


Computer Hardware Tuning and Cooling Technologies

Olinda Braden

Hank Cullen

First Edition, 2012

ISBN 978-81-323-1471-4

Published by:
College Publishing House
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1 - Performance Tuning

Chapter 2 - Dynamic Frequency Scaling and Jumper (Computing)

Chapter 3 - Case Modding

Chapter 4 - Memory Divider and Modchip

Chapter 5 - Modding

Chapter 6 - Overclocking

Chapter 7 - Power Management and SpeedFan

Chapter 8 - SpeedStep and Tweaking

Chapter 9 - Underclocking

Chapter 10 - Computer Cooling

Chapter 11 - Computer Fan

Chapter 12 - Heat Sink

Chapter 13 - Laptop Cooler

Chapter 14 - Thermal Management of Electronic Devices and Systems

Chapter 15 - Fan Control

Chapter 1

Performance Tuning

Performance tuning is the improvement of system performance. This is typically a computer application, but the same methods can be applied to economic markets, bureaucracies or other complex systems. The motivation for such activity is called a performance problem, which can be real or anticipated. Most systems will respond to increased load with some degree of decreasing performance. A system's ability to accept higher load is called scalability, and modifying a system to handle a higher load is synonymous to performance tuning.

Systematic tuning follows these steps:

1. Assess the problem and establish numeric values that categorize acceptable behavior.
2. Measure the performance of the system before modification.
3. Identify the part of the system that is critical for improving the performance. This is called the bottleneck.
4. Modify that part of the system to remove the bottleneck.
5. Measure the performance of the system after modification.

This is an instance of the measure-evaluate-improve-learn cycle from quality assurance.

A performance problem may be identified by slow or unresponsive systems. This usually occurs because high system loading, causing some part of the system to reach a limit in its ability to respond. This limit within the system is referred to as a bottleneck.

A handful of techniques are used to improve performance. Among them are code optimization, load balancing, caching strategy, distributed computing and self-tuning.

Performance analysis

Performance analysis, commonly known as profiling, is the investigation of a program's behavior using information gathered as the program executes. Its goal is to determine which sections of a program to optimize.

A profiler is a performance analysis tool that measures the behavior of a program as it executes, particularly the frequency and duration of function calls. Performance analysis tools existed at least from the early 1970s. Profilers may be classified according to their output types, or their methods for data gathering.

Performance engineering

Performance engineering is the discipline encompassing roles, skills, activities, practices, tools, and deliverables used to meet the non-functional requirements of a designed system, such as increase business revenue, reduction of system failure, delayed projects, and avoidance of unnecessary usage of resources or work.

Several common activities have been identified in different methodologies:

- Identification of critical business processes.
- Elaboration of the processes in use cases and system volumetrics.
- System construction, including performance tuning.
- Deployment of the constructed system.
- Service management, including activities performed after the system has been deployed.

Code optimization

Enhancing performance by rewriting specific portions of a program to run faster is one form of code optimization. The term code optimization can refer to improving the implementation of a particular algorithm for performing a task (code tuning). It can also refer to utilizing a better algorithm. Examples of code optimization include improving the code so that work is done once before a loop rather than inside a loop or replacing a call to a simple selection sort with a call to the more complicated algorithm for a quicksort.

Caching strategy

Caching is a fundamental method of removing performance bottlenecks that are the result of slow access to data. Caching improves performance by retaining frequently used information in high speed memory, which reduces access time and thus improves performance. Caching is an effective manner of improving performance in situations where the principle of locality of reference The methods used to determine which data is stored in progressively faster storage are collectively called **caching strategies.**

Load balancing

A system can consist of independent components, each able to service requests. If all the requests are serviced by one of these systems (or a small number) while others remain idle then time is wasted waiting for used system to be available. Arranging so all systems are used equally is referred to as load balancing and can improve over-all performance.

Load balancing is often used to achieve further gains from a distributed system by intelligently selecting which machine to run an operation on based on how busy all potential candidates are, and how well suited each machine is to the type of operation that needs to be performed.

Distributed computing

Distributed computing is used to increase the performance of operations that can be performed in parallel, by concurrently executing multiple operations. Operations may be distributed across multiple processes on a single CPU, taking advantage of multitasking, multiple processes across multiple CPUs, or across multiple machines. As operations are executed concurrently, ensuring synchronization between processes is essential to ensure correct results.

As the trend of increasing the potential for parallel execution on modern CPU architectures continues, the use of distributed systems is essential to achieve performance benefits from the available parallelism. High performance cluster computing is a well known use of distributed systems for performance improvements.

Distributed computing and clustering can negatively impact latency while simultaneously increasing load on shared resources, such as database systems. To minimize latency and avoid bottlenecks, distributed computing can benefit significantly from distributed caches.

Self-tuning

A self-tuning system is capable of optimizing its own internal running parameters in order to maximize or minimize the fulfillment of an objective function; typically the maximization of efficiency or error minimization. Self-tuning systems typically exhibit non-linear adaptive control. Self-tuning systems have been a hallmark of the aerospace industry for decades, as this sort of feedback is necessary to generate optimal multi-variable control for nonlinear processes.

Bottlenecks

The bottleneck is the part of a system which is at capacity. Other parts of the system will be idle waiting for it to perform it's task.

In the process of finding and removing bottlenecks, it is important to prove their existence, such as by sampling, before acting to remove them. There is a strong temptation to *guess*. Guesses, by definition, are often wrong, and investing time in them is itself a bottleneck.

Chapter 2

Dynamic Frequency Scaling and Jumper (Computing)

Dynamic frequency scaling

Dynamic frequency scaling (also known as **CPU throttling**) is a technique in computer architecture whereby the frequency of a microprocessor can be automatically adjusted "on-the-fly", either to conserve power or to reduce the amount of heat generated by the chip. Dynamic frequency scaling is commonly used in laptops and other mobile devices, where energy comes from a battery and thus is limited. It is also used in quiet computing settings and to decrease energy and cooling costs for lightly loaded machines. Less heat output, in turn, allows the system cooling fans to be throttled down or turned off, reducing noise levels and further decreasing power consumption. It is also used for reducing heat in insufficiently cooled systems when the temperature reaches a certain threshold, such as in poorly cooled overclocked systems.

The dynamic power (*switching power*) dissipated by a chip is $C \cdot V^2 \cdot f$, where C is the capacitance being switched per clock cycle, V is voltage, and f is the switching frequency (as a unitless quantity). As frequency changes, the dynamic power will change linearly with it. Dynamic power does not account the total power of the chip however, as there is also static power, which is primarily due to various leakage currents. Leakage current has become more and more important as feature sizes has become smaller and threshold levels lower. In state-of-the-art deep submicrometre technologies in 2008, dynamic power accounts for approximately two-thirds of the total chip power, which limits the effectiveness of frequency scaling.

Dynamic voltage scaling is another power conservation technique that is often used in conjunction with frequency scaling, as the frequency that a chip may run at is related to the operating voltage.

The efficiency of some electrical components, such as voltage regulators, decreases with increasing temperature, so the power used may increase with temperature. Since increasing power use may increase the temperature, increases in voltage or frequency may increase system power demands even further than the CMOS formula indicates, and vice-versa.

Performance impact

Dynamic frequency scaling reduces the number of instructions a processor can issue in a given amount of time, thus reducing performance. Hence, it is generally used when the workload is not CPU-bound.

Dynamic frequency scaling by itself is rarely worthwhile as a way to conserve switching power. Saving the most power requires dynamic voltage scaling too, because of the V^2 component and the fact that modern CPUs are strongly optimized for low power idle states. In most constant-voltage cases it is more efficient to run briefly at peak speed and stay in a deep idle state for longer (called "race to idle"), than it is to run at a reduced clock rate for a long time and only stay briefly in a light idle state. However, reducing voltage along with clock rate can change those tradeoffs.

A related-but-opposite technique is overclocking, whereby processor performance is increased by ramping the processor's (dynamic) frequency beyond the manufacturer's design specifications.

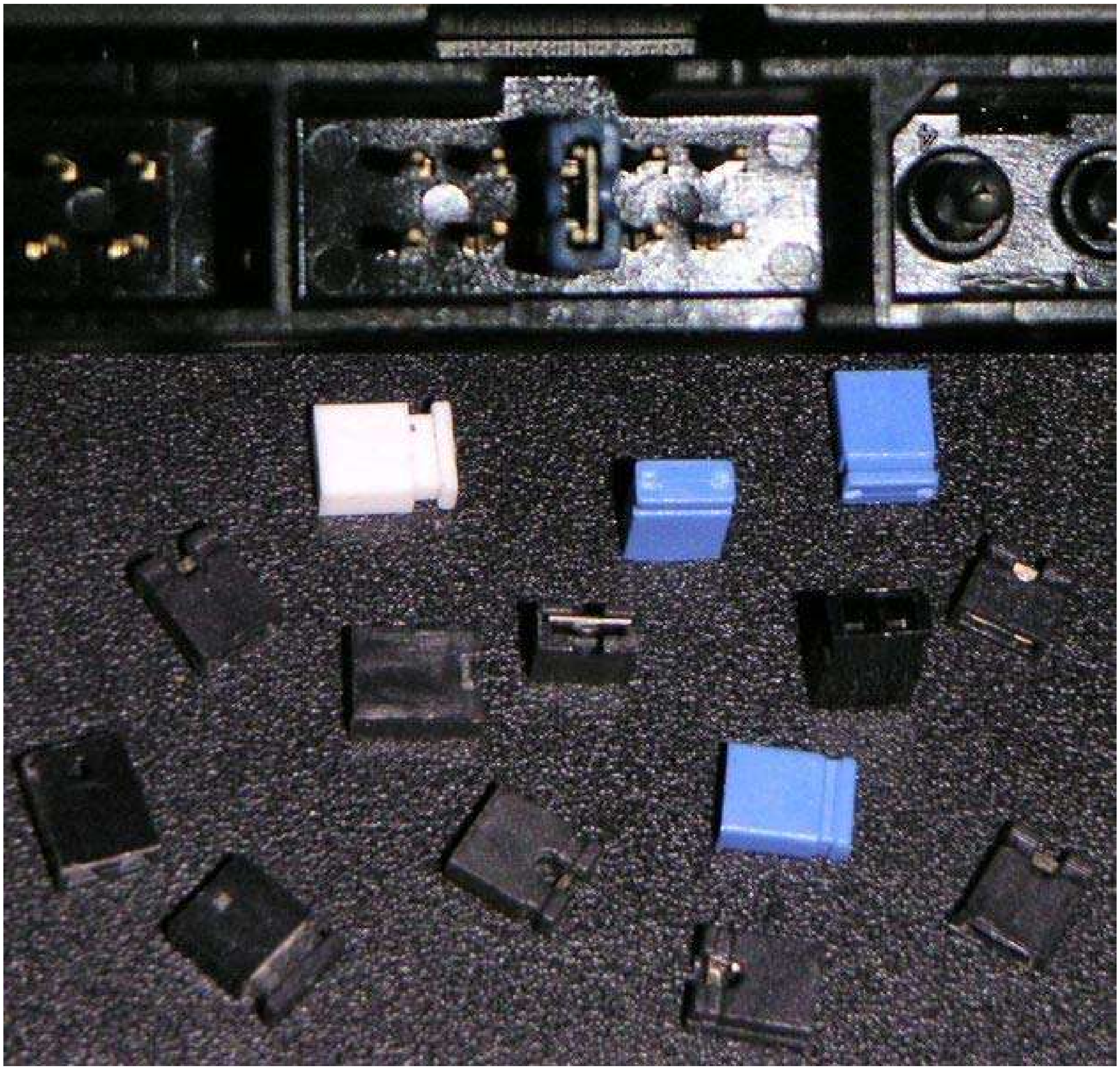
One major difference between the two is that in modern PC systems overclocking is mostly done over the Front Side Bus (mainly because the multiplier is normally locked), but dynamic frequency scaling is done with the multiplier. Moreover, overclocking is often static, while dynamic frequency scaling is always dynamic. Software can often incorporate overclocked frequencies into the frequency scaling algorithm, if the chip degradation risks are allowable.

Implementations

Intel's CPU throttling technology, SpeedStep, is used in its mobile CPU line. AMD employs two different CPU throttling technologies. AMD's Cool'n'Quiet technology is used on its desktop and server processor lines. The aim of Cool'n'Quiet is not to save battery life, as it is not used in AMD's mobile processor line, but instead with the purpose of producing less heat, which in turn allows the system fan to spin down to slower speeds, resulting in cooler and quieter operation, hence the name of the technology. AMD's PowerNow! CPU throttling technology is used in its mobile processor line, though some supporting CPUs like the AMD K6-2+ can be found in desktops as well.

According to the ACPI Specs, the C0 working state of a modern-day CPU can be divided into the so called "P"-states (performance states) which allow clock rate reduction and "T"-states (throttling states) which will further throttle down a CPU (but not the actual clock rate) by inserting STPCLK (stop clock) signals and thus omitting duty cycles.

Jumper (computing)



Top: jumper block on IDE hard drive with jumper; **bottom:** assorted jumpers

In electronics and particularly computing, a **jumper** is a short length of conductor used to close a break in or bypass part of an electrical circuit. Jumpers are typically used to set up or adjust printed circuit boards, such as the motherboards of computers.

Description

Jumper pins (points to be connected by the jumper) are arranged in groups called *jumper blocks*, each group having at least one pair of contact points and often more. Sometimes these groups are referred to as *headers*. In general, each contact in a jumper block terminates in a small metal pin. An appropriately sized conductive sleeve called a jumper, or more technically, a jumper shunt, is slipped over the pins to complete the circuit.

Jumpers must be electrically conductive; they are usually encased in a non-conductive block of plastic for convenience. This also avoids the risk that an unshielded jumper will accidentally short out something critical (particularly if it is dropped on a live circuit).

When a jumper is placed over two or more jumper pins, an electrical connection is made between them, and the equipment is thus instructed to activate certain settings accordingly. For example, with older PC systems, CPU speed and voltage settings were often made by setting jumpers. Informally, technicians often call setting jumpers "strapping". To adjust the SCSI ID jumpers on a hard drive, for example, is to "strap it up".

Jumper blocks and jumpers are also often used on motherboards to clear the CMOS information, resetting the BIOS configuration settings. This allows the computer to boot if a recent BIOS setting made it unable to boot, or if the CMOS boot password was forgotten.

Move to reduce jumpers

A jumper

Early generations of any given computer hardware technology usually have many jumper blocks, often laid out in a way that is poorly documented and difficult to set correctly. Often, designers find ways to streamline and simplify the jumper layout. For example, a typical early model Intel 386 motherboard might have 30 or 40 jumper pairs, while the last production models typically had just a handful, or sometimes only one. Typically, each jumper block is assigned and labeled with a number, which is documented in an instructional list printed on the motherboard or in the manual.

The recent trend has been to try to eliminate jumpers entirely from hardware devices by the use of auto-configuration or software-controlled configuration. Configurations may be stored in NVRAM, loaded by a host processor, or negotiated at system initialization time. In some cases, hot swappable devices may be able to renegotiate their configuration while the system is running. Jumperless designs have the advantage that they are usually fast and easy to set up, often require little technical knowledge, and can be adjusted without having physical access to the circuit. With newer PCs, the most common use of jumpers is in setting the operating mode for ATA drives (master, slave, or cable select).

Systems using boards with physical jumpers, on the other hand, tend to be configured correctly by end users as, in general, non-technical people are less willing to physically alter hardware settings than they are to experiment with settings from the keyboard. They also have the advantage that they usually only need to be set once; while firmware settings can be easily lost or corrupted by a careless user, a virus, or a power failure, the only way to alter a correct jumper setting is to physically change it.

Chapter 3

Case Modding

Modded PC case with CCFL lights and switches, LED fans, perspex window, custom fan hole, and customized Windows XP.

Case modification (commonly referred to as **case modding** where an individual project is referred to as a **case mod**) is the modification of a computer chassis (often just referred to as the case), or a video game console chassis. Modifying a computer case in any non-standard way is considered a case mod. Many people, particularly hardware enthusiasts, use case mods to illustrate a computer's power (by showing off the internal hardware), and also for aesthetic purposes. Cases may also be modified to improve a computer's performance.

History

When personal computers first became available to the public, the majority were produced in simple, beige-colored cases. This design is sometimes referred to be as a *beige box*. Although this met the purpose of containing the components of the personal computer, many users considered their computers as "tacky" or "dull", and some began modifying their existing chassis, or building their own from scratch.

A new market for third-party computer cases and accessories began to develop, and today cases are available in a wide variety of colors and styles. Today the business of "modding" computers and their cases is a hugely profitable endeavor, and modding competitions are everywhere.

Common modifications

A computer modded with lights

- **Window mods**: This refers to a window placed within one of the panels of a computer case. This is most often done to the left hand side panel, and less often to the top panel. This modification is so popular that many of the major case manufacturers now offer cases with the windows pre-installed, or replaceable side panels with a window installed. Some companies even offer entire cases made out of transparent materials. A window kit may be modified to hold an LCD screen. Laser engraving can be done on acrylic windows to add a distinct look to a modded case.

- **Lighting mods**: A lighting mod refers to lighting in or on the computer cases. This is usually achieved with cold cathode lights (CCLs), LED case fans, or electroluminescent wire lights. The lights are sometimes paired with sound controllers that make the lights pulse in time to sound. CCLs come in long tubes and generally produce a little bit of heat. LEDs come in many sizes and forms, most often seen in bars similar to CCLs or within fans, called LED fans. Electroluminescent wire, which takes the form of a small light rope, is often embedded in cables such as SATA cables.

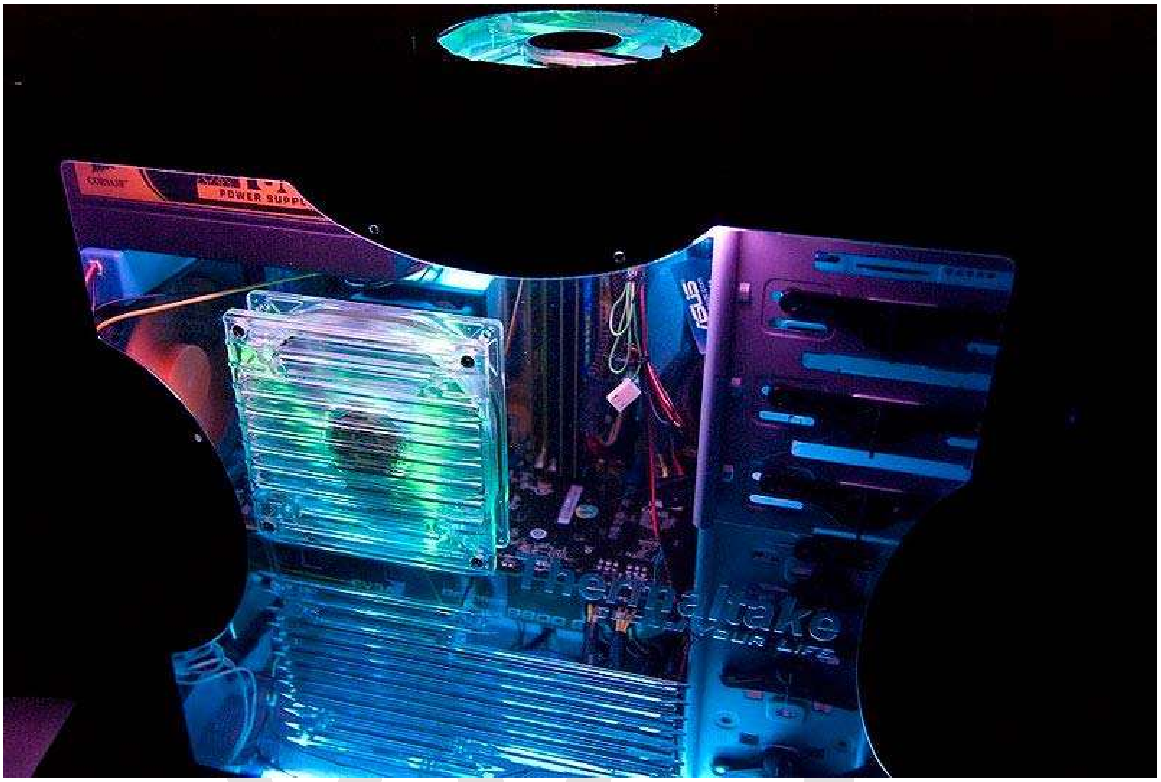
Lighting modifications are often paired with window mods to help show off the components. Although not as common, they are also placed in cases without a window, allowing light to shine through any holes or gaps of the case which add subtle aesthetics to an otherwise plain looking case; this is also done for practical purposes such as to make a computer double as a night light. Internal components such as case fans, CPU heatsink fans, and power supplies are often available with built in lighting.

- **Cooling mods**: There are many modifications that can fall into this category. The most common one is simply drilling a mount for a new fan, or removing a restrictive fan grill. Others include air ducts, water cooling, filtering, sealing openings to promote better air flow, or even the adding of a tank of pressurized carbon dioxide or liters of mineral oil to the case. These modifications are often performed by overclockers either looking for better cooling for hot components or sound reduction. Some fan modifications are merely a show of modding skill or talent and have no true functional purpose.

- **Spray paint**: Painting a case is another method of distinguishing your work from others. Spray paint is the common method preferred among amateur modders. There are many spray painting guides for amateur modders. The finish cannot be compared to automotive paint or powder coating, but is a simple way to change the look of a case.

- **Vinyl dye**: Re-colouring the plastics of a case, keyboard, mouse and speakers is another method of highlighting your system and making it different than the rest. Vinyl dye has the advantages of ease of use like spray paint, but is much more durable as it doesn't chip or scratch off, it also doesn't create a 'layer' like paint.

Less common modifications



A computer modded with different coloured CCFLs and LED fans

- **Automotive paint & other finishes**: Automotive paint refers to the paint typically seen on cars and trucks. This type of finish requires a compressed air source, such as an air compressor or CO2 tank, and a spray gun. It is more expensive than a finish using spray cans, but when done skillfully it can be better looking and much more durable. Other methods of painting can include powder coating which is highly durable though not quite as aesthetically pleasing to many modders as automotive paint. Electroplating can also be done on steel computer cases and parts. Aluminum cases can be plated or anodized as well, and processes are available to plate plastic cases. Plated coatings can range from nickel to chrome and even gold. More elaborate finishes can be crafted by using a combination of techniques, such as chrome plating with a transparent color coat.
- **Body filler**: Body filler (or Bondo) is a two-part putty often used to fix dents in automobiles. Case modders use it to fill and sculpt their own creations. When mixed with a paste catalyst the filler hardens in a short period of time and can be sanded, ground or cut to a desired shape. An alternative system uses fiberglass resin (catalyzed with liquid hardener) and either fiberglass cloth or mat to fill holes and form shapes. Lacquer based Spot Putty is often used to fill smaller imperfections before the application of primer. Typically, a case modder uses a combination of these materials to obtain the desired result. This method is usually used on the front plastic bezel of a computer case to give it a new look.

Types of case mods

Computer built into a Microwave oven

- **Peripheral mods**: Peripherals like the keyboard, mouse, and speakers are sometimes painted or otherwise modified to match the computer. Some system builders, in an effort to make their system more portable and convenient, install speakers and small LCD screens into the case.
- **Unusual cooling mods**: Hardcore overclockers often install cooling systems for the sole purpose of achieving performance records. Such systems frequently include phase change, thermoelectric/Peltier and liquid nitrogen. However, some of these systems are noisy and expensive. They are rarely used for extended periods of time.
- **Case building**: Sometime modders even build entire cases from scratch. Some may attempt to treat the case as a work of art. Others make it appear to be something else, like a teddy bear, wooden cabinet, or a shelf mounted on a wall. Still others pursue a retro look, like a Macintosh Plus or an old Atari 2600 video game console. Case modders (or case builders) who create their computer cases from scratch are few and far between. These people sometimes put hundreds of hours into their work. The WMD case, Project Nighthawk, and Dark Blade case are a few examples of professional cases built from scratch.
- **Component modding**: This type of modding, as the name suggests, involves modifying the PC components themselves. An example is the relocation of buttons on optical drives. This is often done in combination with "stealthing", which hides the drive's visibility by masking it with a blank face. A riskier modification involves installing hard drive windows. This is done in a clean room

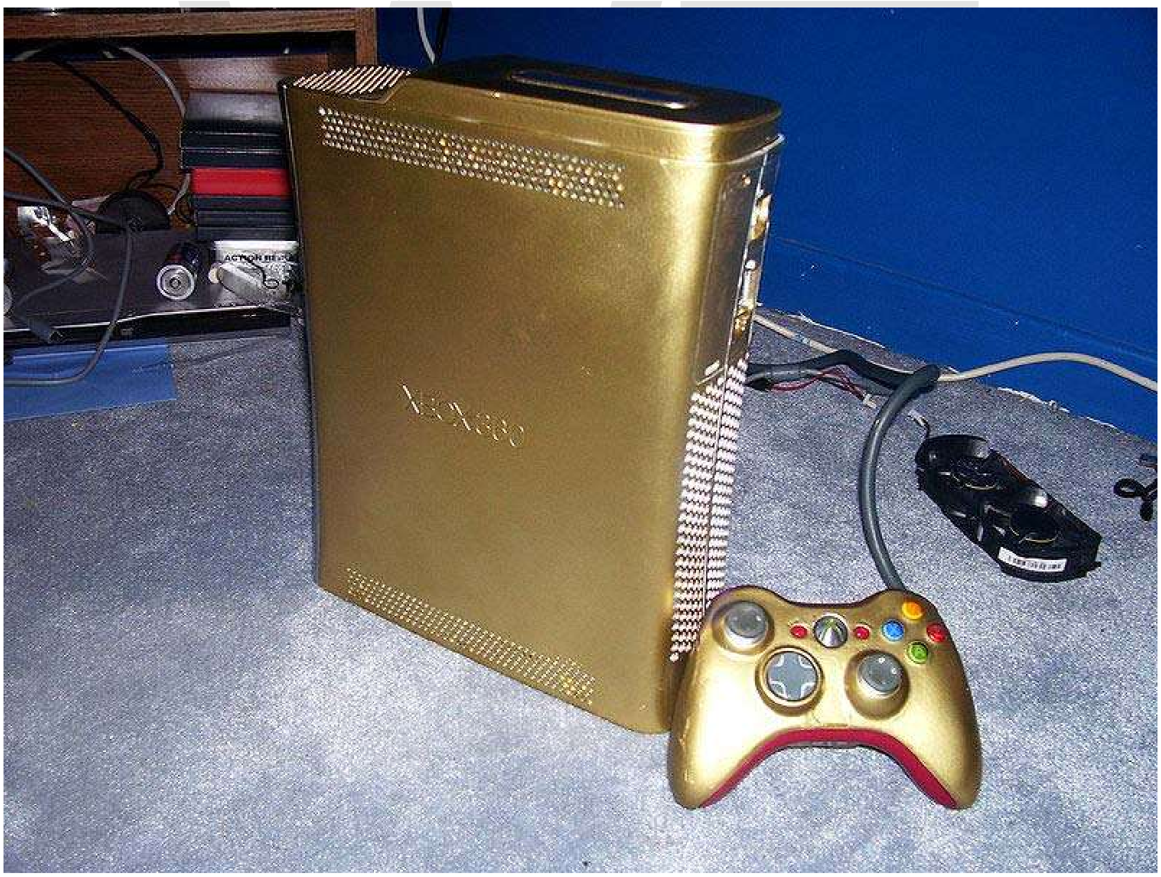
where there is little to no dust. Few people have attempted it and results seem to vary. Some hard drives, including the WD Raptor, now come with a window as standard.

- **Laptop modding**: Laptops can be modified much like a typical computer case. While most laptop mods consist of new paint or other finishes, others have chosen to engrave or cut out designs into their laptop cover (behind the screen). Laptops may also be turned into digital photo frames. These types of mods will typically void the warranty of the device. To avoid warranty issues, skins or stickers can be purchased that are easily removable from the casing.

Case modding contests

Many websites and companies run contests for case modders, awarding prizes and accolades to the winners. Some of these contests are sponsored by computer enthusiast magazines and others by computer retailers.

Console case modding



An example of an Xbox 360 case mod

Console case modding refers to gamers who modify the case of their game consoles. The most common consoles to modify are the Xbox and Xbox 360, because there is much more room inside to customize them with items such as lights, fans, etc. These consoles and their controllers are also relatively easy to take apart. For those who do not wish to scratch-build mods, there are several companies that sell transparent Xbox cases and various cooling/lighting equipment for them.

Console case modding started in the late 1980s when the NES and Sega Genesis, were released; many customers simply put pictures or stickers on them until the PlayStation was released. Many case modders started to change hardware, for example by altering them to play copied games (known as 'chipping' the games console). The most common modification for the PlayStation was the 'chipping' process (mentioned above). When the Nintendo 64, Dreamcast and PlayStation 2 was released, many chipped them, styled them, and added additional cooling; some went as far as changing the hardware itself. When the Xbox and Xbox 360 were released, many modders personally customized them further, using neon lights, transparent cases, fans, and PC hard drives (as opposed to Xbox-branded drives). Many modders found that altering the interior of Xbox 360s was difficult due to absence of a power cable (normally in a PC, this cord attaches the hard disk drive to its motherboard). Despite shortcomings, modders also found a way to power neon lighting and other powered equipment by using the DVD-ROM power supply; however, due to insufficient power to the hard disk drive, it often caused freezing during disk access. Fortunately, the most recent products available for internal case modding uses the power outlet for the internal fan by splitting the cord with a "Y" connector.

