Understanding Power Management of Intel® Processors for Mil/Aero Applications

Introduction

With the advent of the Intel® Core™ i7 processors, the 2nd Generation of which feature 32nm silicon geometry and up to four cores, it is now possible for board designers to fit new levels of functionality and performance into traditional 3U and 6U form factor boards. And while the efficiency of these new processors (performance/Watt) is improved over previous generations, the amount of power that can be consumed, along with the associated heat, has also risen. Add to that the additional thermal challenge posed by extracting this heat from increasingly small silicon surface area.

For high performance multi-processor designs, where the aim is to provide the maximum possible compute performance in a given volume, we are entering an era where such boards may need to be operated at a de-rated performance level when operated at the industry standard 85°C card-edge, and 71°C inlet temperature conditions. This introduces a new element of power management for system designers to consider.

This paper discusses the behavior of the Intel Core i7 processor, as implemented on the Curtiss-Wright Controls Embedded Computing multi-processor CHAMP-AV series products at elevated temperatures, and how to take advantage of board features to maximize the performance potential of the system.

The Challenge

Designers of thermal solutions for high performance processor boards are contending with two technology trends. Processor power continues to increase, while the size of the processor die continues to decrease. For example the power density of the Freescale™ 8640 processor is about 11W/cm² (see note 1). The 2nd Generation Intel Core i7 is greater than 20W/cm² and is close to 50W/cm² in localized areas of the CPU. A conduction-cooled card typically has an 85°C card-edge temperature specification. CPUs have maximum die temperatures in the 100°C to 105°C range. Therefore the thermal design has a 15 to 20°C total temperature rise budget from die to card edge. This increased power density means that a greater proportion of this temperature budget is expended in the interface between the CPU and the contact point of the heat sink, making the design of the cooling solution critical to the performance that can be attained at elevated temperatures.

The good news is that these new processors can do more work per unit of energy, but the bad news is that under environmental conditions of extreme heat, their operating frequency may have to be de-rated in order to maintain die temperatures within specification.
Intel Core i7 Thermal Management Explained

The latest Intel processors have a number of features related to thermal management that need explanation (and perhaps for some readers, correction of common misconceptions).

First, some definitions of terms:

**Adaptive Thermal Monitor**
The name for the collective set of sensors and circuits that operate to control the maximum temperature of the processor.

**Digital Thermal Sensor (DTS)**
The processors feature a number of thermal sensors with internal analog to digital conversion. The sensors report a relative temperature from a reference temperature. There are multiple sensor locations on the processor die.

**Platform Environmental Control Interface (PECI)**
PECI is an external interface signal from the processor to report to external chipset components the state of the DTS sensors. The PECI circuit monitors all of the DTS sensors and reports a rolling average of the highest sensed temperature.

**TjMax**
The maximum specified die temperature for the device. The TjMax value is a property of the specific processor part number (SKU). Typically TjMax is 100°C to 105°C.

**Thermal Control Circuit**
The feature of the processor that reduces the operating frequency in response to over-temperature conditions.

First and 2nd Generation Intel Core i7 processors have a feature that will optionally reduce the operating frequency of the CPU in response to over-temperature conditions. Intel refers to the collective operation of sensors, micro-controller and clock control as the Adaptive Thermal Monitor. When enabled, an autonomous controller within the CPU will dynamically adjust (“throttle”) the operating frequency of the CPU, with the objective of keeping the CPU die temperature at or below the maximum (see note 2). The minimum and maximum clock rate is referred to as Low Frequency Mode (LFM) and High Frequency Mode (HFM) respectively. The range from HFM to LFM is CPU model-specific but is approximately 2:1 for some embedded Intel Core i7 models used by Curtiss-Wright. In concert with the clock frequency changes, the Adaptive Thermal Monitor will also adjust the voltage of the CPU power supply (it can be lower voltage at lower frequency). Since CPU power approximately correlates with V2F, lowering both has a significant effect on reducing power. If the frequency/voltage reduction does not resolve the over-temperature condition, the processor will employ a more aggressive clock modulation strategy, where the clock will be active for a 37.5% duty cycle over a 32 microsecond period.

**Understanding Intel Processor Clock Throttling**
The clock throttling feature can be disabled. In the case of a Curtiss-Wright board this is controlled via settings in the BIOS (see note 3). Disabling the throttling will allow the CPU to run at higher temperatures, however doing so will be operating the CPU beyond its guaranteed operational limits. For most systems, operating in this region would only be appropriate for emergency situations. A correctly functioning system should not have occasion to operate the processors above their rated temperature. Even with the clock throttling feature disabled, the Intel Core i7 processor will shut down at a highly elevated temperature of approximately 130°C. This hard wired behavior is to protect the device from catastrophic failure. This should not be a concern for system designers, since at such a temperature, the board is far above its specifications with a reasonable likelihood that other functions of the card have failed due to over-temperature.
than the large majority of actual applications. In the example shown, the CPU never has excursions to the minimum clock frequency, and was running at approximately 85% of maximum on average, while running at 85°C card-edge.

