DYNAMIC POWER MANAGEMENT WITH REAL-TIME THERMAL CALCULATIONS FOR PROCESSORS

APEC 2013

Maurice Steinman
Senior Fellow
Client SoC Architecture & Modeling
March 20, 2013
AGENDA

- SOC High-level Architecture
- SOC Power Budget Allocation and TDP Headroom
- Dynamic Power Management Evolution:
  - State-based Frequency Boost
  - Digital Activity Monitoring and Power Calculation
  - Real-time Thermal Calculations
- Managing infrastructure current limits
- Summary
- Q & A
AMD SOC ARCHITECTURE EXAMPLE- A-SERIES "TRINITY"

Dual-channel DDR3 Memory Controller: up to DDR3-1866

Dual-channel DDR3 Memory Controller: up to DDR3-1866

AMD HD Media Accelerator (UVD, AMD Accelerated Video Converter)

Up to 4 AMD "Piledriver" Cores with 2MB L2

PCI Express® I/O — 24 lanes, optional digital display interfaces

HDMI™, DisplayPort 1.2, DVI controllers

L2 Cache

D2 Cache

L2 Cache

Dual-core x86 Module

Dual-core x86 Module

PCIe®

Disp. PLL

DP / HDMI

Display Controller

AMD Radeon™ GPU
(up to 384 AMD Radeon cores 2.0)
SOC POWER BUDGET ALLOCATION

x86 Processor Power

GPU Power

MM Power

I/O Power

Display PHY Power

Memory PHY Power

ROC Power

SOC Thermal Design Power (TDP)

Cooling Solution

Fan

Heat Sink

Lid

Microprocessor Die

Socket
TDP HEADROOM – HOW TO DETECT AND EXPLOIT

- Power and performance varies greatly by workload
- AMD Turbo CORE power estimation evolution:
  - Version 1.0: Count active cores
  - Version 2.0: Calculate power using digital monitors
  - Version 3.0: Calculate die temperatures in real time
- High frequency is used when power or thermal limit allows – high frequency for high performance within the same envelope
- Power budget can be dynamically allocated to different compute units (in processor and GPU)
- Over-thermal design current (over-TDC) protection
AMD Turbo CORE 1.0 – STATE-BASED BOOST

- Count the number of cores in active state
- If less than some threshold (in this case, 3), boost the voltage and frequency of the remaining cores

**Active Core Residency**

- **All cores active**
  - 10W 10W 10W
  - 10W 10W 10W

- **3 or fewer cores active**
  - 1W 1W 10W
  - 1W 10W 10W

<table>
<thead>
<tr>
<th>Cores Active</th>
<th>Example Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-6</td>
<td>3.1 GHz</td>
</tr>
<tr>
<td>1-3</td>
<td>3.6 GHz</td>
</tr>
</tbody>
</table>

**Client Workload Suite**
AMD Turbo CORE 2.0 – ACTIVITY MONITORING AND POWER CALCULATION

Digitally measure activity to estimate power.
Then dither p-state to stay within the selected chip TDP.
Enabled higher frequencies even when all cores active due to ability to measure workload power consumption.

Power headroom that can be utilized per core.
**AMD Turbo CORE 2.0 – REDUCED CORE COUNT BOOST**

- Low activity in one core enables it to be a thermal sink for a more active core
  - The heat transfer isn’t 1:1 (i.e., 5W less in one core does not enable another to consume a full 5W)
- The power management controller applies the power density multiplier when one or more cores are in the power-gated state
  - The GPU can borrow power credit from the CPU in GPU-centric scenarios
- Time constant of power transfer between cores done with a simple moving average algorithm
- Enables smooth increases in frequency for varying reductions in active core count as opposed to the single threshold of AMD Turbo CORE 1.0
- Frequency opportunity from this capability is large: 25% to 30% higher frequency for reduced-thread apps
Rather than using power numbers as a proxy for temperature (only true for sustained DC activity), “Trinity” adds firmware and logic to calculate on-die a deterministic estimate of the temperature of GPU and each CPU.

– Relies on a calculated electrical power dissipated by each core and GPU
– Calculates thermal influence of die and total cooling solution

- Modeled by a thermal resistance and capacitance network

\[
T(t) = \sum_{n=1}^{k} R_n \cdot (1 - e^{-\frac{t}{R_n C_n}})
\]
AMD Turbo CORE 3.0 TECHNOLOGY: OVERVIEW

UTILIZE CALCULATED AVAILABLE DYNAMIC THERMAL HEADROOM TO IMPROVE PERFORMANCE

- SoC divided into “thermal entities” (TE)
  - A TE represents a computational unit that can report its calculated power and a calculated thermal density
    - TEs include CPU0/1, 2/3, GPU, and I/O components

