

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 July 2006 (06.07.2006)

PCT

(10) International Publication Number
WO 2006/072097 A2

- (51) International Patent Classification: Not classified
- (21) International Application Number:
PCT/US2005/047677
- (22) International Filing Date:
29 December 2005 (29.12.2005)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
11/027,433 29 December 2004 (29.12.2004) US
- (71) Applicant (for all designated States except US): **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95052 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **SAMSON, Eric** [CA/US]; 453 Williams St, Folsom, California 95630 (US). **HORIGAN, John** [US/US]; 1209 Villa Street, Mountain View, California 94041 (US). **JACKSON, Robert** [US/US]; 419-15 Camiller Circle, San Jose, California 95134 (US). **THAKKAR, Shreekant** [US/US]; 150 Sw Moonridge Place, Portland, Oregon 97225 (US).
- (74) Agents: **O'DOWD, Shawn W.** et al.; KENYON & KENYON, 1500 K Street, Suite 700, Washington, District Of Columbia 20005 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: METHOD AND APPARATUS FOR EXTERNAL PROCESSOR THERMAL CONTROL

(57) Abstract: A system and method for throttling a slave component of a computer system to reduce an overall temperature of the computing system upon receiving a first signal is disclosed. The first signal may be from a master component indicating that a temperature for the master component has exceeded its threshold temperature. The slave component or the master component may be a central processing unit, a graphics memory and controller hub, or a central processing unit memory controller hub. The slave component may send a second signal to indicate that a temperature for the slave component has exceeded its temperature. The master component would then initiate throttling of the master component to reduce the overall temperature of the computing system. The master component may be throttled to a degree less than the slave component. A first component may be designated the master component and the second component may be designated the slave component based on a selection policy. The selection policy may be received from a user through a graphical user interface. The selection policy may be based on an action being performed by the computing system.



WO 2006/072097 A2

METHOD AND APPARATUS FOR EXTERNAL PROCESSOR THERMAL CONTROL

Background of the Invention

5 [0001] Embodiments of the invention pertain to cooling systems for computer systems. More particularly, embodiments of the invention pertain to throttling a component of a computer system based on a criterion.

[0002] The movement of electrons within the electrical components of a computer system causes a great deal of heat to be generated. Unless the heat is dissipated, it will
10 accumulate, causing damage to the system. Such damage may include the warping of the electrical components and possible fire hazards.

[0003] Currently, thermal sensors are attached to a die to read the actual temperature of the die hot spots. When the hot spot temperatures are exceeded on a particular die, that die reduces its temperature independently of the other die using some
15 form of reduction in work per unit time, also called throttling. This throttling prevents a die from reaching its maximum working temperature and damaging the system. Throttling may be performed by clock gating and clock frequency reduction.

[0004] The throttling may be triggered if the thermal sensors read a throttling threshold temperature up to some maximum tolerable temperature. To ensure safety, this
20 maximum temperature may be set well below a temperature that causes actual catastrophic damage.

[0005] Usually, different components in a system, such as the central unit and the graphics memory and controller hub (GMCH), may share a cooling system for a more efficient design to the computer system. However, these different components often have
25 different cooling needs.

Brief Description of the Drawings

[0006] **Figure 1** illustrates one embodiment of a computing system according to the present invention.

[0007] **Figure 2** illustrates in a diagram one embodiment of the shared cooling
5 system according to the present invention.

[0008] **Figure 3** illustrates in a flow chart one method for throttling a component to reduce the temperature by using a PROCHOT pin according to an embodiment of the present invention.

[0009] **Figure 4** illustrates in a flow chart one method for throttling a component
10 to reduce the temperature by using the FSB according to an embodiment of the present invention.

[0010] **Figure 5** illustrates in a flow chart one method for using a selection policy in throttling a component to reduce the temperature according to an embodiment of the present invention.

15 [0011] **Figure 6** illustrates in a flow chart one of a method for using an action-based selection policy according to an embodiment of the present invention.

Detailed Description

[0012] A system and method for throttling a slave component of a computer system to reduce an overall temperature of the computing system upon receiving a first
20 signal is disclosed. The first signal may be from a master component indicating that a temperature for the master component has exceeded its threshold temperature. The slave component or the master component may be a central processing unit (CPU), a graphics memory and controller hub (GMCH), or a CPU memory controller hub. The slave component may send a second signal to indicate that a temperature for the slave

component has exceeded its temperature. The master component may then initiate throttling of the master component to reduce the overall temperature of the computing system. The master component may be throttled to a degree less than the slave component. A first component may be designated the master component and the second component may be designated the slave component based on a selection policy. The selection policy may be received from a user through a graphical user interface. The selection policy may be based on an action being performed by the computing system.