Chapter 4

Memory Divider and Modchip

Memory divider

A **memory divider** is a ratio which is used to determine the operating clock frequency of computer memory in accordance with front side bus (FSB) frequency, if the memory system is dependent on FSB clock speed. Along with memory latency timings, memory dividers are extensively used in overclocking memory subsystems to find stable, working memory states at higher FSB frequencies. A memory divider is also commonly referred to as "DRAM:FSB ratio".

Memory dividers are only applicable to those chipsets in which memory speed is dependent on FSB speeds. Certain chipsets like nVidia 680i have separate memory and FSB lanes due to which memory clock and FSB clock are asynchronous and memory dividers are not used there. Setting memory speeds and overclocking memory systems in such chipsets are different issues which do not use memory dividers.

Overview

Memory Dividers allows system memory to run slower than or faster than the actual FSB (Front Side Bus) speed. Ideally, Front Side Bus and system memory should run at the same clock speed because FSB connects system memory to the CPU. But, it is sometimes desired to run the FSB and system memory at different clock speeds. It is possible to run FSB and memory clock at different clock speeds, within certain limits of the motherboard and corresponding chipset. So, settings termed as Memory Divider or FSB/DRAM settings are available and are expressed in a "ratio" which control the difference in memory clock rate and FSB speed.

Entry Level motherboards usually do not provide memory dividers to be changed and the memory dividers are managed by Memory Controller (if chipset supports memory dividers). High end motherboards meant for overclocking provide facilities to change memory dividers (if chipset support memory dividers). However, in certain chipsets memory dividers are not used, because in those systems memory speed is independent of FSB speed.

Description & Application

Usually (Memory Divider) × (Front Side Bus Frequency) gives I/O Bus clock of the memory. Memory clock then determines the final operating frequency or effective clock speed of memory system depending upon DRAM types (DDR, DDR2 and DDR3 SDRAM).

By default, FSB speed and memory are usually set to a 1:1 ratio, meaning that increasing FSB speed (by overclocking) increases memory speed by the same amount. Normally system memory is not built for overclocking and thus may not be able to take the level of overclocking that the processor or motherboard can achieve. The memory divider allows users to mitigate this problem by reducing the speed increase of the memory relative to that of the FSB and the processor.

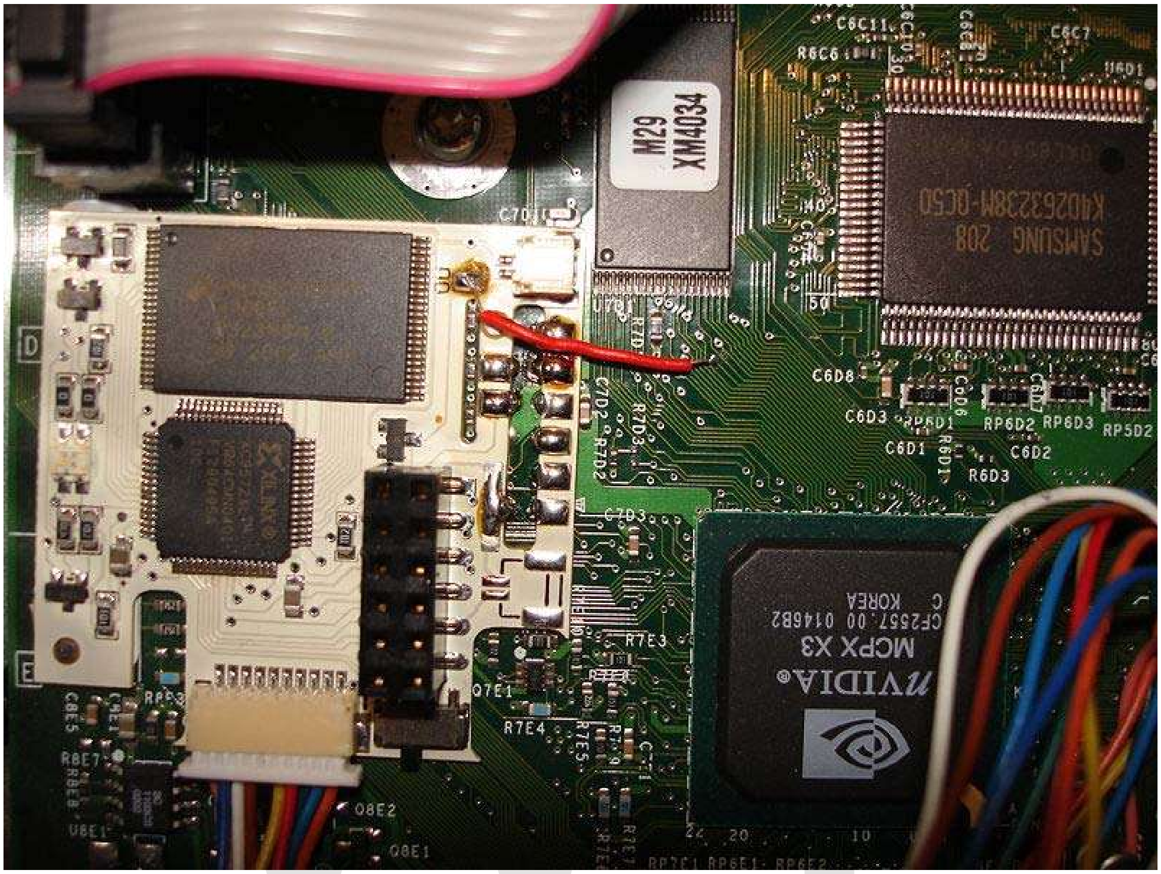
Example

Suppose a computer system has DDR memory, a Memory Divider of 1:1, a FSB operating at 200 MHz and a CPU multiplier of 10x. Then, the base memory clock will operate at (Memory Divider) × (FSB) = 1 × 200 = 200 MHz and the effective memory clock would be 400 MHz since it's a DDR system ("DDR" stands for Double Data Rate; the effective memory clock speed is double the actual clock speed). The CPU will operate at 10 × 200 MHz = 2.0 GHz. Using I/O bus frequency of 200 MHz various types of DRAM will operate as:

```
DDR SDRAM at  400 MHz (DDR -400 or PC -3200)
DDR2 SDRAM at  800 MHz (DDR2-800 or PC2-6400)
DDR3 SDRAM at  1600 MHz (DDR3-1600 or PC3-12800)
```

Now suppose that we overclock FSB to 250 MHz so that CPU operates at 10 × 250 MHz = 2.5 GHz and memory clock operates at 250 MHz (Memory Divider × FSB). Since DDR-400 RAM is used then effective memory clock (Actual Memory Frequency) will be 500 MHz. A normal DDR-400 SDRAM will fail to work at 500 MHz since it is designed to work at 400 MHz and system will become unstable. But a modern CPU (having overclocking potential) can work at 2.5 GHz (even if it is designed to work at 2 GHz) flawlessly without giving any problem of stability. To keep running overclocked CPU at 2.5 GHz or even at higher speeds (by increasing FSB) we need to slow down memory clock so as to achieve a stable system. For this if we decrease DRAM:FSB ratio to say 4:5 then resulting memory clock speed is (4/5) × 250 MHz = 200 MHz resulting effective clock speed of 400 MHz on DDR-400. So, we are able to operate upon a stable overclocked CPU at 2.5 GHz from 2 GHz without increasing effective memory clock.

Modchip



A Modchip, the white circuit board, attached to an Xbox's PCB

A **modchip** (short for *modification chip*) is a small electronic device used to modify or disable built-in restrictions and limitations of computers, specifically videogame consoles. It introduces various modifications to its host system's function, including the circumvention of region coding, digital rights management, and copy protection checks for the purpose of running software intended for other markets, copied game media, or unlicensed third-party (homebrew) software.

Modchips are mainly used in CD/DVD-based videogame systems due to the availability and low cost of user-writable media. In addition to games consoles, modchips are also available for some DVD players to circumvent region code enforcement and user operation prohibitions.

Function and construction

Modchips operate by replacing or overriding a system's protection hardware or software. They achieve this by either exploiting existing interfaces in an unintended or undocumented manner, or by actively manipulating the system's internal communication, sometimes to the point of re-routing it to substitute parts provided by the modchip.

Most modchips consist of one or more integrated circuits (microcontrollers, FPGAs, or CPLDs), often complemented with discrete parts, usually packaged on a small PCB to fit within the console system it is designed for. Although there are modchips that can be reprogrammed for different purposes, most modchips are designed to work within only one console system or even only one specific hardware version.

Modchips typically require some degree of technical acumen to install. Modchips must be connected to a console's circuitry, most commonly by soldering wires to select traces or chip legs on a system's circuit board. Some modchips allow to be installed by directly soldering the modchip's contacts to the console's circuit ("quicksolder"), by the precise positioning of electrical contacts ("solderless"), and, in rare cases, by plugging it into a system's internal or external connector.

Memory cards or cartridges that offer functions similar to modchips work on a completely different concept, namely by exploiting flaws in the system's handling of media. Such devices are not referred to as modchips, even if they are frequently traded under this umbrella term.

History

Cartridge-based console systems did not have modchips produced for them. They usually implemented copy protection and regional lockout with game cartridges, both on physical and software level. Converters or passthrough devices have been used to circumvent the restrictions, while flash memory devices (game backup devices) were widely adopted in later years to copy game media. Early in the transition from solid-state to optical media, CD-based console systems did not have regional market segmentation or copy protection measures due to the rarity and high cost of user-writeable media at the time.

Modchips started to surface with the PlayStation system, due to its popularity and the increasing availability and affordability of CD writers. At the time, a modchip's sole purpose was to allow the use of imported and copied game media.

Today, modchips are available for practically every current console system, often in a great number of variations. In addition to circumventing regional lockout and copy protection mechanisms, modern modchips may introduce more sophisticated modifications to the system, such as allowing the use of user-created software (homebrew), expanding the hardware capabilities of its host system, or even installing an alternative operating system to completely re-purpose the host system (*e.g.* for use as a home theater PC).

Anti-modchip measures

Most modchips open the system to copied media, therefore the availability of a modchip for a console system is undesirable for console manufacturers. They react by removing the intrusion points exploited by a modchip from subsequent hardware or software versions, changing the PCB layout the modchips are customized for, or by having the firmware or software detect an installed modchip and refuse operation as a consequence. Since modchips oftentimes hook into fundamental functions of the host system that cannot be removed or adjusted, these measures may not completely prevent a modchip from functioning but only prompt an adjustment of its installation process or programming, e.g. to include measures to make it undetectable ("stealth") to its host system.

With the advent of online services to be used by video game consoles, some manufacturers have executed their possibilities within the service's license agreement to ban consoles equipped with modchips from using those services.

Some console manufacturers included the option to run homebrew software or even an alternative operating system on their consoles, however, these features have mostly been withdrawn again. An argument can be made that a console system remains largely untouched by modchips as long as their manufacturers provide a legitimate way of running unlicensed third-party software.

Legality

One of the functions of many modchips—the circumvention of copy protection mechanisms—is outlawed by many countries' copyright laws such as the DMCA in the USA, the EUCD and its various implementations by the EU member countries, and the Australian Copyright Act. However, due to the many diversified functions of a modchip, other laws may apply to a modchip as well, allowing specific functions (e.g. the circumvention of region coding under Australian law).

The nonuniform interpretation of applicable law by the courts and constant profound changes and amendments to copyright law do not allow for a definitive statement on the legality of modchips. Because of the ambiguity of applicable laws, a modchip's legality under a country's legislature may only be individually asserted in court.

Most of the very few cases that have been brought before a court have ended with the conviction of the modchip merchant or the manufacturer under the countries' anti-circumvention laws. A small number of cases in the United Kingdom and Australia have been dismissed under the argument that a system's copy protection mechanism would not be able to prevent the actual infringement of copyright (which would be the actual process of copying game media), therefore it cannot be considered an effective TPM protected by anti-circumvention laws. Australian copyright law has since been amended to effectively close this legal loophole.

Chapter 5

Modding

Modding is a slang expression that is derived from the verb "modify". Modding refers to the act of modifying a piece of hardware or software or anything else for that matter, to perform a function not originally conceived or intended by the designer. The term modding is often used within the computer game community, particularly in regard to creating new or altered content and sharing that via the web. It may also be applied to the overclocking of computers in order to increase the frequency at the which the CPU operates. Case modding is also a popular activity amongst many computer enthusiasts which involves the customization of a computer chassis or the installation of water cooling technology. In connection with automobiles, modding often refers to engine tuning, remapping of a vehicle's engine control unit or customization of the bodywork.

Computers and digital equipment

Legal issues

Modding may sometimes infringe the legal rights of the copyright owner. Some nations have laws prohibiting modding and accuse modders of attempting to overcome copy protection schemes. In the United States, the DMCA has set up stiff penalties for mods that violate the rights of intellectual property owners. In the European Union, member states have agreed the EU Copyright Directive and are transposing it into national law. A man was convicted in the United Kingdom in July 2005 for selling a modded Xbox with built in software and games. However it is also worthy of note that some other European countries have not interpreted the legal issues in the same way. In Italy a judge threw out a Sony case saying it was up to owners of a console what they did with it. Similarly in Spain, mod chips are seen as legal despite the EU copyright legislation. Modding may be an unauthorized changed made to a software or hardware to a platform in gaming. Case mods are modifications to a device with the altering of certain styles. For example, people who mod a Microsoft Xbox 360 can alter the led lights on the controller to glow different colors.

Multi-user licensing

Computer systems, hardware, and software are often sold or licensed to one home machine, to a business that has many computers, a government agency, or a non-profit organization. When the software license says that it is for a specific person, then it is not legal for that software to be used by some other person on that same computer, even a member of the same family, or another employee of the same company. But this strict licensing is only one approach. In this form of licensing, for more than one person to be using that software or hardware, they need to have a multi-user license that usually dictates how many different people may use it.

Derivative software

Some software is licensed with a copy of the program source code supplied along with the executable code in which the license specifically authorizes changes to the supplied software. This is a common standard in business software packages. Hundreds of thousands of computer programmers in some nations have jobs because businesses want the purchased software tailored to the specific needs of the individual businesses. Most every major city has want ads in the newspaper where there are job openings for people to modify some company's computer systems, where the ad specifies what programming languages or operating systems the applicant needs to know.

Derivative software is licensed copyrighted software used in whole or part in new programs written by programmers at businesses using the licensed software. Programmers copy the copyright notices into the source code where the code was copied, and track all such places, because if the license is permitted to expire, then the business loses software use rights, including any place to which it was copied. An annual fee is typically paid to keep the license in effect, and over time, the software supplier can increase the fees to the point that the business chooses to convert to some other commercial software that seems to be more cost effective. It is not unusual in business software to find programs that have many different copyright notices, each referring to different sources of the derived source code.

Video game consoles

A common example of one kind of modding is video game console mod chips, which can allow users to play homemade games, games legitimately purchased in other regions, or legal backup copies, but can also allow illegal unauthorized copies by allowing the player to play personally recorded CD or DVD copies of video games. Modchips, in their current form, were first available for the Sony PlayStation (and later the PlayStation 2). Various other types of copyright circumvention systems also existed for the Nintendo 64 and the older Game Boy consoles (though neither include actual modding, but instead backup devices).

Types of modding

There are two different ways of running unsigned code on a game console. One is through soft modding (modifying software, normally using a softmod) to allow the user to change data contained on its hard drive in the case of the Xbox. Another type of modding, known as hard modding, exploits the BIOS of the console to run unsigned code, or games. This form of 'modding' (more correctly termed as hacking) is very popular as it is able to 'run' many different types of software. But soft modding is even more popular because of its ease of installation and its relatively low price (it can even be done for free with the right tools).

Another type of console modding is about appearances, much like computer case modification. Which includes, adding lights (most likely LEDS, cathodes or other electro-luminescent lighting). Cutting the game system case, to fit hardware and/or expose the internal systems. Cooling is a large part of console hard 'modding', including: heat sink upgrades, more powerful or quieter fans, some even go so far as to abandon common heat exchange to air all together by liquid a console (most notably in the Xbox 360, which initially had some heat problems).

Game software

On the other side, some companies actively encourage modding of their products. In cases such as TiVo and Google, there has been an informal agreement between the modders and the company in which the modders agree not to do anything that destroys the company's business model and the company agrees to support the modding community by providing technical specifications and information. Some commercial video games thrive through a modding community. In the case of *Half-Life*, a mod called *Counter-Strike* drove sales of the original software for years.

Many games, such as *The Elder Scrolls series*, come with a mod editing tool that allows users to create original content for themselves and others. Other games provide the source code for users to use in experimenting and creating. Still others, like *18 Wheels of Steel*, will provide the non-programmed data (images, small codepieces and the like) in a simple archive, which can often be opened by renaming it to a .zip file. Often modders will take the game in directions that the developers never anticipated or didn't have time/funding to include. Generally, a small percent of game players will spend much time mod making, but those who do usually develop communities around modding a particular game. Communities are generally connected via a web forum where new modders can ask questions of more experienced ones, and everyone can find inspiration in the work of others. Some games, like *Neverwinter Nights*, could never have been as successful as they are without a thriving user community. And as more people have been more connected via the web, this has become a vital and dynamic creative phenomena where users become content creators not just content consumers.

Skilled computer users who are able to crack data formats and reverse engineer a game can modify them to their heart's content, because the creator of the software has copyright authority over who may use it, or change it. Software is sold with a license that spells out what guarantees, if any, come with the software, and what rights the purchaser has to change the software. Many people do not read these contracts, or store them in such a way to be able to prove what contract came with what purchase, so some computer users are ignorant to what their rights are with respect to backing up software, modifying software, and sharing it with other potential users.

On August 5, 2009 Matthew Crippen, a student at California State University, was arrested for modifying game consoles such as Xbox, Playstation, and Nintendo Wii for profit. According to him it was so that the owners could play their backup discs of games they legally own. However, according to the DMCA, it is illegal to circumvent copyright protection software, even for non-pirating uses such as backing up legally owned games.

Device drivers

Modded drivers are made for improved performance which official versions of drivers do not offer or in cases where there are no official versions of drivers for new hardware designed for older operating systems such as Windows 98.

Computer hardware

Case modding may range from simple case painting to full blown case mod with cooling mods and fabricated pieces.

Overclocking may also be termed as 'modding', and the overclocking of a graphics card using driver software to gain the performance of a more expensive model is known as 'soft-modding'. Volt modding is a term in which jumpers and rheostats are used to mod a hardware's voltages to yield better overclocks.

Cars and vehicles

Orthopedic

Ortho-modding is the car adaptation (seats, pedals, etc.) to help drivers to prevent, correct and diminish light orthopedic and backbone/spine problems.

Eco-modding

Eco-modding is the reduction of drag, petroleum car adaptation to use renewable energy (generally, changing or adding a new engine or motor), generally hydrogen or electricity. Occasionally, it has been known to run a Diesel engine on plant and animal oils.

Performance tuning

Car and engine tuning are modification of a car for enhanced performance, in engine power output, handling, aerodynamics, or cosmetic purposes.

Industrial machines

Factories get rather expensive machines that are used to mass produce specialized parts. These machines can be altered to make parts other than how the manufacturers of the machines designed or intended them. The legality of doing this depends on who owns the machines, and whether the agreement, that supplied the machines to the factories, said anything about this, and what the laws are in the nation where this is being done.

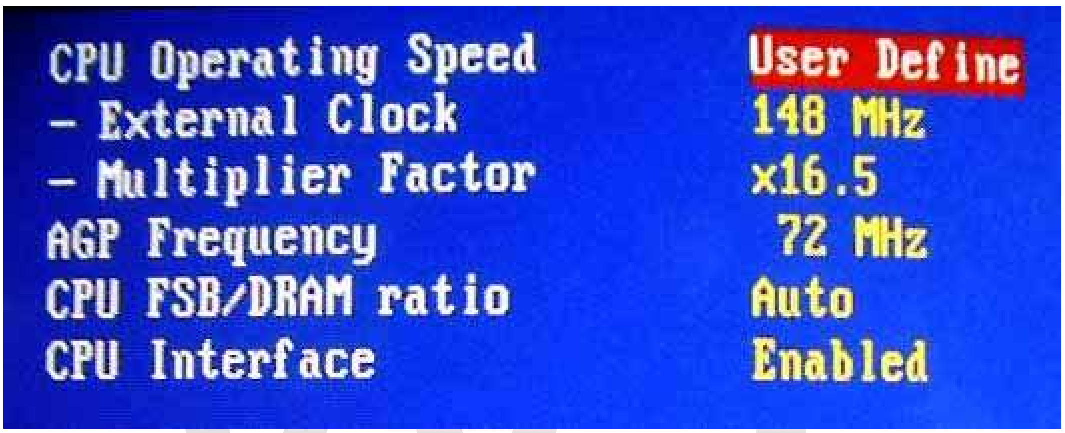
For example, the machines might be leased from the manufacturer of the machines. If they are ever to be returned, they need to go back in the same kind of condition and engineering shape as when they were first delivered. There is an annual physical inventory to make sure the factory has everything that they are leasing. This audit might be done by representatives of the leasing company, who are looking to see recognizable machines, that match their models and safety rules.

Pens

Pen-modding is the act of combining many pen parts either to help with pen spinning, in which a perfect balance is desired to create an ideal spinning pen, or simply for decoration. These pen mods can either be made by combining parts from different pens and/or mechanical pencils, or by buying modded pens online. In some cases, pen mods can exceed over $30–40 USD per pen. Recently, the practice of pen modding has grown dramatically in popularity, with several mod brands appearing, and multiple online communities dedicated to pen modding and spinning.

Chapter 6

Overclocking



AMD Athlon XP overclocking BIOS setup on ABIT NF7-S. Front side bus frequency (external clock) has increased from 133 MHz to 148 MHz, and the clock multiplier factor has changed from 13.5 to 16.5

Overclocking is the process of running a computer component at a higher clock rate (more clock cycles per second) than it was designed for or was specified by the manufacturer, usually practiced by enthusiasts seeking an increase in the performance of their computers. Some purchase low-end computer components which they then overclock to higher clock rates, or overclock high-end components to attain levels of performance beyond the specified values. Others overclock outdated components to keep pace with new system requirements, rather than purchasing new hardware.

People who overclock their components mainly focus their efforts on processors, video cards, motherboard chipsets, and RAM. It is done through manipulating the CPU multiplier and the motherboard's front-side bus (FSB) clock rate until a maximum stable operating frequency is reached, although with the introduction of Intel's new X58 chipset and the Core i7 processor, the front side bus has been replaced with the QPI (Quick Path Interconnect); often this is called the Baseclock (BCLK). While the idea is simple, variation in the electrical and physical characteristics of computing systems complicates the process. CPU multipliers, bus dividers, voltages, thermal loads, cooling techniques and several other factors such as individual semiconductor clock and thermal tolerances can affect it.

FSB × Multiplier = Frequency

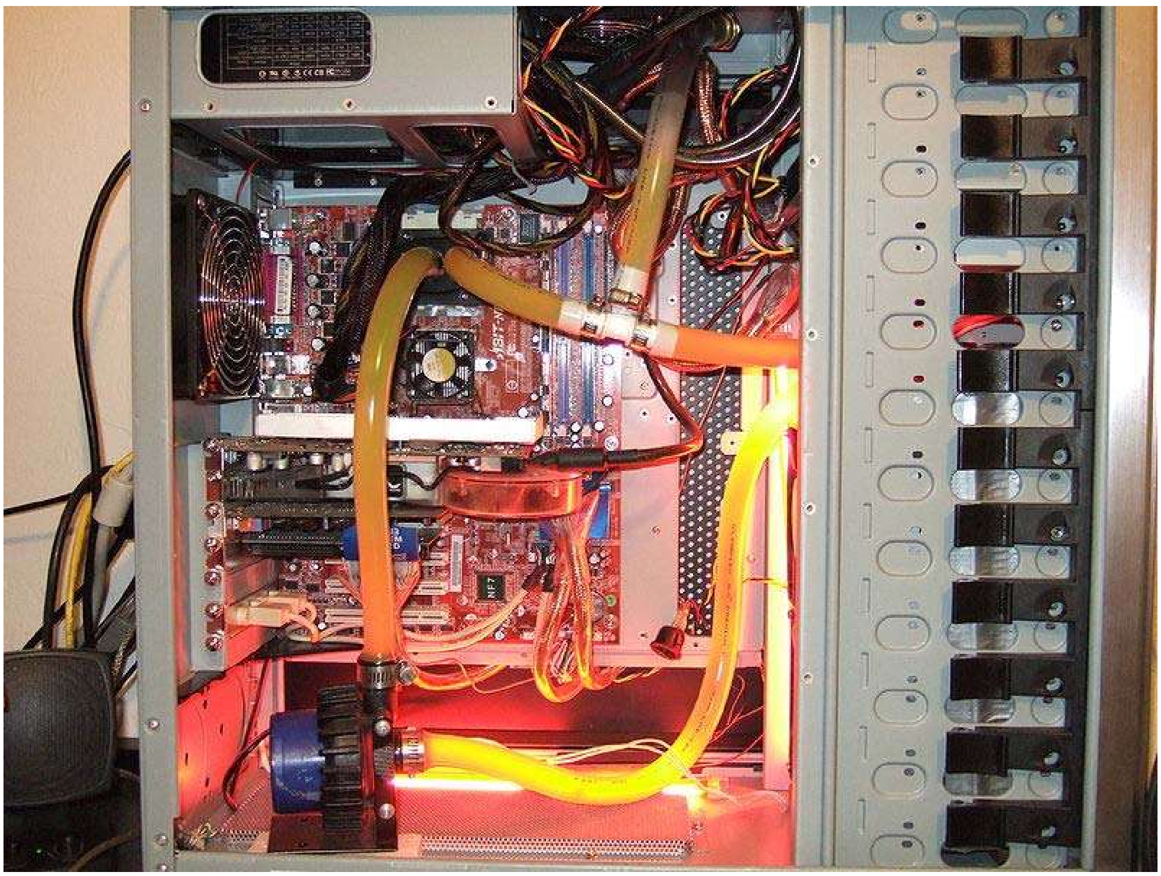
Considerations

There are several considerations when overclocking. First is to ensure that the component is supplied with adequate power to operate at the new clock rate. However, supplying the power with improper settings or applying excessive voltage can permanently damage a component. Since tight tolerances are required for overclocking, only more expensive motherboards—with advanced settings that computer enthusiasts are likely to use—have built-in overclocking capabilities. Motherboards with fewer features, such as those found in original equipment manufacturer (OEM) systems, often do not support overclocking.

Cooling

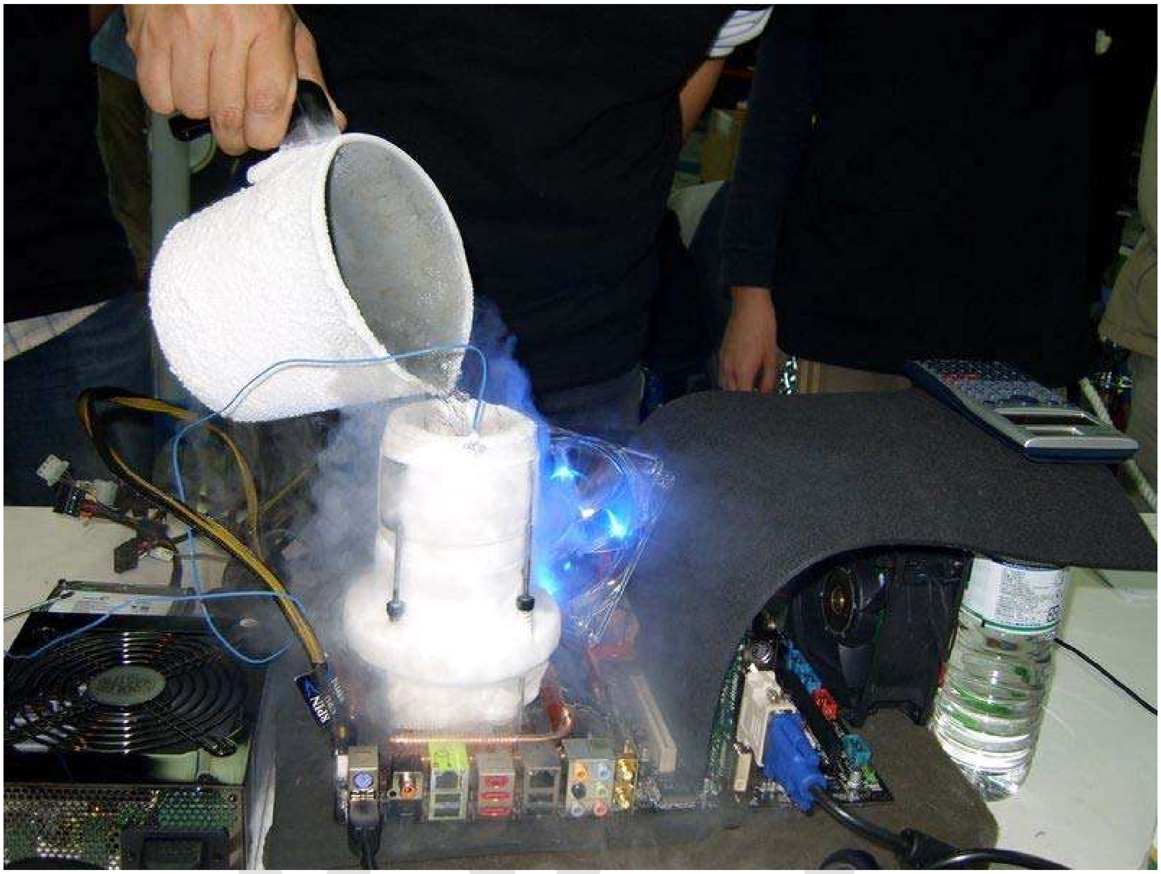
High quality heat sinks are often made of copper

All electronic circuits produce heat generated by the movement of electrical current. As clock frequencies in digital circuits and voltage applied increase, the heat generated by components running at the higher performance levels also increases. The relationship between clock frequencies and thermal design power (TDP) are linear. However, there is a limit to the maximum frequency which is called a "wall". To overcome this issue, overclockers raise the chip voltage to increase the overclocking potential. The relationship between chip voltage and TDP is exponential due to the fact that as the chip warms, the resistance increases. This increased heat requires effective cooling to avoid damaging the hardware. In addition, some digital circuits slow down at high temperatures due to changes in MOSFET device characteristics. Because most stock cooling systems are designed for the amount of power produced during non-overclocked use, overclockers typically turn to more effective cooling solutions, such as powerful fans, larger heat sinks, heat pipes and water cooling. Size, shape, and material all influence the ability of a heatsink to dissipate heat. Efficient heatsinks are often made entirely of copper, which has high thermal conductivity, but is expensive. Aluminium is more widely used; it has poorer thermal conductivity, but is significantly cheaper than copper. Heat pipes are commonly used to improve conductivity. Many heatsinks combine two or more materials to achieve a balance between performance and cost.



