



# Propel<sup>®</sup> Accelerator

## Testing Procedures Guide

## Legal Notices

Reproduction of this document in whole or in part in any form or medium without express permission is prohibited. No part of this document may be photocopied, reproduced, or translated to another language without the prior consent of Propel Software Corporation (“Propel”).

### **Warranty Disclaimer**

This document is supplied on an “as is” basis with no warranty and no support. Propel makes no express warranty, whether written or oral with respect to this document. PROPEL SOFTWARE CORPORATION DISCLAIMS ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.

LIMITATION OF LIABILITY: IN NO EVENT SHALL PROPEL BE LIABLE FOR ERRORS CONTAINED HEREIN OR FOR ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES WHETHER BASED ON WARRANTY, CONTRACT, TORT, OR ANY OTHER LEGAL THEORY IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS MATERIAL.

### **Trademarks**

Propel®, Propel® Accelerator, Propel® Complete, Propel® Self-hosted™, Propel® Acceleration Server and the Propel logos are trademarks and “Our faster is better” is a service mark, of Propel Software Corporation in the United States and/or other countries. All other trademarks shown are the property of their respective owners.

Copyright © 2009. Propel Software Corporation. All right reserved.

### **Confidentiality Notice**

This document contains Propel confidential information. Do not read this document if you are not the intended recipient. If you are not the intended recipient, or a person responsible for delivering it to the intended recipient, you are hereby notified that any disclosure, copying, distribution or use of any of the information contained in or attached to this document is strictly prohibited. If you have received this document in error, please immediately notify Propel at (408) 571-6300 and destroy the original copy without reading or saving it in any manner.

The information contained in this document is subject to change without notice.

Propel Software Corporation  
1010 Rincon Circle  
San Jose, CA 95131-1325  
USA

## Table of Contents

Introduction	3
Propel Testing Considerations	4
Step-by-Step Guide to Testing	6
Comparing and Analyzing Test Results	9
Quick Reference Summary	12
Sample Bumper Pages	14

## Introduction

Evaluating the end user performance benefit of web acceleration by effectively comparing the speed of Propel Accelerator with the speed of a non-accelerated user connection is a complex experimental process. It is easy to arrive at erroneous conclusions unless the nature of this complexity is understood and controlled when testing.

This paper will guide you through the step-by-step procedure we've found effective for measuring performance. You will see how important the careful measurement process described in this paper is to obtaining accurate performance testing results.

Before you get started, Propel has found the items below are needed to perform an accurate test.

- A pair of matched systems/machines and modems<sup>1</sup> - one running with Propel Accelerator and one as a control system running without acceleration.
- It is critical that all systems connect to the Internet at the same speed. Modem connections can be speed limited with initialization strings and the speed verified with a tool such as the [www.numion.com/YourSpeed3/Run.php?QuickStart=MaxSpeed](http://www.numion.com/YourSpeed3/Run.php?QuickStart=MaxSpeed) max-speed utility.
- A set of at least 50 URLs that are representative of your end user's web browsing patterns.
- A minimum of 5 valid data observations per URL for meaningful analysis (we recommend that you plan to execute 6-8 test runs to insure 5 valid observations).
- We recommend that you use a tool that automatically browses and records time the page load time such as the Propel developed tool we call PropelIRT. We make PropelIRT available for you free use for 90 days. Please request it from your sales representative.
- You can concurrently evaluate more than one Accelerator level or version by using multiple matched pairs of systems.<sup>2</sup>
- You should expect to spend 2 to 5 days for set-up and testing. Each test run consists of two rounds. The initial page visit round and the repeat page visit round. Thus, each run provides 2 sets of data. Your actual test time will be determined by the number and page complexity of the URL list, the number of alternate configurations tested, etc.

---

<sup>1</sup> This document focusses on testing with dial-up connections. If you are testing with low-speed broadband, the same principle of controlling the experimental environment must be applied. In particular:

- 1) Each test system should have a dedicated connection which will not interfere with other test systems. This means two test systems must not share the same cable or DSL modem.
- 2) The quality, capacity, Internet routing, etc. should be equivalent to what your end users would have installed in their home or office.
- 3) The nominal speed of all connections over all test runs should be the same.
- 4) Bandwidth limiting tools such as NotEm and NistNet have questionable algorithms for delay insertion and may not accurately simulate real network conditions.

