be cooler than the area of a given remote thermal sensor 330, the digital controller 340 reduces the operating frequency or the operating voltage of the processor core of where that remote thermal sensor 330 is located. Similarly, if the area of the reference thermal sensor 320 is determined to be hotter than the area of the remote thermal sensor 330, the digital controller 340 increases the operating frequency or the operating voltage of the processor core of where the remote thermal sensor 330 is located.

[0021] In the foregoing specification the invention has been described with reference to specific exemplary embodiments thereof. It will, however, be evident that various modification and changes may be made thereto without departure from the broader spirit and scope of the invention as set forth in the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than restrictive sense.

What is claimed is:

1. A system comprising: a first thermal sensor sensing a first temperature of the system and generating a first digital signal having a frequency representing the first temperature; a second thermal sensor sensing a second temperature of the system and generating a second digital signal having a frequency representing the second temperature, wherein the second temperature is a reference temperature; a first counter coupled to the first thermal sensor and enabled by the first digital signal, the first counter counts in a first direction in response to a transition of the first digital signal; a second counter coupled to the second thermal sensor and enabled by the second digital signal, the second counter counts in a second direction in response to a transition of the second digital signal; and a digital controller coupled to the first counter and the second counter, wherein the digital controller compares the count values of the first counter and the second counter, wherein if the second counter reaches a target value before the first counter, the digital controller decreases an operating voltage of the system.
2. The system of claim 1, wherein the operating voltage decrease is at a step value.
3. The system of claim 1, wherein the digital controller adjusts an operating frequency of the system.
4. A method comprising: monitoring a reference thermal sensor; monitoring a plurality of thermal sensors placed in a processor having a first and a second processor core; and increasing an operating frequency of the first processor core to improve throughput of the first processor core if a temperature at the reference thermal sensor is greater than a temperature at a thermal sensor in the first processor core.
5. The method of claim 4, further comprising: decreasing an operating voltage of the first processor core if the temperature at the thermal sensor exceeds the temperature at the reference thermal sensor.
6. The method of claim 5, wherein the operating voltage adjustment is made according to a look-up table, wherein the greater the temperature at the reference thermal sensor is exceeded, the greater the operating voltage is adjusted.

23 February 2007

7. The method of claim 4, wherein the plurality of thermal sensors sense the temperature of a floating point unit of the first processor core.

8. The method of claim 4, wherein the plurality of thermal sensors sense the temperature of an integer execution unit of the first processor core.