Q1. What is SPEC?

SPEC is the Standard Performance Evaluation Corporation. SPEC is a non-profit organization whose members include computer hardware vendors, software companies, universities, research organizations, systems integrators, publishers and consultants.

SPEC’s goal is to establish, maintain and endorse a standardized set of relevant benchmarks for computer systems.

Although no one set of tests can fully characterize overall system performance, SPEC believes that the user community benefits from objective tests which can serve as a common reference point.
Q2. What is a benchmark?

A benchmark is "a standard of measurement or evaluation" (Webster’s II Dictionary).

A computer benchmark is typically a computer program that performs a strictly defined set of operations - a workload - and returns some form of result - a metric - describing how the tested computer performed.

Computer benchmark metrics usually measure speed: how fast was the workload completed; or throughput: how many workload units per unit time were completed.

Running the same computer benchmark on multiple computers allows a comparison to be made.

Q3. Should I benchmark my own application?

Ideally, the best comparison test for systems would be your own application with your own workload.

Unfortunately, it is often impractical to get a wide base of reliable, repeatable and comparable measurements for different systems using your own application with your own workload.

Problems might include generation of a good test case, confidentiality concerns, difficulty ensuring comparable conditions, time, money, or other constraints.
Q5. What does SPEC CPU2006 measure?

SPEC CPU2006 focuses on compute intensive performance, emphasizing:
- the computer processor (CPU),
- the memory architecture, and
- the compilers.

SPEC CPU2006 contains two components that focus on two different types of compute intensive performance:
- CINT2006 suite measures compute-intensive integer performance,
- CFP2006 suite measures compute-intensive floating point performance.

SPEC CPU2006 is not intended to stress other computer components such as networking, the operating system, graphics, or the I/O system.

For single-CPU tests, the effects from such components on SPEC CPU2006 performance are usually minor. For large rate runs, operating system services may affect performance, and the I/O system - number of disks, speed, striping - can have an effect.

Many Other SPEC Benchmarks

http://www.spec.org/benchmarks.html
Q6. Why use SPEC CPU2006?

SPEC CPU2006 provides a comparative measure of integer and/or floating point compute intensive performance.

Other advantages to using SPEC CPU2006 include:

• The benchmark programs are developed from actual end-user applications, as opposed to being synthetic benchmarks.
• Multiple vendors use the suite and support it.
• SPEC CPU2006 is highly portable.
• A wide range of results are available at http://www.spec.org
• The benchmarks are required to be run and reported according to a set of rules to ensure comparability and repeatability.

Q8. What is included in the SPEC CPU2006 package?

SPEC provides the following on the SPEC CPU2006 media (DVD):

• Source code for the CINT2006 benchmarks
• Source code for the CFP2006 benchmarks
• A tool set for compiling, running, validating and reporting on the benchmarks
• Pre-compiled tools for a variety of operating systems
• Source code for the SPEC CPU2006 tools, for systems not covered by the pre-compiled tools
• Run and reporting rules defining how the benchmarks should be used to produce SPEC CPU2006 results.
• Documentation
Q10. What are the basic steps in running the benchmarks?

- Ensure that you meet the system requirements.
- Install SPEC CPU2006 from the DVD
- Determine which metric you wish to run.
- Locate a configuration file as a starting point.
- Use runspec to build the benchmarks.
- If the above steps are successful, use runspec to run, validate, and create a report on the performance of the benchmarks.

CINT 2006

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Language</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>perlbench</td>
<td>C</td>
<td>PERL Programming Lang</td>
</tr>
<tr>
<td>bzip2</td>
<td>C</td>
<td>Compression</td>
</tr>
<tr>
<td>Gcc</td>
<td>C</td>
<td>C Compiler</td>
</tr>
<tr>
<td>Mcf</td>
<td>C</td>
<td>Combinatorial Optimization</td>
</tr>
<tr>
<td>Gobmk</td>
<td>C</td>
<td>Artificial Intelligence: go</td>
</tr>
<tr>
<td>Hmmer</td>
<td>C</td>
<td>Search Gene Sequence</td>
</tr>
<tr>
<td>Sjeng</td>
<td>C</td>
<td>Artificial Intelligence: chess</td>
</tr>
<tr>
<td>Libquantum</td>
<td>C</td>
<td>Physics: Quantum Computing</td>
</tr>
<tr>
<td>h264ref</td>
<td>C</td>
<td>Video Compression</td>
</tr>
<tr>
<td>Omnetpp</td>
<td>C++</td>
<td>Discrete Event Simulation</td>
</tr>
<tr>
<td>Astar</td>
<td>C++</td>
<td>Path-finding Algorithms</td>
</tr>
<tr>
<td>Xalanbmk</td>
<td>C++</td>
<td>XML Processing</td>
</tr>
</tbody>
</table>
Q13. What metrics can be measured?