<sup>2</sup> A single control system may be utilized to collect comparison data for multiple accelerated systems so long as the only difference between the systems is the accelerator configuration. It is not appropriate to use different ISPs, POPs, etc., within data sets to be compared.

## Propel Testing Considerations

### Overview

A wide of variables including line quality, network connections, rotating website content, and origin server performance can affect test results. Tests performed on both accelerated and control non-accelerated configurations show broad variations due to these types of issues. The following section describes techniques designed to reduce the analysis impact of conditions outside of your control.

### Testing Considerations and URL Lists

1. The more samples you can collect, the lower the impact of anomalies caused by variables noted above. Thus, it is important to collect at least five valid samples.<sup>3</sup>
2. Categorized page characteristics so that you can analyze results for similar pages. Users have different expectations for page load times based on page content complexity. Your analysis should attempt to mimic user expectations.
3. Create URL list that represents the types of sites your customers will be visiting. In addition to the common mainstream such as Ebay, CNN, Amazon, etc., include corporate web sites and sites that are advertisement or graphics-heavy. Propel has observed that a significant percentage of web traffic involves visual content which would be awkward in a business context. To evaluate the acceleration impact of such content, we utilize sites such as those which feature travel or family snapshot thumbnails linked to full size pictures. We also utilize the Sports Illustrated site which includes tolerable pictures with acceleration characteristics similar to the more offensive sites. Note: You will find analysis is easier if you group URLs in your list based on characteristics.
4. You should exclude unreliable URLs from your list. Major public servers or reliable test servers you control are best.
5. Simulate browsing within a site by using URLs a user would reach by clicking links from the home page. Depending on the total URL list size, use 3-6 URLs from each domain.
6. Test for first visits and same-day repeat visits to measure the impact of browser and accelerator caching of content.<sup>4</sup>
7. Your URL list should be appropriate for your region. We would be happy to review your list of URLs for benchmarking during your planning phase.
8. Your final URL list should begin and end with trivial BUMPER URLs (see the samples at the end of this document) which are eliminated from your analysis. This ensures that the acceleration network is fully enabled prior to making measurement.

### Traffic Distribution

The following table summarizes the distribution of different web content types based on the total data bytes in each category. For your measurements to be realistic, you URL should approximate this distribution. This data was obtained by analyzing log records from Propel-hosted acceleration servers, which accelerate millions of individual user requests per day.

---

<sup>3</sup> As noted, generally 6-8 test runs will be required.

<sup>4</sup> Subsequent day repeat visits are interesting to evaluate, but the required testing methodology is very complex and beyond the scope of this paper.

## **Primary Web Traffic Content Type Distribution**

Ideally, your overall mix of URLs will approximate at least the first few rows of the above distribution.

## **Timing and PropelRT**

PropelRT, Propel's web browsing tool, automates the testing process by automatically loading and timing web page loads, working in conjunction with a real browser, and recording all data in a form that can be exported easily to programs such as Microsoft Excel. Timing data collected by PropelRT is equivalent to what an end user would observe by watching the browser's busy/done indicators.

Using lower level tools that make http GET requests directly will not be representative of an actual user's experience since much of the web is based on dynamic content decisions implemented by browser JavaScript, Flash Objects etc. Each page requested by a user click event may contain as many as 100 references to objects which must be retrieved to display the page. The Numion Stopwatch is an alternative web page based approach which can be used to time page loads. The primary shortcoming of Numion Stopwatch is the lack of automation.

## Step-by-Step Guide to Testing

CONTENT TYPE	% OF TOTAL TRAFFIC (BYTES)
HTML and other forms of Text	44.14
JPEG	30.47
GIF	17.25
Flash	1.46
PNG	0.14
BMP	<0.01

The objective of these procedures is to obtain multiple timing measurements for the initial page visits and repeat page visits under the same acceleration settings. These procedures should be repeated for each acceleration setting to be evaluated.

### Initial Page Visits (aka Initial Visits) / Same Day Repeat Visits (aka Repeat Visits)

These timings are important as they are intended to reflect real-world user browsing habits. Users generally tend to visit the same sites every day. In a recent Propel Survey, less than 5% of respondents stated that they surf mostly new sites each day, while 47% say they generally visit the same sites each day including email, news, auction sites, etc. Given that browsers and some accelerators use client-side caching, repeat visits should always see a significant decrease in page load times.