Interior of a water-cooled computer, showing CPU water block, tubing and pump

Water cooling carries waste heat to a radiator. Thermoelectric cooling devices, also known as Peltier devices, are recently popular with the onset of high thermal design power (TDP) processors made by Intel and AMD. Thermoelectric cooling devices create temperature differences between two plates by running an electric current through the plates. This method of cooling is highly effective, but itself generates significant heat. For this reason, it is often necessary to supplement thermoelectric cooling devices with a convection-based heatsink or a water-cooling system.



Liquid nitrogen may be used for cooling an overclocked system, when an extreme measure of cooling is needed.

Other cooling methods are forced convection and phase change cooling which is used in refrigerators and can be adapted for computer use. Liquid nitrogen, liquid helium, and dry ice are used as coolants in extreme cases, such as record-setting attempts or one-off experiments rather than cooling an everyday system. In June 2006, IBM and Georgia Institute of Technology jointly announced a new record in silicon-based chip clock rate above 500 GHz, which was done by cooling the chip to 4.5 K (−268.7 °C; −451.6 °F) using liquid helium. These extreme methods are generally impractical in the long term, as they require refilling reservoirs of vaporizing coolant, and condensation can be formed on chilled components. Moreover, silicon-based junction gate field-effect transistors (JFET) will degrade below temperatures of roughly 100 K (−173 °C; −280 °F) and eventually cease to function or "freeze out" at 40 K (−233 °C; −388 °F) since the silicon ceases to be semiconducting so using extremely cold coolants may cause devices to fail.

Submersion cooling, used by the Cray-2 supercomputer, involves sinking a part of computer system directly into a chilled liquid that is thermally conductive but has low electrical conductivity. The advantage of this technique is that no condensation can form on components. A good submersion liquid is Fluorinert made by 3M, which is expensive and can only be purchased with a permit. Another option is mineral oil, but impurities such as those in water might cause it to conduct electricity.

Stability and functional correctness

As an overclocked component operates outside of the manufacturer's recommended operating conditions, it may function incorrectly, leading to system instability. Another risk is silent data corruption by undetected errors. Such failures might never be correctly diagnosed and may instead be incorrectly attributed to software bugs in applications, device drivers, or the operating system. Overclocked use may permanently damage components enough to cause them to misbehave (even under normal operating conditions) without becoming totally unusable.

In general, overclockers claim that testing can ensure that an overclocked system is stable and functioning correctly. Although software tools are available for testing hardware stability, it is generally impossible for any private individual to thoroughly test the functionality of a processor. Achieving good fault coverage requires immense engineering effort; even with all of the resources dedicated to validation by manufacturers, faulty components and even design faults are not always detected.

A particular "stress test" can verify only the functionality of the specific instruction sequence used in combination with the data and may not detect faults in those operations. For example, an arithmetic operation may produce the correct result but incorrect flags; if the flags are not checked, the error will go undetected.

To further complicate matters, in process technologies such as silicon on insulator (SOI), devices display hysteresis—a circuit's performance is affected by the events of the past, so without carefully targeted tests it is possible for a particular sequence of state changes to work at overclocked rates in one situation but not another even if the voltage and temperature are the same. Often, an overclocked system which passes stress tests experiences instabilities in other programs.

In overclocking circles, "stress tests" or "torture tests" are used to check for correct operation of a component. These workloads are selected as they put a very high load on the component of interest (e.g. a graphically-intensive application for testing video cards, or different math-intensive applications for testing general CPUs). Popular stress tests include Prime95, Everest, Superpi, OCCT, IntelBurnTest/Linpack/LinX, SiSoftware Sandra, BOINC, Intel Thermal Analysis Tool and Memtest86. The hope is that any functional-correctness issues with the overclocked component will show up during these tests, and if no errors are detected during the test, the component is then deemed "stable". Since fault coverage is important in stability testing, the tests are often run for long periods of time, hours or even days. An overclocked computer is sometimes described using the number of hours and the stability program used, such as "prime 12 hours stable".

Factors allowing overclocking

Overclockability arises in part due to the economics of the manufacturing processes of CPUs and other components. In most cases components with different rated clock rates are manufactured by the same process, and tested after manufacture to determine their actual ratings. The clock rate that the component is rated for is at or below the clock rate at which the CPU has passed the manufacturer's functionality tests when operating in worst-case conditions (for example, the highest allowed temperature and lowest allowed supply voltage). Manufacturers must also leave additional margin for reasons discussed below. Sometimes manufacturers produce more high-performing parts than they can sell, so some are marked as medium-performance chips to be sold for medium prices. Pentium architect Bob Colwell calls overclocking an "uncontrolled experiment in better-than-worst-case system operation".

Measuring effects of overclocking

Benchmarks are used to evaluate performance. The benchmarks can themselves become a kind of 'sport', in which users compete for the highest scores. As discussed above, stability and functional correctness may be compromised when overclocking, and meaningful benchmark results depend on correct execution of the benchmark. Because of this, benchmark scores may be qualified with stability and correctness notes (e.g. an overclocker may report a score, noting that the benchmark only runs to completion 1 in 5 times, or that signs of incorrect execution such as display corruption are visible while running the benchmark). A widely used test of stability is Prime95 as this has in-built error checking and the computer fails if unstable.

Given only benchmark scores it may be difficult to judge the difference overclocking makes to the overall performance of a computer. For example, some benchmarks test only one aspect of the system, such as memory bandwidth, without taking into consideration how higher clock rates in this aspect will improve the system performance as a whole. Apart from demanding applications such as video encoding, high-demand databases and scientific computing, memory bandwidth is typically not a bottleneck, so a great increase in memory bandwidth may be unnoticeable to a user depending on the applications used. Other benchmarks, such as 3DMark attempt to replicate game conditions.

Variance

The extent to which a particular part will overclock is highly variable. Processors from different vendors, production batches, steppings, and individual units will all overclock with varying degrees of success.

Manufacturer and vendor overclocking

Commercial system builders or component resellers sometimes overclock to sell items at higher profit margins. The retailer makes more money by buying lower-value components, overclocking them, and selling them at prices appropriate to a non-overclocked system at the new clock rate. In some cases an overclocked component is functionally identical to a non-overclocked one of the new clock rate, however, if an overclocked system is marketed as a non-overclocked system (it is generally assumed that unless a system is specifically marked as overclocked, it is not overclocked), it is considered fraudulent.

Overclocking is sometimes offered as a legitimate service or feature for consumers, in which a manufacturer or retailer tests the overclocking capability of processors, memory, video cards, and other hardware products. Several video card manufactures now offer factory overclocked versions of their graphics accelerators, complete with a warranty, which offers an attractive solution for enthusiasts seeking an improved performance without sacrificing common warranty protections. Such factory-overclocked products may cost a little more than standard components, but may be more cost-effective than a product with a higher specification.

Naturally, manufacturers would prefer enthusiasts to pay additional money for profitable high-end products, in addition to concerns of less reliable components and shortened product life spans affecting brand image. It is speculated that such concerns are often motivating factors for manufacturers to implement overclocking prevention mechanisms such as CPU locking. These measures are sometimes marketed as a consumer protection benefit, which typically generates a negative reception from overclocking enthusiasts.

Advantages

- The user can, in many cases, purchase a lower performance, cheaper component and overclock it to the clock rate of a more expensive component.
- Higher performance in games, encoding, video editing applications, and system tasks at no additional expense, but at an increased cost for electrical power consumption. Particularly for enthusiasts who regularly upgrade their hardware, overclocking can increase the time before an upgrade is needed.
- Some systems have "bottlenecks," where small overclocking of a component can help realize the full potential of another component to a greater percentage than the limiting hardware is overclocked. For instance, many motherboards with AMD Athlon 64 processors limit the clock rate of four units of RAM to 333 MHz. However, the memory performance is computed by dividing the processor clock rate (which is a base number times a CPU multiplier, for instance 1.8 GHz is most likely 9×200 MHz) by a fixed integer such that, at a stock clock rate, the RAM would run at a clock rate near 333 MHz. Manipulating elements of how the processor clock rate is set (usually lowering the multiplier), one can often overclock the processor a small amount, around 100–200 MHz (less than 10%), and gain a RAM clock rate of 400 MHz (20% increase), releasing the full potential of the RAM.
- Overclocking can be an engaging hobby in itself and supports many dedicated online communities. The PCMark website is one such site that hosts a leader-board for the most powerful computers to be bench-marked using the program.
- A new overclocker with proper research and precaution or a guiding hand can gain useful knowledge and hands-on experience about their system and PC systems in general.

Disadvantages

Many of the disadvantages of overclocking can be mitigated or reduced in severity by skilled overclockers. However, novice overclockers may make mistakes while overclocking which can introduce avoidable drawbacks and which are more likely to damage the overclocked components (as well as other components they might affect).

General

- The lifespan of a processor may be reduced by higher operating frequencies, increased voltages and heat, although processors rapidly become obsolete in performance due to technological progress.
- Increased clock rates and/or voltages result in higher power consumption.
- While overclocked systems may be tested for stability before use using programs that "burn" the computer, these programs create an artificial strain that pushes individual or many components to their maximum (or beyond it). Some common stability programs are Prime95, Super PI (32M), Intel TAT, LinX, PCMark, FurMark and OCCT. Stability problems may surface after prolonged usage due to new workloads or untested portions of the processor core. Aging effects previously discussed may also result in stability problems after a long period of time. Even when a computer appears to be working normally, problems may arise in the future. For example, Windows may appear to work with no problems, but when it is re-installed or upgraded, error messages may be received such as a “file copy error" during Windows Setup. Microsoft says this of errors in upgrading to Windows XP: "Your computer [may be] over-clocked. Because over-clocking is very memory-intensive, decoding errors may occur when files are extracted from the Windows XP CD-ROM".
- High-performance fans used for extra cooling can be noisy. Older popular models of fans used by overclockers can produce 50 decibels or more. However, nowadays, manufacturers are overcoming this problem by designing fans with aerodynamically optimized blades for smoother airflow and minimal noise (around 20 decibels at approximately 1 metre). The noise is not always acceptable, and overclocked machines are often much noisier than stock machines. Noise can be reduced by utilizing strategically-placed larger fans, which are inherently less noisy than smaller fans; by using alternative cooling methods (such as liquid and phase-change cooling); by lining the chassis with foam insulation; and by installing a fan-controlling bus to adjust fan speed (and, as a result, noise) to suit the task at hand. Now that overclocking is of interest to a larger target audience, this is less of a concern as manufacturers have begun researching and producing high-performance fans that are no longer as loud as their predecessors. Similarly, mid- to high-end PC cases now implement larger fans (to provide better airflow with less noise) as well as being designed with cooling and airflow in mind.
- Even with adequate CPU cooling, the excess heat produced by an overclocked processing unit increases the ambient air temperature of the system case; consequently, other components may be affected. Also, more heat will be

expelled from the PC's vents, raising the temperature of the room the PC is in - sometimes to uncomfortable levels.

- Overclocking has the potential to cause component failure ("heat death"). Most warranties do not cover damage caused by overclocking. Some motherboards offer safety measures that will stop this from happening (e.g. limitations on FSB increase) so that only voltage control alterations can cause such harm.
- Some motherboards are designed to use the airflow from a standard CPU fan in order to cool other heatsinks, such as the northbridge. If the cpu heatsink is changed on such boards, other heatsinks may receive insufficient cooling.
- Overclocking a PC component may void its warranty (depending on the conditions of sale).
- Changing the Heatsink on a Graphics Card often voids its warranty

Incorrectly performed overclocking

- Increasing the operation frequency of a component will usually increase its thermal output in a linear fashion, while an increase in voltage usually causes heat to increase quadratically. Excessive voltages or improper cooling may cause chip temperatures to rise almost instantaneously, causing the chip to be damaged or destroyed.
- More common than hardware failure is functional incorrectness. Although the hardware is not permanently damaged, this is inconvenient and can lead to instability and data loss. In rare, extreme cases entire filesystem failure may occur, causing the loss of all data.
- With poor placement of fans, turbulence and vortices may be created in the computer case, resulting in reduced cooling effectiveness and increased noise. In addition, improper fan mounting may cause rattling or vibration.
- Improper installation of exotic cooling solutions like liquid cooling may result in failure of the cooling system, which may result in water damage.
- With sub-zero cooling methods such as phase-change cooling or liquid nitrogen, extra precautions such as foam or spray insulation must be made to prevent water from condensing upon the PCB and other areas. This can cause the board to become "frosted" or covered in frost. While the water is frozen it is usually safe, however once it melts it can cause shorts and other malignant issues.
- Sometimes products claim to be intended specifically for overclocking and may be just decoration. Novice buyers should be aware of the marketing hype surrounding some products. Examples include heat spreaders and heat sinks designed for chips (or components) which do not generate enough heat to benefit from these devices (capacitors, for example).

Limitations

The utility of overclocking is limited for a few reasons:

- Personal computers are mostly used for tasks which are not computationally demanding, or which are performance-limited by bottlenecks outside of the local machine. For example, web browsing does not require a high performance computer, and the limiting factor will almost certainly be the bandwidth of the Internet connection of either the user or the server. Overclocking a processor will also do little to help increase application loading times as the limiting factor is reading data off the hard drive. Other general office tasks such as word processing and sending email are more dependent on the efficiency of the user than on the performance of the hardware. In these situations any performance increases through overclocking are unlikely to be noticeable.

- It is generally accepted that, even for computationally-heavy tasks, clock rate increases of less than ten percent are difficult to discern. For example, when playing video games, it is difficult to discern an increase from 60 to 66 frames per second (FPS) without the aid of an on-screen frame counter. Overclocking of a processor will rarely improve gaming performance noticeably, as the frame rates achieved in most modern games are usually bound by the GPU at resolutions beyond 1024×768. One exception to this rule is when the overclocked component is the bottleneck of the system, in which case the most gains can be seen.

Graphics cards

The BFG GeForce 6800GSOC ships with higher memory and clock rates than the standard 6800GS.

Graphics cards can also be overclocked, with utilities such as EVGA's Precision, RivaTuner, ATI Overdrive (on ATI cards only), MSI Afterburner, Zotac Firestorm on Zotac cards, or the PEG Link Mode on Asus motherboards. Overclocking a GPU will often yield a marked increase in performance in synthetic benchmarks, and usually will improve game performance too. Sometimes, it is possible to see that a graphics card is pushed beyond its limits before any permanent damage is done by observing on-screen distortions known as artifacts. Two such discriminated "warning bells" are widely understood: green-flashing, random triangles appearing on the screen usually correspond to overheating problems on the GPU itself, while white, flashing dots appearing randomly (usually in groups) on the screen often mean that the card's RAM is overheating. It is common to run into one of those problems when overclocking graphics cards. Showing both symptoms at the same time usually means that the card is severely pushed beyond its heat/clock rate/voltage limits. If seen at normal clock rate, voltage and temperature, they may indicate faults with the card itself. However, if the video card is simply clocked too high and doesn't overheat then the artifacts are a bit different. There are many different ways for this to show up and any irregularities should be considered but usually if the core is pushed too hard black circles or blobs appear on the screen and overclocking the video memory beyond its limits usually results in the application or the entire operating system crashing. Luckily though, after the computer is restarted the settings is reset to stock (stored in the video card firmware) and the maximum clock rate of that specific card has been found.

Some overclockers use a hardware voltage modification where a potentiometer is applied to the video card to manually adjust the voltage. This results in much greater flexibility, as overclocking software for graphics cards is rarely able to freely adjust the voltage. Voltage mods are very risky and may result in a dead video card, especially if the voltage modification ("voltmod") is applied by an inexperienced individual. A pencil volt mod refers to changing a resistor's value on the graphics card by drawing across it with a graphite pencil. This results in a change of GPU voltage. It is also worth mentioning that adding physical elements to the video card immediately voids the warranty.

Alternatives

Flashing and Unlocking are two popular ways to gain performance out of a video card, without technically overclocking.

Flashing refers to using the firmware of another card, based on the same core and design specs, to "override" the original firmware, thus effectively making it a higher model card; however, 'flashing' can be difficult, and sometimes a bad flash can be irreversible. Sometimes stand-alone software to modify the firmware files can be found, i.e. NiBiTor, (GeForce 6/7 series are well regarded in this aspect). It is not necessary to acquire a firmware file from a better model video card (although it should be said that the card in which firmware is to be used should be compatible, i.e. the same model base, design and/or manufacture process, revisions etc.). For example, video cards with 3D accelerators (the vast majority of today's market) have two voltage and clock rate settings - one for 2D and one for 3D - but were designed to operate with *three* voltage stages, the third being somewhere in the middle of the aforementioned two, serving as a fallback when the card overheats or as a middle-stage when going from 2D to 3D operation mode. Therefore, it could be wise to set this middle-stage prior to "serious" overclocking, specifically because of this fallback ability - the card can drop down to this clock rate, reducing by a few (or sometimes a few dozen, depending on the setting) percent of its efficiency and cool down, without dropping out of 3D mode (and afterwards return to the desired high performance clock and voltage settings).

Some cards also have certain abilities not directly connected with overclocking. For example, Nvidia's GeForce 6600GT (AGP flavor) features a temperature monitor (used internally by the card), which is invisible to the user in the 'vanilla' version of the card's firmware. Modifying the firmware can allow a 'Temperature' tab to become visible in the card driver's advanced menu.

Unlocking refers to enabling extra pipelines and/or pixel shaders. The 6800LE, the 6800GS and 6800 (AGP models only) and Radeon X800 Pro VIVO were some of the first cards to benefit from unlocking. While these models have either 8 or 12 pipes enabled, they share the same 16x6 GPU core as a 6800GT or Ultra, but may not have passed inspection when all their pipelines and shaders were unlocked. In more recent generations, both ATI and Nvidia have laser cut pipelines to prevent this practice..

It is important to remember that while pipeline unlocking sounds very promising, there is absolutely no way of determining if these 'unlocked' pipelines will operate without errors, or at all (this information is solely at the manufacturer's discretion). In a worst-case scenario, the card may not start up ever again, resulting in a 'dead' piece of equipment. It is possible to revert to the card's previous settings, but it involves manual firmware flashing using special tools and an identical but original firmware chip.

Chapter 7

Power Management and SpeedFan

Power management

Power management is a feature of some electrical appliances, especially copiers, computers and computer peripherals such as monitors and printers, that turns off the power or switches the system to a low-power state when inactive. In computing this is known as PC power management and is built around a standard called ACPI. This supersedes APM. All recent (consumer) computers have ACPI support.

Motivations

PC power management for computer systems is desired for many reasons, particularly:

- Reduce overall energy consumption
- Prolong battery life for portable and embedded systems
- Reduce cooling requirements
- Reduce noise.
- Reduce operating costs for energy and cooling.

Lower power consumption also means lower heat dissipation, which increases system stability, and less energy use, which saves money and reduces the impact on the environment.

Processor level techniques

The power management for microprocessors can be done over the whole processor, or in specific areas.

With dynamic voltage scaling and dynamic frequency scaling, the CPU core voltage, clock rate, or both, can be altered to decrease power consumption at the price of potentially lower performance. This is sometimes done in real time to optimize the power-performance tradeoff.

Examples:

- AMD Cool'n'Quiet
- AMD PowerNow!
- IBM EnergyScale
- Intel SpeedStep
- Transmeta LongRun and LongRun2
- VIA LongHaul (PowerSaver)

Additionally, processors can selectively power off internal circuitry (power gating). For example:

- Newer Intel Core processors support ultra-fine power control over the functional units within the processors.
- AMD CoolCore technology get more efficient performance by dynamically activating or turning off parts of the processor.

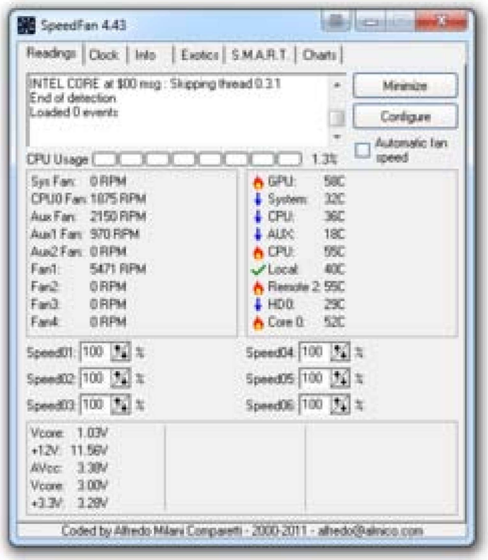
Intel VRT technology split the chip into a 3.3V I/O section and a 2.9V core section. The lower core voltage reduces power consumption.

Operating system level: Hibernation

When a computer system hibernates it saves the contents of the RAM to disk and powers down the machine. On startup it reloads the data. This allows the system to be completely powered off while in hibernate mode. This requires a file the size of the installed RAM to be placed on the hard disk, potentially using up space even when not in hibernate mode. Hibernate mode is enabled by default in some versions of Windows and can be disabled in order to recover this disk space.

SpeedFan

SpeedFan



SpeedFan 4.43 in Windows 7

Original author(s)	Alfredo Milani Comparetti
Developer(s)	Alfredo Milani Comparetti
Initial release	?
Stable release	4.43 (March 17, 2011; 22 days ago) [+/−]
Preview release	4.44 beta 4 (April 7, 2011; 1 day ago) [+/−]
Written in	Delphi, C++, C
Operating system	Windows 95 and later
Available in	Multilanguage
Type	System hardware monitor
License	Freeware

SpeedFan is software that can read temperatures, voltages and fan speeds of computer components. It can change computer fan speeds depending on the temperature of various components. The program can display system variables as a chart and as an indicator in the system tray. Fully configurable user events can be defined to execute specific actions based on system status. As of version 4.37, temperatures in the event section must be specified in °C no matter if the display setting is for °F.

Hard disk support

SpeedFan also monitors S.M.A.R.T. readings for EIDE, SATA and SCSI hard disks. Starting with version 4.35, SpeedFan fully supports Areca RAID controllers. Version 4.38 added full support for AMCC/3ware SATA and RAID controllers.

Hard disk in-depth online analysis

SpeedFan offers a feature named "in-depth online analysis" that compares your hard disk's S.M.A.R.T. data to a comprehensive database with statistical models of every known hard disk allowing early detection of potentially degraded hard disks. Full text messages inform the user of specific situations and problems, as if a human expert had looked at the data.

Risks

SpeedFan crashes some systems when it is launched, possibly causing registry corruption or loss of recent changes to the registry. This can cause recently-installed programs to disappear from Add/Remove Programs, programs to no longer function correctly, or loss of operating-system stability.

Chapter 8

SpeedStep and Tweaking

SpeedStep

SpeedStep is a trademark for a series of dynamic frequency scaling technologies (codenamed **Geyserville** and including SpeedStep, SpeedStep II, and SpeedStep III) built into some Intel microprocessors that allow the clock speed of the processor to be dynamically changed (to different *P-states*) by software. This allows the processor to meet the instantaneous performance needs of the operation being performed, while minimizing power draw and heat dissipation. *Enhanced Intel SpeedStep* is sometimes abbreviated as *EIST*.

Explanation

Running a processor at high clock speeds allows for better performance. However, when the same processor is run at a lower frequency (speed), it generates less heat and consumes less power. In many cases, the core voltage can also be reduced, further reducing power consumption and heat generation. This can conserve battery power in notebooks, extend processor life, and reduce noise generated by variable-speed fans. By using SpeedStep, users can select the balance of power conservation and performance that best suits them, or even change the clock speed dynamically as the processor burden changes.

The power consumed by a CPU with a capacitance C, running at frequency f and voltage V is approximately

$$P = CV^2f.$$

For a given processor, C is a fixed value. However, V and f can vary considerably. For example, for a 1.6 GHz Pentium M, the clock frequency can be stepped in 200 MHz increments over the range from 0.6 to 1.6 GHz. At the same time, the voltage requirement decreases from 1.484 V to 0.956 V. The result is that the power consumption theoretically goes down by a factor of 6.4. In practice, the effect may be smaller because some CPU instructions use less energy per tick of the CPU clock than others. For example, when an operating system is not busy, it tends to issue halt instructions, which suspend operation of parts of the CPU for a time period, so it uses less energy per tick of the CPU clock than when executing productive instructions in its normal state. For a given rate of work, a CPU running at a higher clock rate will execute a greater proportion of HLT instructions. The simple equation which relates power, voltage and frequency above also does not take into account the static power consumption of the CPU. This tends not to change with frequency, but does change with temperature and voltage. Hot electrons, and electrons exposed to a stronger electric field are more likely to migrate across a gate as "gate leakage" current, leading to an increase in static power consumption.

Older processors, using older versions of the SpeedStep technology, have fewer increments, such as the Pentium 4-M. For example, a 1.7 GHz Pentium 4M can run at 1.6 GHz, at 1.2 GHz, and at 786 MHz.

SpeedStep technology is partly responsible for the reduced power consumption of Intel's Pentium M processor, part of the Centrino brand.

Problems when using SpeedStep

Microsoft reports there may be problems previewing video files when SpeedStep (or the AMD equivalent PowerNow!) is enabled. It also may decrease reliability when overclocking.

Operating system support

Older versions of Microsoft Windows, Windows 2000 and earlier, need a special driver and dashboard application to access the SpeedStep feature. Intel's website specifically states that such drivers *must* come from the computer manufacturer; there are no generic drivers supplied by Intel which will enable SpeedStep for older Windows versions if one cannot obtain a manufacturer's driver.

Under Microsoft Windows XP, SpeedStep support is built into the power management console under the control panel. In Windows XP a user can regulate processor speed indirectly by changing power schemes. The "Home/Office Desk" setting disables SpeedStep, the "Portable/Laptop" power scheme enables SpeedStep, and the "Max Battery" uses SpeedStep to slow the processor to minimal power levels as the battery weakens. The SpeedStep settings for power schemes, either built-in or custom, cannot be modified from the control panel's GUI, but can be modified using the POWERCFG.EXE command-line utility.

In contrast, AMD continues to supply and support drivers for its competing PowerNow! technology that will work on Windows 2000, ME, 98, and NT.

Linux has full SpeedStep support integrated into the kernel version 2.6.

Mac OS also has SpeedStep built into the kernel, since the release of the Intel version of Mac OS X 10.4 and is already enabled. It cannot be controlled in the System Preference

"Energy Saver." To disable this feature, and set a specific clock speed (full speed or reduced) requires a third party application, such as coolbook

Solaris has supported SpeedStep since OpenSolaris SXDE 9/07.

The BSD kernels have full SpeedStep support integration.

Versions

V1.1 is used by second generation Pentium III processors. It enables the CPU to switch between two modes: high and low frequency. This is done by modifying the CPU's multiplier. A 1 GHz Pentium III consuming about 20 watts could be reduced to 600 MHz which reduces the power consumption to about 6 watts.

V2.1 (Enhanced SpeedStep) is used in Pentium III-Mobile processors and is similar to the previous version, but in the low frequency mode the CPU also uses a different voltage than the high frequency mode.

V2.2 is adapted for Pentium 4-Mobile processors. With this, a 1.8 GHz Pentium 4-M consuming about 30 watts can lower its frequency to 1.2 GHz, thus reducing power consumption to about 20 watts.

V3.1 (EIST) is used with the first and second generation of Pentium M processors (Banias and Dothan cores, used in Centrino platforms). With this technology, the CPU varies its frequency (and voltage) between about 40% and 100% of its base frequency in increments of 100 MHz (for Banias core) or 133 MHz (for Dothan core). With this technology, Intel also introduces realtime Level 2 cache capacity variation, further improving power savings.