Sensors Are Key to Thermal Management

We have established that high power cards, working at the extremes of their temperature, will need to be monitored and controlled to ensure they are operating within specification. Curtiss-Wright’s Intel Core i7 board products provide a large number of sensors to assist with thermal management strategy. In Figure 3, the top side of the CHAMP-AV5 dual Intel Core i7 610e processor board is shown. There are two types of temperature sensors, internal and external. Many large scale ICs have internal temperature sensing. In the case of the AV5, there are a total of five such sensors for the CPU (2), PCH (2) and the Core Functions FPGA. The CPU sensor is depicted as a single device but in fact there is a more elaborate implementation in the CPU, with multiple sensing points and an internal controller that reports a rolling average of the hottest.

In Figure 1 and Figure 2, the clock throttling feature can be observed in action. These graphs show the clock frequency, expressed as percentage of maximum, versus card edge temperature. This test was done with software designed to heavily burden the processors, memory and PCI Express® switch components, resulting in power consumption that should be greater
The other temperature sensors of interest are external sensors, located to approximately measure the card edge temperature and the temperature of the center of the board. These are implemented with discrete devices that connect to the CPU via an SMBus interface. Even though there are four different types of sensors, there is a single sensor reading software API that hides this complexity from the application programmer.

The Intel DTS implementation adds some complexity to the task of monitoring temperature. It is the CPU that is the single largest consumer of power and is therefore the device that needs most careful monitoring. The DTS sensor will not report temperatures in excess of $T_{j\text{Max}}$ (see note 4). If during system testing the readings from the DTS sensor stay below $T_{j\text{Max}}$, there is no issue of over-temperature. For a system that is both demanding of the CPU (driving power up) and is subjected to high temperatures, we may find a situation where at full clock rate the CPU will either throttle (if enabled) or exceed the allowable temperature. There are means to detect both. The AV5 Board Support Package (BSP) provides a function to read the instantaneous operating frequency of the CPU. This was used to plot the data in Figure 1 and Figure 2. Any reading less than HFM indicates the CPU has entered into throttling. If the CPU has the throttling feature disabled, then it demands that the system designer use the thermal sensors. By reading both the CPU temperature sensor (which tops out at $T_{j\text{Max}}$) and the card-edge sensors, one can correlate the temperature drop between these two points. For a constant power dissipation, the delta between the CPU and the card edge sensor will remain constant (see Figure 4). So if the CPU is measured to be within a few degrees of maximum, but the card edge is 10 degrees away, then one can safely assume the CPU will rise at least another 10 degrees at full card-edge temperature. In this scenario, one would have to take steps to reduce the CPU power (or reduce XMC/PMC power which will raise the amount of power that can be dissipated by the CPUs).
Reducing Power Consumption

A very useful feature of the Intel Core i7 processors is SpeedStep which provides for a range of different operating clock rates of the processor. This is the primary method that a system designer can use to reduce the power consumption of an Intel-powered board. Implementations may differ. In the case of the CHAMP-AV5, it is possible to dynamically program the clock speed from an application program. Intel refers to the different available clock frequencies as P-states. In the case of the 2.53 GHz Core i7 610e, the CPU is capable of operating at speeds between 1.20 GHz and 2.53 GHz in 133 MHz increments. Combined with the on-board sensors, it is a relatively simple task to test one’s software at different clock speeds while verifying the operating temperature of the key components on the board.

Conclusions

High performance multi-processor boards may need de-rating at maximum inlet/card-edge temperatures. This will depend on the nature of the software, utilization of memory and other I/O functions on the board, as well as the use of mezzanine modules. Intel processors can optionally throttle their clock rates to control temperature. However, many real-time system designers may prefer to disable this feature and manually choose an operating frequency to manage the CPU temperatures and/or minimize the board power consumption for a given workload.

Curtiss-Wright can provide test data showing the thermal characteristics of these boards under a high utilization regime. This data can be used for planning purposes to estimate cooling and performance. Since the user application has a significant impact on actual power consumption, designers are advised to measure the consumption and temperature sensor readings for their particular use of the card. Although not mentioned in the paper, the CHAMP-AV5 also features a board power sensor which can also assist in understanding the power/thermal characteristics of a system.

The SpeedStep dynamic clock frequency feature can be used to minimize power consumption in systems with variable workloads. For instance a radar system may have a high-power attack mode that would normally not be engaged. A system could be engineered to operate in a low power mode most of the time with occasional usage in a high power mode. The transition between p-states is very fast.

Notes

1. Based on the low-voltage 1.06 GHz dual-core Freescale MPC8640 with a 20 W power consumption. This is the most power efficient speed-grade for the 8640 family.
2. The maximum temperature for an Intel CPU is model number specific. Core i7 processors in Curtiss-Wright CHAMP-AV boards are rated for 100°C and 105°C.
3. On Curtiss-Wright CHAMP-AV5 and CHAMP-AV8 products the default setting for clock throttling is enabled for Level 0 air-cooled cards and disabled for rugged cards.
4. The Intel DTS sensor expresses temperature as the delta below TjMax. (i.e. 0=TjMax, 10=TjMax-10) so once the CPU die is above this temperature, the DTS ceases to indicate by how much. Note also that the DTS sensor is designed to be most accurate near TjMax and bottoms out at a relatively warm temperature.

Intel, Intel Core, Intel SpeedStep and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.
*Other names and brands may be claimed as the property of others.
Copyright © 2011, Intel Corporation. All rights reserved.