- Thermal RC network
  - Coefficients that describe thermal transfer between TEs, substrate, and package are characterized.
  - Firmware on the management processor calculates per-TE temperatures using heat transfer coefficients and calculated TE power.
  - TE voltage/frequency adjusted according to defined temperature throttle points and workload heuristics.
AMD Turbo CORE 3.0 TECHNOLOGY: TEMPERATURE CALCULATION

- **CPU/GPU temperature**
  - Firmware regularly calculates instantaneous temperature for each TE based on the power estimate and prior temperature
  - Uses a 5-stage thermal RC ladder

- **Other silicon contributors**
  - High-speed I/O interfaces and Northbridge are modeled as power and/or temperature offsets to simplify calculations
  - This has limited impact on accuracy

- **Measured error of +/-5°C on 3DMark® analysis**

- **Algorithm provides deterministic operation and reproducibility of results**
**AMD Turbo CORE 3.0 TECHNOLOGY: CALCULATED VS. MEASURED TEMPERATURE**

Experimental results for engineering review, no observable product functional operational difference results from thermal differences. No claims made to accuracy.

Estimated +/- 3-5°C difference in calculated hotspot vs. measured hotspot temperature at steady thermal state.
**AMD Turbo CORE 3.0 TECHNOLOGY: DYNAMIC FINE-GRAINED POWER TRANSFERS**

- AMD Turbo CORE 2.0 incorporates a simple binary power transfer from GPU->CPU if GPU activity is low.

- With AMD Turbo CORE 3.0, the dynamically calculated temperature of each core and the GPU enables the operating point of each is dynamically balanced to maximize performance within temperature limits.
TDP WORKLOAD SCENARIOS WITH AMD Turbo CORE 3.0

- GPU LightWeight
- GPU Idle
- 3D Game
- 3D Bench Mark
- OpenCL™

Thermally sustainable

CPU
P0, P1, P2

GPU

ROC

Avg. Power

Avg. Power

Avg. Power

Avg. Power

Avg. Power

Avg. Power
AMD Turbo CORE 3.0: CURRENT-LIMITING

Electrical design current (EDC) tracking logic enforces maximum duration of continuous high current draw.

Thermal design current (TDC) tracking logic forces hardware P-states to switch between HWP2 and HWP3 to keep moving average within longer thermal time constant limits.
AMD Turbo CORE 3.0: MANAGING TO INFRASTRUCTURE LIMITS
(NOT DRAWN TO SCALE)

Electrical design power (EDP)

- 3: EDC tracking algorithm enforces maximum duration of operation at EDC limit.

Maximum boost power (MBP)

- 4: CPU cores in HW P2 state.

Traditional TDP specification

- 6: CPU Cores in HW P3 state.

System idle (<1W)

- 1: Power Virus launched: Initial "cool" condition allows all 4 cores to boost to maximum allowed frequency for 4 threads.

- 2: All CPU cores in P-Boost (HW P1).

- 7: Calculated high temp. Cores reduced to HW P3. Calc die temp (°C)

- 5: Calculated moving average current exceeds TDC spec. Cores reduced to HW P3 state.

- 8: Moving average current

- 9: Calculated temp

- ~50 °C
AMD Turbo CORE 3.0 TECHNOLOGY: RESIDENCY AT DIFFERENT CPU PERFORMANCE STATES

- Workloads of moderate activity have high residency at maximum frequency
  - Thermal headroom allows hotspot to remain below maximum control temperature

- Higher-activity workloads offer fewer opportunities to raise frequency and benefit from algorithms to bias power levels between CPU and GPU
  - Collaborative or compute CPU/GPU applications
  - Multi-threaded workloads

Configuration:
AMD A10-4600M APU with Radeon(tm) HD Graphics, 4GB DDR3-1600, on Pumori Reference Board with Hitachi 5400 RPM HDD.
SUMMARY

- Power and performance vary greatly by workload
- Exploit TDP headroom
  - Requires monitoring/measurement scheme to measure headroom
  - Evolution of AMD Turbo CORE technology:
    • Version 1.0: Count active cores
    • Version 2.0: Calculate power using digital monitors
    • Version 3.0: Calculate die temperatures in real time
- Current delivery infrastructure limits must be honored
- Q & A
THANK YOU
Disclaimer

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD makes no representations or warranties with respect to the contents hereof and assumes no responsibility for any inaccuracies, errors or omissions that appear in this information.

AMD specifically disclaims any implied warranties of merchantability or fitness for any particular purpose. In no event will AMD be liable to any person for any direct, indirect, special or other consequential damages arising from the use of any information contained herein, even if AMD is expressly advised of the possibility of such damages.

Trademark Attribution

AMD, the AMD Arrow logo, AMD Radeon and combinations thereof are trademarks of Advanced Micro Devices, Inc. HDMI is a trademark of HDMI licensing, LLC. PCIe and PCI Express are registered trademarks of PCI-SIG. OpenCL is a trademark of Apple, Inc. used by permission from Khronos

©2013 Advanced Micro Devices, Inc. All rights reserved.