Tweaking

Tweaking refers to fine-tuning or adjusting a complex system, usually an electronic device. Tweaks are any small modifications intended to improve a system.

In electronics, it is a synonym for "trimming." Analog circuit boards often have small potentiometers or other components on them that are used to calibrate or adjust the board as a service procedure: the small insulated screwdriver used to turn them is often called a "tweaker."

This use was echoed in the name of the product *Tweek*, a popular but controversial audio product during the 1980s, which was claimed to improve the electrical characteristics of audio switch contacts.

Hardware

Hardware tweaking is a process of modifying certain parts of a hardware such as changing the cables, cleaning the heads of a VHS player with a branded cleaning fluid or oiling the moving parts of an engine with the best possible oil.

Computer hardware

Computer hardware tweaking is an extension of hardware tweaking, specifically geared towards the components of a PC. They include: changing voltage and clock rates of processing units, modifying RAM unit timing, improving cooling systems to reduce chance of overheating, etc.

Tweaks specifically designed to allow a processor to operate at a higher clock speed than normal are known as overclocking.

Modifications of computer systems not aimed at increasing performance, such as quieter fans, external controls, and decorations such as lights or windows, are known as modding.

Software

Software tweaking is the process of improving the performance of an application or the quality of its output. There can be two ways of accomplishing this: manually (that is, if one is familiar with programming; though it may be irrelevant if the source of the application is closed, and there are no built-in means to adjust its performance) or using another piece of software specialized for that purpose (such as Red Button and SysTool, semi-automatic tweaking utilities used to improve the performance of Microsoft Windows XP. Tweaking of this kind generally increases usability, in terms of personal configuration preferences, rather than objective performance of the system overall). Linux and other open source products are designed to facilitate the tweaking process as much as possible, as opposed to Microsoft Windows which limits tweaking but allows it with provision, and Mac OS which largely prohibits or strongly discourages tweaking.

Some very precise applications need constant and thorough tweaking to stay up to date and deliver best possible results. One of the most obvious examples of such a fine tuning is LAME MP3 encoder, whose 3.9x branch is not only considered as the state-of-the-art MP3 encoder, but also continues to shape the boundaries of the MP3 codec and stay competitive with its successors.

Chapter 9

Underclocking

"**Underclocking**" also known as "**Downclocking**" is the practice of modifying a synchronous circuit's timing settings to run at a lower clock rate than it was specified to operate at. It may be said to be the computer equivalent to drive a car at a speed below the speed limit. Usually, underclocking is used to reduce a computer's power consumption and heat emission, sometimes also to increase the system's stability and compatibility. Underclocking may be implemented by the factory, but many computers and components are end user underclockable.

Types of underclocking

CPU underclocking

For microprocessors, the purpose is generally to decrease the need for heat dissipation devices or decrease the electrical power consumption. This can provide increased system stability in high-heat environments, or can allow a system to run with a lower airflow (and therefore quieter) cooling fan or without one at all. For example, a Pentium 4 processor clocked at 2.4 GHz can be "underclocked" to 1.8 GHz and can then be safely run with reduced fan speeds. This invariably comes at the expense of some system performance. However, the performance usually is reduced less than the reduction of clock speed because the performance often is limited by other bottlenecks: The hard disk, the disk controller, the Internet, the network, etc.

Graphics cards

"Underclocking" can also be performed on graphics card processors (graphics processing units; GPUs), usually with the aim of reducing heat output. For instance, it is possible to set a GPU to run at lower clock rates when performing everyday tasks (e.g. internet browsing and word processing), thus allowing the card to operate at lower temperature and thus lower, quieter fan speeds. The GPU can then be overclocked for more graphically intense applications, such as games. "Underclocking" a GPU will reduce performance, but this decrease will probably not be noticeable except in graphically intensive applications.

Memory "underclocking"

Newer and faster RAM may be "underclocked" to match older systems as an inexpensive way to replace rare or discontinued memory. This might also be necessary if stability problems are encountered at higher settings, espcially in a PC with several memory modules of different clock speed. If you underclock a PC processor, and do not change the clock factor or multipilier (the ratio between the processor and the memory clock speed), the memory will also be underclocked.

When used

Dynamic frequency scaling (automatic "underclocking") is very common on laptop computers and is beginning to emerge on desktop computers as well. In laptops, the processor is usually "underclocked" automatically whenever the computer is operating on batteries. Most newer notebook and some desktop processors (like AMD's Cool'n'Quiet and PowerNow!) will also underclock themselves automatically when under a light processing load. Intel has also used this method on their Core 2 Duo and later processors, through a feature called SpeedStep.

Some processors "underclock" automatically as a defensive measure, to prevent overheating which could cause permanent damage. When such a processor reaches a temperature level deemed too high for safe operation, the *thermal control circuit* activates, automatically decreasing the clock and CPU core voltage until the temperature has returned to a safe level. In a properly cooled environment, this mechanism should trigger rarely (if ever).

There are several different "underclocking" competitions similar in format to overclocking competitions, except the goal is to have the lowest clocked computer, as opposed to the highest.

Advantages

- Reduced electrical power consumption, especially when combined with undervolting (i.e., reduce the component's voltage below the nominal). For instance, by underclock an Athlon XP 1700+ processor from 1466 to 1000 MHz and reduce the core voltage from 1.75 to 1.15V, a computer user reduced the power consumption from 64.0 to 21.6W, i.e. 66 % power reduction and only 26 % less performance. The same is true for newer processors: When a single-core Intel CPU was 20 % underclocked, the PC's performance was down only 13 % with a 49 % power reduction.

In general, the power consumed by a CPU with a capacitance C, running at frequency f and voltage V is approximately

$$P = CV^2f.$$

- Reduced heat generation, which is exactly proportional to the power consumption.
- Less noise because the cooling fans may be slowed down, or even eliminated. A cooling fan's efficiency is proportional to its rotation speed, but it noise grows much more.
- Longer hardware lifespan.
- Increased stability.
- Reduced noise from cooling parts due to reduced heat dissipation requirements.
- Increased battery life.
- Better compability with old applications.

In practice

Linux

Linux kernel supports CPU frequency modulation. In supported processors, using *cpufreq* to gain access to this feature gives the system administrator a variable level of control over the CPU's clock rate. The kernel includes five governors by default: Conservative, Ondemand, Performance, Powersave, and Userspace. The Conservative and Ondemand governors adjust the clock rate depending on the CPU load, but each with different algorithms. The Ondemand governor jumps to maximum frequency on CPU load and decreases the frequency step by step on CPU idle, whereas the Conservative governor increases the frequency step by step on CPU load and jumps to lowest frequency on CPU idle. The Performance, Powersave and Userspace governors set the clock rate statically: Performance to the highest available, Powersave to the lowest available, and Userspace to a frequency determined and controlled by the user.

Windows

"Underclocking" can be done manually in the BIOS or with Windows applications, or dynamically using features such as Intel's SpeedStep or AMD's Cool'n'Quiet.

Asus Eee PC

Some versions of the Asus Eee PC uses a 900 MHz Intel Celeron M processor underclocked to 630 MHz.

Smartphones and PDAs

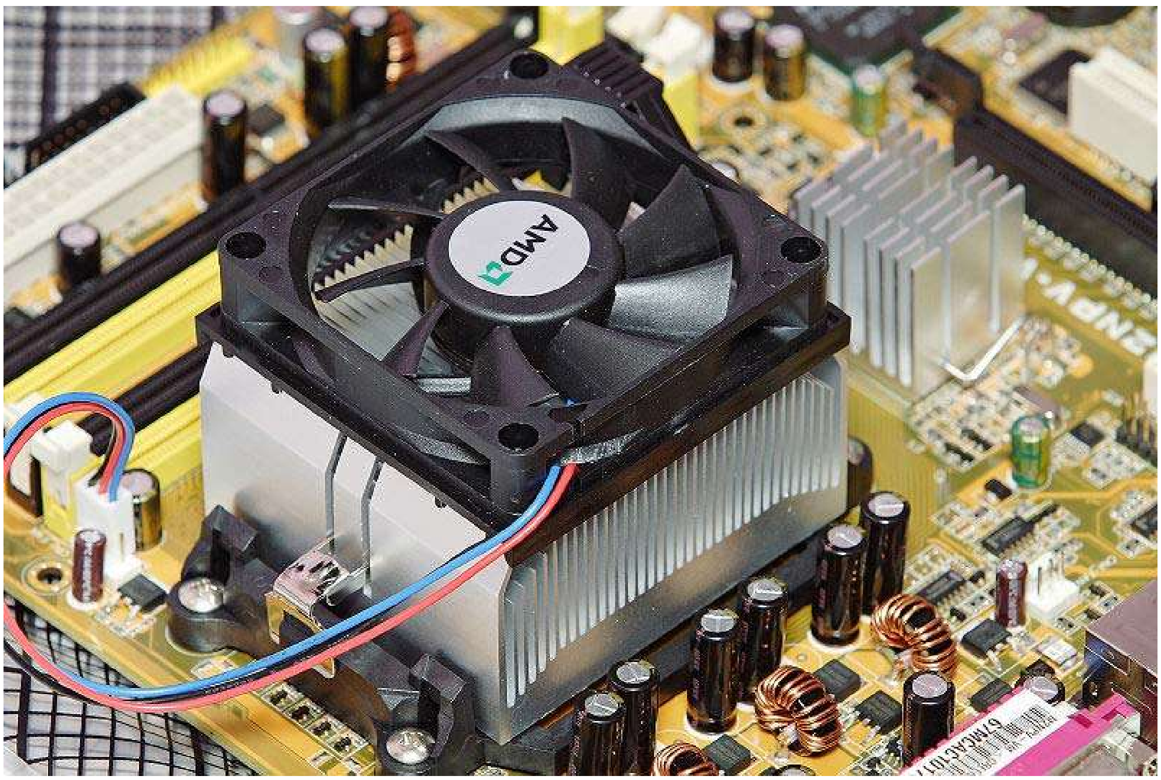
Most smartphones and PDAs such as the Motorola Droid, Palm Pre, and Apple's iPhone, iPhone 3G, and iPhone 3GS use the "underclocking" of a more powerful processor, rather than the full clocking of a less powerful processor, to maximize battery life. The designers for such mobile devices often discover that a slower processor gives worse battery life than a more powerful processor at a lower clock rate. They select a processor on the basis of the performance per watt of the processor.

Performance

The performance of an "underclocked" machine will often be better than might be expected. Under normal desktop use, the full power of the CPU is rarely needed. Even when the system is busy, a large amount of time is usually spent waiting for data from memory, disk, or other devices. Such devices communicate with the CPU through a bus which operates at a much lower bandwidth. Generally speaking, the lower the CPU multiplier (and thus clockrate of a CPU), the closer its performance will be to that of the bus, and the less time it will spend waiting.

Chapter 10

Computer Cooling



An OEM AMD heatsink mounted onto a motherboard.

Computer cooling is required to remove the waste heat produced by computer components, to keep components within their safe operating temperature limits. Various cooling methods help to improve processor performance or reduce the noise of cooling fans.

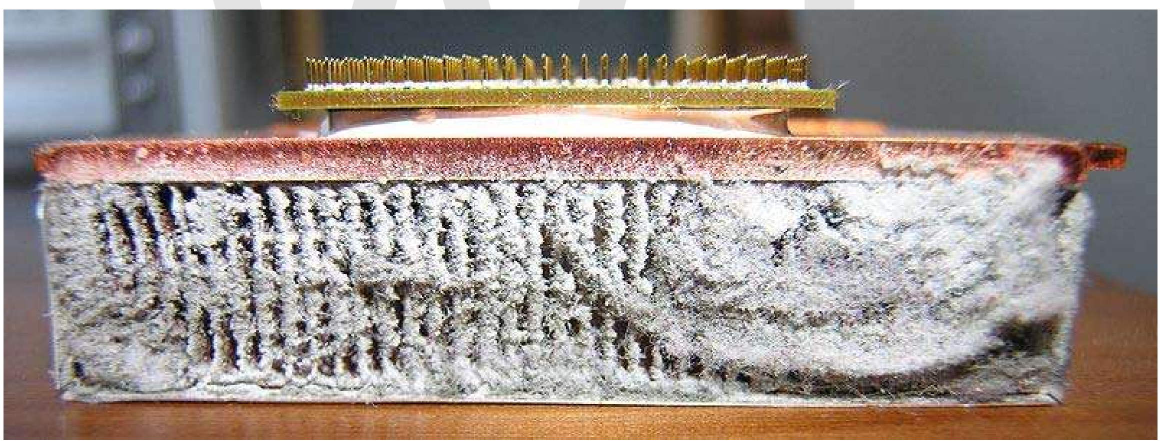
Components which produce heat and are susceptible to performance loss and damage include integrated circuits such as CPUs, chipset and graphics cards, along with hard drives (though excessive cooling of hard drives has been found to have negative effects). Overheated parts fail early and may give sporadic problems resulting in system freezes or crashes.

Both integral and peripheral means are used to keep the temperature of each component at a safe level. With regard to integral means, CPU and GPUs are designed with energy efficiency, including heat dissipation, in mind; though improved efficiency may only allow increased performance instead of reduced heat. Peripheral means include heat sinks to increase the surface area which dissipates heat, fans to speed up the exchange of air heated by the computer parts for cooler ambient air, and in some cases softcooling, the throttling of computer parts in order to decrease heat generation.

As a safety measure, many computers are designed to turn themselves off if the internal temperature exceeds a certain point. Alternatively, some have an option in their BIOS that allows the user to determine if the system emits an alarm beep or shuts itself down when the core temperature reaches the level set by the user. However, setting this incorrectly can result in hardware damage or erratic system behaviour.

Causes of heat build up

The amount of heat generated by an integrated circuit (e.g., a CPU or GPU), the prime cause of heat build up in modern computers, is a function of the efficiency of its design, the technology used in its construction and the frequency and voltage at which it operates.



The dust buildup on this laptop CPU heat sink after three years of use has made the laptop unusable due to frequent thermal shutdowns.

In operation, the temperature of a computer's components will rise until the heat lost to the surroundings is equal to the heat produced by the component, and thus the temperature of the component reaches equilibrium. For reliable operation, the equilibrium temperature must be sufficiently low for the structure of the computer's circuits to survive.

Cooling can be hindered by:

- **Dust** acting as a thermal insulator and impeding airflow, thereby reducing heat sink and fan performance.
- **Poor airflow** including turbulence due to friction against impeding components such as ribbon cables, or improper orientation of fans, can reduce the amount of air flowing through a case and even create localized whirlpools of hot air in the case.
- **Poor heat transfer** due to a lack of, or poor application of thermal compounds and sufficient surface area of heat sinks to radiate off the heat.

Damage prevention

Thermal sensors in some CPUs and GPUs can shut down the computer when high temperatures are detected. However, reliance on such measures may not prevent repeated incidents from permanently damaging the integrated circuit.

An integrated circuit may also shut down parts of the circuit when it is idling, or to scale back the clock speed under low workloads or high temperatures, with the goal of reducing both power use and heat generation.

Air cooling

Fans are most commonly used for air cooling. A *computer fan* may be attached to the computer case, or attached to a CPU, GPU, chipset, PSU, hard drive or PCI slot. Common fan sizes include 40, 60, 80, 92, 120, and 140 mm. Recently, 200mm fans have begun to creep into the performance market, as well as even larger sizes such as 230 and 240mm.

In desktops

Typical airflow through a desktop ATX case

Desktop computers typically use one or more fans for cooling. Almost all desktop power supplies have at least one fan to exhaust air from the case. Most manufacturers recommend bringing cool, fresh air in at the bottom front of the case, and exhausting warm air from the top rear.

If there is more air being forced into the system than is being pumped out (due to an imbalance in the number or strength of fans), this is referred to as a "positive" airflow, as the pressure inside the unit would be higher than outside. A balanced or neutral airflow is the most efficient, although a slightly positive airflow results in less dust build up if dust filters are used. Negative pressure inside the case can create problems such as clogged optical drives due to sucking in air (and dust).

In high density computing

Data centers typically contain many racks of flat 1U servers. Air is drawn in at the front of the rack and exhausted at the rear. Because data centers typically contain such large numbers of computers and other power-consuming devices, they risk overheating of the various components if no additional measures are taken. Thus, extensive HVAC systems are used. Often a raised floor is used so the area under the floor may be used as a large plenum for cooled air and power cabling.

Another way of accommodating large numbers of systems in a small space are blade chassis. In contrast to the horizontal orientation of flat servers, blade chassis are often oriented vertically. This vertical orientation facilitates convection. When the air is heated by the hot components, it tends to flow to the top on its own, creating a natural air flow along the boards. This stack effect can help to achieve the desired air flow and cooling. Some manufacturers expressly take advantage of this effect.

In laptop computing

Most laptops use air cooling in order to keep the CPU and other components within their operating temperature range. Because the fan's air is forced through a small port, the fan and heatsinks can be clogged by dust or be obstructed by objects placed near the port. This can cause overheating, and can be a cause of component failure in laptops. The severity of this problem varies with laptop design, its use and power dissipation. With recent reductions in CPU power dissipation, this problem can be anticipated to reduce in severity.

Liquid submersion cooling

An uncommon practice is to submerge the computer's components in a thermally conductive liquid. Personal computers that are cooled in this manner do not generally require any fans or pumps, and may be cooled exclusively by passive heat exchange between the computer's parts, the cooling fluid and the ambient air. Extreme component density supercomputers such as the Cray-2 and Cray T90 used additional liquid to chilled liquid heat exchangers in order to facilitate heat removal.

The liquid used must have sufficiently low electrical conductivity in order for it not to interfere with the normal operation of the computer's components. If the liquid is somewhat electrically conductive, it may be necessary to insulate certain parts of components susceptible to electromagnetic interference, such as the CPU. For these reasons, it is preferred that the liquid be dielectric.

Liquids commonly used in this manner include various liquids invented and manufactured for this purpose by 3M, such as Fluorinert. Various oils, including but not limited to cooking, motor and silicone oils have all been successfully used for cooling personal computers.

Evaporation can pose a problem, and the liquid may require either to be regularly refilled or sealed inside the computer's enclosure. Liquid may also slowly seep into and damage components, particularly capacitors, causing an initially functional computer to fail after hours or days immersed.

Waste heat reduction

Where full-power, full-featured modern computers are not required, some companies opt to use less powerful computers or computers with fewer features. For example: in an office setting, the IT department may choose a thin client or a diskless workstation thus cutting out the heat-laden components such as hard drives and optical disks. These devices are also often powered with direct current from an external power supply brick which still wastes heat, but not inside the computer itself.

The components used can greatly affect the power consumption and hence waste heat. A VIA EPIA motherboard with CPU typically generates approximately 25 watts of heat whereas a Pentium 4 motherboard typically generates around 140 watts. While the former has considerably less computing power, both types are adequate and responsive for tasks such as word processing and spreadsheets. Choosing a LCD monitor rather than a CRT can also reduce power consumption and excess room heat, as well as the added benefit of increasing available physical desk space.

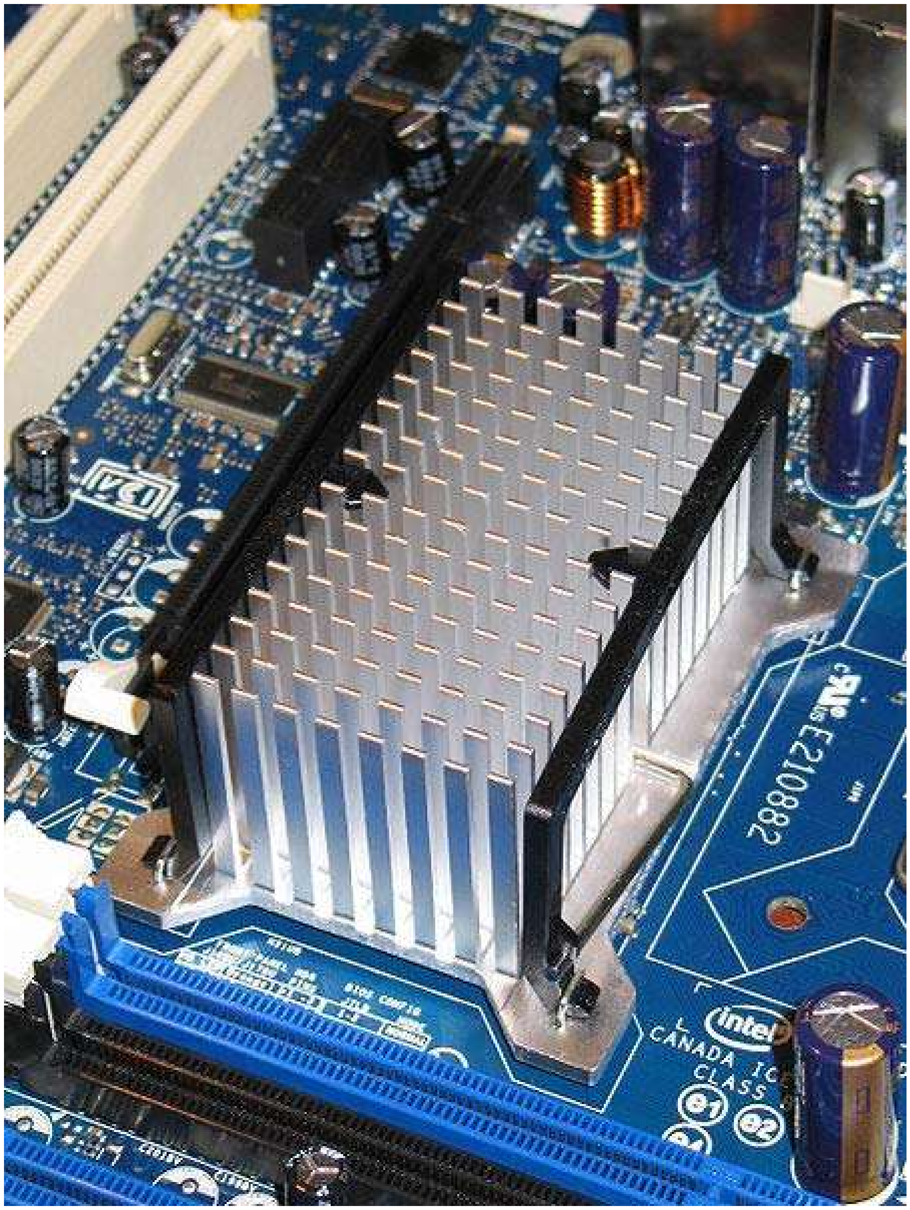
Conductive and radiative cooling

Some laptop components, such as hard drives and optical drives, are commonly cooled by having them make contact with the computer's frame, increasing the surface area which can radiate and otherwise exchange heat.

Spot cooling

In addition to system cooling, various individual components usually have their own cooling systems in place. Components which are individually cooled include, but are not limited to, the CPU, GPU and the Northbridge chip. Some cooling solutions employ one or more methods of cooling, and may also utilize logic and/or temperature sensors in order to vary the power used in active cooling components.

Passive heat-sink cooling



Passive heatsink on an Intel GMA graphics chip

Passive heat-sink cooling involves attaching a block of machined or extruded metal to the part that needs cooling. A thermal adhesive may be used. More commonly for a personal-computer CPU, a clamp holds the heat sink directly over the chip, with a thermal grease or thermal pad spread between. This block usually has fins and ridges to increase its surface area. The heat conductivity of metal is much better than that of air, and it radiates heat better than does the component that it is protecting (usually an integrated circuit or CPU). Until recently, fan-cooled aluminium heat sinks were the norm for desktop computers. Today, many heat sinks feature copper base-plates or are entirely made of copper.

Dust buildup between the metal fins of a heat sink gradually reduces efficiency, but can be countered with a gas duster by blowing away the dust along with any other unwanted excess material.

Passive heat sinks are commonly found on older CPUs, parts that do not get very hot (such as the chipset), and low-power computers.

Usually a heat-sink is attached to the integrated heat spreader (IHS), essentially a large, flat plate attached to the CPU, with conduction paste layered between. This dissipates or spreads the heat locally. Unlike a heat sink, a spreader is meant to redistribute heat, not to remove it. In addition, the IHS protects the fragile CPU.

Passive cooling involves no fan noise.

Active heat-sink cooling

Active heat sink with a fan and heat pipes

Active heat-sink cooling uses the same principle as passive, with the addition of a fan that blows over or through the heat sink. The air movement increases the rate at which the heat sink can exchange heat with the ambient air. Active heat sinks are the primary method of cooling modern processors and graphics cards.

The buildup of dust is greatly increased with active heat-sink cooling, because the fan continually takes in the dust present in the surrounding air.

Peltier cooling or thermoelectric cooling

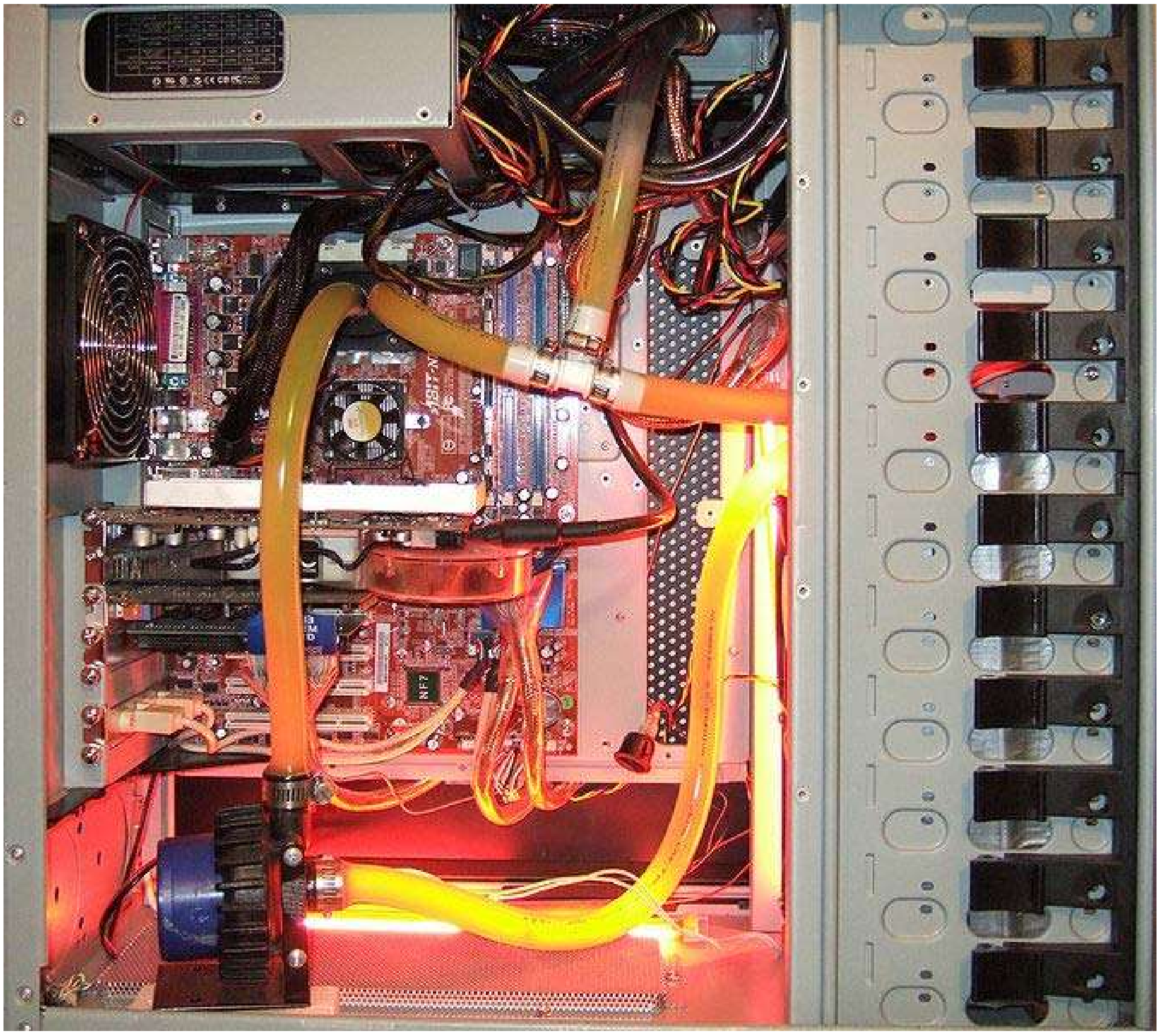
In 1821 T. J. Seebeck discovered that different metals, connected at two different junctions, will develop a micro-voltage if the two junctions are held at different temperatures. This effect is known as the "Seebeck effect"; it is the basic theory behind the TEC (thermoelectric cooling).

In 1834 Jean Peltier discovered the inverse of the Seebeck effect, now known as the "Peltier effect". He found that applying a voltage to a thermocouple creates a temperature differential between two sides. This results in an effective, albeit extremely inefficient heat pump.

Modern TECs use several stacked units each composed of dozens or hundreds of thermocouples laid out next to each other, which allows for a substantial amount of heat transfer. A combination of bismuth and tellurium is most commonly used for thermocouples.

As active heat pumps, TECs can cool the surface of components below ambient temperatures. This is impossible with common radiator cooled water cooling systems and heatpipe HSFs.

