Monitoring and measuring server availability

Supervisor: Dr. Tim Moors
Assessor: Dr. Vijay Sivaraman

web servers, mail servers, file servers and others. It's important that these servers are available and accessible by users. Hence, there is a need for a tool to monitor and measure the availability of these servers so that network administrators can detect such server unavailability and take necessary actions to rectify it.

End user perspective availability
Conventional method of describing server unavailability is using the concept of uptime/downtime. This metric is usually expressed in percentage such as 99.95% or 99.995%. This metric is insufficient to describe server availability since it does not provide server availability from an end-user perspective. Better methods of measuring server availability include:

- Take into consideration partial responses of server availability instead of binary representations of availability.
- Obtain user impact server availability by measuring server availability based on the server connection distribution. E.g. server outage during the peak hours of the day is more severe than during non-peak hours.
- Provide more detailed information on the causes of server unavailability such as failures due to DNS hostname resolve, inability to connect to server or slow/no response from server.

As seen from the server connection distribution above, server outage during the early morning is less severe than during mid-afternoon. Thus, this program will take this into account when computing the server availability metric.

The diagrams above illustrate the comparison between a completely loaded web page and an incompletely loaded web page. An algorithm has been developed to detect partial responses from web server as incompletely loaded web page also indicates server unavailability.

System architecture
The implementation of this system consists of two main parts which are

- Prober module
- Log analysis module

The prober module is responsible for generating requests, send requests and process responses from server while the log analysis module is to analyze the logs generated by the prober module to produce important server availability and performance metrics.

The main log analysis tool developed is the log_analyzer program. Its generated metrics or data are as follows:

- Number of probes
- Connection time (mean, median, SLA min, max)
- Data transfer time (mean, median, SLA min, max)
- Subjective availability (incompletely loaded web pages)
- MTTF availability ratio and user impact availability
- Monitoring details (duration, resolution etc)
- Breakdown of user impact availability metric according to days monitored.