[0013] Embodiments of the present invention also relate to apparatus for

performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, compact disk-read only memories (CD-ROMs), and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), erasable programmable read only memories (EPROMs), electronically erasable programmable read only memories (EEPROMs), magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus. Instructions are executable using one or more devices (*e.g.*, central processing units, etc.). In other embodiments, steps of the present invention might be performed by specific hardware components that contain reconfigurable or hardwired logic for performing the steps, or by any combination of programmed computer components and custom hardware components.

[0014] **Figure 1** illustrates one embodiment of a computing system 100 according to the present invention. A first component, such as a CPU 110, may be coupled to a

second component, such as a GMCH 120, by a front side bus (FSB) 130. While this description will refer specifically to a CPU and a GMCH, it is to be understood that other components may also be used. For example, the component may also be a CPU memory controller hub. The CPU 110 and the GMCH 120 share a cooling system 140. This

5 cooling system 140 may take one of any number of forms known in the art, such as air circulation units, heat exchangers, or other methods. While the cooling system 140 should be able to handle the sum of the thermal design power (TDP) of both the CPU 110 and the GMCH 120 in most computing systems, in some computing systems this is not the case for various reasons. The TDP for a component is defined as the steady state power for

10 which a thermal solution for that component should be designed so that the component will not exceed any reliability temperature threshold, and is generally quoted at a specific ambient temperature. The maximum power for the CPU 110 and GMCH 120 may be more than the TDP of each device. Since the maximum power is more than the TDP power, physical damage due to overheating may occur when operating beyond the TDP

15 power for a sufficiently long time.

[0015] The minimum residual GMCH thermal power budget is the power available to the GMCH 120 when the CPU 110 is at its maximum operating power in steady state.

The minimum residual CPU thermal power budget is the power available to the CPU 110 when the GMCH 120 is at its maximum operating power in steady state.

20 **[0016]** The CPU 110 has a microprocessor 111 to process software instructions. The CPU 110 may have a thermal sensor 112 to detect when the CPU 110 is getting too hot. The thermal sensor 112 may alert a CPU throttling arbiter 113, which may contain throttling control logic to control CPU throttling hardware 114. The throttling hardware 114 then reduces the amount of processing being performed by the microprocessor. For a

computing system 100 that executes graphics, a graphics driver 115 may be used to interact with the GMCH 120 via the FSB 130. Messages may be transmitted via the FSB 130 using the inband message protocol 116.

[0017] The GMCH 120 may have a graphics engine 121 to execute graphics processing. The GMCH 120 may have a thermal sensor 122 to detect when the GMCH 120 is getting too hot. The thermal sensor 122 may alert a GMCH throttling arbiter 123, which may contain throttling control logic to control GMCH throttling hardware 124. The throttling hardware 114 then reduces the amount of graphics execution being performed by the microprocessor. Messages may be transmitted via the FSB 130 using the inband message protocol 125.

[0018] The CPU 110 may have a pin 150, such as a PROCHOT pin, which receives a signal from the GMCH 120. Upon receiving the signal, the CPU throttling arbiter 113 may cause the CPU throttling hardware 114 to throttle the microprocessor 111. Additionally, the GMCH 120 may also have a PROCHOT pin 160, which receives a signal from the CPU 110. Upon receiving the signal, the GMCH throttling arbiter 123 may cause the graphics throttling hardware 124 to throttle the graphics engine 121.

[0019] **Figure 2** illustrates in a simplified diagram one embodiment of the shared cooling system 140. A first junction 210 may couple the CPU 110 to a shared thermal solution 220. The first junction 210 has a heat capacity 212 and a thermal conductivity 214 and the shared thermal solution 220 has a heat capacity 222 and a thermal conductivity 224. A second junction 230 may couple the CPU 110 to the shared thermal solution 220. The second junction 230 also has a heat capacity 232 and a thermal conductivity 234. The shared cooling system may reduce the entire system to the ambient temperature 240 of the surroundings.

[0020] The heat capacity 222 and the thermal conductivity 224 of the shared thermal solution 220 create a heat reduction factor θ_{sa} . The heat capacity 212 and the thermal conductivity 214 of the first junction 210 create a heat reduction factor θ_{js1} . The heat capacity 232 and the thermal conductivity 234 of the second junction 230 create a heat reduction factor θ_{js2} . The temperature for the CPU 110 and the GMCH 120 may be governed by the equations:

$$T_{cpu} = (P_{cpu} + P_{gmch}) * \theta_{sa} + Ta + P_{cpu} * \theta_{js1}$$

$$T_{gmch} = (P_{cpu} + P_{gmch}) * \theta_{sa} + Ta + P_{gmch} * \theta_{js2}$$

where P_{cpu} is the power from CPU 110, P_{gmch} is the power from the GMCH 120, and Ta is the ambient temperature 240. If the temperature of the CPU 110 is greater than its maximum allowed die junction temperature, then the temperature of the CPU 110 must be reduced. If the temperature of the GMCH 120 is greater than its maximum allowed die junction temperature, then the temperature of the GMCH 120 must be reduced.

