Service Level Agreement

E-Commerce Website

# Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Reason** |
| 1.0 | 11/03/2009 | Jesse Squire | Initial draft |

Table of Contents

[Revision History 2](#_Toc309398552)

[Executive Summary 4](#_Toc309398553)

[Overview 4](#_Toc309398554)

[Scope of Agreement 4](#_Toc309398555)

[Audience Segmentation 5](#_Toc309398556)

[Approach: Consumer Segment 5](#_Toc309398557)

[Approach: In-Store Segment 5](#_Toc309398558)

[Consumer Segment Execution 6](#_Toc309398559)

[Performance Monitoring Execution 6](#_Toc309398560)

[Page Segment Identification 6](#_Toc309398561)

[Performance Metrics 7](#_Toc309398562)

[Reporting 8](#_Toc309398563)

[In-Store Segment Execution 8](#_Toc309398564)

[Performance Monitoring Execution 8](#_Toc309398565)

[Page Segment Identification 9](#_Toc309398566)

[Performance Metrics 10](#_Toc309398567)

[Reporting 11](#_Toc309398568)

[Remediation Responsibilities 11](#_Toc309398569)

[Overview 11](#_Toc309398570)

[Standard Monitoring and Response 12](#_Toc309398571)

[Production Deployment Monitoring and Response 12](#_Toc309398572)

# Executive Summary

This document defines a Service Level Agreement (SLA) between [DEVELOPER] and [SITE OWNER] with respect to the performance aspects of the e-commerce website. It outlines the metrics, goals, and monitoring approach to be taken to ensure that the site is performing optimally.

Monitoring will be performed on key pages of the website from multiple locations spread across the country. The response time of each webpage is collected and the average across all locations is measured against the expected performance.

Service levels are differentiated by the type of audience, page, and monitoring being considered. The defined levels are:

|  |  |  |  |
| --- | --- | --- | --- |
| **Audience** | **Page Segment** | **Monitoring Scan Type** | **Expected Performance** |
| Consumer | Landing | HTML Only | < 1 second |
| Consumer | Landing | Full Object Set | < 2.5 seconds |
| Consumer | Browse and Search | HTML Only | < 1.5 seconds |
| Consumer | Browse and Search | Full Object Set | < 3 seconds |
| Consumer | Product Detail | HTML Only | < 1.5 seconds |
| Consumer | Product Detail | Full Object Set | < 3 seconds |
| In-Store | Landing | HTML Only | < 1 second |
| In-Store | Landing | Full Object Set | < 2.5 seconds |
| In-Store | Browse and Search | HTML Only | < 1 seconds |
| In-Store | Browse and Search | Full Object Set | < 4.5 seconds |
| In-Store | Product Detail | HTML Only | < 2.5 seconds |
| In-Store | Product Detail | Full Object Set | < 5 seconds |
| In-Store | Transaction | HTML Only | < 2 seconds |
| In-Store | Transaction | Full Object Set | < 5 seconds |

Service level deviations will be addressed via the current support ticket workflow. For each new production deployment, a one week window for escalated response will be observed. During this time, performance will be monitored daily and measured against the SLA metrics and thirty days of historical data. In the event that performance is not compliant with the SLA or has degraded by more than thirty percent, a priority ticket will be created and expedited. Scheduling of the priority ticket will be the collaborative responsibility of the [DEVELOPER] and [SITE OWNER] leadership teams.

# Overview

## Scope of Agreement

This SLA pertains to the public-facing [SITE OWNER] website at the address [www.ecomsite.com](http://www.google.com/), defining sets of unique goals and requirements with respect to the identified target audience segments. This SLA should be considered a living document and may be revised by mutual agreement between [DEVELOPER] and [SITE OWNER].

By definition, this document outlines an agreement, not a contract. It is understood that while diligent efforts will be made to fulfill the performance goals of this SLA by the stakeholders, there may be exceptional circumstances when the SLA will not be realized. Examples of these circumstances are network outages by service providers, hardware failures, and site maintenance windows.

## Audience Segmentation

The e-commerce website was designed to accommodate two primary audiences with different usage patterns, performance expectations, and access methods. In order to ensure the best possible user experience, the target audience has been segmented with each having a unique set of goals and requirements.