Water cooling



DIY Water cooling setup showing 12v pump, CPU Waterblock and the typical application of a T-Line

While originally limited to mainframe computers, water cooling has become a practice largely associated with overclocking in the form of either manufactured kits, or in the form of do-it-yourself setups assembled from individually gathered parts. The past few years has seen water cooling increasing its popularity with pre-assembled, moderate to high performance, desktop computers. Water has the ability to dissipate more heat from the parts being cooled than the various types of metals used in heatsinks, making it suitable for overclocking and high performance computer applications.

Advantages to water cooling include the fact that a system is not limited to cooling one component, but can be set up to cool the central processing unit, graphics processing unit, and/or other components at the same time with the same system. As opposed to air cooling, water cooling is also influenced less by the ambient temperature. Water cooling's comparatively low noise-level compares favorably to that of active cooling, which can become quite noisy. One disadvantage to water cooling is the potential for a coolant leak. Leaked coolant can damage any electronic components it comes in contact with. Another drawback to water cooling is the complexity of the system; an active heat sink is much simpler to build, install, and maintain than a water cooling solution.

Computing folklore holds that users of the Sinclair ZX81, one of the first home computers, had to balance a carton of milk on top of the case to cool it down – perhaps an early form of water cooling.

Heat pipe

A graphics card with a heatpipe cooler design

A heat pipe is a hollow tube containing a heat transfer liquid. As the liquid evaporates, it carries heat to the cool end, where it condenses and then returns to the hot end (under capillary action, or, in earlier implementations, under gravitation). Heat pipes thus have a much higher effective thermal conductivity than solid materials. For use in computers, the heat sink on the CPU is attached to a larger radiator heat sink. Both heat sinks are hollow as is the attachment between them, creating one large heat pipe that transfers heat from the CPU to the radiator, which is then cooled using some conventional method. This method is expensive and usually used when space is tight (as in small form-factor PCs and laptops), or absolute quiet is needed (such as in computers used in audio production studios during live recording). Because of the efficiency of this method of cooling, many desktop CPUs and GPUs, as well as high end chipsets, use heat pipes in addition to active fan-based cooling to remain within safe operating temperatures.

A new design wrinkle is known as HDT for Heatpipe Direct Touch. In this usage, the heat pipe is in direct contact with the CPU chip skin. Heatpipe Direct Touch was first introduced in the Zaward ZikaRay ZIKA-01 heatsink on February 2007, using a patent obtained from Golden Sun News Techniques Corporation in Taiwan.

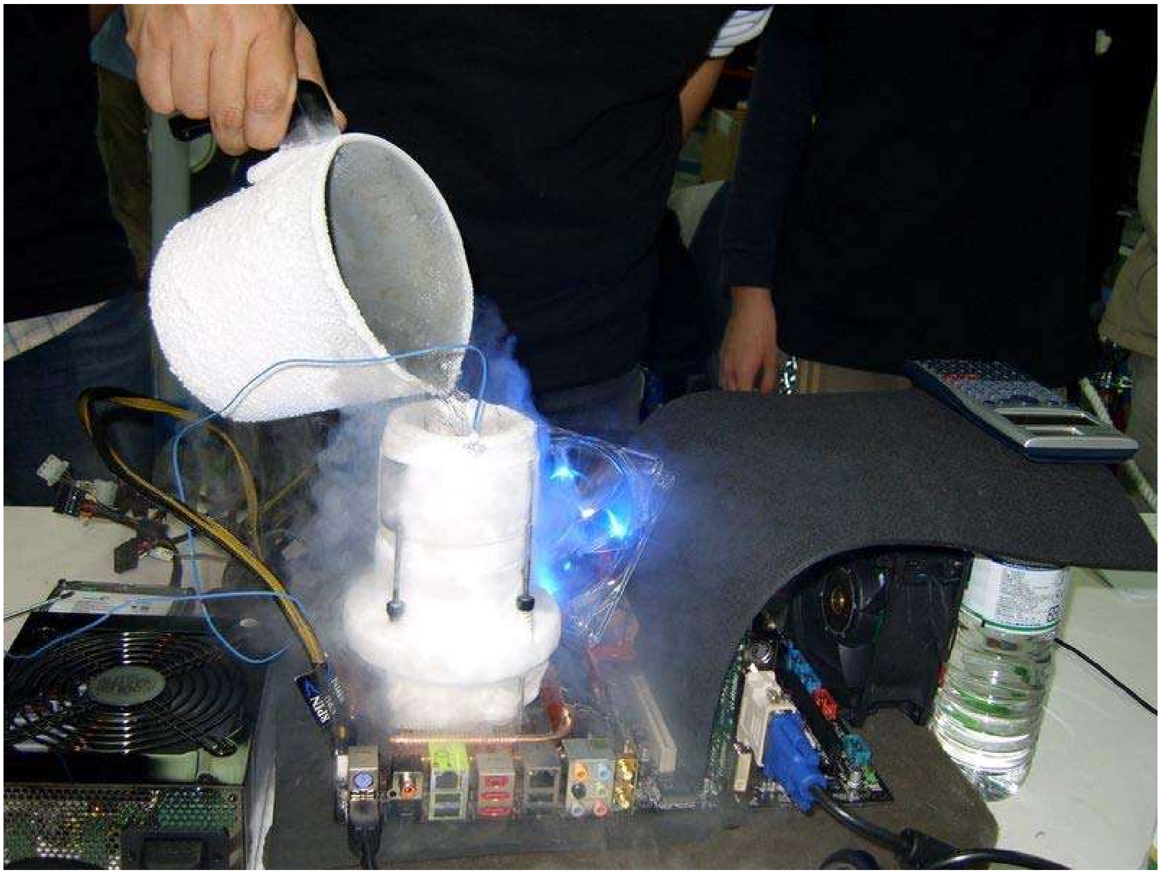
Phase-change cooling

Phase-change cooling is an extremely effective way to cool the processor. A vapor compression phase-change cooler is a unit which usually sits underneath the PC, with a tube leading to the processor. Inside the unit is a compressor of the same type as in a window air conditioner. The compressor compresses a gas (or mixture of gases) which condenses it into a liquid. Then, the liquid is pumped up to the processor, where it passes through an expansion device, this can be from a simple capillary tube to a more elaborate thermal expansion valve. The liquid evaporates (changing phase), absorbing the heat from the processor as it draws extra energy from its environment to accommodate this change. The evaporation can produce temperatures reaching around −15 to -150 degrees Celsius. The gas flows down to the compressor and the cycle begins over again. This way, the processor can be cooled to temperatures ranging from −15 to −150 degrees Celsius, depending on the load, wattage of the processor, the refrigeration system and the gas mixture used. This type of system suffers from a number of issues but mainly one must be concerned with dew point and the proper insulation of all sub-ambient surfaces that must be done (the pipes will sweat, dripping water on sensitive electronics).

Alternately a new breed of cooling system is being developed inserting a pump into the thermo siphon loop. This adds another degree of flexibility for the design engineer as the heat can now be effectively transported away from the heat source and either reclaimed or dissipated to ambient. Junction temperature can be tuned by adjusting the system pressure; higher pressure equals higher fluid saturation temperatures. This allows for smaller condensers, smaller fans and/or the effective dissipation of heat in a high ambient environment. These systems are in essence the next generation liquid cooling paradigm as they are approximately 10 times more efficient than single phase water. Since the system uses a dielectric as the heat transport media, leaks do not cause a catastrophic failure of the electric system.

This type of cooling is seen as a more extreme way to cool components, since the units are relatively expensive compared to the average desktop. They also generate a significant amount of noise, since they are essentially refrigerators, however the compressor choice and air cooling system is the main determinant of this, allowing for flexibility for noise reduction based on the parts chosen.

Liquid nitrogen



Liquid nitrogen may be used to cool an overclocked PC

As liquid nitrogen boils at -196 °C, far below the freezing point of water, it is valuable as an extreme coolant for short overclocking sessions.

In a typical installation of liquid nitrogen cooling, a copper or aluminum pipe is mounted on top of the processor or graphics card. After being heavily insulated against condensation, the liquid nitrogen is poured into the pipe, resulting in temperatures well below -100°C.

Evaporation devices ranging from cut out heat sinks with pipes attached to custom milled copper containers are used to hold the nitrogen as well as to prevent large temperature changes. However, after the nitrogen evaporates, it has to be refilled. In the realm of personal computers, this method of cooling is seldom used in contexts other than overclocking trial-runs and record-setting attempts, as the CPU will usually expire within a relatively short period of time due to temperature stress caused by changes in internal temperature.

Although liquid nitrogen is non-flammable, it can condense oxygen directly from air. Mixtures of liquid oxygen and flammable materials can be dangerously explosive.

Liquid helium

Liquid helium, colder than liquid nitrogen, has also been used for cooling. Liquid helium evaporates at -269 °C, and temperatures ranging from -230 to -240 °C have been measured from the heatsink.

Soft cooling

Softcooling is the practice of utilizing software to take advantage of CPU power saving technologies to minimize energy use. This is done using halt instructions to turn off or put in standby state CPU subparts that aren't being used or by underclocking the CPU.

Undervolting

Undervolting is a practice of running the CPU or any other component with voltages below the device specifications. An undervolted component draws less power and thus produces less heat. The ability to do this varies by manufacturer, product line, and even different production runs of the same exact product (as well as that of other components in the system), but modern processors are typically shipped with voltages higher than strictly necessary. This provides a buffer zone so that the processor will have a higher chance of performing correctly under sub-optimal conditions, such as a lower quality mainboard (motherboard). However, too low a voltage will not allow the processor to function correctly, producing errors, system freezes or crashes, or the inability to turn the system on. Undervolting too far does not typically lead to hardware damage, though in worst-case scenarios, program or system files can be corrupted.

This technique was generally employed by those seeking low-noise systems, as less cooling is needed because of the reduction of heat production. Since the popularity of hand-held or remote computers (Unmanned vehicles, mobile & cordless phones/camera/viewers, etc), undervolting is used to prolong battery endurance.

Integrated chip cooling techniques

Conventional cooling techniques all attach their “cooling” component to the outside of the computer chip, or via IHS and/or heat sinks. This “attaching” technique will always exhibit some thermal resistance, reducing its effectiveness. The heat can be more efficiently and quickly removed by directly cooling the local hot spots. At these locations, power dissipation of over 300W/cm^2 (typical CPU are less than 100W/cm^2, although future systems are expected to exceed 1000W/cm^2) can occur. This form of local cooling is essential to developing high power density chips. This ideology has led to the investigation of integrating cooling elements into the computer chip. Currently there are two techniques: micro-channel heat sinks, and jet impingement cooling.

In micro-channel heat sinks, channels are fabricated into the silicon chip (CPU), and coolant is pumped through them. The channels are designed with very large surface area which results in large heat transfers. Heat dissipation of $3000W/cm^2$ has been reported with this technique. In comparison to the Sun power density of around $7400W/cm^2$. The heat dissipation can be further increased if two-phase flow cooling is applied. Unfortunately the system requires large pressure drops, due to the small channels, and the heat flux is lower with dielectric coolants used in electronic cooling. Another local chip cooling technique is jet impingement cooling. In this technique, a coolant is flown through a small orifice to form a jet. The jet is directed toward the surface of the CPU chip, and can effectively remove large heat fluxes. Heat dissipation of over $1000W/cm^2$has been reported. The system can be operated at lower pressure in comparison to the micro-channel method. The heat transfer can be further increased using two-phase flow cooling and by integrating return flow channels (hybrid between micro-channel heat sinks and jet impingement cooling).

Cooling and overclocking

Extra cooling is usually required by those who run parts of their computer (such as the CPU and GPU) at higher voltages and frequencies than manufacturer specifications call for, called overclocking. Increasing performance by this modification of settings results in a greater amount of heat generated and thus increasing the risk of damage to components and/or premature failure.

The installation of higher performance, non-stock cooling may also be considered modding. Many overclockers simply buy more efficient, and often, more expensive fan and heat sink combinations, while others resort to more exotic ways of computer cooling, such as liquid cooling, Peltier effect heatpumps, heat pipe or phase change cooling.

There are also some related practices that have a positive impact in reducing system temperatures:

Heat sink lapping

Heat sink lapping is the smoothing and polishing of the contact (bottom) part of a heat sink to increase its heat transfer efficiency. The desired result is a contact area which has a more even surface, as a less even contact surface creates a larger amount of insulating air between the heat sink and the computer part it is attached to. Polishing the surface using a combination of fine sandpaper and abrasive polishing liquids can produce a mirror-like shine, an indicator of a very smooth metal surface. Even a curved surface can become extremely reflective, yet not particularly flat, as is the case with curved mirrors; thus heat sink quality is based on *overall flatness*, more than optical properties. Lapping a high quality heat sink can damage it, because, although the heat sink may become shiny, it is likely that more material will be removed from the edges, making the heat sink less effective overall.

If attempted, a piece of float glass should be used, as it self-levels as it cools and offers the most economical solution to producing a perfectly flat surface.

Use of exotic thermal conductive compounds

Some overclockers use special thermal compounds whose manufacturers claim to have a much higher efficiency than stock thermal pads. Heat sinks clean of any grease or other thermal transfer compounds have a very thin layer of these products applied, and then are placed normally over the CPU. Many of these compounds have a high proportion of silver as their main ingredient due to its high thermal conductivity. The resulting difference in the temperature of the CPU is measurable (several celsius degrees), so the heat transfer does appear to be superior to stock compounds. Some people experience negligible gains and have called to question the advantages of these exotic compounds, calling the style of application more important than the compound itself. Also note that there may be a 'setting' or 'curing' period and negligible gains may improve over time as the compound reaches its optimum thermal conductivity.

Use of rounded cables

Most older PCs use flat ribbon cables to connect storage drives (IDE or SCSI). These large flat cables greatly impede airflow by causing drag and turbulence. Overclockers and modders often replace these with rounded cables, with the conductive wires bunched together tightly to reduce surface area. Theoretically, the parallel strands of conductors in a ribbon cable serve to reduce crosstalk (signal carrying conductors inducing signals in nearby conductors), but there is no empirical evidence of rounding cables reducing performance. This may be because the length of the cable is short enough so that the effect of crosstalk is negligible. Problems usually arise when the cable is not electromagnetically protected and the length is considerable, a more frequent occurrence with older network cables.

These computer cables can then be cable tied to the chassis or other cables to further increase airflow.

This is less of a problem with new computers that use Serial ATA which has a much narrower cable.

Airflow optimization

The colder the cooling medium (the air), the more effective the cooling. Cooling air temperature can be improved with these guidelines:

- Supply cool air to the hot components as directly as possible. Examples are air snorkels and tunnels that feed outside air directly and exclusively to the CPU or GPU cooler. For example, the BTX case design prescribes a CPU air tunnel.
- Expel warm air as directly as possible. Examples are: Conventional PC (ATX) power supplies blow the warm air out the back of the case. Many dual-slot graphics card designs blow the warm air through the cover of the adjacent slot. There are also some aftermarket coolers that do this. Some CPU cooling designs blow the warm air directly towards the back of the case, where it can be ejected by a case fan.
- Air that has already been used to spot-cool a component should not be reused to spot-cool a different component (this follows from the previous items). The ATX case design can be said to violate this rule, since the power supply gets its "cool" air from the inside of the case, where it has been warmed up already. The BTX case design also violates this rule, since it uses the CPU cooler's exhaust to cool the chipset and often the graphics card.
- Prefer cool intake air, avoid inhaling exhaust air (outside air above or near the exhausts). For example, a CPU cooling air duct at the back of a tower case would inhale warm air from a graphics card exhaust. Moving all exhausts to one side of the case, conventionally the back, helps to keep the intake air cool.
- Hiding cables behind motherboard tray or simply apply ziptie and tucking cables away to provide unhindered airflow.

Fewer fans strategically placed will improve the airflow internally within the PC and thus lower the overall internal case temperature in relation to ambient conditions. The use of larger fans also improves efficiency and lowers the amount of waste heat along with the amount of noise generated by the fans while in operation.

There is little agreement on the effectiveness of different fan placement configurations, and little in the way of systematic testing has been done. For a rectangular PC (ATX) case, a fan in the front with a fan in the rear and one in the top has been found to be a suitable configuration. However, AMD's (somewhat outdated) system cooling guidelines notes that "A front cooling fan does not seem to be essential. In fact, in some extreme situations, testing showed these fans to be recirculating hot air rather than introducing cool air." It may be that fans in the side panels could have a similar detrimental effect—possibly through disrupting the normal air flow through the case. However, this is unconfirmed and probably varies with the configuration.

Chapter 11

Computer Fan

A 3D illustration of four 80 mm fans, a type of fan commonly used in personal computers (sometimes as a set, or mixed with other fan sizes).

A **computer fan** is any fan inside, or attached to, a computer case used for cooling purposes, and may refer to fans that draw cooler air into the case from the outside, expel warm air from inside, or move air across a heatsink to cool a particular component. The use of fans to cool a computer is an example of active cooling.

Usage

As processors, graphics cards, RAM and other components in computers that have increased in speed and power consumption, the amount of heat produced by these components as a side-effect of normal operation has also increased. These components need to be kept within a specified temperature range to prevent overheating, instability, malfunction and damage leading to a shortened component lifespan.

While in earlier personal computers it was possible to cool most components using natural convection (passive cooling), many modern components require more effective active cooling. To cool these components, fans are used to move heated air away from the components and draw cooler air over them. Fans attached to components are usually used in combination with a heatsink to increase the area of heated surface in contact with the air, thereby improving the efficiency of cooling.

In the IBM compatible PC market, the computer's power supply unit (PSU) almost always uses an exhaust fan to expel warm air from the PSU. Active cooling on CPUs started to appear on the Intel 80486, and by 1997 was standard on all desktop processors. Chassis or case fans, usually one exhaust fan to expel heated air from the rear and optionally an intake fan to draw cooler air in through the front, became common with the arrival of the Pentium 4 in late 2000. A third vent fan in the side of the PC, often located over the CPU, is also common. The graphics processing unit (GPU) on many modern graphics cards also requires a heatsink and fan. In some cases, the northbridge chip on the motherboard has another fan and heatsink. Other components such as the hard drives and RAM may also be actively cooled, though as of 2007 this remains relatively unusual. It is not uncommon to find five or more fans in a modern PC.

Cooling fan applications

An 80x80x25mm computer fan

Case mount

Used to aerate the case of the computer. The components inside the case cannot dissipate heat efficiently if the surrounding air is too hot. Case fans move air through the case, usually drawing cooler outside air in through the front (where it may also be drawn over the internal hard drive racks) and expelling it through the rear. There may be a third fan in the side or top of the case to draw outside air into the vicinity of the CPU, which is usually the largest single heat source. Standard case fans are 80 mm, 92 mm or 120 mm in width and length. As case fans are often the most readily visible form of cooling on a PC, decorative fans are widely available and may be lit with LEDs, made of UV-reactive plastic, and covered with decorative grilles. Decorative fans and accessories are popular with case modders. Air filters are often used over intake fans, to prevent dust from entering the case.

A power supply (PSU) fan often plays a double role, not only keeping the PSU itself from overheating, but also removing warm air from inside the case. PSUs with two fans are also available, which typically have a fan on the inside to supply case air into the PSU and a second fan on the back to expel the heated air.

CPU fan

Used to cool the CPU (central processing unit) heatsink.

Graphics card fan

Used to cool the graphics processing unit or the memory on graphics cards. These fans were not necessary on older cards because of their low power dissipation, but most modern graphics cards, especially those designed for 3D graphics and gaming, need their own dedicated cooling fans. Some of the higher powered cards can produce more heat than the CPU (up to 289 watts), so effective cooling is especially important. Passive coolers for new video cards, however, are not unheard of, such as the Thermalright HR-03.

Chipset fan

Used to cool the northbridge of a motherboard's chipset, which may be necessary for system bus overclocking.

Other types of fans

Other less commonly encountered fans may include:

- **PCI slot fan:** A fan mounted in one of the PCI slots, usually to supply additional cooling to the PCI and/or graphics cards.
- **Hard disk fan:** A fan mounted next to or on a hard disk drive. This may be desirable on faster-spinning (e.g. 10,000 RPM) hard disks with greater heat production.
- **CD burner fan:** Some internal CD and/or DVD burners included cooling fans.

Physical characteristics

The width and height of these usually square fans are measured in millimeters; common sizes include 60 mm, 80 mm, 92 mm and 120 mm. Fans with a round frame are also available; these are usually designed so that one may use a larger fan than the mounting holes would otherwise allow (i.e., a 120 mm fan with 90 mm holes). The amount of airflow which fans generate is typically measured in cubic feet per minute (CFM), and the speed of rotation is measured in revolutions per minute (RPM). Often, computer enthusiasts choose fans which have a higher CFM rating, but produce less noise (measured in decibels, or dB), and some fans come with an adjustable RPM rating to produce less noise when the computer does not require additional airflow. Fan speeds may be controlled manually (a simple potentiometer control, for example), thermally, or by the computer hardware or by software. It is also possible to run many 12V fans from the 5 V supply, at the expense of airflow, but with reduced noise levels.

The other consideration when choosing a computer fan is static pressure. A fan with high static pressure is more effective at forcing air through restricted spaces, such as the gaps between a radiator or heatsink. Therefore, enthusiasts often prioritize static pressure over CFM when choosing a fan for use with a heatsink. The relative importance of static pressure depends on the degree to which the airflow is restricted by geometry (i.e. static pressure becomes more important as the spacing between heatsink blades decreases). Static pressure is usually measured in either mm Hg or mm H2O.

The type of bearing used in a fan can affect its performance and noise output. Most computer fans use one of the following bearing types:

- **Sleeve bearing** fans use two surfaces lubricated with oil or grease as a friction contact. Sleeve bearings are less durable as the contact surfaces can become rough and/or the lubricant dry up, eventually leading to failure. Sleeve bearings may be more likely to fail at higher temperatures, and may perform poorly when mounted in any orientation other than vertical. The lifespan of a sleeve bearing fan may be around 40,000 hours at 50 °C. Fans that use sleeve bearings are generally cheaper than fans that use ball bearings, and are quieter at lower speeds early in their life, but can grow considerably noisier as they age.
- **Rifle bearing** fans are similar to sleeve bearing, but are quieter and have almost as much lifespan as ball bearings. The bearing has a spiral groove in it that pumps fluid from a reservoir. This allows them to be safely mounted horizontally (unlike sleeve bearings), since the fluid being pumped lubricates the top of the shaft. The pumping also ensures sufficient lubricant on the shaft, reducing noise, and increasing lifespan.
- **Ball bearing** fans use ball bearings. Though generally more expensive, ball bearing fans do not suffer the same orientation limitations as sleeve bearing fans, are more durable especially at higher temperatures, and quieter than sleeve bearing fans at higher rotation speeds. The lifespan of a ball bearing fan may be around 63,000 hours at 50 °C.
- **Fluid bearing** fans have the advantages of near-silent operation and high life expectancy (comparable to ball bearing fans). However, these fans tend to be the most expensive. The **enter bearing** fan is a variation of the fluid bearing fan, developed by Everflow.
- **Magnetic bearing** or **maglev** fans, in which the fan is repelled from the bearing by magnetism.

Fan Sizing

Fans are available in wide variety of sizes and capacities. In general the faster the fan, the more noise it produces. Within a given physical size capacity is roughly proportional to current draw. For a given flow a larger fan will be quieter than a smaller fan.

Fan connector

The standard connectors for computer fans are

3-pin Molex connector KK Family

This connector is used when connecting a fan to the motherboard or other circuit board. It is a small thick rectangular in-line female connector with two tabs on the outer-most edge of one long side. The size and spacing of the pin sockets is identical to a standard 3-pin female IC connector. The three pins are used for ground, +12 V power, and a tachometer signal. Molex Part number of receptacle is 22-01-3037. Molex Part number of individual crimp contacts is 08-55-0101.

4-pin Molex connector KK Family

This is a special variant of the Molex KK connector with four pins but with the locking/polarisation features of a 3-pin connector. The additional pin is used for a pulse-width modulation signal to provide variable speed control. These can be plugged into 3-pin headers, but will lose their fan speed control. Molex Part number of receptacle is 22-01-3047. Molex Part number of individual crimp contacts is 08-55-0101.

4-pin Molex connector

This connector is used when connecting the fan directly to the power supply. It consists of two wires (red/12V and black/ground) leading to and splicing into a large in-line 4-pin male-to-female Molex connector.

Dell, Inc. proprietary

This connector is an expansion of a simple 3-pin female IC connector by adding two tabs to the middle of the connector on one side and a lock-tab on the other side. The size and spacing of the pin sockets is identical to a standard 3-pin female IC connector and 3-pin Molex connector. Some models have the wiring of the white wire (speed sensor) in the middle, whereas the standard 3-pin Molex requires the white wire as pin #3, thus compatibility issues may exist.

Alternatives

If a fan is not desirable, because of noise, reliability, or environmental concerns, there are some alternatives:

- Very Rarely, such as ultra silent home theatre machines, can rely on passive cooling alone and do not require a case fan to keep computer components at ordinary operating temperatures. More commonly (such as in simple business and home machines) a power supply fan alone is sufficient to cool the machine.
- Undervolting and/or underclocking to reduce power dissipation
- Larger heatsinks (for example, some motherboards have northbridge fans; others have larger, more costly heatsinks)
- Natural convection cooling: carefully designed, correctly oriented, and sufficiently large CPU coolers can dissipate up to 100 W by natural convection alone
- More unusual solutions, *e.g.* heatpipes bonded to the metal case, water cooling, or refrigeration
- Motherboards sunk in liquid oil provides excellent convection cooling and protects from humidity and water without the need for heatsinks or fans. Special care must be taken to ensure compatibility with adhesives and sealants used on

the motherboard and ICs. This solution is used in some external environments like wireless equipments located in the wild.

- Ionic wind cooling is being researched, whereby air is moved by ionizing air between 2 electrodes. this replaces the fan and has the advantage of no moving parts.

Chapter 12

Heat Sink

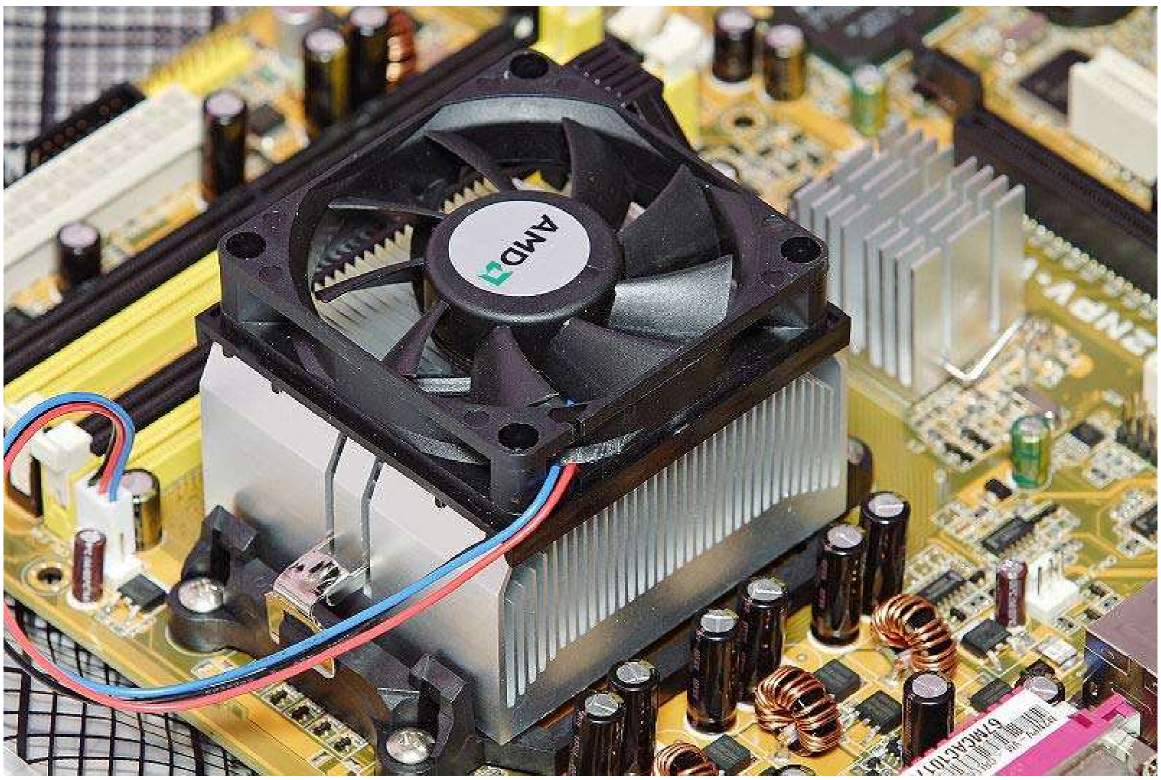


Figure 1: An active (fan cooled) heat sink used for the processor cooling on a PC motherboard. To its right is a smaller pin fin heat sink used to cool the Northbridge of the motherboard.

A **heat sink** is a term for a component or assembly that transfers heat generated within a solid material to a fluid medium, such as air or a liquid. Examples of heat sinks are the heat exchangers used in refrigeration and air conditioning systems and the radiator (also a heat exchanger) in a car. Heat sinks also help to cool electronic and optoelectronic devices, such as higher-power lasers and light emitting diodes (LEDs).

