Performance and Drupal – Using Caching as an Example

Gerhard Killesreiter

August 28, 2008
Computer performance

Computer performance is characterized by the amount of useful work accomplished by a computer system compared to the time and resources used. Depending on the context, good computer performance may involve one or more of the following:
Computer performance

- Short response time for a given piece of work
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: Wikipedia, the free encyclopedia
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: Wikipedia, the free encyclopedia

For Drupal this translates to:
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: *Wikipedia, the free encyclopedia*

For Drupal this translates to:

- Short computation time per page
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: *Wikipedia, the free encyclopedia*

For Drupal this translates to:

- Short computation time per page
- Many pages per minute
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: Wikipedia, the free encyclopedia

For Drupal this translates to:

- Short computation time per page
- Many pages per minute
- Use RAM and database sparingly
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: Wikipedia, the free encyclopedia

For Drupal this translates to:

- Short computation time per page
- Many pages per minute
- Use RAM and database sparingly
- Don’t break
Computer performance

- Short response time for a given piece of work
- High throughput (rate of processing work)
- Low utilization of computing resource(s)
- High availability of the computing system or application

From: *Wikipedia, the free encyclopedia*

For Drupal this translates to:

- Short computation time per page
- Many pages per minute
- Use RAM and database sparingly
- Don’t break
In software engineering, performance analysis, more commonly profiling, is the investigation of a program’s behavior using information gathered as the program runs [...]. The usual goal of performance analysis is to determine which parts of a program to optimize for *speed* or *memory* usage.

From: *Wikipedia, the free encyclopedia*
In software engineering, performance analysis, more commonly profiling, is the investigation of a program’s behavior using information gathered as the program runs [...]. The usual goal of performance analysis is to determine which parts of a program to optimize for *speed* or *memory* usage.

From: *Wikipedia, the free encyclopedia*

Translates to:
Performance analysis

In software engineering, performance analysis, more commonly profiling, is the investigation of a program’s behavior using information gathered as the program runs [...]. The usual goal of performance analysis is to determine which parts of a program to optimize for *speed* or *memory* usage.

From: *Wikipedia, the free encyclopedia*

Translates to:

Monitoring!
If you don’t monitor the health of your Drupal install, you are a fool!
Monitoring

If you don’t monitor the health of your Drupal install, you are a fool!

From: Gerhard Killesreiter
Monitoring

Programms to use:
Monitoring

Programms to use:

- watchdog
Monitoring

Programms to use:

- watchdog
- rrdtool (efficient data collection)
Monitoring

Programms to use:

- watchdog
- rrdtool (efficient data collection)
- Cacti (Graphical display of collected data)
Monitoring

Programms to use:

- watchdog
- rrdtool (efficient data collection)
- Cacti (Graphical display of collected data)
- Munin (does pretty much the same)
Monitoring

Programms to use:

- watchdog
- rrdtool (efficient data collection)
- Cacti (Graphical display of collected data)
- Munin (does pretty much the same)

Result:
Performance tuning

Performance tuning is the improvement of system performance. [...] The motivation for such activity is called a performance problem, which can be real or anticipated. [...] 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.

From: Wikipedia, the free encyclopedia
Performance tuning

Systematic tuning follows these steps:
Performance tuning

Systematic tuning follows these steps:

1. Assess the problem and establish numeric values that categorize acceptable behavior.
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.
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.
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.
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.
Measurement

Measurement is the process of estimating the magnitude of some attribute of an object. [...] The act of measuring usually involves using a measuring instrument. [...] Measurements always have errors and therefore uncertainties. [...] Measurement errors are often assumed to be normally distributed about the true value of the measured quantity. Under this assumption, every measurement has three components: the estimate, the margin of error or uncertainty or error bound, and the confidence level — that is the probability that the actual magnitude lies within the margin of error. For example, a measurement of the length of a plank might result in an estimate of 2.53 meters plus or minus 0.01 meter, with a level of confidence of 99%.

From: Wikipedia, the free encyclopedia
Measurement

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- Drupal’s Caches
- Efficiency of Caches
- Top Secret!
Measurement

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- Drupal’s Caches
- Efficiency of Caches
- Top Secret!