The first segment is consumers, recognized as individuals who browse and/or purchase goods from the website, accessing it via web-browser from a general purpose internet service provider. The remaining segment is in-store sales representatives, recognized as individuals who browse and/or place orders on a customer’s behalf, accessing the website from within a store using a thin client appliance.

## Approach: Consumer Segment

In order to verify performance goals, an external service provider will be used to measure and record response times for a set of page segments from disparate geographic locations and using a varied set of network service providers.

The pages chosen for monitoring will each be part of a central hub within the site with a high traffic volume and will be representative in terms of size and complexity to pages of a similar type. Where possible, a complex candidate and simple candidate will be chosen for a page type in order to ensure coverage across edge cases.

## Approach: In-Store Segment

In order to observe in-store performance, a custom monitoring application will be utilized within a set of [SITE OWNER] stores to measure and record response times for a set of page segments. The selected stores will be from different geographic regions of the United States in order to illustrate performance expected performance characteristics for stores throughout the country.

The pages chosen for monitoring will each be part of a central hub within the site or will belong to part of the core ordering workflow. Pages will be representative in terms of size and complexity to pages of a similar type. Where possible, a complex candidate and simple candidate will be chosen for a page type in order to ensure coverage across edge cases.

# Consumer Segment Execution

## Performance Monitoring Execution

Monitoring will be performed by Compuware’s Gomez Networks service. Each target webpage will be monitored from three different geographic regions of the United States, and will each utilize a unique network service backbone in order to ensure accurate measurement of performance consistency for consumers in different regions. The following nodes will be implemented to perform the monitoring:

* New York (Sprint)
* Chicago (Qwest)
* Los Angeles (Verizon)

Monitoring of each target will be applied using two distinct scans. The first scan, referred to as “Full Object Set” going forward, is indicative of an end-user’s experience when visiting the website for the first time when none of the site’s assets are available in the browser cache. This scan calculates the total response time for DNS resolution, establishing a connection to the webserver, download of the HTML content of the page, and download of all assets linked from that page including images, script files, and style sheets. This test is performed continuously once every five minutes for each of the target pages.

The second scan, referred to as “HTML Only” going forward, is indicative of an end-user’s experience when visiting the website for a repeat visit, when the site’s assets are available in the browser cache. This scan calculates the total response time for DNS resolution, establishing a connection to the webserver, and download of the HTML content of the page. This test is performed continuously once every five minutes for each of the target pages.

## Page Segment Identification

To ensure that performance metrics are met for an experience that is characteristic of an end-user, pages have been chosen from three different segments of the site. The first segment is comprised of landing pages, for which the content is typically driven by both product and merchandising data. These pages serve as a central hub from which a user would branch into locating products or viewing of details for a specific product.

The next segment is comprised of browse and search pages, for which content is typically driven by search engine results as well as product and merchandising data. These pages provide the user with operations related to location of products, offering the ability to drill into the details for a specific product.

The final segment is comprised of product detail pages, for which the content is typically driven by an expanded set of product data from multiple sources including heavy use of imagery. These pages provide the user with a comprehensive view of a product including its full set of details, store inventory, user reviews, and what other products may compliment or replace it.

The following pages have been selected to represent these segments for monitoring. Wherever possible, two candidates were chosen; one representing the simplest set of content, the other representing the most complex. The pages to be monitored, by segment, are –

**Landing Pages**

* Home

[www.ecomsite.com](http://www.google.com/)

**Browse and Search Pages**

* Widget Category

[http://www.ecomsite.com/Widgets](http://www.google.com/)

* Snowshovel Category

[http://www.ecomsite.com/Snowshovels](http://www.google.com/)

**Product Detail Pages**

* Confabulator Widget 3000

[http://www.ecomsite/Widgets/confabulator-widget-3000-106739](http://www.google.com/)

* Snowmaster Plastic Shovel

[http://www.ecomsite/Showshovels/snowmaster-plastic-889643](http://www.google.com/)

## Performance Metrics

The following metrics will be applied as goals for each of the page segment and monitoring scan type combinations. Scans in which these metrics are not met will be considered failures and may imply an anomaly in the scan itself, an exceptional circumstance being encountered, or that an issue exists with the website. Repeated failures indicate that an action is needed on the part of [DEVELOPER] to diagnose and correct the cause of these performance issues, ensuring that the performance goals continue to be met.