- SPECint2006: The geometric mean of twelve normalized ratios - one for each integer benchmark - when the benchmarks are compiled with peak tuning.
- SPECint_base2006: The geometric mean of twelve normalized ratios when the benchmarks are compiled with base tuning.
- SPECint_rate2006: The geometric mean of twelve normalized throughput ratios when the benchmarks are compiled with peak tuning.
- SPECint_rate_base2006: The geometric mean of twelve normalized throughput ratios when the benchmarks are compiled with base tuning.

Similar for FP
Q14. What is the difference between a "base" metric and a "peak" metric?

- The base metrics (e.g. SPECint_base2006) are required for all reported results and have stricter guidelines for compilation. For example, the same flags must be used in the same order for all benchmarks of a given language. This is the point closer to those who might prefer a relatively simple build process.

- The peak metrics (e.g. SPECint2006) are optional and have less strict requirements. For example, different compiler options may be used on each benchmark, and feedback-directed optimization is allowed. This point is closer to those who may be willing to invest more time and effort in development of build procedures.

Q15. What is the difference between a "rate" and a "speed" metric?

There are several ways to measure computer performance.
- One way is to measure how fast the computer completes a single task; this is a speed measure.
- Another way is to measure how many tasks a computer can accomplish in a certain amount of time; this is called a throughput, capacity or rate measure.

For the rate metrics, multiple copies of the benchmarks are run simultaneously. Typically, the number of copies is the same as the number of CPUs on the machine, but this is not a requirement.
Q17: SPEC CPU2000 is already available. Why create SPEC CPU2006?

Technology is always improving. As the technology improves, the benchmarks should improve as well. SPEC needed to address the following issues:

- **Run-time:**
  As of summer, 2006, many of the CPU2000 benchmarks are finishing in less than a minute on leading-edge processors/systems. Small changes or fluctuations in system state or measurement conditions can therefore have significant impacts on the percentage of observed run time. SPEC chose to make run times for CPU2006 benchmarks longer to take into account future performance and prevent this from being an issue for the lifetime of the suites.

- **Application size:**
  As applications grow in complexity and size, CPU2000 becomes less representative of what runs on current systems. For CPU2006, SPEC included some programs with both larger resource requirements and more complex source code.

- **Application type:**
  SPEC felt that there were additional application areas that should be included in CPU2006 to increase variety and representation within the suites. For example, video compression and speech recognition have been added, and molecular biology has been significantly expanded.

- **Moving target:**
  CPU2000 has been available for six years and much improvement in hardware and software has occurred during this time. Benchmarks need to evolve to keep pace with improvements.

Q19: Is there a way to translate SPEC CPU2000 results to SPEC CPU2006 results or vice versa?

There is no formula for converting CPU2000 results to CPU2006 results and vice versa; they are different products.

There probably will be some correlation between CPU2000 and CPU2006 results (i.e., machines with higher CPU2000 results often will have higher CPU2006 results), but there is no universal formula for all systems.
Q20: What criteria were used to select the benchmarks?

- portability to a variety of CPU architectures (32- and 64-bit including AMD64, Intel IA32, Itanium, PA-RISC, PowerPC, SPARC, etc.)
- portability to various operating systems, particularly UNIX and Windows
- nearly all of the time is spent compute bound
- little time spent in IO and system services
- benchmarks should run in about 1GB RAM without swapping or paging
- no more than five percent of benchmarking time should be spent processing code not provided by SPEC
- well-known applications or application areas
- available workloads that represent real problems

Q23: Why does SPEC use a reference machine? What machine is used for SPEC CPU2006?

- SPEC uses a reference machine to normalize the performance metrics used in the CPU2006 suites. Each benchmark is run and measured on this machine to establish a reference time for that benchmark. These times are then used in the SPEC calculations.
- SPEC uses a historical Sun system, the "Ultra Enterprise 2" which was introduced in 1997, as the reference machine. The reference machine uses a 296 MHz UltraSPARC II processor, as did the reference machine for CPU2000. But the reference machines for the two suites are not identical: the CPU2006 reference machine has substantially better caches, and the CPU2000 reference machine could not have held enough memory to run CPU2006.
- It takes about 12 days to do a rule-conforming run of the base metrics for CINT2006 and CFP2006 on the CPU2006 reference machine.
Q26: Where are SPEC CPU2006 results available?