



keynote™

The Mobile & Internet Performance Authority™

Solution Brief

Measuring and Monitoring Web 2.0 Applications

Measuring & Monitoring Web 2.0 Applications with Keynote

Executive Summary

The Internet and the Web continue to evolve to deliver new customer experiences and increased application utility. The label “Web 2.0,” while imprecise, signifies the newest and best examples of this evolutionary process.

Many organizations are now adopting these Web 2.0 technologies and design methods to enable the creation of richer and more responsive interactions. But to be effective, the resulting applications must also be significantly more complex than traditional Web sites, complicating performance management and imposing new requirements on performance measurement tools. This paper reviews the challenges posed by the evolution to Web 2.0 and explains how Keynote is addressing them in its suite of Test and Measurement solutions.

Table of Contents

1. The Web Application Landscape	3
1.1. The Trend to Rich Internet Applications	3
2. The Importance of Web Performance Management	4
2.1. How Effective Is Your Web Application?	4
2.2. Common Goals, Divided Responsibilities	5
2.3. Measurement Challenges of Web 2.0	6
2.4. Getting the Right Data	6
3. Keynote Web Performance Management Solutions	7
3.1. Measuring the Diverse Customer Community	7
3.2. Measuring Upload and Download Performance	8
3.3. Measuring a Richer Variety of Media	9
3.4. Measuring Logical Web Pages	9
3.5. Reporting More Than Web Page Metrics	11
Conclusion	12
References / Resources	13

Characteristics of Web 2.0 Applications

The network as a platform.

Applications are delivered entirely through a browser, or even a microbrowser that resides on a handheld device such as a PDA or mobile phone.

A collaborative environment

Whereas a traditional Web site publishes content, many Web 2 sites act as a hub for sharing content. Site users may supply, even control, the published content.

Social networking. Many Web 2.0 applications include features that connect people with common interests.

Hybrid applications

(**mashups**). Standard Web protocols and interfaces allow third-party applications to integrate both data and functions from existing ones.

Rich media interfaces.

Compared to the original concept of the Web as a collection of HTML hypertext documents, user interfaces are now more dynamic and more interactive, incorporating a variety of media including audio and video streams and voice interactions.

1. The Web Application Landscape

The Internet and the Web have become the primary vehicle for business communications, evolving to subsume and replace older technologies. As software technologies exploit steady advances in the Internet hardware platform, the Web continues to evolve to deliver new user experiences and increased application utility.

In recent years, the term “Web 2.0” has emerged as an inclusive label signifying both the current state of this technological evolution and the newest and best examples of its application.ⁱ

Web 2.0 Technologies and Applications

The exact definition of Web 2.0 is debatable, but all Web 2.0 applications exhibit one or more of the features shown in the sidebar “Characteristics of Web 2.0 Applications.”ⁱⁱ [WEB1]

Highlighting its interactive nature, Web 2.0 has been described by Richard Monson-Haefel of the Burton Group as *an architecture of participation built on the World Wide Web*.ⁱⁱⁱ [HAE1]

1.1. The Web Application Landscape

The most advanced example of the Web becoming a platform is the rich Internet application (RIA). The RIA architecture reflects the gradual transition of Web applications from the simple thin-client Web browser to a richer distributed-function paradigm that behaves more like the desktop in a client/server model.

These richer user experiences are being implemented using “client-engine” technologies like Flash, Ajax, and Java, using standard Internet and Web protocols. Figure 1 illustrates the relationships among these concepts and technologies and the larger Web 2.0 landscape.

Measuring RIA Performance

This architecture complicates performance measurement, whose goal is to understand the customer’s experience. In an RIA, the time to complete a Web page download may no longer correspond to something a user perceives as important, because (for example) the client engine may be prefetching some of the downloaded content for future use. Standard tools that measure the time for Web page downloads to complete can record misleading data for RIAs. In some cases, the client engine may not even invoke any standard Web page downloads once the application is launched, using asynchronous background communications for all server requests.