[0021] The temperatures of the CPU 110 and the GMCH 120 may be reduced in a number of ways. **Figure 3** illustrates in a flow chart one embodiment of a method 300 for throttling a component to reduce the temperature by using a PROCHOT pin. The process starts (Block 302) when a first component, designated the slave component (SCOMP), receives a first signal via the first PROCHOT pin from a second component, designated the master component (MCOMP) (Block 304). SCOMP and MCOMP may be either the CPU 110 or the GMCH 120, depending on the circumstances. Further, the CPU 110 or the GMCH 120 may be a master component at one moment and a slave component at the next moment. Additionally, the master-slave relationship of the components need not extend past the cooling situation described herein. MCOMP is indicating with the first signal that the temperature of MCOMP (MCT) has exceeded the threshold temperature of

MCOMP (MCTT). The throttling arbiter then has the throttling hardware throttle the performance of SCOMP (Block 306). SCOMP may also receive a temperature reading of SCOMP (SCT) from its thermal sensor (Block 308). If SCT is not greater than the threshold temperature of SCOMP (SCTT) (Block 310), then the process ends (Block 312).

5 If SCT is greater than SCTT (Block 310), then a second signal may optionally be sent to the PROCHOT pin of MCOMP (Block 314), ending the process (Block 312). This second signal indicates to the throttling arbiter of MCOMP to throttle MCOMP.

[0022] Figure 4 illustrates in a flow chart one embodiment of a method 400 for throttling a component to reduce the temperature by using the FSB 130. The process starts (Block 402) when SCOMP receives a first signal via the FSB from MCOMP (Block 404). Again, MCOMP is indicating with the first signal that MCT has exceeded MCTT. The throttling arbiter then has the throttling hardware throttle the performance of SCOMP (Block 406). SCOMP receives SCT from its thermal sensor (Block 408). If SCT is not greater than SCTT (Block 410), then the process ends (Block 412). If SCT is greater than SCTT (Block 410), then a second signal is sent to MCOMP via the FSB (Block 414), ending the process (Block 412). This second signal indicates to the throttling arbiter of MCOMP to throttle MCOMP.

[0023] In a further embodiment, a selection policy may be used to designate which component is throttled. Figure 5 illustrates in a flow chart one embodiment of a method 500 for using a selection policy in throttling a component to reduce the temperature. The selection policy may be devised in a number of ways. In one embodiment, the process starts (Block 502) when the computing system 100 receives a selection policy by a user through a graphical user interface (GUI) or other method (Block 504). The selection policy may also be already present in the system or received by some other method. The

throttling arbiter of a first component (COMP1) registers a first component temperature (CT1) received from the thermal sensor exceeding a first threshold temperature for that component (CTT1) (Block 506). The throttling arbiter refers to the selection policy (Block 508). If the selection policy indicates COMP1 is the slave component and should
5 be throttled (Block 510), then the throttling arbiter has the throttling hardware throttle COMP1 (Block 512). At the same time, a second component (COMP2) receives a second component temperature (CT2) from its thermal sensor. If CT2 is not greater than the second component threshold temperature (CTT2) at this point (Block 514), the process ends (Block 516). If CT2 is still greater than CTT2 (Block 514), then the throttling arbiter
10 of COMP2 has the throttling hardware of COMP2 throttle COMP2 (Block 518), ending the process (Block 516). If the selection policy indicates COMP2 is the slave component and should be throttled (Block 510), then the throttling arbiter of COMP2 has the throttling hardware of COMP2 throttle COMP2 (Block 520). The throttling arbiters of COMP1 and COMP2 may communicate using the methods described in **Figure 3** and
15 **Figure 4**. If CT1 is not greater than CTT1 at this point (Block 522), the process ends (Block 516). If CT1 is still greater than CTT1 (Block 522), then the throttling arbiter of COMP1 has the throttling hardware of COMP1 throttle COMP1 (Block 524), ending the process (Block 516). The second throttling may be to a lesser degree than the first throttling.

20 **[0024]** In a further embodiment, a selection policy may be based on the actions being performed by the computing system at that time. **Figure 6** illustrates in a flow chart one embodiment of a method 600 for using an action-based selection policy. In one embodiment, the process starts (Block 602) when the throttling arbiter of COMP1 registers a temperature received from the thermal sensor exceeding CTT1 (Block 604).

