



Search (Underlying Metrics) December 2007

Table of Contents

- Purpose and Methodology
- Ranking Methodology
- Key Performance Factors
 - High Speed Response
 - Dial-Up Response
 - Response Time Consistency
 - Geographic Uniformity
 - Load Handling
 - Availability
 - Outage Hours





keynote™

The Mobile & Internet Performance Authority™

Purpose and Methodology

Competitive Research

- In 1996 Keynote published the first Web performance indices which are now considered in the industry to be the 'gold standard' for benchmarking page download and transaction speed and reliability against competitors.
- At the core of these services is the Keynote Transaction Perspective measurement agent, the most accurate and the highest fidelity measurement technology available. The **Keynote Transaction Agent uses the IE 6.0 browser to ensure measurement accuracy.** Just like a user, the agent uses two concurrent threads to fetch page components and allows the use of persistent connections. This browser also handles SSL, JavaScript, Flash, ActiveX, and other interactive languages that can have a significant impact on both performance and availability. Measurement data includes rendering time of all components to ensure that the data gathered represents the complete customer experience.
- Keynote's Service Level Analysts have deep experience in comparative analysis and have performed studies for some of the world's largest financial and retail companies over the last 5 years. **Our team includes statisticians, application tuning experts, and Internet performance analysts.** Together this team has measured and diagnosed the most prominent Web sites providing independent, unbiased assessments.

Purpose

- **Keynote Service Level Rankings:**
 - Provide an objective comparison of a site's service level, focusing on response time and reliability factors (key components of a user's overall experience that are tied to user satisfaction and site success)
 - Answer questions like:
 - **Which sites offer the quickest transactions (fewest steps, fastest pages)?**
 - **Which sites are the most stable, offering consistent levels of service?**
 - **Which sites offer well-designed Web pages?**
 - **Which sites provide the best service for dial-up users? Broadband?**
 - **Which sites are the most reliable (fewest outages)?**
 - Are based on quantifiable metrics that site owners can use to improve rankings and overall customer/site experience
 - Create a baseline for the industry and a way to systematically track that baseline (this study will be done bi-annually)
 - Provide insight into which sites are doing well and identify industry-wide areas for improvement
 - Help site owners set appropriate goals for their specific industry
 - Help industry analysts keep abreast of the ever-changing technology underlying the industry

Sites Profiled

- AOL
- Ask
- Google
- MSN
- Yahoo

Transaction Compared



Home Page



Search Results
(General and
Image)



Rotating search terms used randomly (5 terms for general, 3 image terms)

Image Results only page in image search path used for ranking

Methodology

1. Investigation

- Which sites are the largest in their industry?
- Which transactions are most commonly performed on these sites?



2. Deployment

- Transactions are defined across all sites
- For each transaction to be tracked, a script is created for each site to step through the appropriate pages for that site
- Transactions are deployed to the Keynote agents for measurement



3. Data Collection

- Data is collected over a one-month period



4. Analysis

- Data sample undergoes statistical analysis
- Sites are investigated individually
- Industry trends are noted