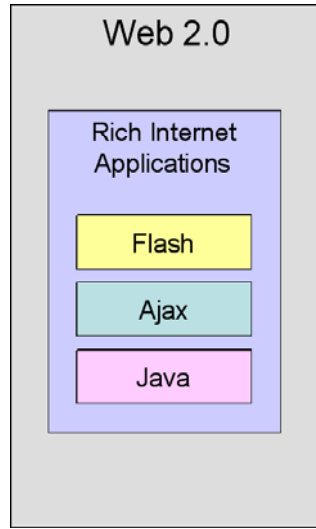


Figure 1. The Web 2.0 and RIA landscape

To implement RIAs successfully, enterprises must re-evaluate their approach to performance management. Instead of relying on the definition of physical Web pages to drive the subdivision of application response times, RIA developers or tool users must break the application into *logical pages*. Measurement tools must recognize meaningful application *milestones* or *markers* that signal logical boundaries of interest for reporting, and subdivide the application's response time accordingly. For more detailed discussion of RIA architecture, see the paper *Rich Internet Applications: Design, Measurement, and Management Challenges*.^{iv}

2. The Importance of Web Performance Management

What does Web application *performance* mean to you? Most business executives would evaluate the success of a Web application by looking at business performance metrics such as revenue, costs, and customer satisfaction. Because an application may be created to serve customers, partners, members of an organization, or employees, the relative importance of those metrics may vary. But no matter what its audience or purpose may be, the success of every Web site or application can (and should) be judged by a common criterion—how effectively does it achieve its goals?

2.1. How Effective Is Your Web Application?

For any Web application, *effectiveness* means simply fulfilling the planned design and delivery objectives, delivering online experiences that lead to satisfied customers, and so meeting the intended business performance goals.

If the mission of a Web site is “*turning triers into buyers*,” the two keys to site effectiveness can be described as “*first attracting enough triers, then turning enough triers into buyers*.”

In this paper, our focus is on the second key—a site’s *conversion rate* — which depends on delivering an effective service to those customers who do enter the site. To achieve this, a site or application must meet the four distinct needs shown in Figure 2: *availability*, *responsiveness*, *clarity* and *utility*, each of which is essential to its success. The Web 2.0 environment offers new ways to achieve these goals, but also adds new challenges.

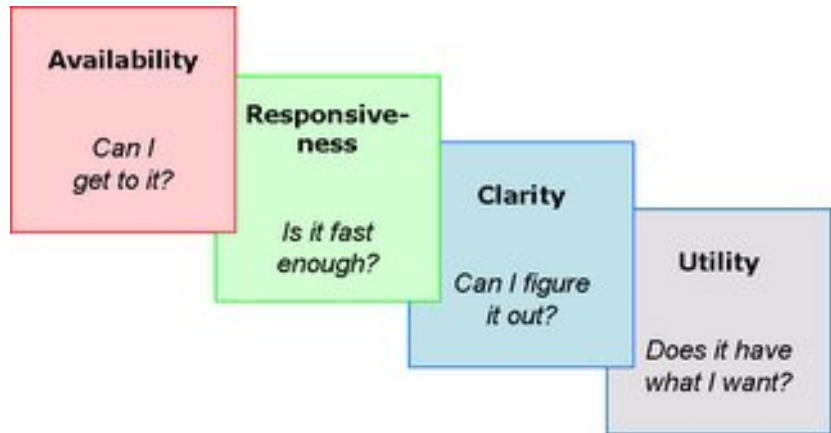


Figure 2. Four dimensions of application effectiveness (RIA WP Figure 5)

2.2. Common Goals, Divided Responsibilities

In many organizations, responsibility for these four aspects of effectiveness is divided among four groups or departments. (See the sidebar “Divide and Conquer.”)

Such an approach allows people to specialize—an essential division of labor, since few have the necessary skills in all four areas. But effectiveness depends on all four aspects, and on their relationships. Application *utility* is tied to user behavior, which is driven by the *clarity* of the site, the design of the user interface, and the *responsiveness* of interactions between client and server components. The design and implementation of those components ultimately also determines application *availability*.

Thus the division of labor and responsibilities introduces the need for more coordination while the application is being developed, more measurement and testing as a precursor to deployment, and more monitoring and tracking after deployment to ensure its success.

Divide and Conquer

Business experts provide the content or specify the behaviors that are the site’s purpose (Utility).

Design and usability professionals specialize in making it easy for customers to navigate the site (Clarity).

Site developers build the site in ways that determine download speed (Responsiveness).

IT staff manage the systems that keep the site up and running (Availability) and maintain a consistent level of performance (Responsiveness).

Measuring Web Effectiveness

Web analytics (WA) tools focus on utility. Business analysts use these tools to count and segment site visitors by the frequency of different actions, providing data to answer questions about behavior and conversion rates.

User experience (UX) tools measure clarity. Site designers use UX tools for behavioral and attitudinal analysis, user experience testing, and online surveys.

Measurement and testing tools measure responsiveness and availability. Web developers, testers, and performance analysts use these tools to obtain data about application performance during deployment and under production conditions.