The throttling arbiter refers to the selection policy (Block 606). If a processing intensive action is being performed (Block 608), then the GMCH throttling arbiter 123 has the graphics throttling hardware 124 throttle the graphics engine 121 (Block 610), ending the process (Block 612). If a graphics intensive action is being performed (Block 608), then
5 the CPU throttling arbiter 113 has the CPU throttling hardware 114 throttle the microprocessor 111 (Block 610), ending the process (Block 612).

[0025] In the above description, for purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention.

10 It will be apparent, however, to one skilled in the art that the present invention can be practiced without these specific details.

What is claimed is:

1. A method comprising:

receiving in a slave component a first signal from a master component indicating that a temperature for the master component has exceeded a master threshold temperature;

5 and

throttling the slave component to reduce an overall temperature of the computing system.

2. The method of claim 1, further comprising:

10 sending to the master component a second signal from the slave component indicating that a temperature for the slave component has exceeded a slave threshold temperature to initiate throttling of the master component to reduce the overall temperature of the computing system.

15 3. The method of claim 2, further comprising throttling the master component to a degree less than the slave component.

4. The method of claim 1, further comprising:

20 selecting a first component to be the master component based on a selection policy; and

selecting a second component to be the slave component based on a selection policy.

5. The method of claim 4, further comprising allowing a user to create the selection

policy.

6. The method of claim 5, wherein the selection policy is determined by an action
5 being performed by the computing system.

7. A set of instructions residing in a storage medium, said set of instructions to be
executed by a processor to implement a method for processing data, the method
comprising:

10 receiving at a pin of a slave component a first signal from a master component
indicating that a temperature for the master component has exceeded a master threshold
temperature; and
throttling the slave component to reduce an overall temperature of the computing
system.

15

8. The set of instructions of claim 7, further comprising:
sending to the master component a second signal from the slave component
indicating that a temperature for the slave component has exceeded a slave threshold
temperature to initiate throttling of the master component to reduce the overall
20 temperature of the computing system.

9. The set of instructions of claim 7, further comprising:
selecting a first component to be the master component based on a selection
policy; and

selecting a second component to be the slave component based on a selection policy.

10. The set of instructions of claim 9, wherein the selection policy is determined by an
5 action being performed by the computing system.

11. A slave component of a computing system comprising:
a slave throttling hardware to throttle the slave component; upon receiving; and
a throttling control logic to activate the slave throttling hardware of the slave
10 component upon receiving a first signal from a master component that shares a cooling
system with the slave component, the signal indicating that a temperature for the master
component has exceeded a master threshold temperature.

12. The slave component of claim 11, wherein the master component and the slave
15 component are each one of a central processing unit, a graphics memory and controller
hub, or a central processing unit memory controller hub.

13. The slave component of claim 11, further comprising a first slave pin to receive the
first signal.

20

14. The slave component of claim 13, further comprising:
a thermal sensor to read a temperature for the slave component; and
a second slave pin to send a second signal from the slave component indicating
that a temperature for the slave component has exceeded a slave threshold temperature to

induce throttling of the master component.

15. A computing system comprising:

5 a first component including:

a first component throttling hardware to throttle the first component; and

a first throttling control logic to activate the first throttling hardware of the first component upon receiving a first signal;

a second component including:

10 a second component thermal sensor to sense a temperature of the second component; and

a second throttling control logic to send the first signal indicating that the temperature of the second component has exceeded a second threshold temperature; and

15 a shared cooling solution to cool the first component and the second component.

16. The computing system of claim 15, wherein the first component and the second component are each one of a central processing unit, a graphics memory and controller hub, or a central processing unit memory controller hub.

20

17. The computing system of claim 15, wherein

the first component further includes a first component thermal sensor to read a temperature for the slave component;

the second component further includes a second component throttling hardware to

throttle the second component; and

the first throttling control logic sends a second signal to the second throttling control logic to throttle the second component.

5 18. The computing system of claim 17, wherein the first throttling control logic and the second throttling control logic select the first component to be a master component and the second component to be a slave component based on a selection policy.

19. The computing system of claim 18, further comprising a graphical user interface to
10 allow a user to create the selection policy.

20. The computing system of claim 18, wherein the selection policy is determined by an action being performed by the computing system.

15 21. The computing system of claim 17, wherein
the first component further includes a first receiving pin to receive the first signal and a first transmitting pin to send the second signal; and
the second component further includes a second receiving pin to receive the second signal and a second transmitting pin to send the first signal.