The metrics are listed below. A successful scan is defined as one in which the scan is completed within the defined time goal. The metrics are based on the average of all monitoring nodes for a given snapshot in time.

**Full Object Set Scans**

* Landing Pages - less than two and a half seconds
* Browse and Search Pages - less than three and a half seconds
* Product Detail Pages - less than three seconds

**HTML Only Scans**

* Landing Pages - less than one second
* Browse and Search Pages - less than one and a half seconds
* Product Detail Pages - less than one and a half seconds

## Reporting

In order to observe compliance with the performance service levels outlined in this document, monitoring reports will be provided on a weekly basis. These reports will be comprised of month-to-date performance data in a summary format with each test and scan type represented as an individual line item. Reporting data will be normalized, with any outlying data anomalies not considered.

Reports will be generated through the Gomez Networks monitoring service and delivered via email to a to-be-specified email alias.

# In-Store Segment Execution

## Performance Monitoring Execution

Monitoring will be performed by a performance monitoring application running as a service on the thin client machines in select retail stores. The application will execute its performance measurements in the background, causing no disruption to the sales representative during normal usage of the machine. Upon completion of the test suite, the results will be electronically transmitted to an online service for collection and aggregation of the metrics. To ensure that performance measures are an accurate representation across retail stores, monitoring will be performed from fourdifferent geographic regions. The following stores have been selected to perform the monitoring:

* Times Square Location, NY, NY
* Goose Island Location, Chicago, IL
* Santa Fe Location, Santa Fe, NM
* Rose Quarter Location, Portland, OR

Monitoring of each target will be applied using two distinct scans. The first scan, referred to as “Full Object Set” going forward, is indicative of a sales representative’s experience when visiting the website for the first time when none of the site’s assets are available in the browser cache. This scan calculates the total response time for DNS resolution, establishing a connection to the webserver, sales representative authentication, SPO session creation, and download of the HTML content of the page, and download of all assets linked from that page including images, script files, and style sheets. This test is performed continuously once every ten minutes for each of the target pages.

The second scan, referred to as “HTML Only” going forward, is indicative of a sales representative’s experience when visiting the website for a repeat visit, when the site’s assets are available in the browser cache. This scan calculates the total response time for DNS resolution, establishing a connection to the webserver, sales representative authentication, thin client session creation, and download of the HTML content of the page. This test is performed continuously once every ten minutes for each of the target pages.

## Page Segment Identification

To ensure that performance metrics are met for an experience characteristic of a common workflow for a sales representative, pages have been chosen from four different segments of the site. The first segment is comprised of landing pages, for which the content is typically driven by both product and merchandising data. These pages serve as a central hub from which a user would branch into locating products or viewing of details for a specific product.

The next segment is comprised of browse and search pages, for which content is typically driven by search engine results as well as product and merchandising data. These pages provide the user with operations related to location of products, offering the ability to drill into the details for a specific product.

The third segment is comprised of product detail pages, for which the content is typically driven by an expanded set of product data from multiple sources including heavy use of imagery. These pages provide the user with a comprehensive view of a product including its full set of details, store inventory, user reviews, and what other products may compliment or replace it.

The final segment consists of pages which service a sales transaction. These pages provide the user with the means to add products to a shopping cart, review the items for an order, and to complete the ordering workflow.

The following pages have been selected to represent these segments for monitoring. Wherever possible, two candidates were chosen; one representing the simplest set of content, the other representing the most complex. The pages to be monitored, by segment, are:

**Landing Pages**

* Home

[www.ecomsite.com](http://www.google.com/)

**Browse and Search Pages**

* Free Text Search

[https://www.ecomsite.com/Search?internal=1&q=blue+widget](http://www.google.com/)

* Widget Category

[http://www.ecomsite.com/Widgets](http://www.google.com/)

* Snowshovel Category

[http://www.ecomsite.com/Snowshovels](http://www.google.com/)

**Product Detail Pages**

* Confabulator Widget 3000

[http://www.ecomsite/Widgets/confabulator-widget-3000-106739](http://www.google.com/)

* Snowmaster Plastic Shovel

[http://www.ecomsite/Showshovels/snowmaster-plastic-889643](http://www.google.com/)