Performance monitoring and reporting tools track availability and consistency. IT and operations staff use these tools to monitor and manage service levels after applications are deployed.

2.3. Measurement Challenges of Web 2.0

Just as development skills and responsibilities are divided, so too are the capabilities of measurement tools. (See the sidebar “Measuring Web Effectiveness.”)

The characteristics of Web 2.0 applications highlighted earlier—the network as a platform, collaborative environment, social networking, mashups, and rich media interfaces—create several additional challenges for all measurement tools. In increasing degree of complexity, these are:

- Measuring an increasingly diverse customer community
- Measuring the quality of collaborative interactions, in which customers both upload and download information.
- Measuring applications that incorporate a much richer variety of media and interfaces.
- Measuring application phases that no longer correspond to one (or more) HTML pages
- Reporting data in ways that provide insight into Web 2.0 applications.

2.4. Getting the Right Data

To sum up the importance of these issues, consider the popular saying coined by Tom DeMarco: *You can't control what you don't measure.*^v Measuring the wrong things, or basing key management decisions on reports that contain incomplete data, is as bad, if not worse, than not measuring at all.

So it's not wise to measure Web 2.0 applications using only the tools and approaches developed for traditional Web sites. Inaccurate data undermines the effectiveness of any program of systematic performance management and causes performance-tuning skills and resources to be applied in ways that are not optimal. It can also lead to unproductive interdepartmental conflicts and disputes over service-level agreements with internal or external service providers when staff question the accuracy of the data, or discover discrepancies in data from different sources.

Five Performance Management Challenges

Setting targets. Before adopting and committing to service-level targets, do you know what level of performance your customers are likely to expect? How can you match, even stay ahead of, the competition?

Testing. As you prepared for the peak season or a sales event, will your Web site and applications scale to meet the projected user demand?

Measurement. How will you measure site and application responsiveness as your customers experience it? Can you measure and trend the online performance and availability for all your applications, and for the newest Internet technologies? Will you know if your customers experience a sudden drop, or even a downward trend, in service levels?

Monitoring. How do you identify, diagnose, and fix problems quickly—before they impact your customers? Do you get alerts if your Web applications are unavailable from anywhere outside the firewall?

Reporting. How will you monitor, quantify, and report on progress against your goals? Do you have a complete view of the Web infrastructure and application performance from major cities around the globe?

3. Keynote Web Performance Management Solutions

To maintain effective and optimal online business performance you need to address the Five Performance Management Challenges presented in the sidebar.

Success in every one of these five performance management activities depends crucially on an organization's ability to gather and report meaningful, timely, and accurate measurement data with the focus on the right metrics. Since a key idea of Web 2.0 is enhancing the user's experience, it is vital to measure actual customer experience proactively.

Keynote is the only vendor offering a complete set of performance monitoring solutions that meet the many new challenges posed by Web 2.0 applications. This section describes how Keynote's products address Web 2.0 performance management requirements.

3.1. Measuring the Diverse Customer Community

On Web 2.0 sites, personalization options allow customers to tailor their experience of a site to their individual preferences, and sites are carefully designed to download and display contents efficiently and successfully in all major browsers. Because customers' experience depends on their Internet connectivity, sites may even adjust their content based on the browser's connection speed. Measurement data must reflect this diversity.

Keynote support for all types of connectivity

Keynote's global agent network reflects this, collecting measurements from a wide range of environments. For Keynote users who want to measure and track their site's upload or download experience for a diverse population of business customers and consumers, Keynote agents can measure sites over fiber, broadband (cable, DSL), dial-up, or 3G cellular networks.

Keynote support for browser diversity

Although statistics show that the vast majority of Web users still use Internet Explorer on Windows, usage of other browsers (Firefox, Safari, Opera, etc.) and other operating system platforms (Mac OS X, Linux) is increasing.^{vi} For customers who wish to measure an application's performance on multiple platforms, Keynote's agents can identify themselves as any desired browser version or release.

In keeping with its commitment to providing an accurate measurement of user experience on the predominant platform, Keynote plans to include Windows Vista OS platform support in the near future.

Keynote support for cached and uncached measurement

Browsers offer a variety of caching options. For example, some users may elect to disable caching for security or other reasons. For any site visited regularly, disabling the browser cache causes additional downloads and longer transaction response times.

Most sites are designed with caching in mind to optimize responsiveness. However, Keynote can measure cached and uncached experiences separately for organizations wanting to understand and track both.

Keynote's test and measurement products allow customers to track the fully cached user experience, and they provide the capability to exclude completely any subset of the cacheable content elements from the reported measurements. Keynote test and measurement services can be also configured to enable or disable caching between transactions.