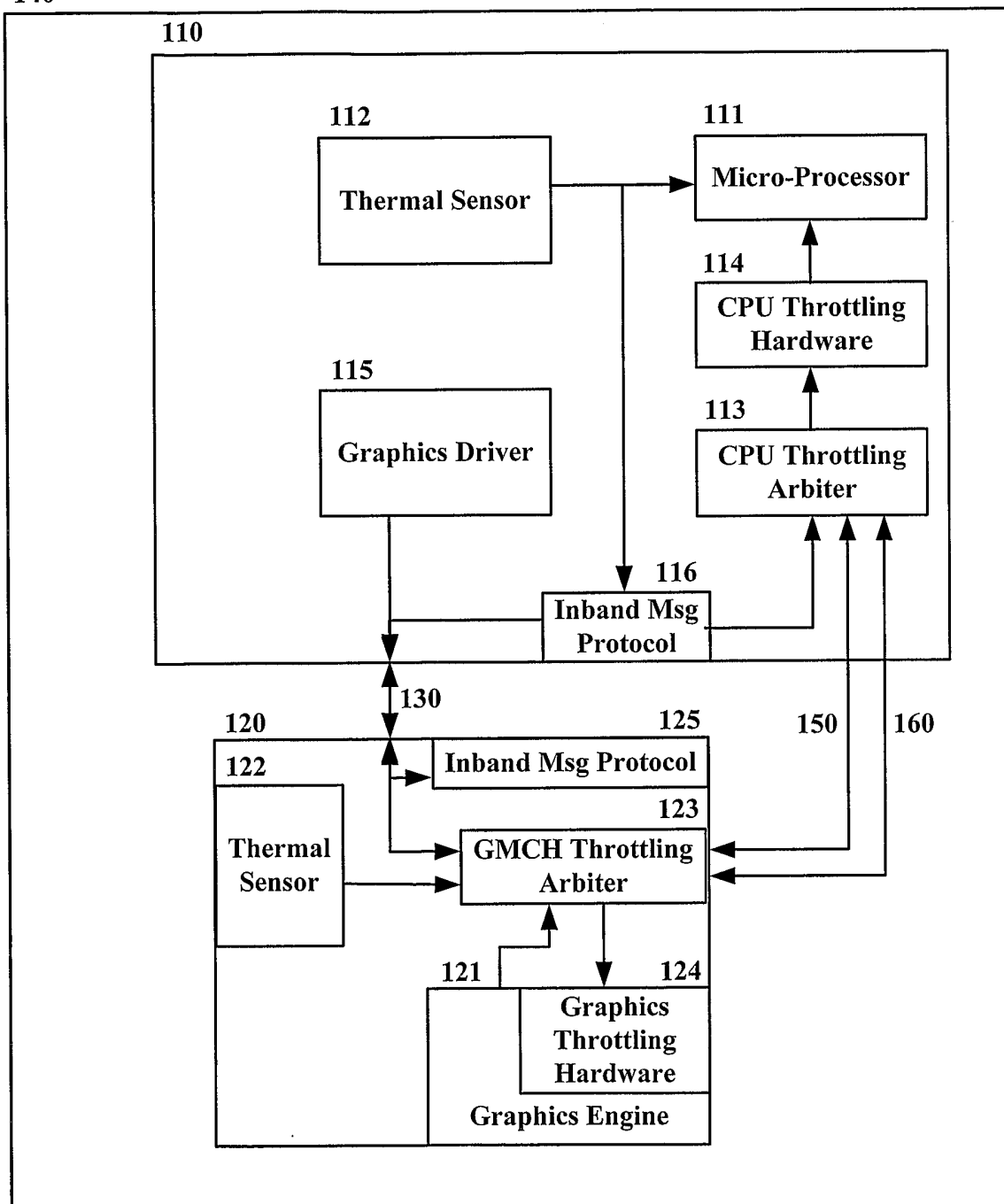
20

22. The computing system of claim 17, further comprising a front side bus coupling the first component to the second component to communicate the first signal and the second signal.