**Sales Transaction Pages**

* Selection and Add To Cart

[https://www.ecomsite.com/Account/Cart.aspx?qty=1&itemno=1385902](http://www.google.com/)

* View Cart

[https://www.ecomsite.com/Account/Cart.aspx?view=instore](http://www.google.com/)

* Order Checkout

[https://www.ecomsite.com/Account/Checkout.aspx](http://www.google.com/)

## Performance Metrics

The following metrics will be applied as goals for each of the page segment and monitoring scan type combinations. Scans in which these metrics are not met will be considered failures and may imply an anomaly in the scan itself, an exceptional circumstance being encountered, or that an issue exists with the website. Repeated failures indicate that an action is needed on the part of [DEVELOPER] to diagnose and correct the cause of these performance issues, ensuring that the performance goals continue to be met.

The metrics are listed below. A successful scan is defined as one in which the scan is completed within the defined time goal. The metrics are based on the average of all monitoring nodes for a given snapshot in time.

**Full Object Set Scans**

* Landing Pages - less than two and a half seconds
* Browse and Search Pages - less than four and a half seconds
* Product Detail Pages - less than five seconds
* Sales Transaction Pages – less than five seconds

**HTML Only Scans**

* Landing Pages - less than one second
* Browse and Search Pages - less than one second
* Product Detail Pages - less than two and a half seconds
* Sales Transaction Pages – less than two seconds

## Reporting

In order to observe compliance with the performance service levels outlined in this document, monitoring reports will be available on-demand from within the standard administration portal. Two reports will be available, one at the summary level and one at the detail level. Data at the summary level will allow drilling-down on individual line items for inspection on a more granular level.

Each report will allow for multiple criteria to be specified in order to best conform the results to the desired target time window, store, and monitoring scan dates. Reporting data will be normalized, with any outlying data anomalies not considered

# Remediation Responsibilities

## Overview

In the event performance does not comply with the defined service levels, remediation efforts will be needed to identify and correct the underlying cause. In order to ensure that performance issues are addressed, scheduled, and prioritized appropriately, monitoring and response procedures will be defined according to two scenarios.

The most likely time for critical performance issues to be discovered is shortly after a new release of the website is deployed into the production environment. While every reasonable attempt is made during the development and testing phases of a release to ensure that performance is properly tuned, the possibility exists that the differences in hosting environment and overall load could expose bottlenecks that could not be detected earlier. In order to ensure that any performance issues impacting a new release are discovered and addressed as quickly as possible, a period of enhanced monitoring and response will be observed.

Most performance issues that impact stable codebases are degenerative in nature. Typically, they begin to present as minor performance erosions and grow at a slow and steady pace over time. In the majority of cases, these types of issues can be corrected as part of the normal development process, allowing them to be fully vetted during testing.

## Standard Monitoring and Response

The majority of monitoring will occur by observation of the weekly reports and will be the responsibility of the [SITE OWNER] team. If additional monitoring attention is needed, reports may be generated on demand through the appropriate monitoring system.

In the event that monitoring indicates that the SLA is not being met, a work item ticket will be opened. The ticket will be scheduled and addressed in accordance with the existing process, including the need for a production patch, inclusion in a target release, and its priority with respect to ongoing work items.

## Production Deployment Monitoring and Response

For a period of one week after the deployment of a major release or patch into the production environment, a window of enhanced monitoring and response will be in effect. It is expected that any critical performance issues resulting from the release will be discovered in this period of time. During this window, the development team will monitor performance on a daily basis, measuring it against the SLA metrics and thirty days of historical data.

In the event that performance is not compliant with the SLA or that a performance degradation of greater than thirty percent is observed, a priority work item ticket will be opened and escalated in the work item queue. It is the responsibility of the leadership teams from [DEVELOPER] and [SITE OWNER] to collaborate on reprioritization of other outstanding work items to ensure that the performance issue can be addressed in an expedited fashion. Should a patch be deployed to address the performance issue, it will be considered a release and therefore subject to the one week enhanced window.

If a performance degradation of less than thirty percent is observed and performance is still complaint with SLA metrics, a work item ticket of normal priority will be opened and scheduled in accordance with the existing process.

Once the week has elapsed and the production website proven to be compliant with the SLA, monitoring and response actions will adhere to the standard procedures previously detailed.