### Testing Procedures

1. We recommend gathering at least six samples per URL for each configuration and visit.
2. Set-up all test machines ("clean machines") with the appropriate software, timing tools and the accelerator client installed. (No acceleration client on the control machine).
3. Establish a dial-up connection from each machine, using the same access number, pop-up blocking, ad blocking etc. Note: PropelRT provides moderately effective pop-up blocking so it may be appropriate to disable accelerator pop-up blocking.
4. Use a bandwidth meter such as the [www.numion.com/YourSpeed3/Run.php?QuickStart=MaxSpeed](http://www.numion.com/YourSpeed3/Run.php?QuickStart=MaxSpeed) max-speed utility to health-check the connection and verify its data transfer speed. (Use modem INIT strings to set all modems to a common speed if you experience difficulties matching modem connector speeds. Propel generally forces modems to 33 kbps as that speed is usually possible and reasonable for dial-up connections.)
5. Configure the acceleration level you plan to test (Low, Medium, High, etc.)

### Page Visit Timing Samples

1. Clear the browser's temporary file cache (PropelRT has a Clear Browser Cache Button, or you can use Microsoft IE's Internet Options command, click on the Delete Files button under Accelerator Temporary Files.)
2. Clear the Accelerator's client-side cache files. (Go to the Propel icon in the task bar and Click: Options/temporary files/Delete Files.)

3. Launch PropelRT and load the URL list you have created. (See Step 6)
4. Select “2 Times” in the PropelRT dialog. This will result in collecting data for initial and repeat visits in one run.
5. It is ideal if you perform these comparative tests in parallel, so that all configurations load at the same time and experience the same variable factors. In that case, only one control (unaccelerated) machine is required for each physical dial-up network.
6. Save your test data. PropelRT collects data in a spreadsheet-like matrix. We prefer to save each run’s data (1 initial and 1 repeat visit) in a unique file with test conditions coded in the file name. An alternative is to skip Step 3 on subsequent runs and collect all data from one machine in a single file.
7. Repeat Steps 1-6 as required to gather your planned number of timing samples. We recommend at least six initial and size repeat visits to each URL in the test suite.

### Page Visit Results Summary

This is an example of how you might summarize your test results to compare non-accelerated control to Propel Accelerator at different settings:

	NON-ACCELERATED 48K CONNECTION	MIDDLE SETTING PROPEL LEVEL 3	MAXIMUM SETTING PROPEL LEVEL 5
Mainstream Popular URLs			
Average Page Load (sec)	43.51	17.28	14.47
X-Factor	1.0x	2.52x	3.01x
Graphics Heavy URLs			
Average Page Load (sec)	44.68	10.07	5.78
X-Factor	1.0x	4.43x	7.74x
Ad Heavy URLs			
Average Page Load (sec)	59.51	22.75	21.96
X-Factor	1.0x	2.62x	2.71x
% of Pages Loaded			
Less Than - 10 sec	8%	30%	46%
Less Than - 15 sec	11%	55%	64%
Less Than - 20 sec	19%	73%	82%

**Analysis of the Information Presented**

1. Separate data logically into initial visit and repeat visit for each configuration. The analysis process must be repeated for each visit type.
2. Collect all initial page runs for accelerated and non-accelerated runs.
3. Identify any URL that have fewer than five observations and delete these from all test data. Also delete BUMPER URLs.
4. Gather all data for each accelerator/control configuration in one worksheet.
5. Compute average time for each URL for each machine with error data ignored (sample spreadsheet).
6. X-factor can be computed by dividing the control system average for each URL by the corresponding system average. (See sample graph titled *Initial Page Load Acceleration Factor Graph*.)
7. The lift chart can be created by sorting page load times for each configuration in ascending sequence. Then use spreadsheet formulas, etc. to count the number of URLs completed within each time arrival. Graph the results for all of the same visit types (initial and repeat) on a single graph for ease of comparison. (See sample graph under section called *Lift Factor Analysis*.)

Here is a sample spreadsheet that represents how you might show your data. If you are viewing an online document, the actual Excel spreadsheet can be examined if you double-click the chart. If you are viewing this as a pdf file, please ask your salesperson for the Excel workbook.