3.2. Measuring Upload and Download Performance

In traditional Web applications, customers *consume* content, so all performance measurement efforts have focused on download times as the key metric. But as Web 2.0 applications add collaboration and social networking features, customers also supply content. To ensure the quality of a customer's experience, it's therefore necessary to measure and report upload performance as well.

For example, the user of a collaborative application might navigate to a product or member directory, complete a login or authentication dialog, search for a certain subject or interest area, browse the results, select a particular area of interest, proceed to an upload page or dialog, complete a browse dialog to select content to be uploaded from their laptop or desktop, entering some additional descriptive metadata as appropriate for the application, and click the Upload button. Throughout this interaction download activities are minimal, and may respond rapidly. But if, after all this work, the upload stage is painfully slow, or fails altogether, because of congestion at the server end, that customer may be lost forever.

For an organization needing to measure and ensure the quality of its customers' experience when contributing content to a collaborative or social networking application, Keynote's agents can measure both downloads and uploads.

3.3. Measuring a Richer Variety of Media

Keynote support for streamed content

The performance of streamed audio and video must be considered separately from that of other Web application content, because its delivery infrastructure and the Internet protocols it uses are different. Keynote has dedicated streaming agents that can measure the customer experience for applications having streamed content as a significant component.

Keynote support for mobile applications

A growing Web 2.0 trend is for companies to create mobile versions of some Web applications to support the growing handheld device community. Because of the platform and content differences involved, the performance of such a mobile application must be tracked and managed separately from that of Web users. If mobile users are a significant customer segment for any application, Keynote has a network of mobile agents that can measure and track their experience.

3.4. Measuring Logical Web Pages

Traditional Web applications can be viewed as a series of Web pages, each of which requires a distinct download, initiated by an HTTP GET request. This model has been characterized as the *Web page paradigm*. Rich Internet applications deviate from this model, moving some processing to the client and introducing additional asynchronous server communications to support a more responsive user interface. In a hybrid or “mashup” application, some of those communications may be with third-party sites.

To understand Web 2.0 application performance, an analyst may need to measure the performance of transactions (page sequences), single pages, individual page elements, or selected groups or subsets of those page elements that comprise a “logical page” experienced by the customer.

Keynote support for AJAX transaction recording

Keynote measurement agents measure a Web transaction by invoking predefined scripts that represent usage scenarios (“clickpaths”) for a Web application and timing every detail of the resulting server communication for every element downloaded (or uploaded). Keynote customers create measurement scripts using a script recorder application, stepping through the application and capturing the desired navigation path and user actions.

Because scripts for RIAs that use AJAX technologies may trigger downloads that do not affect the response time experienced by customers, a second phase of the recording process shows the user a complete list of page elements, letting them specify the subset of components to be included in “logical pages” and the boundaries between those logical pages.

Complementing this feature, in addition to the former options of waiting until all downloads complete, or waiting for a specified number of seconds after a specified object is downloaded, the customer can specify complex selection logic to determine page completion during measurement. For example, a

user visiting the first page of a site may cause four video files to be downloaded. One is needed immediately; the others will be played later as the user navigates. Script logic can be used to record the user's response time when the first file download completes, even though the Web page download is not complete.

For AJAX-based RIAs, these features let customers measure and report on logical pages that are entirely different from the underlying physical page or pages. To confirm that the measurements produced by the recorded script will properly reflect a customer's intent, a Transaction Tester feature permits interactive recording and testing of AJAX scripts.

Keynote support for logic within transaction scripts

Transaction scripts for Keynote's measurement agents can include JavaScript logic, generating client inputs that include data derived from prior server responses. This capability may be needed (for example) when responding to complex application interactions that ask a variety of questions to authenticate a user during a login step.

Keynote support for transaction tagging

One special use for JavaScript within a script is to inspect server responses and detect application-specific "tags"—distinct content fields inserted as markers. Site developers may choose to embed such tags in application pages solely to pass data to Keynote measurement agents. Keynote agent scripts can also recognize tags that developers have inserted for other purposes, such as to permit tracking by a Web Analytics tool or service.

Regardless of their original purpose, such tags or markers can be recognized by Keynote's transaction recorder and agents and used to define logical page boundaries. The content of a tag (or other HTML statement) can also be saved within the Keynote measurement data, then used to segment measurement results during reporting. Such data might, for example, be the identity of the server that served a page or content element.