1/6

100

140

**Figure 1**

2/6

140

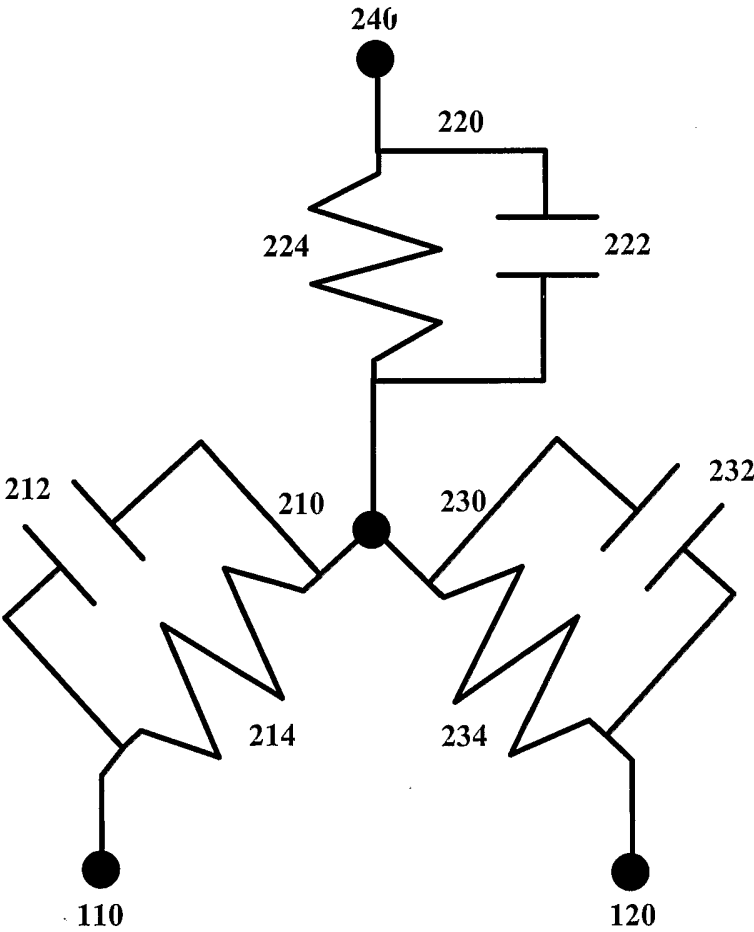