ACCELERATED				NON-ACCELERATED				
Page Load Time				Page Load Time				
Test Case	Samples (#)	Average (Sec)	Total (Sec)	Samples (#)	Average (Sec)	Total (Sec)	Overall X-Factor	Average URL X-Factor
Initial Page Loads				Initial Page Loads				
Level 4	224	18.73	4196	224	45.17	10119	2.4x	2.9x
Level 5	192	15.61	2998	192	45.63	8762	2.9x	3.9x
Repeated Page Loads				Repeated Page Loads				
Level 4	224	4.79	1073	222	13.4	2975	2.8x	3.4x
Level 5	192	4.40	845	192	13.52	2596	3.1x	3.7x

**About Next Day Page Visits**

If you wish to test the loading times for Next Day page visits, please request our comprehensive Performance Benchmarking White Paper from your sales representative. We have omitted Next Day page visits from this guide because of the complexity of the process.

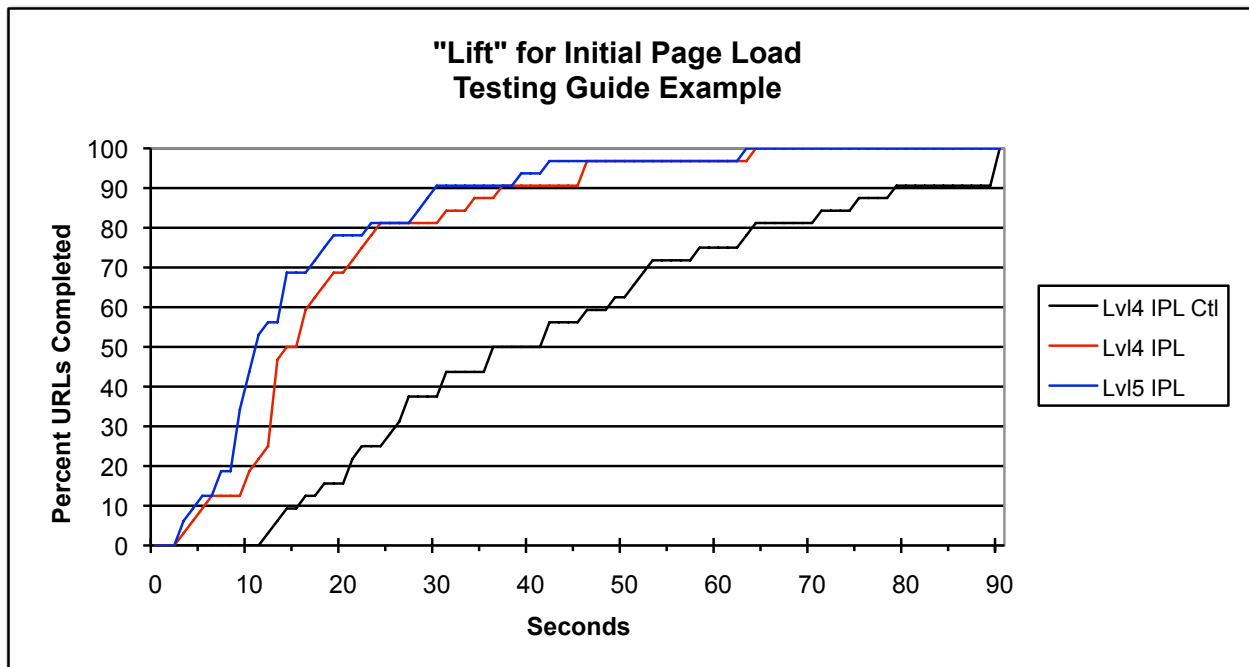
## Comparing and Analyzing Test Results

### Initial Page Load Acceleration Factor Graph

This graph compares the acceleration X-factor for each URL as measured at *Very High Acceleration* (Lvl 4) and *Maximum Acceleration* (Lvl 5). As you can see, the higher acceleration level has a differing impact depending on the characteristics of the page load.

### Lift Factor Analysis

A more effective method consists of sorting pages by the amount of time it takes to load, and the percentage of pages loaded within set times. This offers a real-world view of how quickly a page appears to the end user and provides an instantly recognizable profile of the benefits of acceleration. By plotting the data in graph from the table, it's easy to see that Propel clearly exhibits faster loading times than the non-accelerated connections. Propel uses the term "lift" for this increased acceleration.



