White Paper

Memory and Storage Solutions

Introduction to Composite Temperature

How Composite Temperature affects SSD thermal management and reporting.

Authors

Justin Boldrini
Mechanical Engineer

Executive Summary

Modern solid state drives (SSDs) contain numerous internal components with a varied range of thermal limits. To provide a simplified method of reporting SSD temperature to the host, Intel® SSDs report Composite Temperature. Composite Temperature is a single, normalized temperature value that considers input from multiple temperature sensors. This enables each Intel SSD to monitor multiple temperature points while reporting to the host a single Composite Temperature value for thermal management. The Composite Temperature is calculated based on multiple temperature inputs and does not necessarily represent the actual temperature of any physical point within the SSD; it is intended to determine the thermal margin of the SSD. It is not abnormal for a Composite Temperature to be lower than the local ambient air temperature.

Background and Introduction to Composite Temperature

Data centers support a wide range of platforms, hard disk drives (HDDs), and SSDs. Compute, Storage, and hybrid server platforms present unique thermal environments that require comprehensive thermal management to enable safe operating temperatures for all onboard components. The platform Board Management Control (BMC) monitors the temperature of components such as CPUs, PCIe switches, DIMM, as well as storage and boot drives. A platform's primary thermal management is a bank of fans adjusted to provide airflow at a rate that will satisfy the thermal requirements of all platform components, while operating at a level that is cost-effective for the data center. This level of control requires feedback from all thermally sensitive components of the platform.

The SSD exists as an NVM subsystem within the server platform. SSDs are made up of hundreds of passive components, as well as NAND, DRAM, specialized integrated circuits, and at least one controller. Each component in an SSD has an inherent thermal limit that may differ from other components' limits. The Composite Temperature considers multiple points of measurement and computes a single, normalized temperature value that the platform BMC uses to adjust fan speed to an appropriate level. Composite Temperature works by monitoring multiple temperature sensors at regular intervals and comparing each current sensor value to a pre-set maximum value. Once complete, the sensor value that is closest to its corresponding

Table of Contents

$$T_{Composite} = 70 - \min(Limit_A - CurTemp_A, Limit_B - CurTemp_B, Limit_C - CurTemp_C, \ldots, Limit_N - CurTemp_N)$$

Figure 1. Composite Temperature Algorithm

maximum is subtracted from a secondary value—typically 70° C per NVMe specifications. This provides the single, normalized Composite Temperature value which may used by the platform BMC to initiate an asynchronous event[1]. This NVMe recommendation of 70° C is referred to as WCTEMP or Warning Composite Temperature Threshold[2]. Using the Composite Temperature for thermal monitoring and reporting enables all SSDs in a platform to be assigned the same thermal limit regardless of make, model, or form factor. While this technique works very well in practice, there are nuances that may cause concern if not fully understood.

By understanding how the Composite Temperature is calculated, you can see that it's not a temperature in the traditional sense, but a reference that informs the host system or user of the temperature margin of an SSD before it reaches a thermal limit. Because of this calculation method, it is quite common for an SSD to report a Composite Temperature that is less than the ambient air temperature. This can be alarming but is not necessarily indicative of a problem. This simply means that the thermal margin of the SSD is sufficiently high to cause the Composite Temperature to report a value less than the ambient air temperature. This is more common for SSDs with advanced thermal solutions that are not artificially limited by an enclosure temperature (i.e. non hot-pluggable SSDs) [3]. Example 1 shows the algorithm used to calculate the Composite Temperature and provides some additional scenarios.

Example 1

In Example 1, we assume the following theoretical sensor values and limits:

Sensor Name	Sensor Abbreviation	Current Value	Thermal Limit
ASIC	T_A	89	115
NAND	T_N	62	85
Board SMT	T_B	58	70
PMIC	T_P	64	105

Composite Temperature calculation:

Step 1:

$$T_{Comp} = 70 - \min\,[(115 - T_A), (85 - T_N), (70 - T_B), (105 - T_P)]$$

Step 2:

$$T_{Comp} = 70 - \min\,[(115 - 89), (85 - 62), (70 - 58), (105 - 64)]$$

Step 3:

$$T_{Comp} = 70 - \min\,[(26), (23), (12), (41)]$$

Step 4:

$$T_{Comp} = 70 - 12$$

Step 5:

$$T_{Comp} = 58$$

Though the thermal limits above are theoretical, the result is consistent with expectations of a form factor that is artificially limited by an enclosure temperature. This is common with form factors that are hot-pluggable such as U.2 or EDSFF. In this case all thermal sensors report values that are less than their respective thermal limits. The result is a Composite Temperature less than the 70° C WCTEMP. Example 2 simulates a scenario where the Composite Temperature is outside of the recommended operating temperature range.