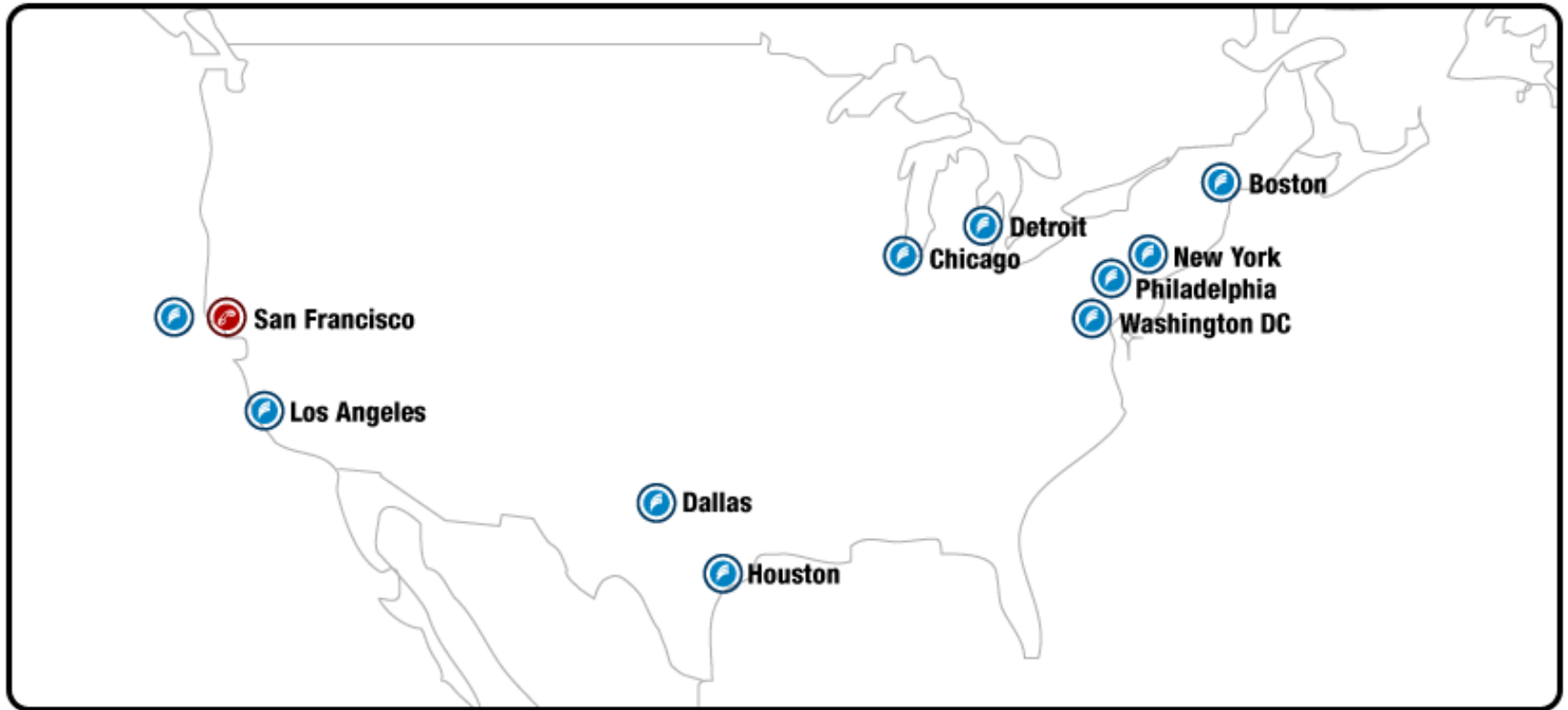


5. Rankings

- Rankings are created and presented
- Best sites are named



Measurement Agents



Data Collection

- **More than 6000 total samples were collected for each site**
- Data was collected from **November 22nd, 2007 – December 13th, 2007**
- The transaction was run against each site every hour throughout the day from each of Keynote's US10 agents as well as a Dial-Up agent. In other words, 11 transactions were run each hour against each site, **or a total of 264 runs each day against each site**
- For many of the metrics presented a peak period was used. This period was defined as daily **8 am to Midnight EST**. Using a peak period was especially important when calculating reliability as many sites have planned maintenance over the weekends that would negatively impact their ranking