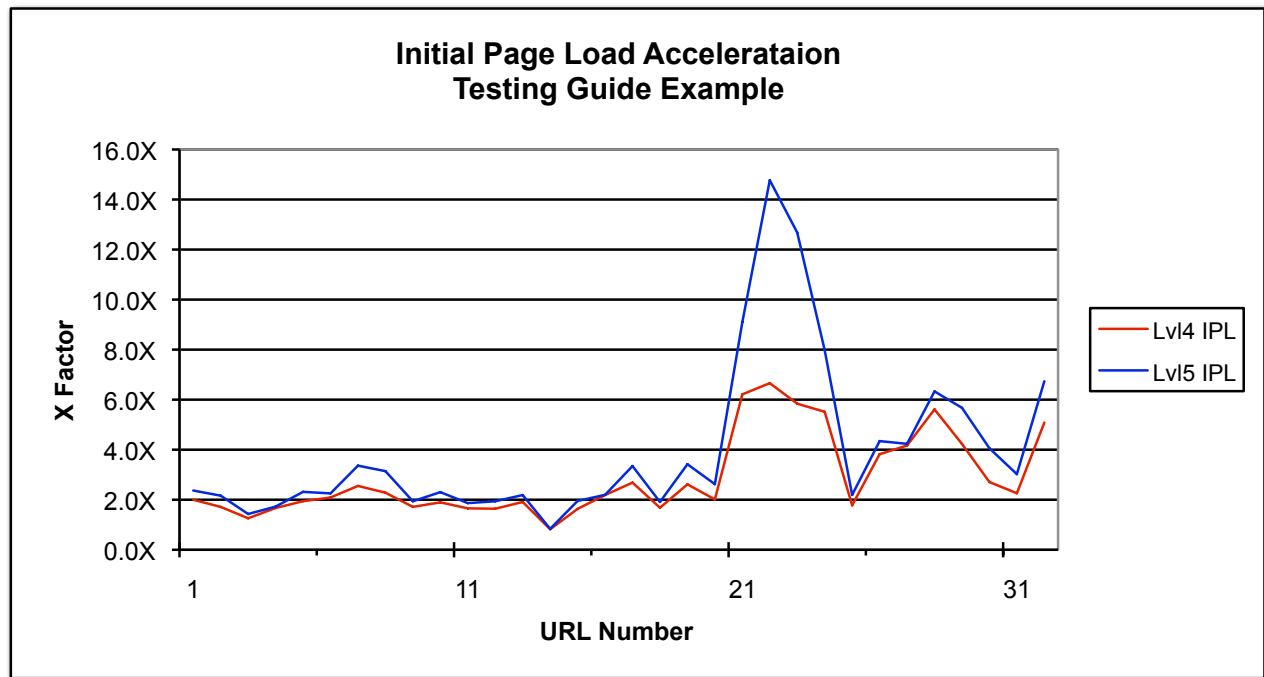
### Real World Comparison Summary

The graphs throughout this document show examples similar to real-world acceleration benchmarking test results performed by Propel. Compare your results to these tables to see how closely they correlate with the benchmark results. If there are discrepancies, please contact Propel sales or support so that we can assist in finding the cause of these differences. These graphs clearly show why non-accelerated users are frustrated by slow performance.

### Other Analytical Metrics - Time-to-Text

Some web acceleration vendors base their load timings on T2T (time-to-text), the time it takes the browser to display the non-graphic text on a page. As the text usually loads faster and more completely than the graphics, these published times can be quite misleading. Propel believes that the user experience is best represented by completion of the page load as it is the rare user who would not want at least partial graphical content before moving to the next page. We encourage all

ISP partners to perform tests and gather their own results rather than rely on vendors' claims. If you would like to pursue your evaluation based on an alternative to the methodology described in this paper, please ask your sales representatives for technical assistance designing your methodology. Based on our experiences, we'd like to help you avoid measurement



pitfalls.

### The Fallacy of Measuring Image Quality with MSE

Some competing vendors use a metric known as Mean Square Error (MSE) to determine equivalent graphic quality client settings. However, it has been well understood for decades in academic and professional communities which analyze and measure human-perceived image quality that MSE returns severely misleading results. The following graphics demonstrate this fallacy. According to MSE scores, the image on the right is supposed to have a better image quality than the one on the left.



Please refer to the paper at

[www.cns.nyu.edu/~zwang/files/papers/icassp02a.pdf](http://www.cns.nyu.edu/~zwang/files/papers/icassp02a.pdf) for more information on MSE.