Keynote support for Adobe Flash applications

For Web applications created using Adobe's Flash technology, Keynote's transaction recorder and script logic can be used to reproduce observed client-side behavior and inputs. Because Keynote's agents invoke and measure the behavior of a standard IE browser, they render and measure all externally visible Flash application interactions. As Adobe creates and delivers more open Web 2.0 interfaces to its Flash technology, Keynote will exploit those APIs to capture more internal details (such as background server interactions) of RIAs implemented using Flash.

3.5. Reporting More Than Web Page Metrics

Keynote support for analysis and reporting at all levels

Using Keynote's reporting tools (MyKeynote and Keynote Performance Scoreboard) the performance of a Web application can be viewed at many levels of detail. Overall transaction response can be broken down into pages (logical or physical). Below the page level, the download times of content elements and server interactions are available, and these times in turn can be broken down into many distinct response-time components (see Table 1).

The *User Experience Time* component helps customers identify and trend client-side processing delays in Web 2.0 applications. It records client-side processing and rendering time and the time required to download client-side application code or browser plugins.

DNS lookup time
TCP connect time
SSL time
Redirection time
Request time
First-byte download time
Base-page download time
Content download time
Client time
User experience time

Table 1. Components of response time

Keynote support for component-level reporting and alerting

To further support detailed performance tracking and problem isolation, every measured component can be reported and graphed separately for use as input to Keynote's alerting service.

Keynote support for selective exclusion of page content from reports

When reporting at the page level or transaction level, Keynote's tools permit selective exclusion of content elements. A common use is for pages that include embedded advertisements. When the measurement goal is to assess the overall customer experience, reported times can include the ad delivery time. But if the goal is to monitor performance of the host Web server infrastructure, the response time of third-party ad content is irrelevant and can be excluded. This feature can be used for any similar purpose, but is particularly useful when monitoring hybrid Web 2.0 applications or "mashups" that include third-party content.

Custom components for reporting

When reporting transactions, Keynote's tools also permit aggregate response time components to be defined at the transaction, page, or object level (see Figure 3). Keynote can report and generate alerts on these custom aggregates.

Statistical analysis of measurement results

For the performance analyst investigating, tracking, or reporting on any of Keynote's Web measurement data at any level of detail, Keynote's tools offer several statistical analysis options (see Table 2).

Raw measurement data can also be downloaded regularly for inclusion in local analysis using Keynote's data feed and data pulse APIs.

Statistics:

- Average
- Geometric mean
- Standard deviation
- Median
- Inter-percentile range
- Trimming of large or small outliers

Displays:

- Time histories
- Histograms
- Dashboard view

Table 2. Statistical analysis options

Conclusion

When considering the need for Web testing and measurement, it can be tempting to view a Web *application* simply as a collection of Web *pages*. Adopting that simplifying view makes performance measurement and reporting tasks easier. But doing so ignores the dimension of Web performance most affected by Web 2.0, and maybe the most important feature of a Web application—the paths taken by Web users.

For every Web application, a user's ability to navigate smoothly through a relatively small set of business transactions—the "money paths"—is the key to business success. It is only by completing these key transactions that visitors are converted into customers, *triers into buyers*. Therefore, in addition to tracking key performance indicators like *conversion rate*, most organizations consider it essential to identify, track, and measure their key business transactions. By doing so, they can better understand and improve each application's conversion rate—the second key to site effectiveness and therefore the health of the business.

Keynote's test and measurement products are designed to capture accurate and actionable information about transaction performance in the latest Web 2.0 technology environments. They measure and report at every possible level of detail you may need to test, measure, track, investigate, and manage Web application effectiveness.

ⁱ Tim O'Reilly, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software,"

<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.

ⁱⁱ Wikipedia: http://en.wikipedia.org/wiki/Web_2.0.

ⁱⁱⁱ Richard Monson-Haefel, "Web 2.0: A Simple Definition,"

http://rmh.blogs.com/weblog/2007/04/web_20_a_simple.html.

^{iv} Chris Loosely, *Rich Internet Applications: Design, Measurement, and Management Challenges* (San Mateo, Calif.: Keynote Systems, 2006).

<http://www.webperformancematters.com/papers-and-talks/web-20/RIA%20Challenges%20WP%202006.pdf>.

^v Tom DeMarco, *Controlling Software Projects: Management, Measurement and Estimation* (Upper Saddle River, N.J.: Prentice Hall, 1986).

^{vi} See http://en.wikipedia.org/wiki/Usage_surveys_of_web_browsers for a survey of trends in Web browser usage. A similar survey at

http://en.wikipedia.org/wiki/Usage_share_of_operating_systems covers desktop computer operating systems.