A heat sink is physically designed to increase the surface area in contact with the cooling fluid surrounding it, such as the air. Approach air velocity, choice of material, fin (or other protrusion) design and surface treatment are some of the design factors which influence the thermal resistance, i.e. thermal performance, of a heat sink. One engineering application of heat sinks is in the thermal management of electronics, often computer CPU or graphics processors. For these, heat sink attachment methods and thermal interface materials also influence the eventual junction or die temperature of the processor(s). Thermal adhesive (also known as thermal grease) is added to the base of the heatsink to help its thermal performance. Theoretical, experimental and numerical methods can be used to determine a heat sink's thermal performance.

Basic heat sink heat transfer principle

A heat sink is an object that transfers thermal energy from a higher temperature to a lower temperature *fluid medium*. The fluid medium is frequently air, but can also be water or in the case of heat exchangers, refrigerants and oil. If the fluid medium is water, the 'heat sink' is frequently called a cold plate.

To understand the principle of a heat sink, consider Fourier's law of heat conduction. Joseph Fourier was a French mathematician who made important contributions to the analytical treatment of heat conduction. Fourier's law of heat conduction, simplified to a one-dimensional form in the *x*-direction, shows that when there is a temperature gradient in a body, heat will be transferred from the higher temperature region to the lower temperature region. The rate at which heat is transferred by conduction, q_k, is proportional to the product of the temperature gradient and the cross-sectional area through which heat is transferred.

$$q_k = -kA\frac{dT}{dx}$$

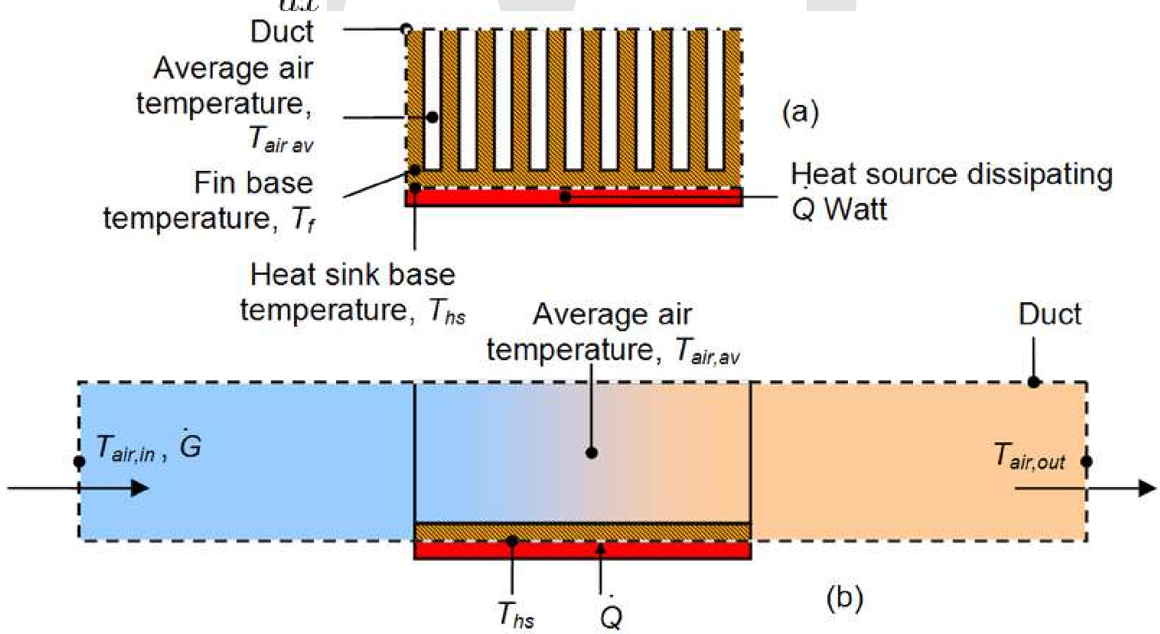


Figure 2: Sketch of a heat sink in a duct used to calculate the governing equations from conservation of energy and Newton's law of cooling.

Consider a heat sink in a duct, where air flows through the duct, as shown in Figure 2. It is assumed that the heat sink base is higher in temperature than the air. Applying the conservation of energy, for steady-state conditions, and Newton's law of cooling to the temperature nodes shown in Figure 2 gives the following set of equations.

$$\dot{Q} = \dot{m}c_{p,in}(T_{air,out} - T_{air,in}) \quad (1)$$

$$\dot{Q} = \frac{T_{hs} - T_{air,av}}{R_{hs}} \quad (2)$$

where $T_{air,av} = \frac{T_{air,out} + T_{air,in}}{2}$ (3)

Using the mean air temperature is an assumption that is valid for relatively short heat sinks. When compact heat exchangers are calculated, the logarithmic mean air temperature is used. $\dot{m}$ is the air mass flow rate in kg/s.

The above equations show that

- When the air flow through the heat sink decreases, this results in an increase in the average air temperature. This in turn increases the heat sink base temperature. And additionally, the thermal resistance of the heat sink will also increase. The net result is a higher heat sink base temperature.
    - The increase in heat sink thermal resistance with decrease in flow rate will be shown in later here.
- The inlet air temperature relates strongly with the heat sink base temperature. For example, if there is recirculation of air in a product, the inlet air temperature is not the ambient air temperature. The inlet air temperature of the heat sink is therefore higher, which also results in a higher heat sink base temperature.
- Therefore, if there is no air or fluid flow around the *heat sink*, the energy dissipated to the air can not be transferred to the ambient air. Therefore, the heat sink functions poorly.
- Furthermore, a heat sink is not a device with the "magical ability to absorb heat like a sponge and send it off to a parallel universe".

Other examples of situations in which a heat sink has impaired efficiency:

- Pin fins have a lot of surface area, but the pins are so close together that air has a hard time flowing through them.
- Aligning a heat sink so that the fins are not in the direction of flow.
- Aligning the fins horizontally for a natural convection heat sink. Whilst a heat sink is stationary and there are no centrifugal forces and artificial gravity, air that is warmer than the ambient temperature ***always*** flows upward, given essentially-still-air surroundings; this is convective cooling.

Design factors which influence the thermal performance of a heat sink

Material

The most common heat sink material is aluminium. Chemically pure aluminium is not used in the manufacture of heat sinks, but rather aluminium alloys. Aluminium alloy 1050A has one of the higher thermal conductivity values at 229 W/m•K. However, it is not recommended for machining, since it is a relatively soft material. Aluminium alloys 6061 and 6063 are the more commonly used aluminium alloys, with thermal conductivity values of 166 and 201 W/m•K, respectively. The aforementioned values are dependent on the temper of the alloy.

Copper is also used since it has around twice the conductivity of aluminium, but is three times as heavy as aluminium. Copper is also around four to six times more expensive than aluminium, but this is market dependent. Copper and aluminium prices can be compared in figures 3 and 4, or on Internet websites, such as the London Metal Exchange. Aluminium has the added advantage that it is able to be extruded, while copper can not. Copper heat sinks are machined and skived. Another method of manufacture is to solder the fins into the heat sink base.

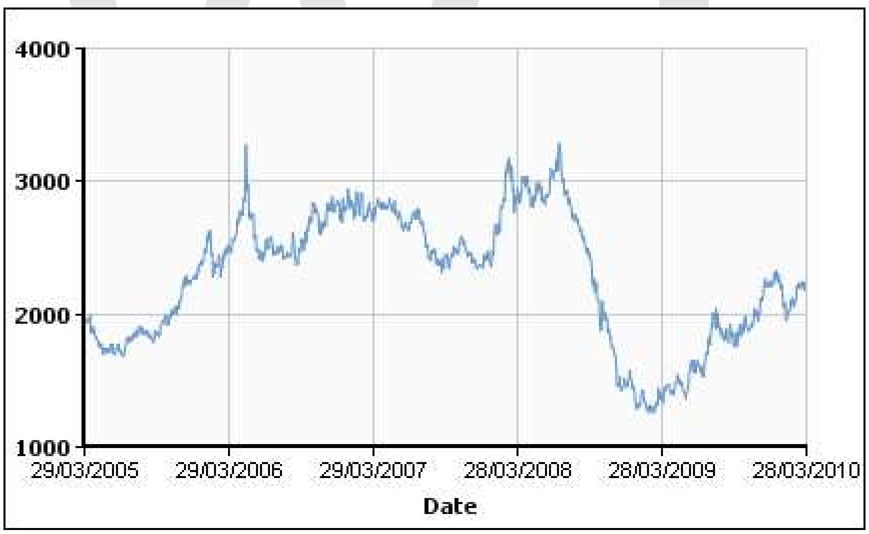


Figure 3: Aluminium cash buyer prices listed in US Dollar per metric tonne. Prices are shown for the period 29 March 2005 to the 29 March 2010.

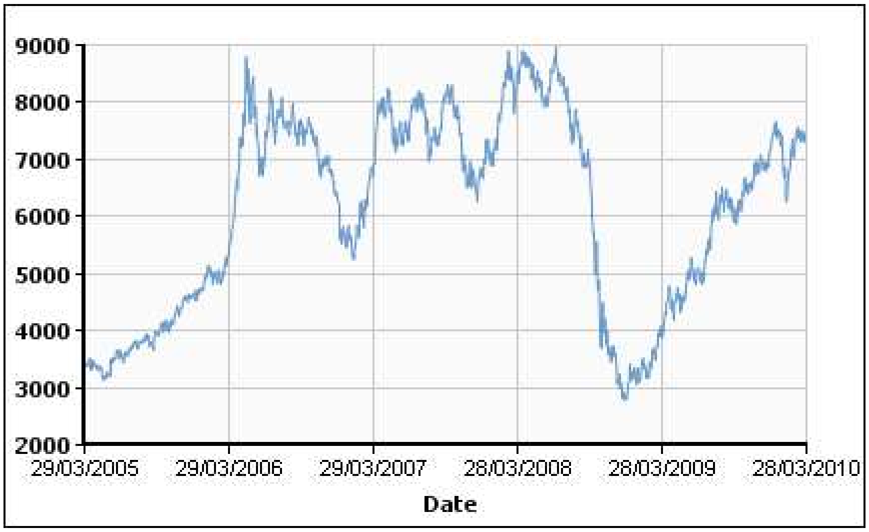


Figure 4: Copper cash buyer prices listed in US Dollar per metric tonne. Prices are shown for the period 29 March 2005 to the 29 March 2010.

Another heat sink material that can be used is diamond. With a value of 2000 W/mK it exceeds that of copper by a factor of five. In contrast to metals, where heat is conducted by delocalized electrons, lattice vibrations are responsible for diamond's very high thermal conductivity. For thermal management applications, the outstanding thermal conductivity and diffusivity of diamond is an essential. Nowadays CVD diamond is used as submounts for high-power integrated circuits and laser diodes.

Composite materials can be used. Examples are a copper-tungsten pseudoalloy, AlSiC (silicon carbide in aluminium matrix), Dymalloy (diamond in copper-silver alloy matrix), and E-Material (beryllium oxide in beryllium matrix). Such materials are often used as substrates for chips, as their thermal expansion coefficient can be matched to ceramics and semiconductors.

Fin efficiency

Fin efficiency is one of the parameters which makes a higher thermal conductivity material important. A fin of a heat sink may be considered to be a flat plate with heat flowing in one end and being dissipated into the surrounding fluid as it travels to the other. As heat flows through the fin, the combination of the thermal resistance of the heat sink impeding the flow and the heat lost due to convection, the temperature of the fin and, therefore, the heat transfer to the fluid, will decrease from the base to the end of the fin. This factor is called the fin efficiency and is defined as the actual heat transferred by the fin, divided by the heat transfer were the fin to be isothermal (hypothetically the fin having infinite thermal conductivity). Equations 6 and 7 are applicable for straight fins.

$$\eta_f = \frac{tanh(mL_c)}{mL_c} \quad (6)$$

$$mL_c = \sqrt{\frac{2h_f}{kt_f}} L_f \quad (7)$$

Where:

- h_f is the convection coefficient of the fin
  - Air: 10 to 100 W/(m^2K)
  - Water: 500 to 10,000 W/(m^2K)
- k is the thermal conductivity of the fin material
  - Aluminum: 120 to 240 W/(m·K)
- L_f is the fin height (m)
- t_f is the fin thickness (m)

To increase the fin efficiency of fins:

- Decrease the fin aspect ratio, by:
  - Increasing the fin thickness, or
  - Decreasing the fin length
- Increase the thermal conductivity of the fins, (e.g. by using copper instead of aluminum)

Spreading resistance

Another parameter that concerns the thermal conductivity of the heat sink material is spreading resistance. Spreading resistance occurs when thermal energy is transferred from a small area to a larger area in a substance with finite thermal conductivity. In a heat sink, this means that heat does not distribute uniformly through the heat sink base. The spreading resistance phenomenon is shown by how the heat travels from the heat source location and causes a large temperature gradient between the heat source and the edges of the heat sink. This means that some fins are at a lower temperature than if the heat source were uniform across the base of the heat sink. This nonuniformity increases the heat sink's effective thermal resistance.

To decrease the spreading resistance in the base of a heat sink:

- Increase the base thickness
- Choose a different material with better thermal conductivity
- Use a vapour chamber or heat pipe in the heat sink base.

Fin arrangements

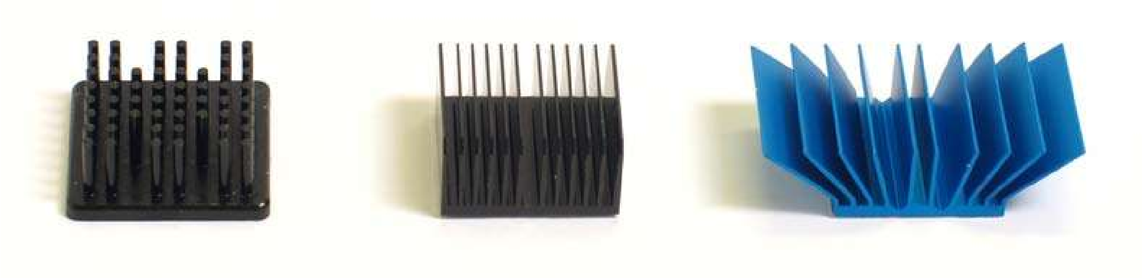


Figure 5: A pin-, straight- and flared fin heat sink types

A pin fin heat sink is a heat sink that has pins that extend from its base. The pins can be cylindrical, elliptical or square. A pin is by far one of the more common heat sink types available on the market. A second type of heat sink fin arrangement is the straight fin. These run the entire length of the heat sink. A variation on the straight fin heat sink is a cross cut heat sink. A straight fin heat sink is cut at regular intervals but at a coarser pitch than a pin fin type.

In general, the more surface area a heat sink has, the better it works. However, this is not always true. The concept of a pin fin heat sink is to try to pack as much surface area into a given volume as possible. As well, it works well in any orientation. Kordyan has compared the performance of a pin fin and a straight fin heat sink of similar dimensions. Although the pin fin has 194 cm^2 surface area while the straight fin has 58 cm^2, the temperature difference between the heat sink base and the ambient air for the pin fin is 50 °C. For the straight fin it was 44 °C or 6 °C better than the pin fin. Pin fin heat sink performance is significantly better than straight fins when used in their intended application where the fluid flows axially along the pins rather than only tangentially across the pins.

Comparison of a pin fin and straight fin heat sink of similar dimensions. Adapted from data of						
Heat sink fin type	**Width [cm]**	**Length [cm]**	**Height [cm]**	**Surface area [cm^2]**	**Volume [cm^3]**	**Temperature difference, T_{case}-T_{air} [°C]**
Straight	2.5	2.5	3.2	58	20	44
Pin	3.8	3.8	1.7	194	24	51

Another configuration is the flared fin heat sink; its fins are not parallel to each other, as shown in figure 5. Flaring the fins decreases flow resistance and makes more air go through the heat sink fin channel; otherwise, more air would bypass the fins. Slanting them keeps the overall dimensions the same, but offers longer fins. Forghan, et al. have published data on tests conducted on pin fin, straight fin and flared fin heat sinks. They found that for low approach air velocity, typically around 1 m/s, the thermal performance is at least 20% better than straight fin heat sinks. Lasance and Eggink also found that for the bypass configurations that they tested, the flared heat sink performed better than the other heat sinks tested.

Surface colour

The heat transfer from the heatsink is mediated by two effects: conduction via the coolant, and thermal radiation. The surface of the heatsink influences its emissivity; shiny metal absorbs and radiates only a small amount of heat, while matte black radiates highly.

In coolant-mediated heat transfer, the contribution of radiation is generally small. A layer of coating on the heatsink can then be counterproductive, as its thermal resistance can impair heat flow from the fins to the coolant. Finned heatsinks with convective or forced flow will not benefit significantly from being colored.

In situations with significant contribution of radiative cooling, e.g. in case of a flat non-finned panel acting as a heatsink with low airflow, the heatsink surface finish can play an important role. Matte-black surfaces will radiate much more efficiently than shiny bare metal.

The importance of radiative vs coolant-mediated heat transfer increases in situations with low ambient air pressure (e.g. high-altitude operations) or in vacuum (e.g. satellites in space).

Engineering applications

Processor/Microprocessor cooling

Heat dissipation is an unavoidable by-product of all but micropower electronic devices and circuits. In general, the temperature of the device or component will depend on the thermal resistance from the component to the environment, and the heat dissipated by the component. To ensure that the component temperature does not overheat, a thermal engineer seeks to find an efficient heat transfer path from the device to the environment. The heat transfer path may be from the component to a printed circuit board (PCB), to a heat sink, to air flow provided by a fan, but in all instances, eventually to the environment.

Two additional design factors also influence the thermal/mechanical performance of the thermal design:

1. The method by which the heat sink is mounted on a component or processor. This will be discussed under the section *attachment methods*.
2. For each interface between two objects in contact with each other, there will be a temperature drop across the interface. For such composite systems, the temperature drop across the interface may be appreciable. This temperature change may be attributed to what is known as the thermal contact resistance. *Thermal interface materials* (TIM) decrease the thermal contact resistance.

Attachment methods for microprocessors and similar ICs

As power dissipation of components increases and component package size decreases, thermal engineers must innovate to ensure components won't overheat. Devices that run cooler last longer. A heat sink design must fulfill both its thermal as well as its mechanical requirements. Concerning the latter, the component must remain in thermal contact with its heat sink with reasonable shock and vibration. The heat sink could be the copper foil of a circuit board, or else a separate heat sink mounted onto the component or circuit board. Attachment methods include thermally conductive tape or epoxy, wire-form z clips, flat spring clips, standoff spacers, and push pins with ends that expand after installing.

- **Thermally conductive tape**

Figure 6: Roll of thermally conductive tape.

Thermally conductive tape is one of the most cost-effective heat sink attachment materials. It is suitable for low-mass heat sinks and for components with low power dissipation. It consists of a thermally conductive carrier material with a pressure-sensitive adhesive on each side.

This tape is applied to the base of the heat sink, which is then attached to the component. Following are factors that influence the performance of thermal tape:

1. Surfaces of both the component and heat sink must be clean, with no residue such as a film of silicone grease.
2. Preload pressure is essential to ensure good contact. Insufficient pressure results in areas of non-contact with trapped air, and results in higher-than-expected interface thermal resistance.
3. Thicker tapes tend to provide better "wettability" with uneven component surfaces. "Wettability" is a term used to describe the percentage area of contact of a tape on a component. Thicker tapes, however, have a higher thermal resistance than thinner tapes. From a design standpoint, it is best to strike a balance by selecting a tape thickness that provides maximum "wettablilty" with minimum thermal resistance.

- **Epoxy**

Epoxy is more expensive than tape, but provides a greater mechanical bond between the heat sink and component, as well as improved thermal conductivity. The epoxy chosen must be formulated for this purpose. Most epoxies are two-part liquid formulations that must be thoroughly mixed before being applied to the heat sink, and before the heat sink is placed on the component. The epoxy is then cured for a specified time, which can vary from 2 hours to 48 hours. Faster cure time can be achieved at higher temperatures. The surfaces to which the epoxy is applied must be clean and free of any residue.

The epoxy bond between the heat sink and component is semi-permanent/permanent. This makes re-work very difficult and at times impossible. The most typical damage caused by rework is the separation of the component die heat spreader from its package.

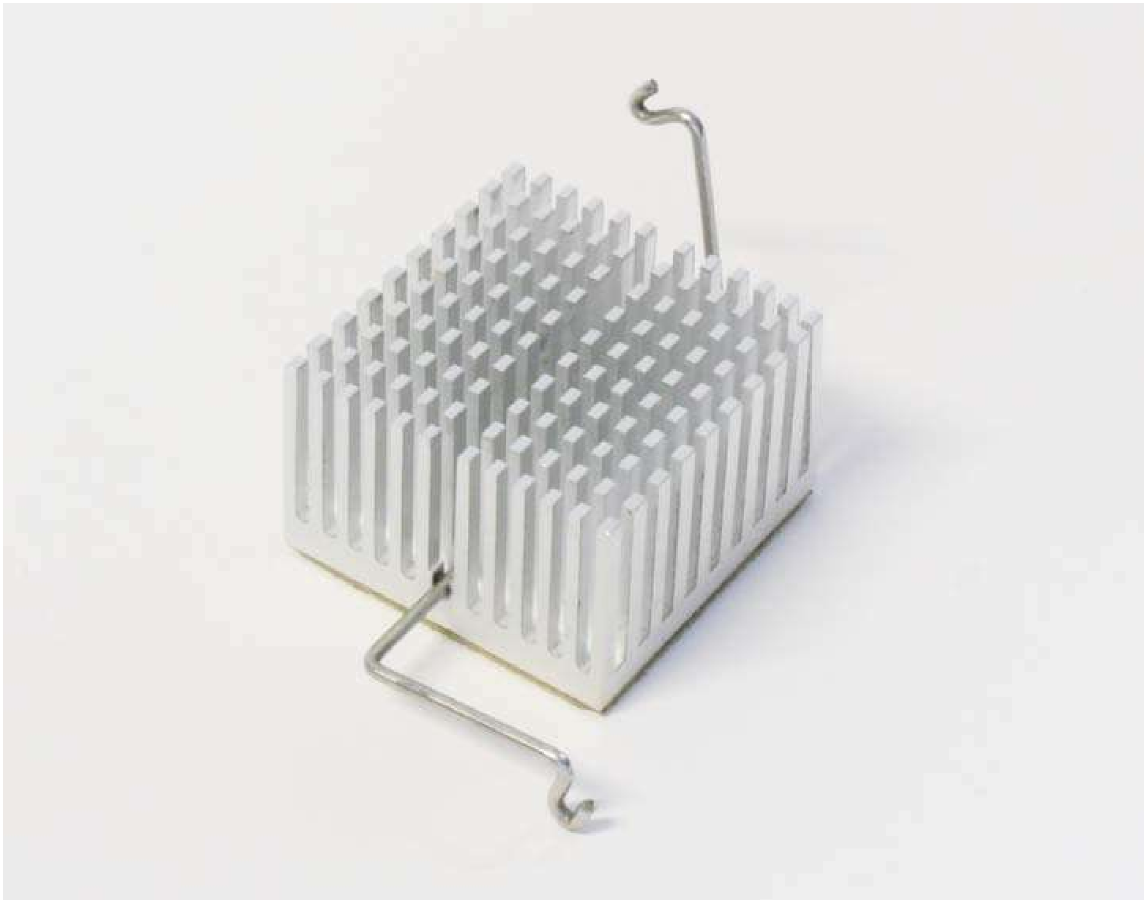


Figure 7: A pin fin heat sink with a z-clip retainer

- **Wire form Z-clips**

More expensive than tape and epoxy, wire form z-clips attach heat sinks mechanically. To use the z-clips, the printed circuit board must have anchors. Anchors can be either soldered onto the board, or pushed through. Either type requires holes to be designed into the board. The use of RoHS solder must be allowed for because such solder is mechanically weaker than traditional Pb/Sn solder.

To assemble with a z-clip, attach one side of it to one of the anchors. Deflect the spring until the other side of the clip can be placed in the other anchor. The deflection develops a spring load on the component, which maintains very good contact. In addition to the mechanical attachment that the z-clip provides, it also permits using higher-performance thermal interface materials, such as phase change types.

Figure 8: Two heat sink attachment methods, namely the maxiGRIP (left) and Talon Clip

- **Clips**

Available for processors and ball grid array (BGA) components, clips allow the attachment of a BGA heat sink directly to the component. The clips make use of the gap created by the ball grid array (BGA) between the component underside and PCB top surface. The clips therefore require no holes in the PCB. They also allow for easy rework of components. Examples of commercially available clips are the maxiGRIP™ and superGRIP™ range from Advanced Thermal Solutions (ATS) and the Talon Clip™ from Malico. The three aforementioned clipping methods use plastic frames for the clips, but the ATS designs uses metal spring clips to provide the compression force. The Malico design uses the plastic "arm" to provide a mechanical load on the component. Depending on the product requirement, the clipping methods will have to meet shock and vibration standards, such as Telecordia GR-63-CORE, ETSI 300 019 and MIL-STD-810.

Figure 9: Push pins

- **Push pins with compression springs**

For larger heat sinks and higher preloads, push pins with compression springs are very effective. The push pins, typically made of brass or plastic, have a flexible barb at the end that engages with a hole in the PCB; once installed, the barb retains the pin. The compression spring holds the assembly together and maintains contact between the heat sink and component. Care is needed in selection of push pin size. Too great an insertion force can result in the die cracking and consequent component failure.

- **Threaded standoffs with compression springs**

For very large heat sinks, there is no substitute for the threaded standoff and compression spring attachment method. A threaded standoff is essentially a hollow metal tube with internal threads. One end is secured with a screw through a hole in the PCB. The other end accepts a screw which compresses the spring, completing the assembly. A typical heat sink assembly uses two to four standoffs, which tends to make this the most costly heat sink attachment design. Another disadvantage is the need for holes in the PCB.

Summary of heat sink attachment methods			
Method	**Pros**	**Cons**	**Cost**
Thermal tape	Easy to attach. Inexpensive.	Cannot provide mechanical attachment for heavier heat sinks or for high vibration environments. Surface must be cleaned for optimal adhesion. Moderate to low thermal conductivity.	$
Epoxy	Strong mechanical adhesion. Relatively inexpensive.	Makes board rework difficult since it can damage component. Surface must be cleaned for optimal adhesion.	$$
Wire form Z-clips	Strong mechanical attachment. Easy removal/rework. Applies a preload to the thermal interface material, improving thermal performance.	Requires holes in the board or solder anchors. More expensive than tape or epoxy. Custom designs.	$$$
Clip-on	Applies a preload to the thermal interface material, improving thermal performance. Requires no holes or anchors. Easy removal/rework.	Must have "keep out" zone around the BGA for the clip. Extra assembly steps.	$$$
Push pin with compression springs	Strong mechanical attachment. Highest thermal interface material preload. Easy removal and installation.	Requires holes in the board which increases complexity of traces in PCB.	$$$$
Stand-offs with compression springs	Strongest mechanical attachment. Highest preload for the thermal interface material. Ideal for large heat sinks.	Requires holes in the board which increases complexity of trace layout. Complicated assembly.	$$$$$

Thermal interface materials

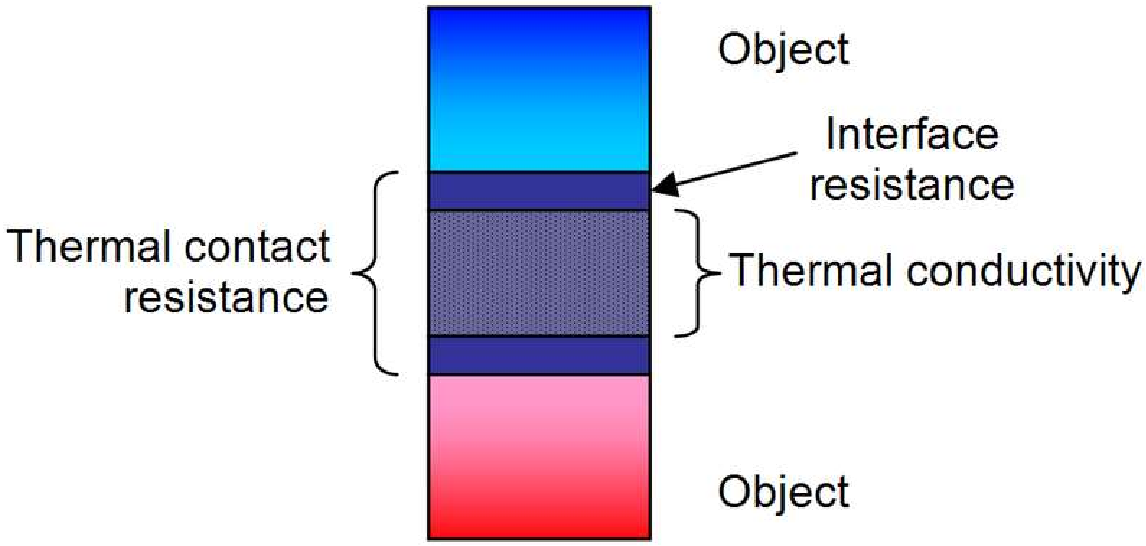


Figure 10: Thermal conductivity and the interface resistance form part of the thermal interface resistance of an thermal interface material.