## Quick Reference Summary

### Testing Do's

1. Test with a pair of matched machines and modems. Configure one machine with acceleration and one without acceleration.
2. Control connection speed and use the same dial-up access numbers.
3. Control runs to minimize the impact of the server load and differing Internet traffic levels.
4. Run at least six test samples of each configuration to create more accurate averages.
5. Test with the default installed acceleration level.
6. Create a list of URLs that accurately reflect your customers' browsing habits. Be sure to evaluate navigating levels within websites by constructing sequences of URLs representing a user visiting a home page followed by other links within the site.
7. Test different types of content acceleration: graphics, text, etc. Use typical website home pages.
8. Clear both browser cache and accelerator cache before starting each test run when taking initial visit test samples.
9. Test with Microsoft Internet Explorer version 5.5 or higher.
10. Use PropelRT page load timing application explained in this document.
11. Use a web-based bandwidth meter (e.g. [www.numion.com](http://www.numion.com) max speed utility) to verify initial connection speeds to confirm all tests are at equivalent connection speeds.

### Testing Don'ts

1. Don't use a bandwidth meter for testing overall accelerator performance.
2. Don't run a different test configuration before completing all times for the first configuration i.e.; if you are planning to test medium graphics quality acceleration settings, run and finish all these tests before starting to test for a different quality setting. (Your data will be higher quality if your measurements minimize changes in the Web content).  
**Exception:** If you are testing with multiple systems, a single control (unaccelerated) machine can be combined with concurrent accelerated runs. Each accelerated run would be at a different quality.
3. Don't run any other programs that might use network bandwidth while testing.
4. Don't clear caches before performing repeat visit tests. This will invalidate your results and create a new set of initial visit samples.
5. Don't use test machines between initial visits and repeat visits. This will affect the caches of the machines and deliver inaccurate results.
6. Don't rely on Mean Square Error as a metric for defining quality of accelerated graphic images; use your own visual comparisons to compare "apples to apples".
7. Don't compare acceleration of single pages with one content type. Such pages do not accurately represent the end user experience.
8. Don't compare raw acceleration results from two different dial-up (or other connection type) infrastructure. This will invalidate your comparison by adding an additional uncontrolled variable.

### **Testing Procedures**

1. Set-up all test machines with the appropriate software, timing tools and accelerator client installed.
2. Establish a dial-up connection from each machine, using the same access number.
3. Use a bandwidth meter to health-check the line and verify its connection rate.
4. Use modem INIT strings to set all modems to a common speed if you experience difficulties matching modem connection speeds.
5. Establish the acceleration settings you plan to test (Low, Medium, High, etc.)

### **Page Visit Timing Samples**

1. Clear the browser's temporary file cache (PropelRT has a Clear Browser Cache Button, or you can use Microsoft IE's Internet Options command and click on the Delete Files button under Temporary Internet Files.)
2. Clear the Accelerator's client-side cache files. (Go to the Propel icon in the message area on the task bar and Click: Options/temporary files/Delete Files.)
3. Launch PropelRT and load the URL list you have created. Set PropelRT to run two times to collect initial visit and repeat visit data in a single PropelRT run.
4. Click: Play. PropelRT will then record the loading times.
5. Save the test data collected by PropelRT.
6. It is recommended that you perform these comparative tests in parallel, so that all configurations load at the same time and experience similar variable factors.
7. Repeat Steps 1-5 multiple times to gather more timing samples. We recommend at least six visit types (initial and repeat) to each URL in the test suite.

## Sample Bumper Pages

The purpose of a BUMPER page is to insure that timing measurements are isolated from the overhead associated with internal accelerator connection management, browser cache clearing issues, etc. We suggest you included pages such as those below on a test server accessible to the Internet. Start your URL list with *first.html* and end it with *last.html*.

**Table 1: Sample first.html**

```
<HTML><HEAD><TITLE>Empty Page To Force a connection before Perf  
Timings</TITLE>  
<BODY bgColor=#d8ffd8>  
<P><P><P><P><P><P>  
<CENTER><H2>Empty Page -- Start of Run</H2></CENTER>  
</BODY></HTML>
```

**Table 2: Sample last.html**

```
<HTML><HEAD><TITLE>Empty Page To Reset Perf Timings</TITLE>  
<BODY bgColor=#ffc8c8>  
<P><P><P><P><P><P>  
<CENTER><H2>Empty Page -- End of Run</H2></CENTER>  
</BODY></HTML>
```