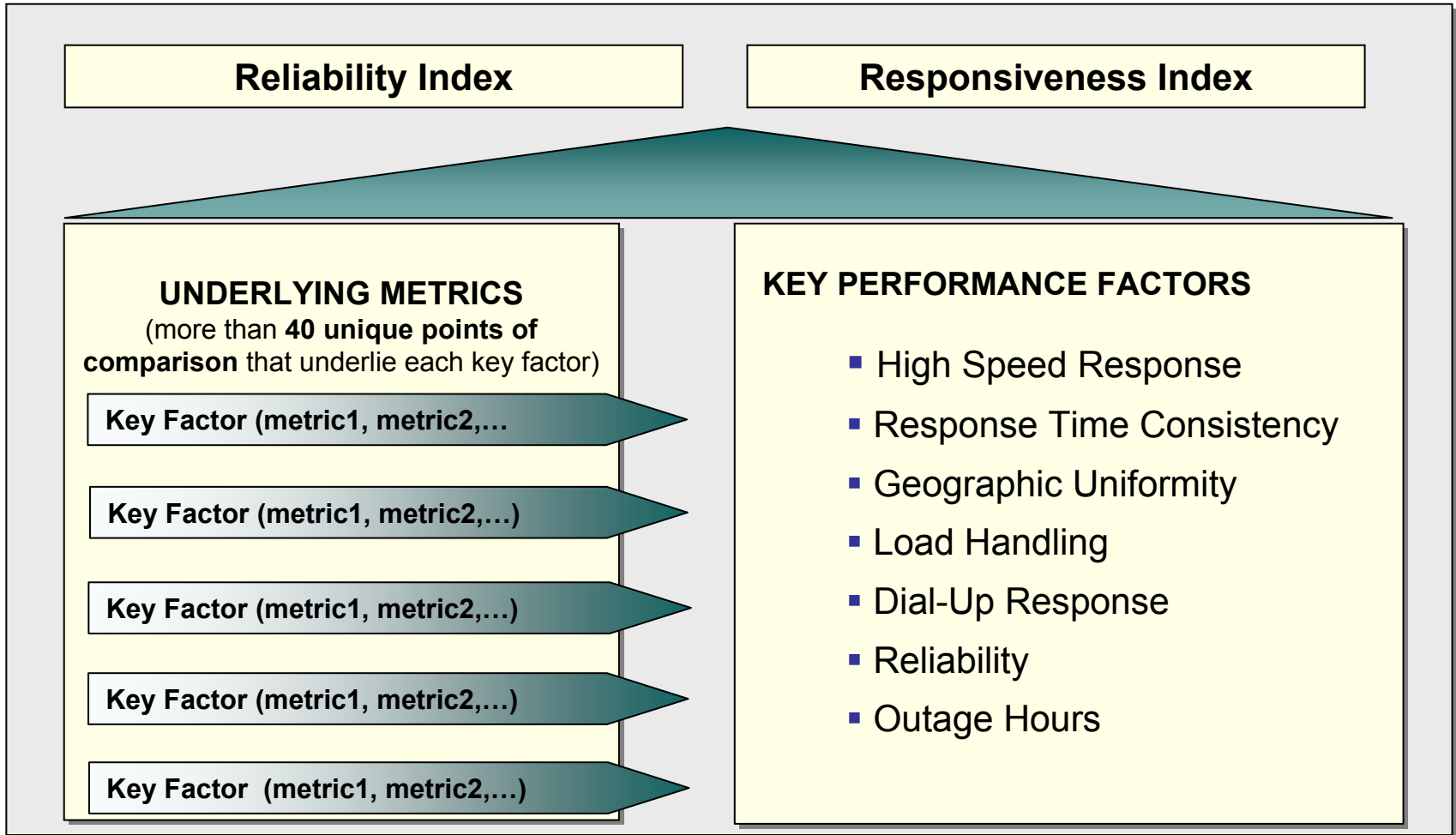


keynote™

The Mobile & Internet Performance Authority™

Ranking Methodology

Ranking Methodology



Ranking Methodology (cont.)

- Rather than simply comparing the site's overall response time and reliability, Keynote chose to rank a number of key areas of each site's performance. This methodology takes a more detailed view of each of the sites' service levels and rewards and penalizes the sites in a more consistent manner.
 - 7 key performance factors were identified. Keynote believes these factors describe a site's overall service level. A site's overall performance score and ranking is based on a combination of these 7 factors. The factors are:
 - **High Speed Response** – How quickly do the transaction and portions of the transaction (individual pages) load when measured from a T1 connection?
 - **Response Time Consistency** – Does the site minimize its response time variation?
 - **Geographic Uniformity** – How well does a site service all portions of the US Internet market?
 - **Load Handling** – Does the site regularly slow down during peak periods?
 - **Dial-Up Response** - How quickly do the transaction and portions of the transaction (individual pages) load when measured from a 56Kbps connection?
 - **Availability** – How often do users on the site receive errors?
 - **Outage Hours** – How many hours is the site effectively unavailable during the measurement period?