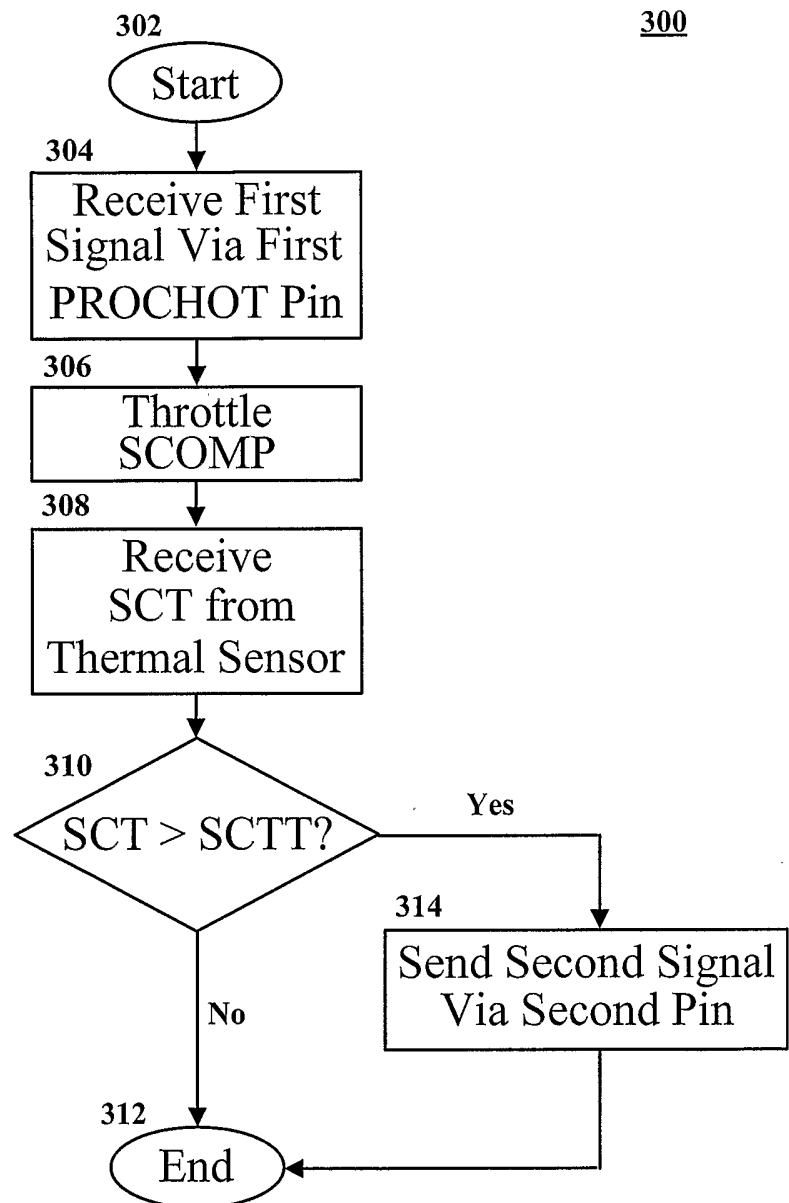
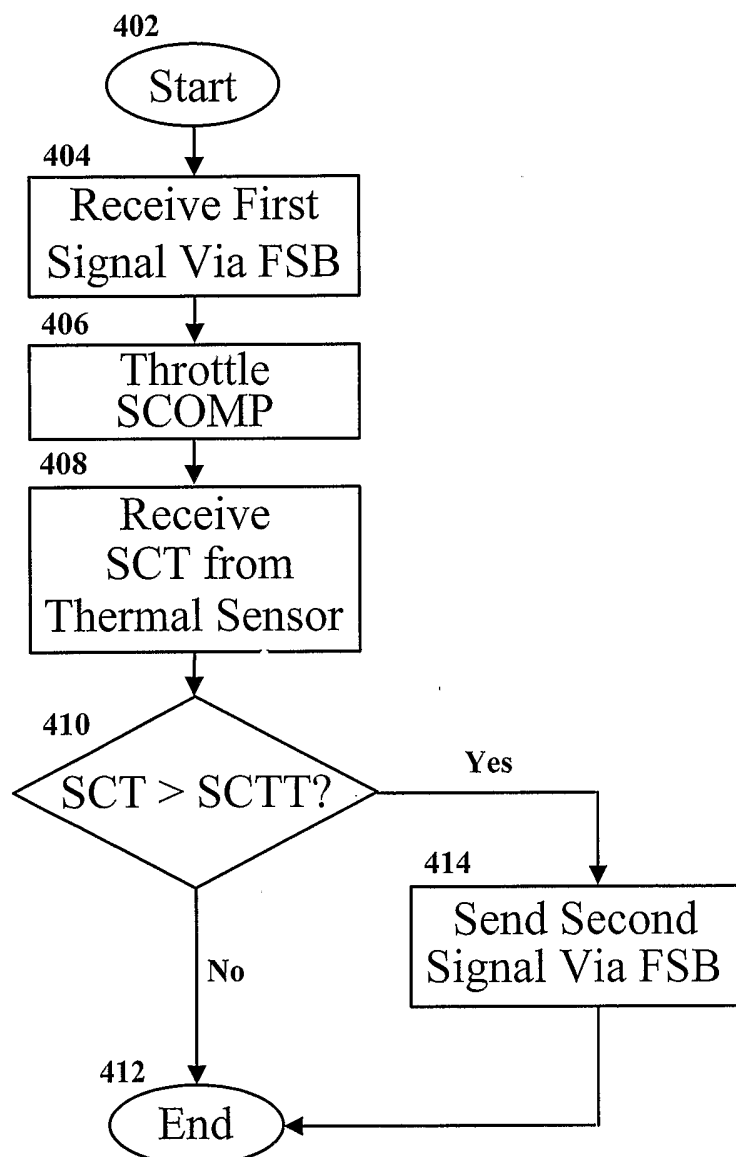


