

## 16.7 Power and Temperature Control on a 90nm Itanium®-Family Processor

Christopher Poirier, Richard McGowen, Christopher Bostak, Samuel Naffziger

Intel, Fort Collins, CO

This paper describes the embedded feedback and control system on a 90nm Itanium®-family processor, code-named Montecito[1], that maximizes performance while staying within a target power and temperature (PT) envelope. This system, referred to as Foxton technology (FT), utilizes on-chip sensors and an embedded microcontroller to measure PT and modulate both voltage and frequency (VF) to optimize performance while meeting PT constraints. Changing both VF, which takes advantage of the cubic relationship of  $P=CV^2F$ , is described for a PDA processor in [2]. Montecito is able to implement FT using only 0.5% of the die area and 0.5% of the die power.

An on-die discrete-time control system measures PT and controls VF as shown in Fig. 16.7.1. The major components of this system are an embedded 16b-instruction/32b-data microcontroller (FTC) and 4 on-die A/D converters. The FTC implements the control system and the A/D converters provide the measurements. Each A/D converter has an analog input MUX that allows measuring system parameters such as voltage, temperature, and calibration-specific signaling.

The embedded firmware within the FTC implements the control system that has a sampling interval (T) of 8 $\mu$ S. Every T, a power, temperature, or calibration measurement is made. Power is measured by utilizing the inherent resistance in the power plane between the edge connector and the die to compute current and thus power. This approach minimizes additional circuitry, wastes no additional power, and provides complete ammeter integration in the package independent of the power supply. The FTC uses the measured power to determine new voltage commands to send to the separate core and cache power supplies. A power error term is computed by subtracting the measured power from a programmable reference power ( $P_{REF}$ ). The error term is then multiplied by a forward gain factor and a new target voltage is calculated. An IIR filter is then applied to maintain system stability. Finally, the new voltages are clipped to remain within the core and cache operating specifications before being sent to the supplies.

System operation is demonstrated in Fig. 16.7.2 which shows that the system can maintain a target  $P_{REF}$  across an activity factor change from min to max by changing VF. The FTC requests a voltage change and a separate, independent sub-system changes the frequency when it detects the voltage change. The clock-generation system includes this sub-system and is described more fully in [3]. FT will only raise the voltage to a point where frequency increases are still possible, preventing wasted power.

The FTC is an on-die custom microcontroller with DSP-like capabilities. It has 4k  $\times$  16b of instruction memory and 4k  $\times$  8b of data memory for firmware support. The embedded firmware is responsible for external interfacing and for scheduling power control, temperature control, and analog calibration routines. Every T, the scheduler is invoked by a real-time counter to call the appropriate power, temperature, or calibration routine. The power routine measures power, performs closed loop power control, and insures system stability using a transfer function  $G_d(Z)$  implemented with an IIR. The temperature routine is responsible for reading the 4 thermal sensors and communicating with the power-control routine, insuring that the junction temperature ( $T_j$ ) never exceeds 90°C. Temperature control is accomplished by limiting the output voltage which the power control algorithm

can request of the power supply. Finally, the calibration routine insures the specified power-measurement accuracy.

The four on-die A/D converters are implemented using high-speed VCO and counters. Typical VCO frequencies are 10GHz with a 1V input that, when counted over an 8 $\mu$ S interval, supports a resolution of 12.5 $\mu$ V. In practice, this is an order of magnitude below the system noise floor of 100 $\mu$ V. Anti-alias filtering is accomplished in part by the averaging that naturally occurs from counting VCO pulses for 8 $\mu$ S. Additional filtering is accomplished by averaging two consecutive counts. Careful design of the VCO ensures monotonicity and linearity as shown in Fig. 16.7.3 and Fig. 16.7.4.

Since the voltage-to-frequency transfer function of a VCO can vary, each VCO is individually calibrated at run time using a band-gap reference voltage and a voltage ladder. By selecting a tap on the voltage ladder, the system generates a count-to-voltage conversion table to utilize in later measurements. The band-gap reference voltage and resistor ladder are shown in Fig. 16.7.5.