Thermal contact resistance occurs due to the voids created by surface roughness effects, defects and misalignment of the interface. The voids present in the interface are filled with air. Heat transfer is therefore due to conduction across the actual contact area and to conduction (or natural convection) and radiation across the gaps. If the contact area is small, as it is for rough surfaces, the major contribution to the resistance is made by the gaps. To decrease the thermal contact resistance, the surface roughness can be decreased while the interface pressure is increased. However, these improving methods are not always practical or possible for electronic equipment. Thermal interface materials (TIM) are a common way to overcome these limitations,

Properly applied thermal interface materials displace the air that is present in the gaps between the two objects with a material that has a much-higher thermal conductivity. Air has a thermal conductivity of 0.022 W/m•K while TIMs have conductivities of 0.3 W/m•K and higher.

When selecting a TIM, care must be taken with the values supplied by the manufacturer. Most manufacturers give a value for the thermal conductivity of a material. However, the thermal conductivity does not take into account the interface resistances. Therefore, if a TIM has a high thermal conductivity, it does not necessarily mean that the interface resistance will be low.

Selection of a TIM is based on three parameters: the interface gap which the TIM must fill, the contact pressure, and the electrical resistivity of the TIM. The contact pressure is the pressure applied to the interface between the two materials. The selection does not include the cost of the material. Electrical resistivity may, or may not, be important, depending upon electrical design details.

Selection Based on Interface Gap	
Interface gap values	**Products types available**
< 2 mil	Thermal grease, epoxy, phase change materials
2 – 5 mil	Phase change materials, polyimide, graphite or aluminium tapes
5 – 18 mil	Silicone coated fabrics
> 18 mil	Gap fillers

Selection Based on Contact Pressure		
Contact pressure scale	**Typical pressure ranges**	**Product types available**
Very low	< 70 kPa	Gap fillers
Low	< 140 kPa	Thermal grease, epoxy, polyimide, graphite or aluminium tapes
High	2 MPa	Silicone coated fabrics

Selection Based on Dielectric Strength			
Electrical insulation	**Dielectric strength**	**Typical values**	**Product types available**
Not required	N/A	N/A	Thermal grease, epoxy, phase change materials, graphite or aluminium tapes.
Required	Low	< 300 V/mil	Silicone coated fabrics, gap fillers
Required	High	> 1500 V/mil	Polyimide tape

TIM Application Notes Based on Product Type		
Product type	**Application notes**	**Thermal performance**
Thermal paste	Messy. Labour intensive. Relatively long assembly time.	++++
Epoxy	Creates ‘permanent’ interface bond.	++++
Phase change	Allows for pre-attachment. Softens and conforms to interface defects at	++++

	operational temperatures. Can be repositioned in field.	
Thermal tapes, including graphite, polyimide, and aluminium tapes	Easy to apply. Some mechanical strength.	+++
Silicone coated fabrics	Provide cushioning and sealing while still allowing heat transfer.	+
Gap filler	Can be used to thermally couple differing-height components to a heat spreader or heat sink. Naturally tacky.	++

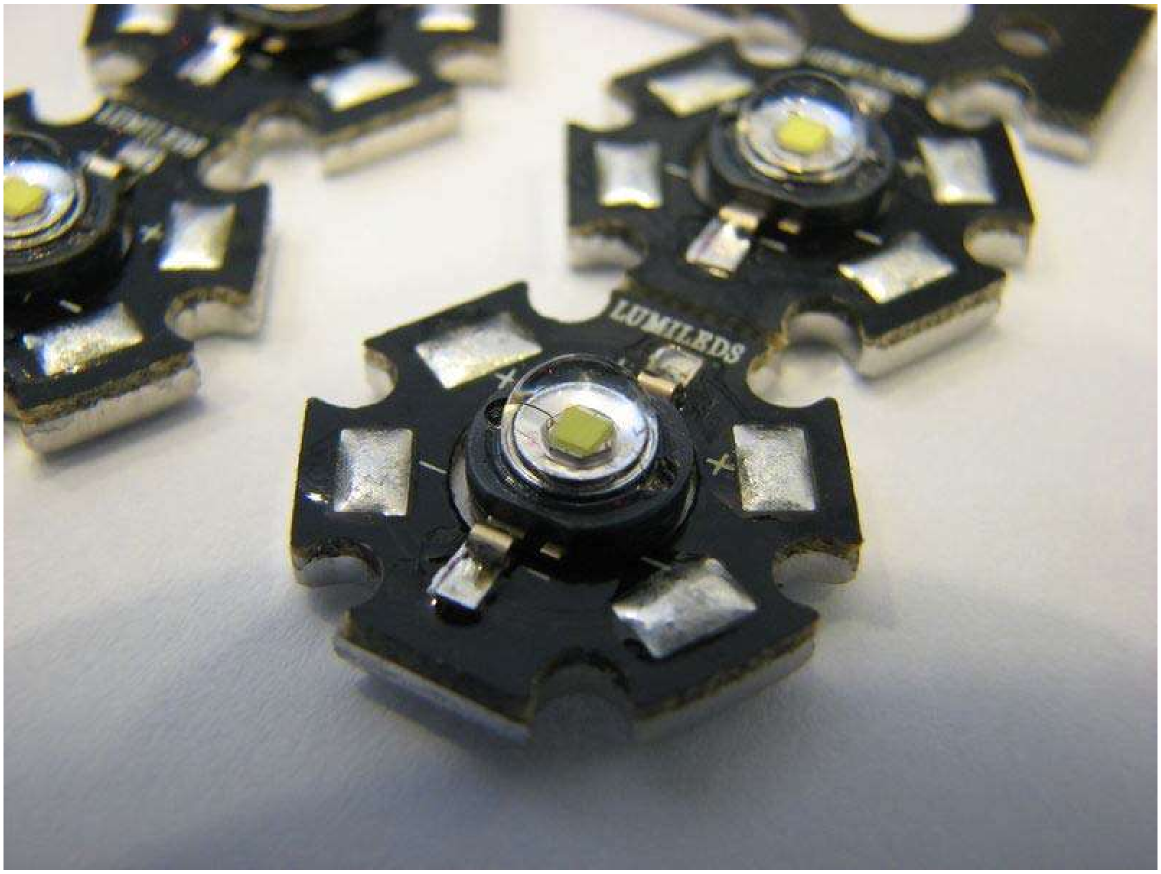


Figure 11: High power LEDs from Philips Lumileds Lighting Company mounted on 21 mm star shaped metal-core PCBs

Light emitting diode lamps

Light emitting diode (LED) performance and lifetime are strong functions of their temperature. Effective cooling is therefore essential. A case study of a LED based downlighter shows an example of the calculations done in order to calculate the required heat sink necessary for the effective cooling of lighting system. The article also shows that in order to get confidence in the results, multiple independent solutions are required that give similar results. Specifically, results of the experimental, numerical and theoretical methods should all be within 10% of each other to give high confidence in the results.

Firestopping and fireproofing

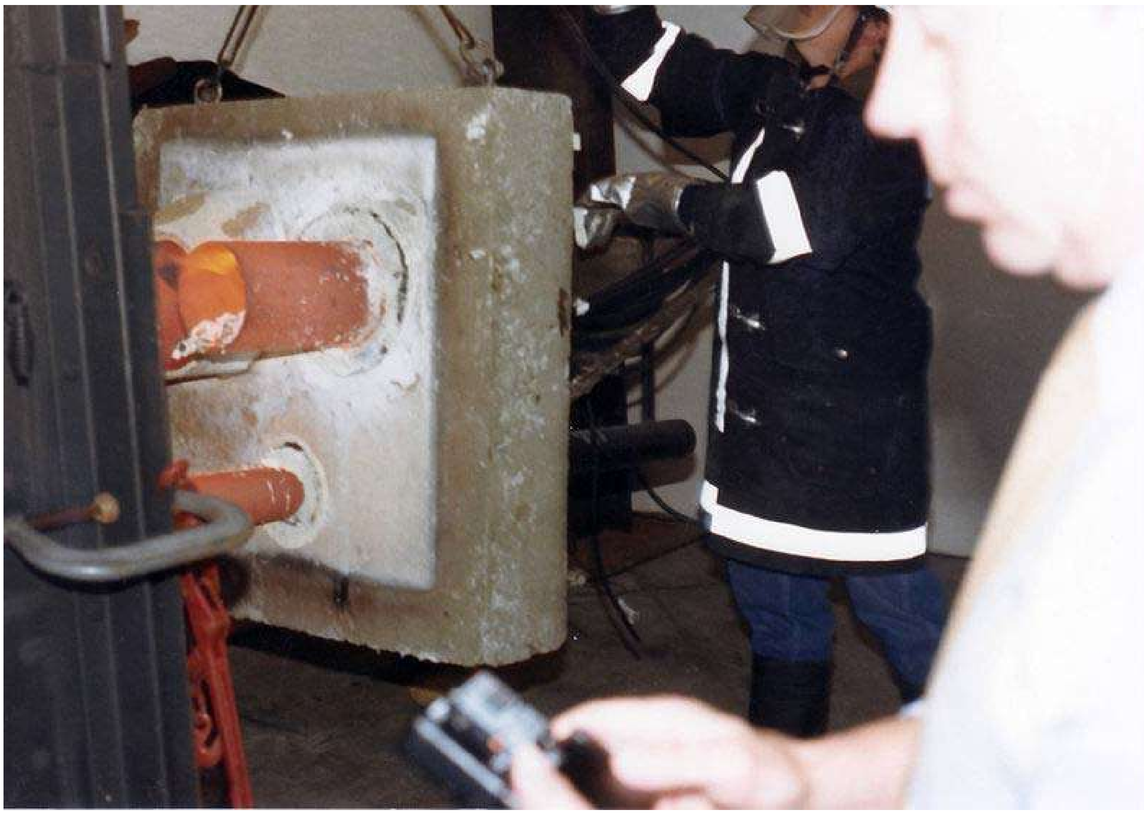


Figure 12: Fire test where the steel pipe penetrant to absorb and conduct heat from the furnace, through to the unexposed side.

A heat sink is rarely a desired thing in passive fire protection. Rather, it is usually a problem that must be overcome to maintain fire-resistance ratings. The proven ability to *overcome* heat sinks in construction is subject to building code and fire code regulations.

Firestopping

- **Problem** – Metallic penetrants and sleeves, at a density of 7.9 kg/L are denser than common firestops or concrete. Consequently, during a fire, they will absorb more heat and conduct it to the unexposed side of a fire barrier (thus *"cooling"* the exposed side at the expense of the unexposed side), such as the cold side of a firewall. This is undesirable. Even if the fire is stopped by the barrier, one must keep the unexposed side cool to prevent autoignition of combustibles on the unexposed side of a fire barrier. The unexposed side may very well be an area of refuge, which must be safeguarded to comply with the building code. Greater penetrant and sleeve conductivity leads to lower T-ratings. Higher density firestops, such as firestop mortars act as heat sinks to absorb heat away from small penetrants, such as cables, thus increasing T-ratings.

- **Benefit** – a rare exception where heat sinks are *beneficial* in firestops is where intumescents must be activated, such as in a firestop containing a plastic pipe. Heat sinks such as wire mesh and extra metallic sleeving may be used to carry heat to intumescents to activate expansion which should choke off a melting plastic pipe or melting pipe covering, such as foamed plastic or fibreglass.

Fireproofing

In fireproofing of structural steel as well as providing circuit integrity to cables, cable trays, junction boxes and electrical conduit, the metallic items that are protected by the fireproofing measures act as heat sinks. Fireproofing methods are used to *defeat* the heat sink properties of the items they protect. In the case of circuit integrity measures, electrical services will fuse and short circuit above 140°C.

In soldering

Temporary heat sinks were sometimes used while soldering circuit boards, preventing excessive heat from damaging sensitive nearby electronics. In the simplest case, this means partially gripping a component using a heavy metal crocodile clip, hemostat or similar clamp. Modern semiconductor devices, which are designed to be assembled by reflow soldering, can usually tolerate soldering temperatures without damage. On the other hand, electrical components such as magnetic reed switches can malfunction if exposed to hotter soldering irons, so this practice is still very much in use.

Methods to determine heat sink thermal performance

In general, a heat sink performance is a function of material thermal conductivity, dimensions, fin type, heat transfer coefficient, air flow rate, duct size. To determine the thermal performance of a heat sink, a theoretical model can be made. Alternatively, the thermal performance can be measured experimentally. Due to the complex nature of the highly 3D flow in present in applications, numerical methods or CFD can also be used. Here we will discuss the aforementioned methods for the determination of the heat sink thermal performance.

A heat transfer theoretical model

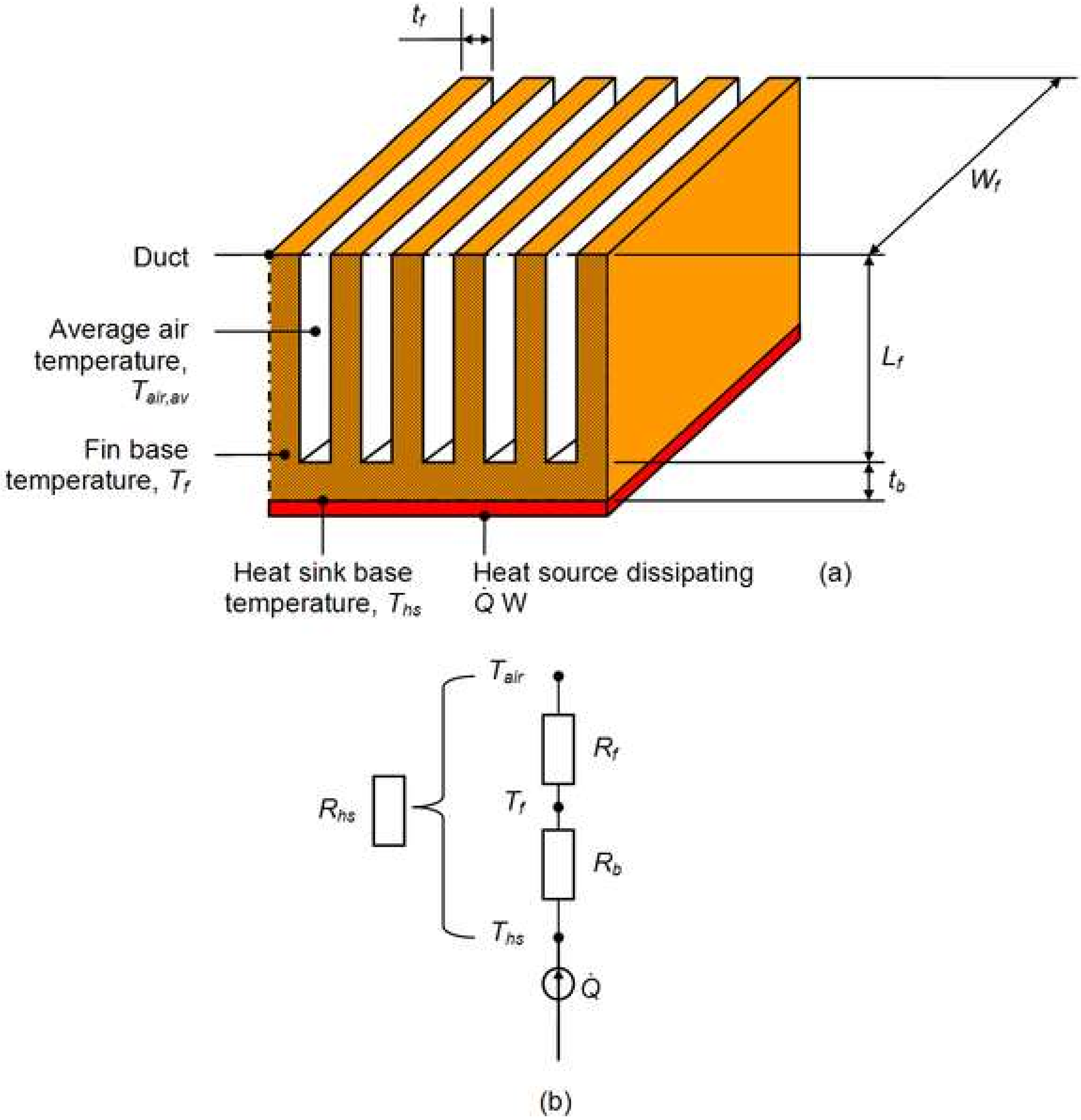


Figure 13: Sketch of a heat sink with equivalent thermal resistances

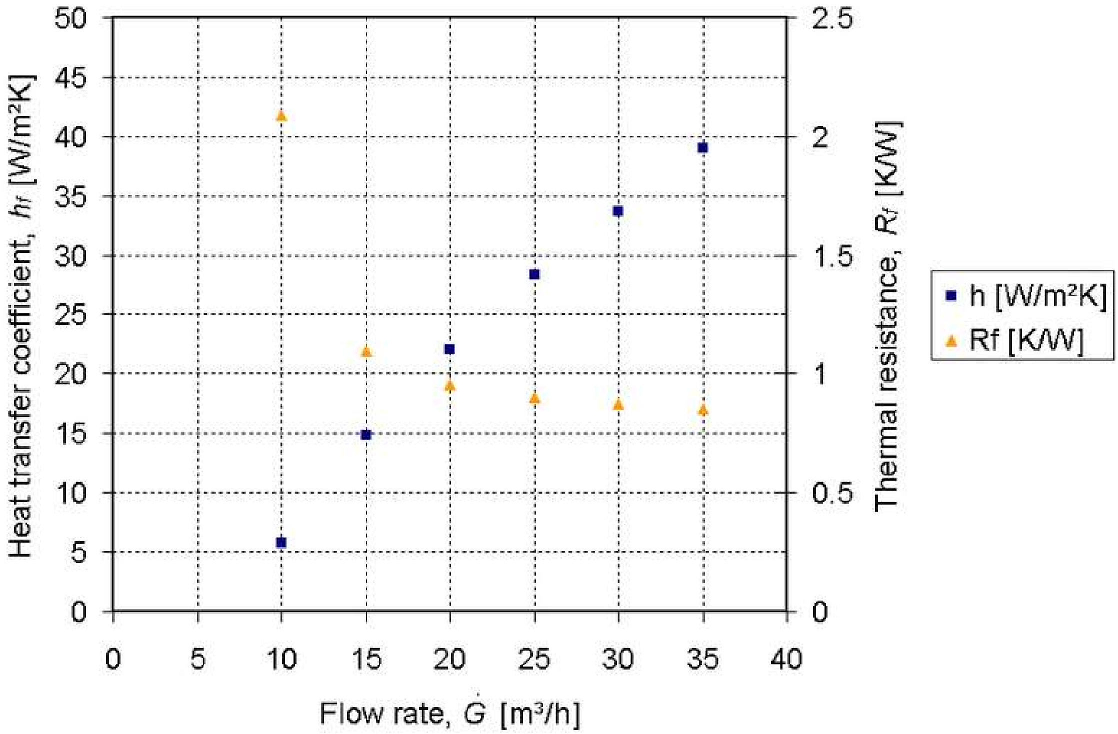


Figure 14: Thermal resistance and heat transfer coefficient plotted against flow rate for the specific heat sink design used in. The data was generated using the equations provided here. The data shows that for an increasing air flow rate, the thermal resistance of the heat sink decreases.

One of the methods to determine the performance of a heat sink is to use heat transfer and fluid dynamics theory. One such method has been published by Jeggels, et al., though this work is limited to ducted flow. Ducted flow is where the air is forced to flow through a channel which fits tightly over the heat sink. This makes sure that all the air goes through the channels formed by the fins of the heat sink. When the air flow is not ducted, a certain percentage of air flow will bypass the heat sink. Flow bypass was found to increase with increasing fin density and clearance, while remaining relatively insensitive to inlet duct velocity.

The heat sink thermal resistance model consists of two resistances, namely the resistance in the heat sink base, R_b, and the resistance in the fins, R_f. The heat sink base thermal resistance, R_b, can be written as follows if the source is a uniformly applied the heat sink base. If it is not, then the base resistance is primarily spreading resistance:

$$R_b = \frac{t_b}{kA_b} \quad (4)$$

where t_b is the heat sink base thickness, k is the heat sink material thermal conductivity and A_b is the area of the heat sink base.

The thermal resistance from the base of the fins to the air, R_f, can be calculated by the following formulas.

$$R_f = \frac{1}{nh_f W_f (t_f + 2\eta_f L_f)} \quad (5)$$

$$\eta_f = \frac{tanh mL_c}{mL_c} \quad (6)$$

$$mL_c = \sqrt{\frac{2h_f}{kt_f}} L_f \quad (7)$$

$$D_h = \frac{4A_{ch}}{P_{ch}} \quad (8)$$

$$Re = \frac{4\dot{G}\rho}{n\pi D_h \mu} \quad (9)$$

$$f = (0.79 \ln Re - 1.64)^{-2} \quad (10)$$

$$Nu = \frac{(f/8)(Re - 1000)Pr}{1 + 12.7(f/8)^{0.5}(Pr^{\frac{2}{3}} - 1)} \quad (11)$$

$$h_f = \frac{Nu k_{air}}{D_h} \quad (12)$$

$$\rho = \frac{P_{atm}}{R_a T_{in}} \quad (13)$$

The flow rate can be determined by the intersection of the heat sink system curve and the fan curve. The heat sink system curve can be calculated by the flow resistance of the channels and inlet and outlet losses as done in standard fluid mechanics text books, such as Potter, et al. and White.

Once the heat sink base and fin resistances are known, then the heat sink thermal resistance, R_{hs} can be calculated as:

$$R_{hs} = R_b + R_f \quad (14)$$

Using the equations 5 to 13 and the dimensional data in, the thermal resistance for the fins was calculated for various air flow rates. The data for the thermal resistance and heat transfer coefficient are shown in Figure 14. It shows that shows that for an increasing air flow rate, the thermal resistance of the heat sink decreases.

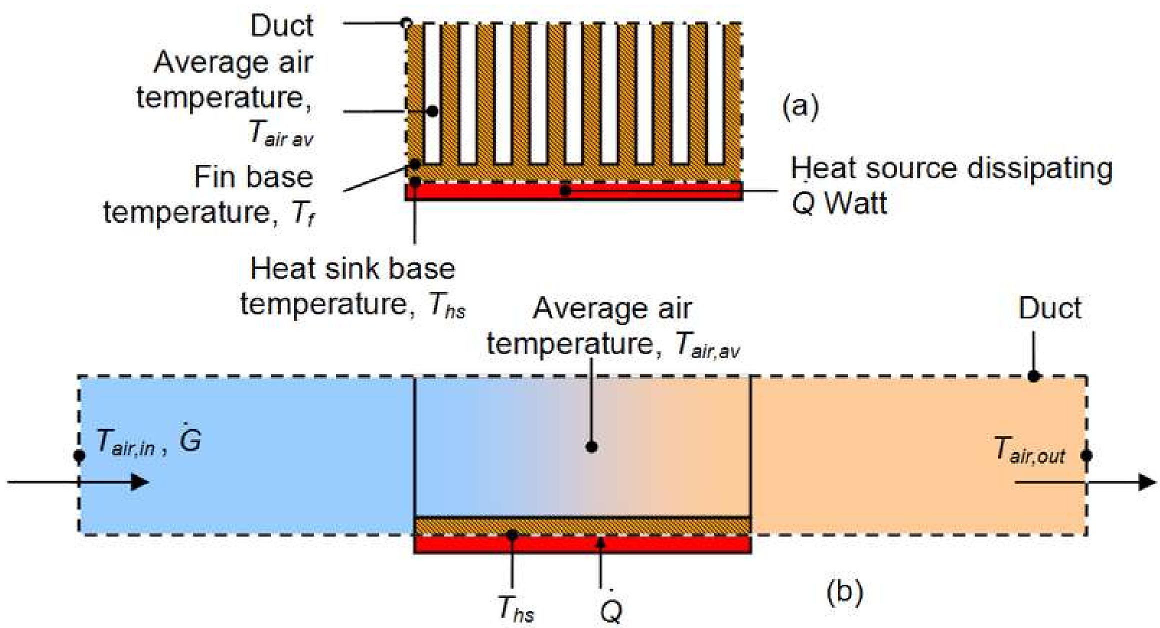


Figure 15: Sketch of a heat sink in a duct used to calculate the governing equations from conservation of energy and Newton's law of cooling.

Experimental methods

Experimental tests are one of the more popular ways to determine the heat sink thermal performance. In order to determine the heat sink thermal resistance, the flow rate, input power, inlet air temperature and heat sink base temperature need to be known, as shown in figure 15. However, figure 15 shows a test setup for a ducted flow heat sink application. Vendor-supplied data is commonly provided for ducted test results. However, the results are optimistic and can give misleading data when heat sinks are used in an unducted application. Another problem with the set up, be it unducted or ducted, is the losses into the board. These must be taken into account. More details on heat sink testing methods and common oversights can be found in Azar, et al.

Numerical methods

In industry, thermal analyses are often ignored in the design process or performed too late — when design changes are limited and become too costly. Of the three methods mentioned here, theoretical and numerical methods can be used to determine an estimate of the heat sink or component temperatures of products before a physical model has been made. A theoretical model is normally used as a first order estimate. Numerical methods or computational fluid dynamics (CFD) provide a qualitative (and sometimes even quantitative) prediction of fluid flows. What this means is that it will give a visual or post-processed result of a simulation, like the images in figures 16 and 17, but the quantitative or absolute accuracy of the result is not guaranteed.

CFD can give an insight into flow patterns that are difficult, expensive or impossible to study using experimental methods. Experiments can give a quantitative description of flow phenomena using measurements for one quantity at a time, at a limited number of points and time instances. If a full scale model is not available or not practical, scale models or dummy models can be used. The experiments can have a limited range of problems and operating conditions. Simulations can give a prediction of flow phenomena using CFD software for all desired quantities, with high resolution in space and time and virtually any problem and realistic operating conditions. However, the results still need to be validated. Another problem with CFD is that the inputs need to be correct. It is the classic case of "Garbage in, garbage out."

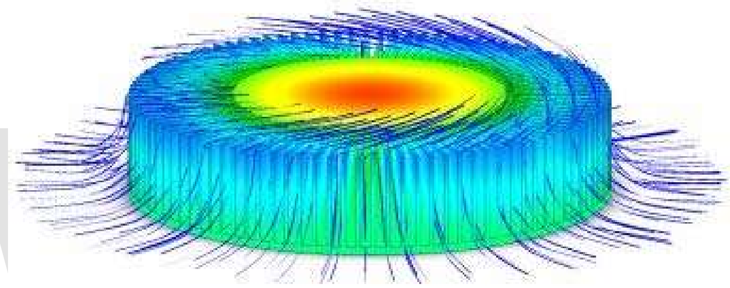


Figure 16: Radial heat sink with thermal profile and swirling forced convection flow trajectories predicted using a CFD analysis package

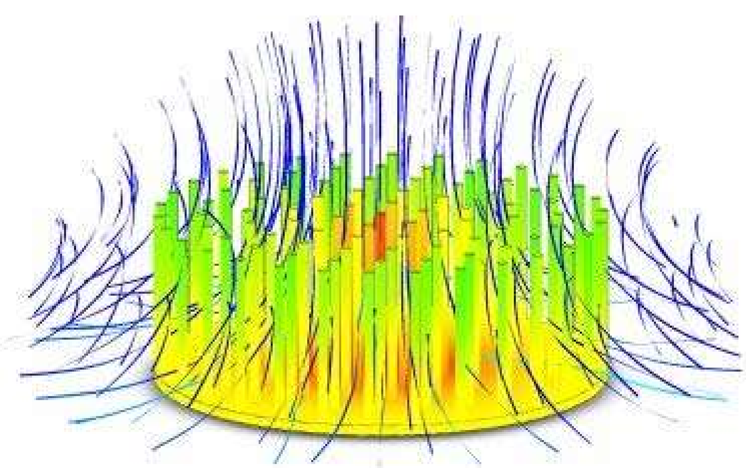


Figure 17: Pin fin heat sink with thermal profile and dione convection flow trajectories predicted using a CFD analysis package

Chapter 13

Laptop Cooler

Passive Laptop cooler

An Active Laptop cooler

A **laptop/notebook cooler**, **cooler pad** or **chill mat** is an accessory for laptop computers that helps reduce their operating temperature. Normally used when the laptop's fan device is unable to sufficiently cool the laptop, a cooling pad may house active or passive cooling methods and rests beneath the laptop. Active coolers move air or liquid to direct heat away from the laptop quickly, while passive methods may rely on thermally conductive materials or increasing passive airflow.

Active coolers

Many cooler pads support the use of a power adapter. They typically run on power drawn through one of the laptop's USB ports. Many cooler pads therefore compensate with a built-in USB hub.

Some of these "powered" coolers are designed to draw heat from the underside of the computer; others work in the opposite way by blowing cool air towards the machine. The fan speed is adjusted manually or automatically on certain models and on others stays at a fixed speed.

The contribution of heat from the energy requirement of the cooler is very minimal impact as USB ports are limited to 2.5 watts of output power, whereas a laptop can easily consume 40 W of electrical power.

Just as there are advantages of having fans in a cooler - there are disadvantages as well. The power drawn by the cooler increases with the wear of the bearings or bushings, or when the vents of fan are blocked by dust, and as friction rises. When the friction becomes significant the power drawn by the cooler increases as well and there have been instances of USB port failures due to overdraw. In a situation like this, although the cooler may appear to work but it may be ineffective or counterproductive by leading to overheating of key components of the mainboard, in rare cases - leading to premature failure of the mainboard.