Example 2

In this example the theoretical sensor values will shift up, but their respective limits will remain the same:

Sensor Name	Sensor Abbreviation	Current Value	Thermal Limit
ASIC	T_A	104	115
NAND	T_N	77	85
Board SMT	T_B	74	70
PMIC	T_P	79	105

Composite Temperature calculation:

Step 1:

$$T_{Comp} = 70 - \min\left[(115 - T_A), (85 - T_N), (70 - T_B), (105 - T_P)\right]$$

Step 2:

$$T_{Comp} = 70 - \min\left[(115 - 104), (85 - 77), (70 - 74), (105 - 79)\right]$$

Step 3:

$$T_{Comp} = 70 - \min\left[(11), (8), (-4), (26)\right]$$

Step 4:

$$T_{Comp} = 70 - (-4)$$

Step 5:

$$T_{Comp} = 74$$

The resulting Composite Temperature in Example 2 is 74° C. This exceeds the WCTEMP of 70° C so the Intel SSD will attempt to reduce the temperature by throttling performance. Depending on the platform's thermal management policy, the BMC may also respond by increasing fan speed.

Example 3

In this example the theoretical sensor values will shift down, and some of the limits will shift up:

Sensor Name	Sensor Abbreviation	Current Value	Thermal Limit
ASIC	T_A	54	115
NAND	T_N	27	85
Board SMT	T_B	24	95
PMIC	T_P	29	105

Composite Temperature calculation:

Step 1:

$$T_{Comp} = 70 - \min\left[(115 - T_A), (85 - T_N), (70 - T_B), (105 - T_P)\right]$$

Step 2:

$$T_{Comp} = 70 - \min\left[(115 - 54), (85 - 27), (95 - 24), (105 - 29)\right]$$

Step 3:

$$T_{Comp} = 70 - \min\left[(61), (58), (71), (76)\right]$$

Step 4:

$$T_{Comp} = 70 - (58)$$

Step 5:

$$T_{Comp} = 12$$

As the individual temperature sensor values show, the lowest temperature recorded is 24° C. This is the raw temperature input as reported by the Board SMT sensor and may be only a few degrees above the ambient air temperature. Though all sensors register temperatures that would be considered "normal" by the end user (no temperatures lower than the ambient air temperature), the Composite Temperature will be lower than the ambient air temperature. This scenario is common among low-power form factors, such as M.2, which are not required to abide by a touch-temperature thermal limit. Though the logical interpretation of this value is a 12° C SSD temperature, a more accurate interpretation of this value is a margin-to-throttle of 58° C.

Conclusion

As form factor offerings, performance requirements, and cooling criteria evolve in the data center, SSD thermal management must also evolve to achieve optimum performance and reliability. Modern SSDs operate at higher power levels and temperatures than their predecessors, this often pushes the thermal capabilities of onboard components to the limit. Composite Temperature plays an integral role in empowering the controller to protect the entire SSD and provides the host platform a single threshold setpoint (WCTEMP) for all SSDs regardless of form factor, SKU, or capacity.

In summary it is important to remember, the Composite Temperature value has been computed based on multiple temperature inputs and does not necessarily represent the actual temperature of any physical point within the SSD; it is intended to convey the thermal margin (in ° C or K) available to the SSD and it is not abnormal for a Composite Temperature to be lower than the local ambient air temperature.

1. Asynchronous events are used to notify host software of the status, error and health information of various events. NVMe controllers or drives report an event to the host software when an error occurs, attributes on the drive change, a SMART change, or a management event is completed.
2. For more information, refer to section 5.14.1.2 SMART / Health Information (Log Identifier 02h) of NVMe specification version 1.4. NVMe specifications can be found at https://nvmexpress.org/wp-content/uploads/NVM-Express-1_4b-2020.09.21-Ratified.pdf
3. For more information, refer to Underwriter's Laboratory standard 60950-1 or 62368-1.