The measurement of current is an integral part of the power calculation. The die current is calculated by measuring the voltage drop across the package's power plane resistance,  $R_{PKG}$ , and dividing by that resistance. Since  $R_{PKG}$  varies with manufacturing process and temperature, the FT system repeatedly measures this resistance. It does this by performing two package voltage drop measurements,  $V_1$  and  $V_2$ . The die is first put into a state where it will draw a constant, but unknown, amount of current. The first measurement,  $V_1$ , measures the drop across  $R_{PKG}$  in this state. The second measurement,  $V_2$ , uses on-die precision current sources ( $I_{SOURCE}$ ), to draw additional current from the supply beyond the ambient current in the measurement of  $V_1$ . With these two measurements,  $R_{PKG}$  can be calculated as follows:

$$R_{PKG} = (V_2 - V_1) / I_{SOURCE}$$

The process of quieting the cores, settling, taking  $V_1$  measurements, firing  $I_{SOURCE}$ , settling, and taking  $V_2$  measurements, takes 56 $\mu$ S and impacts core performance less than 0.1%.

The four on-die thermal sensors are comprised of a precision current source and a forward-biased diode with a -1.6mV/°C temperature coefficient. These sensors are located in the floating point and integer units of each core to detect localized hot spots for varying code workloads. See Fig. 16.7.6 for a TPCC balanced workload thermal map. The FTC uses the A/D converters for measuring these diode voltages and utilizes manufacturing calibrated values for each sensor to calibrate out any offset voltages at the 90°C temperature point.

The FT system allows software to further optimize the power/performance efficiency of a system. An example of this is ACPI P-states and demand based switching [4].

### Acknowledgements:

The authors would like to acknowledge Eric Fetzer, Scott Anderson, Jane Wang, Jim Ignowski, Russ Mason, and Warren Parks.

### References:

- [1] S. Naffziger et al., "The Implementation of a 2-Core, Multi-Threaded Itanium®-Processor," *ISSCC Dig. Tech. Papers*, Paper 10.1, pp. 182-183, Feb., 2005.
- [2] S. Akui et al., "Dynamic Voltage and Frequency Management for a Low-Power Embedded Microprocessor," *ISSCC Dig. Tech. Papers*, pp. 64-65, Feb., 2004.
- [3] T. Fischer et al., "A 90nm Variable-Frequency Clock System for a Power-Managed Itanium®-Family Processor," *ISSCC Dig. Tech. Papers*, Paper 16.2, pp. 294-295, Feb., 2005.
- [4] D. Bodas, "New Server Power-Management Technologies Address Power and Cooling Challenges," *Technology at Intel Magazine*, Sept., 2003.

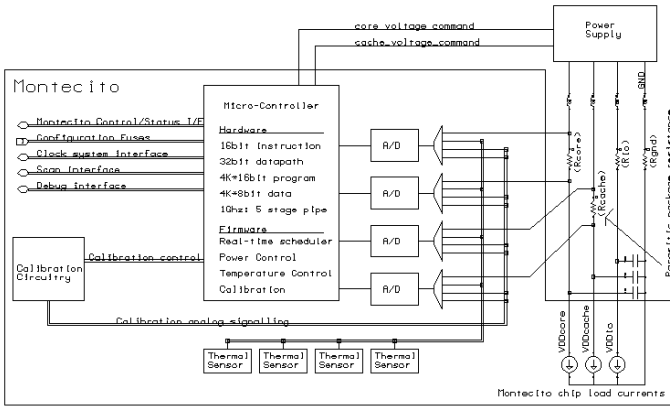


Figure 16.7.1: System block diagram.

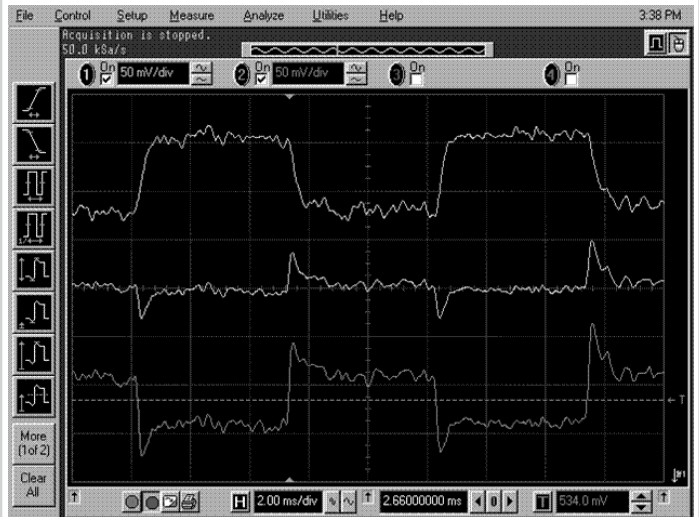


Figure 16.7.2: Voltage, power, and current waveforms.