Passive coolers

Typically, a conductive cooling pad allows for the cooling of a laptop without using any power. These "Pads" are normally filled with an organic salt compound that allows them to absorb the heat from the laptop . They are good for a limited amount of time from around 6–8 hours of cooling. Other designs are simply a pad that elevates the laptop so that the fans in the laptop are allowed greater airflow.

The conductive cooling pads are not advisable for laptops that have fan vents built into the bottom as the cooling pad blocks the vents leading to overheating or premature system failure. The best way to determine if a cooling pad would be suitable for your laptop would be to take a look at the bottom of the laptop and look for air vents or fan vents. If they are on the side and not on the bottom, it is usually safe to use the cooler pad otherwise not.

The other variety that can be used simply as a hard resting surface that provides gap between the cooler and the laptop is normally safer to use. The disadvantages of using these are - they are usually bulky, have to be carried around and sometimes they are too big or too small for a laptop so a bit cumbersome to use.

Multi-Surface Cooler

A Type of Passive Cooler that allows both airflow between the laptop base and cooler, as well as, between the base of the Cooler and Users's lap. These laptop coolers are well suited to laptops that have vents on its base because it prevents these vents from being blocked regardless of what ever surface the laptop is used on. Therefore, these Multi-Surface Coolers are suitable for use on desk, lap and uneven/soft surfaces (couch, bed/duvet, carpet) and the outdoors.

Multipurpose Coolers

Recent advancements have brought forward coolers that are multipurpose. Features include card readers for various forms of media such as keydrives, memory cards, and 2.5" Laptop hard disk drives.

In addition to the above coolers that are a combination of mini work desk with fans are a convenient addition to users that want to use the laptop on a bed or a couch - although they tend to be too heavy and bulky to be carried conveniently everywhere, limiting mobility.

One more variant of this is a cooler with writing pad having an area meant to be used for placing a book or a writing pad – designed with students in mind although the bigger size limits its mobility and the weight usually results in tired legs for the user when used for a prolonged period of time.

One recent addition to the above is an attachable laptop cooler and a comfort pad built into one. It offers combination of cooling, mobility, and comfort as our lives integrate more with mobile computers.

As the laptop gets more powerful and faster it generates more heat. Manufacturers are trying to embed newer technologies to reduce heat generation, improve efficiency, and lower power consumption, which will one day eliminate the use of these devices.

But if Laptop Manufacturer's make smaller and thinner laptops that house higher speed components, then laptops may only be comfortable to use with cooler pads to protect from toasted skin syndrome.

Chapter 14

Thermal Management of Electronic Devices and Systems

Heat generated by electronic devices and circuitry must be dissipated to improve reliability and prevent premature failure. Techniques for heat dissipation can include heatsinks and fans for air cooling, and other forms of computer cooling such as liquid cooling.

In cases of extreme low environmental temperatures, it may actually be necessary to heat the electronic components to achieve satisfactory operation.

Overview

Thermal resistance of devices

This is usually quoted as the thermal resistance from junction to case of the semiconductor device. The units are °C/W. For example, a heatsink rated at 10 °C/W will get 10°C hotter than the surrounding air when it dissipates 1 Watt of heat. Thus, a heatsink with a low °C/W value is more efficient than a heatsink with a high °C/W value.

Thermal time constants

A heatsink's thermal mass can be considered as a capacitor (storing heat instead of charge) and the thermal resistance as an electrical resistance (giving a measure of how fast stored heat can be dissipated). Together, these two components form a thermal RC circuit with an associated time constant given by the product of R and C. This quantity can be used to calculate the dynamic heat dissipation capability of a device, in an analogous way to the electrical case. A specific type of thermal interface material is put between the heat sink and the heat source to increase thermal throughput, such as a microprocessor chip or other power handling semiconductor to stabilise its temperature through increased thermal mass and heat dissipation (primarily by conduction and convection and to a lesser extent by radiation).

Thermal interface material

A **Thermal Interface Material or Mastic** (aka **TIM**) is used to fill the gaps between thermal transfer surfaces, such as between microprocessors and heatsinks, in order to increase thermal transfer efficiency. These gaps are normally filled with air which is a very poor conductor.

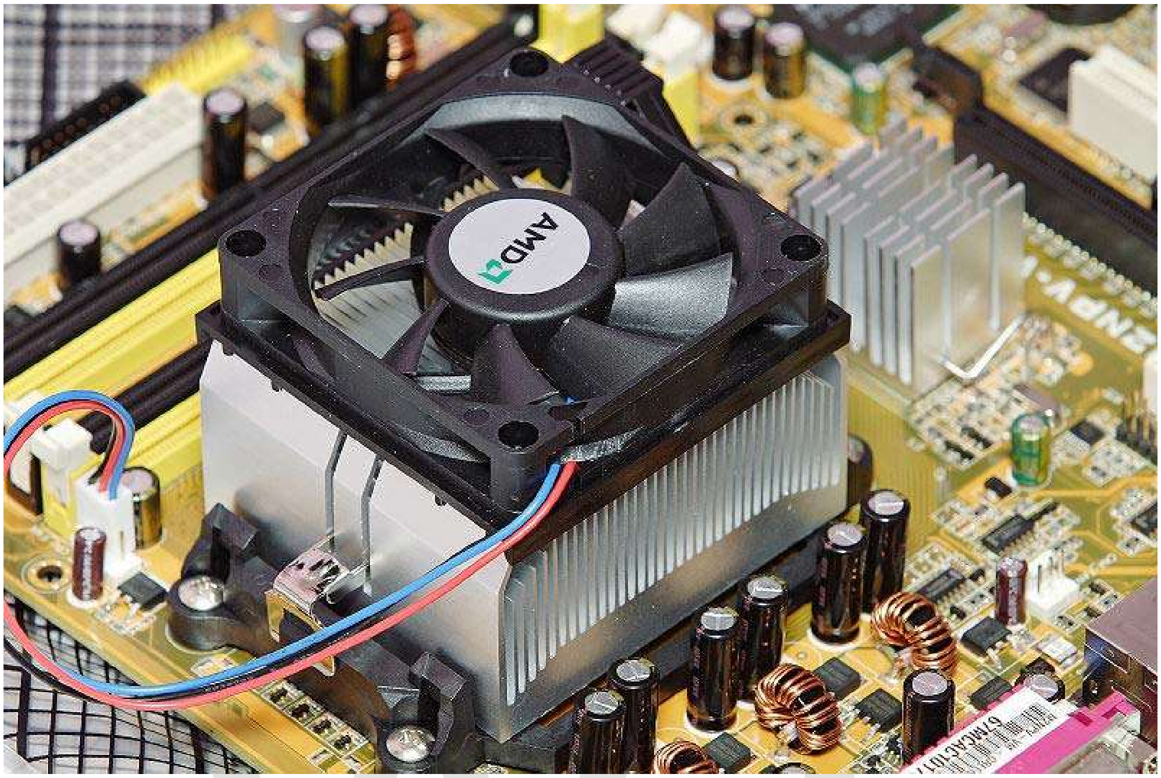
Device heatsinking

Heat sinks

Heat sinks are widely used in electronics, and have become almost essential to modern central processing units. In common use, it is a metal object brought into contact with an electronic component's hot surface — though in most cases, a thin thermal interface material mediates between the two surfaces. Microprocessors and power handling semiconductors are examples of electronics that need a heat sink to reduce their temperature through increased thermal mass and heat dissipation (primarily by conduction and convection and to a lesser extent by radiation). Heat sinks have become almost essential to modern integrated circuits like microprocessors, DSPs, GPUs, and more.

A heat sink usually consists of a metal structure with one or more flat surfaces to ensure good thermal contact with the components to be cooled, and an array of comb or fin like protrusions to increase the surface contact with the air, and thus the rate of heat dissipation.

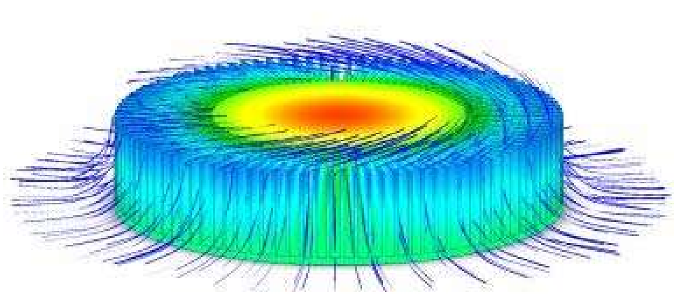
A heat sink is sometimes used in conjunction with a fan to increase the rate of airflow over the heat sink. This maintains a larger temperature gradient by replacing warmed air faster than convection would. This is known as a forced air system.



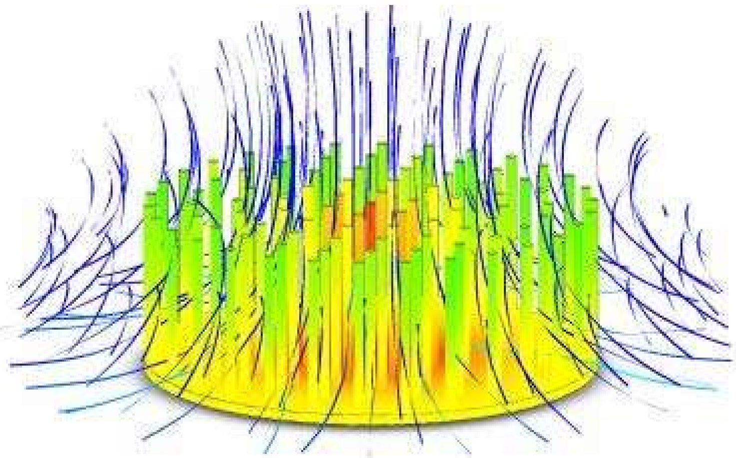
CPU heat sink with fan attached

A **heat sink** (or heatsink) is an environment or object that absorbs and dissipates heat from another object using thermal contact (either direct or radiant). Heat sinks are used in a wide range of applications wherever efficient heat dissipation is required; major examples include refrigeration, heat engines and cooling electronic devices.

Principle

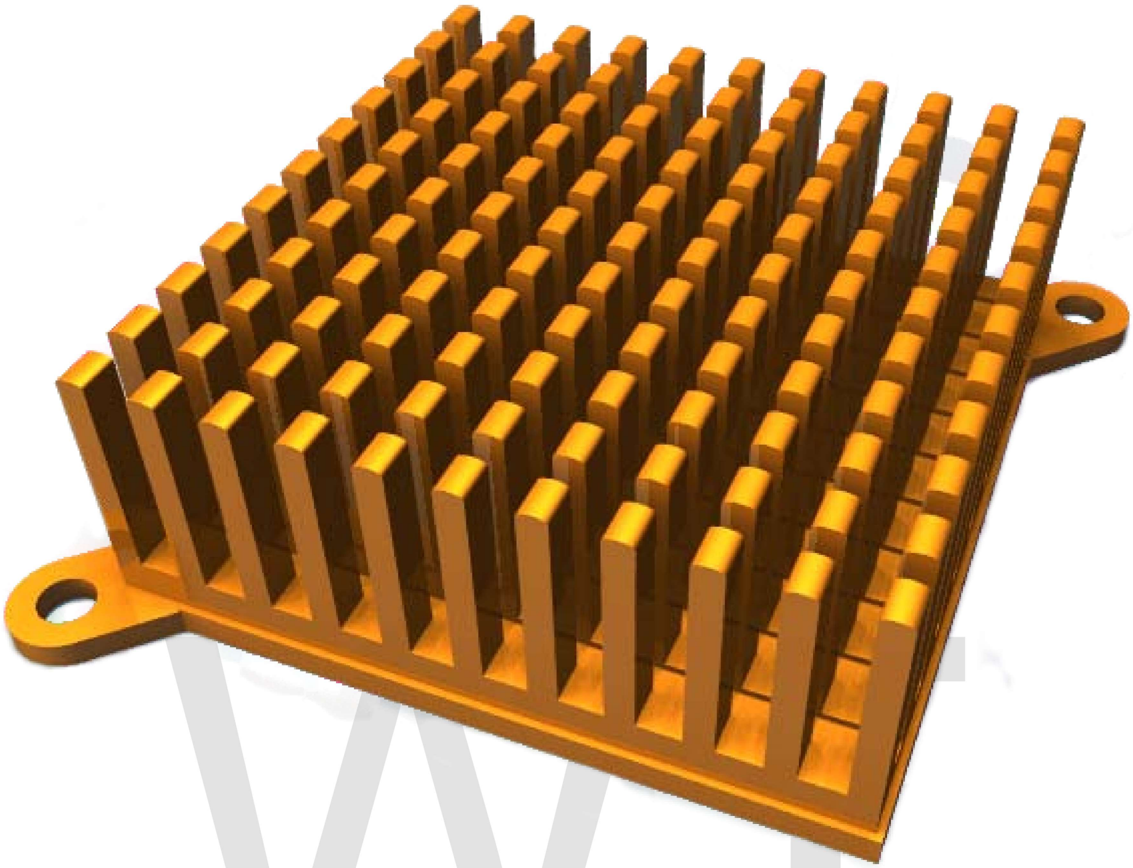


Radial Heat Sink with Thermal Profile and Swirling Forced Convection Flow Trajectories (using CFD analysis)

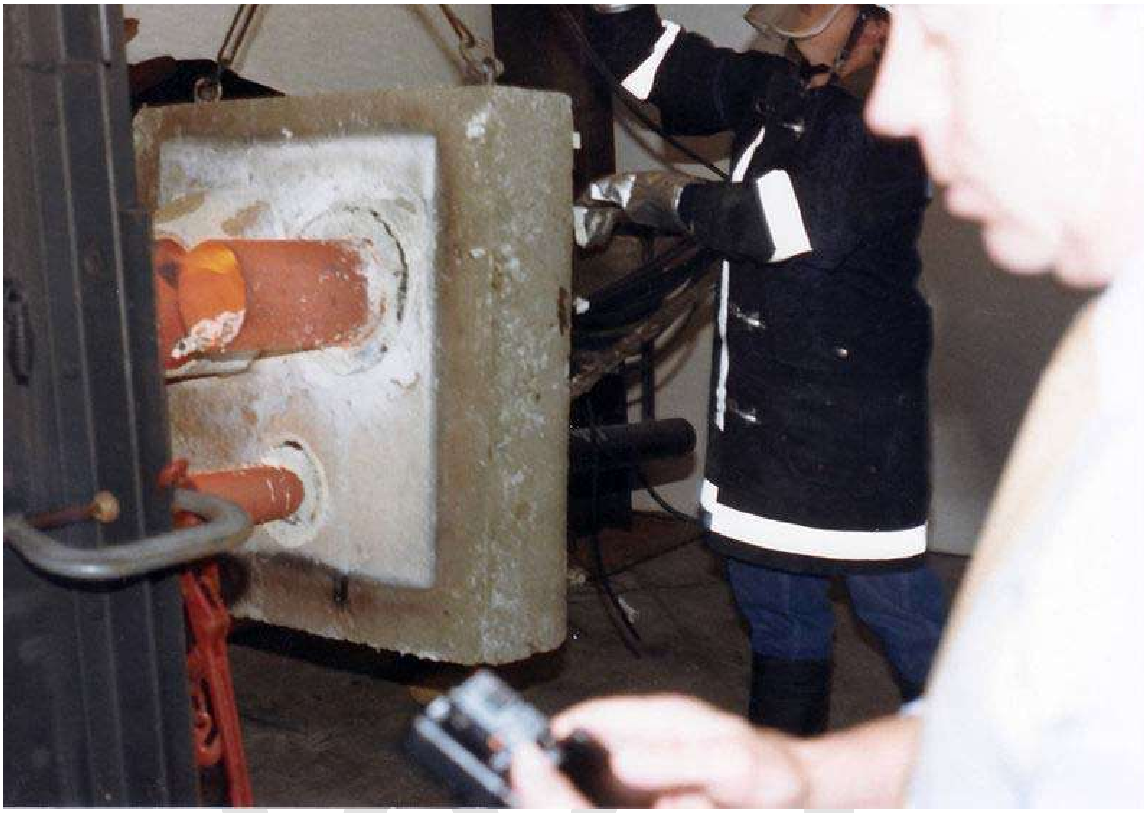


Pin Fin Heat Sink with Thermal Profile and Dione Convection Flow Trajectories (using CFD analysis)

Heat sink in a workstation computer



A motherboard heat sink



Fire test where the steel pipe penetrants clearly act to absorb and conduct heat from the furnace, through to the unexposed side.

Heat sinks function by efficiently transferring thermal energy ("heat") from an object at high temperature to a second object at a lower temperature with a much greater heat capacity. This rapid transfer of thermal energy quickly brings the first object into thermal equilibrium with the second, lowering the temperature of the first object, fulfilling the heat sink's role as a cooling device. Efficient function of a heat sink relies on rapid transfer of thermal energy from the first object to the heat sink, and the heat sink to the second object.

The most common design of a heat sink is a metal device with many fins. The high thermal conductivity of the metal combined with its large surface area result in the rapid transfer of thermal energy to the surrounding, cooler, air. This cools the heat sink and whatever it is in direct thermal contact with. Use of fluids (for example coolants in refrigeration) and thermal interface material (in cooling electronic devices) ensures good transfer of thermal energy to the heat sink. Similarly, a fan may improve the transfer of thermal energy from the heat sink to the air.

Construction and materials

A heat sink usually consists of a base with one or more flat surfaces and an array of comb or fin-like protrusions to increase the heat sink's surface area contacting the air, and thus increasing the heat dissipation rate. While a heat sink is a static object, a fan often aids a heat sink by providing increased airflow over the heat sink — thus maintaining a larger temperature gradient by replacing the warmed air more quickly than passive convection achieves alone — this is known as a forced-air system.

Ideally, heat sinks are made from a good thermal conductor such as silver, gold, copper, or aluminum alloy. Copper and aluminum are among the most-frequently used materials for this purpose within electronic devices. Copper (401 W/(m·K) at 300 K) is significantly more expensive than aluminum (237 W/(m·K) at 300 K) but is also roughly twice as efficient as a thermal conductor. Aluminum has the significant advantage that it can be easily formed by extrusion, thus making complex cross-sections possible. Aluminum is also much lighter than copper, offering less mechanical stress on delicate electronic components. Some heat sinks made from aluminum have a copper core as a trade off. The heat sink's contact surface (the base) must be flat and smooth to ensure the best thermal contact with the object needing cooling. Frequently a thermally conductive grease is used to ensure optimal thermal contact; such compounds often contain colloidal silver. Further, a clamping mechanism, screws, or thermal adhesive hold the heat sink tightly onto the component, but specifically without pressure that would crush the component.

Performance

Heat sink performance (including free convection, forced convection, liquid cooled, and any combination thereof) is a function of material, geometry, and overall surface heat transfer coefficient. Generally, forced convection heat sink thermal performance is improved by increasing the thermal conductivity of the heat sink materials, increasing the surface area (usually by adding extended surfaces, such as fins or foam metal) and by increasing the overall area heat transfer coefficient (usually by increase fluid velocity, such as adding fans, pumps, etc.).

Online heat sink calculators from companies such as Novel Concepts, Inc., can accurately estimate forced convection heat sink performance. For more complex heat sink geometries, and/or heat sinks with multiple materials, and/or heat sinks with multiple fluids, computation fluid dynamics (CFD) analysis is recommended.

Use in electronics

PC marketplace

Due to recent technological developments and public interest, the retail heat sink market has reached an all time high. In the early 2000s, CPUs were produced that emitted more and more heat than earlier, escalating requirements for quality cooling systems.

Overclocking has always meant greater cooling needs, and the inherently hotter chips meant more concerns for the enthusiast. Efficient heat sinks are vital to overclocked computer systems because the higher a microprocessor's cooling rate, the faster the computer can operate without instability; generally, faster operation leads to higher performance. Many companies now compete to offer the best heat sink for PC overclocking enthusiasts. Prominent aftermarket heat sink manufacturers include: Aero Cool, Foxconn, Thermalright, Thermaltake, Swiftech, and Zalman.

In soldering

Temporary heat sinks were sometimes used while soldering circuit boards, preventing excessive heat from damaging sensitive nearby electronics. In the simplest case, this means partially gripping a component using a heavy metal crocodile clip or similar clamp. Modern semiconductor devices, which are designed to be assembled by reflow soldering, can usually tolerate soldering temperatures without damage. On the other hand, electrical components such as magnetic reed switches can malfunction if exposed to higher powered soldering irons, so this practice is still very much in use.

In Batteries

In the battery used for electric vehicles, Nominal battery performance is usually specified for working temperatures somewhere in the + 20°C to +30°C range however the actual performance can deviate substantially from this if the battery is operated at higher or in particular lower temperatures, so some electric cars have heating and cooling for their batteries

Recent developments

More recently, synthetic diamond cooling sinks are being researched to provide better cooling. Also, some heat sinks are constructed of multiple materials with desirable characteristics, such as phase change materials, which can store a great deal of energy due to their heat of fusion.

Convective air cooling

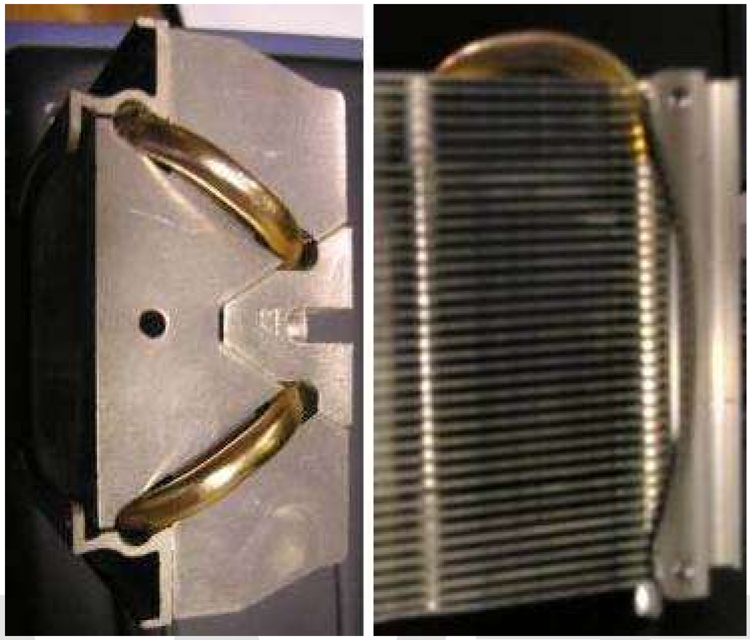
This term describes device cooling by the convection currents of the warm air being allowed to escape the confines of the component to be replaced by cooler air. Since warm air normally rises, this method usually requires venting at the top or sides of the casing to be effective.

Forced air cooling

If there is more air being forced into a system than being pumped out (due to an imbalance in the number of fans), this is referred to as a 'positive' airflow, as the pressure inside the unit is higher than outside.

A balanced or neutral airflow is the most efficient, although a slightly positive airflow can result in less dust build up if filtered properly.

Heat pipes



A heat sink (aluminium) incorporating a heat pipe (copper)

A **heat pipe** is a heat transfer mechanism that can transport large quantities of heat with a very small difference in temperature between the hot and cold interfaces. A typical heat pipe consists of sealed hollow tube made of a thermoconductive metal such as copper or aluminium. The pipe contains a relatively small quantity of a "working fluid" or coolant (such as water, ethanol or mercury) with the remainder of the pipe being filled with the vapour phase of the working fluid, all other gases being excluded. The advantage of heat pipes is their great efficiency in transferring heat. They are actually more "conductive" than a copper bar of equivalent cross-section.

Peltier cooling plates

Peltier cooling plates take advantage of what is known as the Peltier effect to create a heat flux between the junction of two different types of materials. This effect is commonly used for cooling electronic components and small instruments.

There are no moving parts and such a device is maintenance free. Due to the relatively low efficiency, thermoelectric cooling is generally only used in environments where the solid state nature outweighs the poor efficiency. Thermoelectric junctions are generally only around 10% as efficient as the ideal refrigerator (Carnot cycle), compared with 40% achieved by conventional compression cycle systems.

Synthetic Jet Air Cooling

A Synthetic Jet is produced by a continual flow of vortices that are formed by alternating brief ejection and suction of air across an opening such that the net mass flux is zero. A unique feature of these jets is that they are formed entirely from the working fluid of the flow system in which they are deployed can produce a net momentum to the flow of a system without net mass injection to the system.

Synthetic jet air movers have no moving parts and are thus maintenance free. Due to the high heat transfer coefficients, high reliability but lower overall flow rates, Synthetic jet air movers are usually used at the chip level and not at the system level for cooling. However depending on the size and complexity of the systems they can be used for both at times.

Chapter 15

Fan Control

Fan control is the management of the rotational speed of an electric fan. There are many types of electric fans and many types of fan controls.

Need for fan control

As modern PCs grow more powerful so do their requirements for electrical power. Computers convert most of this electrical power into heat generated by all major components.

Some early generation PCs did not need active ventilation. Power supplies eventually needed forced cooling, and soon took up the duty of cooling the rest of the PC with the ATX standard. The byproduct of increased heat generation is that the fan(s) need to move increasing amounts air and thus, need to be more powerful. Since they must move more air through the same area of space, fans will naturally become more noisy.

In fact, if one installs extra fans in a PC case, the noise levels can reach 70 dB. Since fan noise increases with the fifth power of the fan rotation speed, reducing rotations per minute (RPM) by a small amount potentially means a reduction in fan noise. This must be done cautiously, as excessive reduction in speed may cause components to overheat and be damaged. If done properly fan noise can be drastically reduced.

Fan types

Common cooling fans have at least two pins on the connector. These two pins connect to the fan motor, usually a brushless DC type. Extra pins may be present, with either or both of the following features:

- A tachometer that measures the actual speed of the fan as a pulse train, frequency being proportional to speed.
- Control of fan speed using PWM (Pulse-width modulation.)

Two-pin fans operate either as an on/off fan, or can be controlled by varying the voltage.

will have either two, three, or four pins on the connector.

Types of control

No control

The simplest method of fan control is simply to leave the fan on all the time. This creates quite a bit of noise and increases power requirements, but keeps the system the coolest.

Thermostatic

In this style of fan control, the fan is either on or off. A system thermistor checks the temperature inside the chassis, and if it detects a temperature outside of range, it spins the fans up to maximum. When the temperature drops below a threshold again, the fans are turned back off. This control method reduces power requirements during periods of low usage, but when the system is operating at capacity, the fan noise can become a problem again.

Linear voltage regulation

A standard cooling fan is essentially a bladed DC motor. By varying the voltage input across the acceptable range for a fan, the speed of the fan will increase (to added voltage) and decrease (to reduced voltage). A faster fan, obviously, means more air moved, and thus a higher heat exchange rate. There are a few ways to perform this regulation.

Resistors

Resistors are the simplest method of reducing fan noise, but they add to the heat generated inside the computer case. Since the voltage drop is proportional to the current, the fan may not start. They need to be of the appropriate power rating (i.e. higher than the fan). For variable fan control, potentiometers could be used along with a transistor such as a MOSFET whose output voltage is controlled by the potentiometer. It is possible to use a rheostat instead.

Diodes

A diode in series with the fan will reduce the voltage being output to the fan. You can use a zener diode (select one for the desired voltage drop) or a silicon diode (Produce the required voltage drop by connecting multiple diodes in series. Each diode reduces the voltage by approximately 0.75 volts.)

Voltmodding

Voltmodding describes the practice of varying the voltage fed to a component; in this case, a computer fan. This can be achieved by connecting the ground wire of the fan to the +5V rail and the positive wire to the +12V rail of a typical PC power supply to achieve a theoretical +7V (positive seven volts). However, this is a potentially risky method: the parts on the +5V power line might be exposed to overvoltage in case of a short in the fan. A less common variation is to increase the voltage to the fan by connecting the ground wire to a -12V rail (located on the motherboard ATX connector) and the positive wire to the +12V rail, producing a total of +24V. 24V can destroy the fans, other options are 17V and 15.3V.

Pulse-width modulation

Pulse-width modulation (PWM) is a common method of controlling fans. Modern computer motherboard PWM control when used with multi-core CPUs reads data from Digital Temperature Sensors on each core of the CPU.

Unlike the linear methods above that are based on voltage loss, PWM switches the input voltage between (nearly) fully on and fully off. This means there is practically no voltage or power loss and associated heat output. PWM controller can be a relatively small, low-power and cool-running, albeit complex, component that doesn't require heavy duty resistors, diodes or transistors and associated heatsinking.

Fan speed controllers

A fan controller with LEDs to indicate fan status and potentiometer and switch to control fan speed.

Another method, popular with gamers, is the manual fan speed controller. They can be mounted in an expansion slot, a 5.25" or 3.5" drive bay or come built into the computer's case. Using switches or knobs, attached fans can have their speeds adjusted by one of the above methods.

Software

The method by which the software physically controls the fan is usually PWM. Many companies now provide software to control fan speeds on their motherboards under Microsoft Windows.

- AOpen motherboards can use "SilentTEK".

- ASUS motherboards can use "Q-Fan".
- MSI motherboards can use "Core Center".
- Universal abit motherboards can use "µGuru".
- Gigabyte motherboards can use "EasyTune 6".
- Intel desktop boards (older socket 478 etc) use "Active Monitor" and "Desktop Control Centre".
- Intel desktop boards (newer socket 775 etc) use "Desktop Utilities".
- Dell Inspiron/Latitude/Precision computers can use "I8kfanGUI".
- Various computers can use the freeware "SpeedFan".

Computers running Linux can use lm sensors.