From: buytaert.net
Measurement

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- Drupal’s Caches
- Efficiency of Caches
- Top Secret!
Measurement

Server Hostname: drupal.org
Document Path: /
Document Length: 24995 bytes
Concurrency Level: 1
Time taken for tests: 10.937806 seconds
Complete requests: 10
Total transferred: 256061 bytes
Requests per second: 0.91 [#/sec] (mean)
Time per request: 1093.781 [ms] (mean)
Transfer rate: 22.86 [Kbytes/sec] received
Connection Times (ms)

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>mean</th>
<th>[+/-sd]</th>
<th>median</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connect:</td>
<td>208</td>
<td>210</td>
<td>1.9</td>
<td>210</td>
<td>215</td>
</tr>
<tr>
<td>Processing:</td>
<td>859</td>
<td>882</td>
<td>43.5</td>
<td>866</td>
<td>994</td>
</tr>
<tr>
<td>Waiting:</td>
<td>230</td>
<td>252</td>
<td>43.3</td>
<td>234</td>
<td>362</td>
</tr>
<tr>
<td>Total:</td>
<td>1067</td>
<td>1093</td>
<td>43.5</td>
<td>1076</td>
<td>1204</td>
</tr>
</tbody>
</table>
How to measure Drupal?

Standard answer:
How to measure Drupal?

Just run “ab”
How to measure Drupal?

Better:
How to measure Drupal?

Run “ab” and read the manual!
How to measure Drupal?

And know something about measurement: Choose correct parameters for concurrency and number of requests.
How to measure Drupal?

“ab” is rather limited, it is good for testing a particular URL but doesn’t measure the system as a whole.
How to measure Drupal?

Better answers are needed. We want to assess performance improvements involving the whole system.
Siege is an open source stress / regression test and benchmark utility. It can stress a single URL with a user defined number of simulated users or it can read many URLs into memory and stress them simultaneously. The program reports the total number of hits recorded, bytes transferred, response time, concurrency, and return status.

Sproxy is an HTTP proxy server that collects URLs including GET and POST information in a "siege-friendly" manner. It was designed to ease the burden associated with the creation of a siege urls.txt file.

From: http://www.joedog.org/
A Plan

1. Measure important quantities.
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.

Credit: Ethan Fremen
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.

Credit: Ethan Fremen
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.
4. Play back that data on a test system using the snapshot and siege.
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.
4. Play back that data on a test system using the snapshot and siege.
5. Measure
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.
4. Play back that data on a test system using the snapshot and siege.
5. Measure
6. Make changes
A Plan

Credit: Ethan Fremen

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.
4. Play back that data on a test system using the snapshot and siege.
5. Measure
6. Make changes
7. Play back and measure
A Plan

1. Measure important quantities.
2. Take a snapshot of your system.
3. Collect real user data off your website using sproxy.
4. Play back that data on a test system using the snapshot and siege.
5. Measure
6. Make changes
7. Play back and measure
8. Repeat

Credit: Ethan Fremen
Cache

In computer science, a cache [...] is a collection of data duplicating original values stored elsewhere or computed earlier, where the original data is expensive to fetch [...] or to compute, compared to the cost of reading the cache.

From: Wikipedia, the free encyclopedia
Drupal’s Caches

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- **Drupal’s Caches**
- Efficiency of Caches
- Top Secret!
Drupal’s Caches

Drupal stores its caches in the database.
Drupal’s Caches

Drupal stores its caches in the database. Drupal has a swapable cache.inc file ⇒ Define your own caching!
Drupal’s Caches

Drupal stores its caches in the database. Drupal has a swapable cache.inc file ⇒ Define your own caching! Popular example: memcached
Efficiency of Caches

Does caching make sense?

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- Drupal’s Caches
- Efficiency of Caches
- Top Secret!
Efficiency of Caches

Does caching make sense?

![Graph of Memcached Cache Hits and Misses]

**drupal-www4 - Memcached Cache Hits and Misses**


- Hits/sec: Current: 86.79, Average: 58.76, Maximum: 171.85
- Misses/sec: Current: 42.67, Average: 40.47, Maximum: 76.40

**drupal3 - Memcached Cache Hits and Misses**


- Misses/sec: Current: 123.46 m, Average: 168.92 m, Maximum: 7.08

- Computer performance
- Performance analysis
- Monitoring
- Performance tuning
- Measurement
- How to measure Drupal?
- A Plan
- Cache
- Drupal’s Caches
- Efficiency of Caches
- Top Secret!
If a website is called “slow” it turns out that in 80% of all cases not the PHP/SQL–backend is at fault, but that there’s a problem with too many included files (JS/ CSS/ Images). Attend Konstantin’s talk after the break to learn more about this.