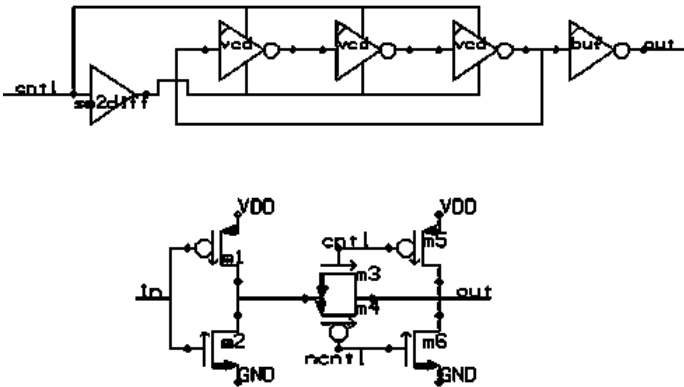


Figure 16.7.3: VCO and VCD schematics.

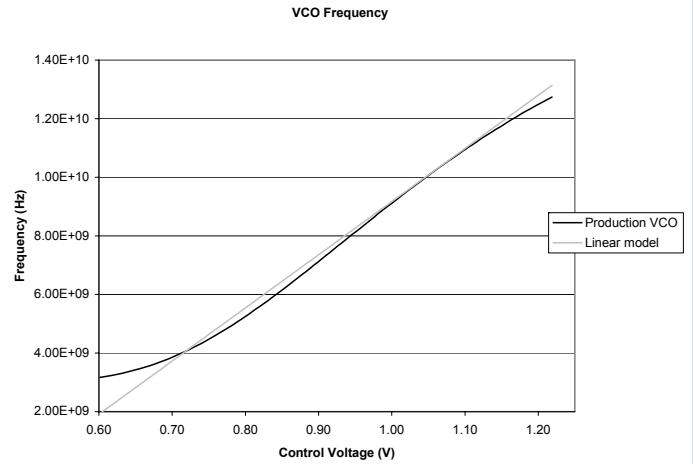


Figure 16.7.4: VCO frequency versus voltage.

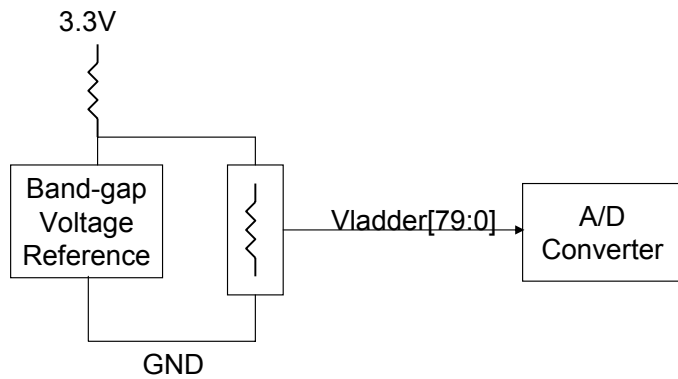


Figure 16.7.5: Band-gap reference and resistance ladder.

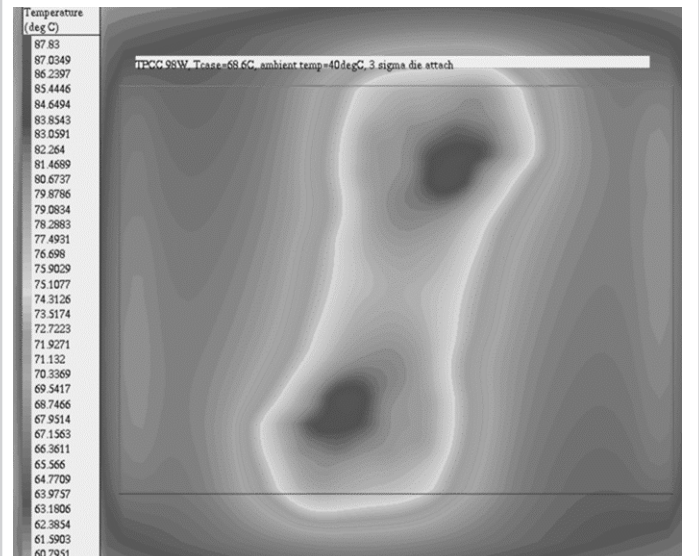


Figure 16.7.6: Montecito thermal map.