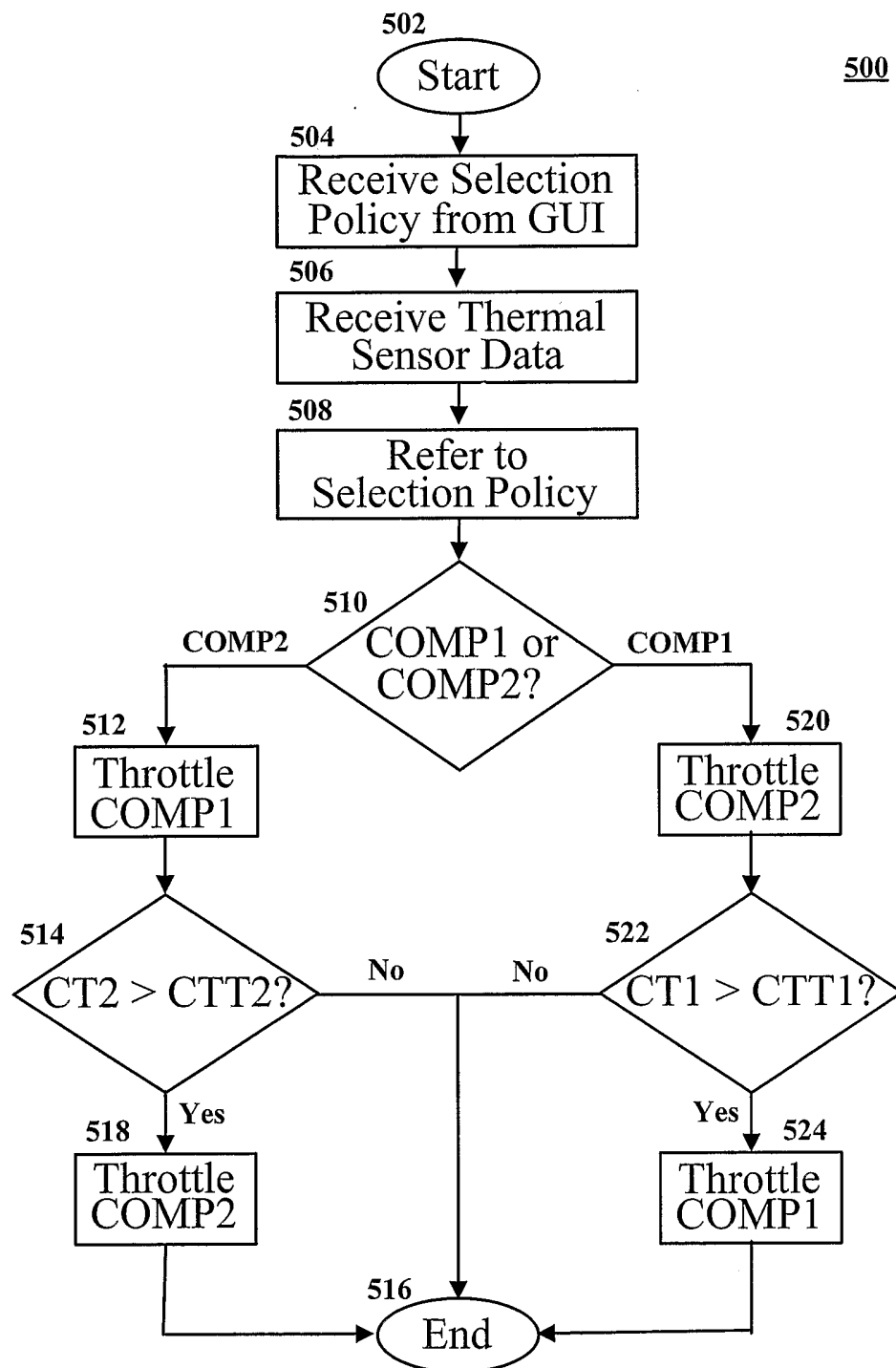
Figure 2

3/6**Figure 3**

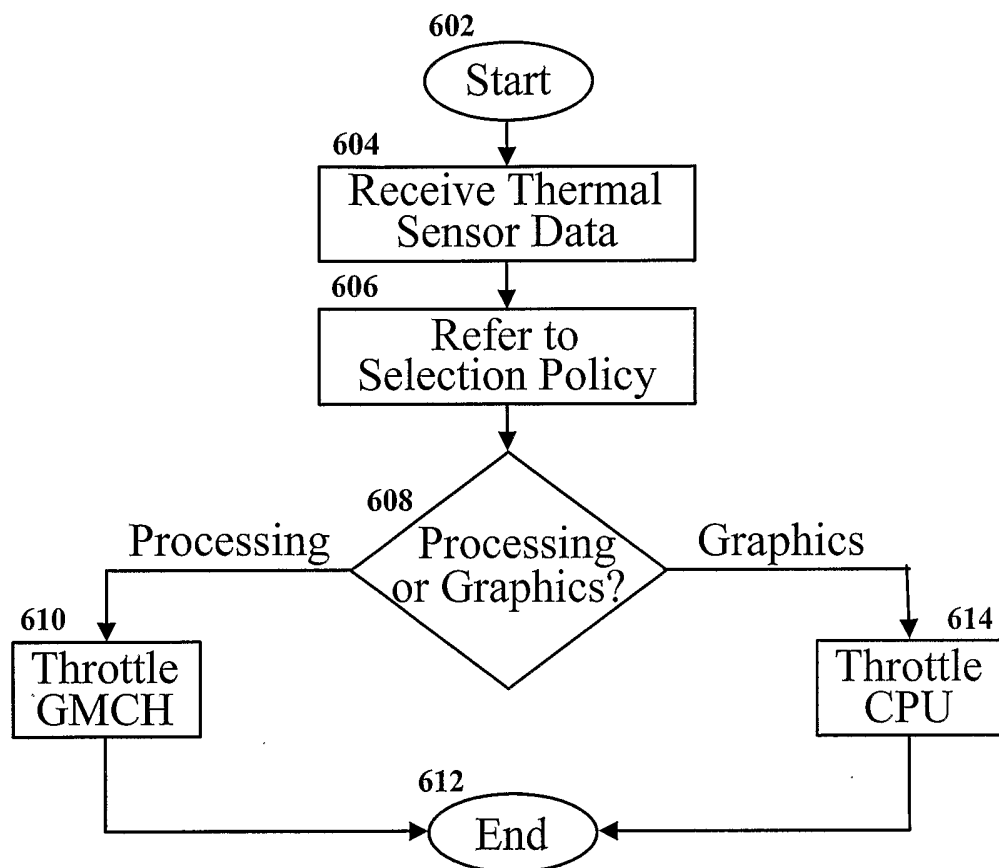
4/6

400**Figure 4**

5/6

**Figure 5**

6/6

600**Figure 6**