Ranking Methodology (cont.)

- **Within each key factor, Keynote identified several underlying metrics. These metrics comprise the final standing in each key area. For example, within the High Speed Response factor these underlying metrics were used:**
- - Overall Transaction Response Time
 - Average Page Download Time
 - Home Page Download Time
 - General Search Page Download Time
 - Image Search Page Download Time



keynote™

The Mobile & Internet Performance Authority™

Overall Results

Responsiveness Ranking

Site	Score	Rank
Google	728	1
Yahoo	710	2
MSN	620	3
AOL	331	4
Ask	211	5

Reliability Ranking

Site	Score	Rank
Ask	1000	1
Google	1000	1
Yahoo	980	3
AOL	387	4
MSN	286	5

Key Factors

High Speed Response	Dial-up Response	Response Time Consistency	Geographic Uniformity	Load Handling	Availability	Outages
Google	Google	Yahoo	MSN	Yahoo	Google	Google
MSN	Yahoo	MSN	Yahoo	AOL	Ask	Ask
Yahoo	Ask	AOL	AOL	Ask	Yahoo	Yahoo
Ask	MSN	Google	Google	MSN	MSN	AOL
AOL	AOL	Ask	Ask	Google	AOL	MSN

“()” indicates tie in key factors



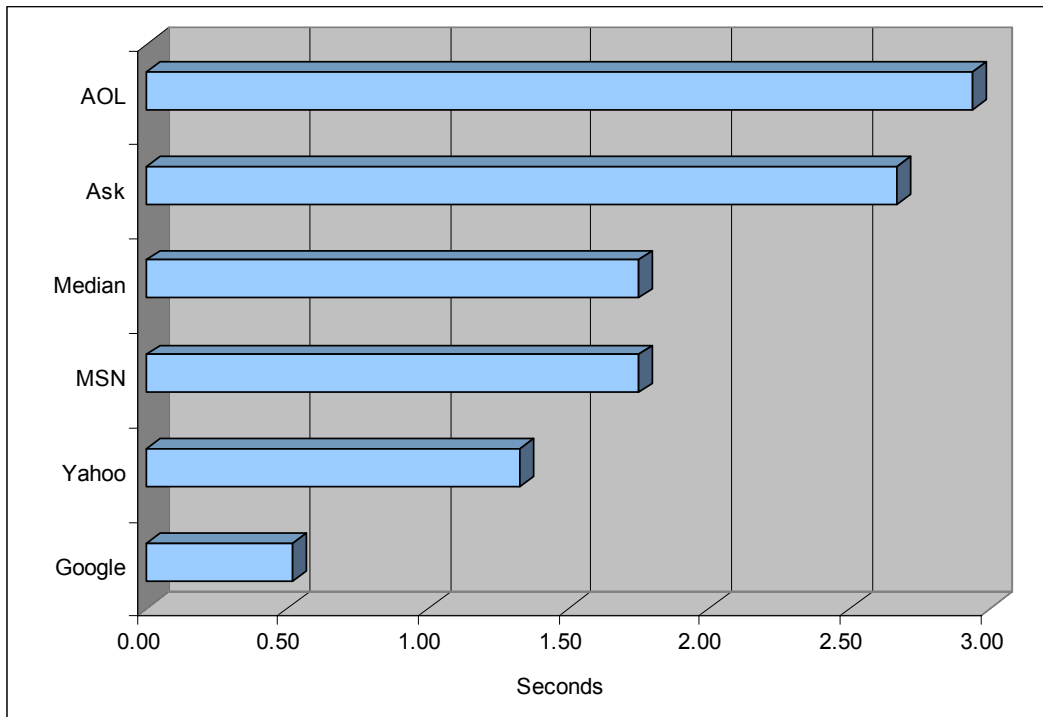
keynote™

The Mobile & Internet Performance Authority™

High Speed Response

High Speed Response

Overall Transaction Response Time



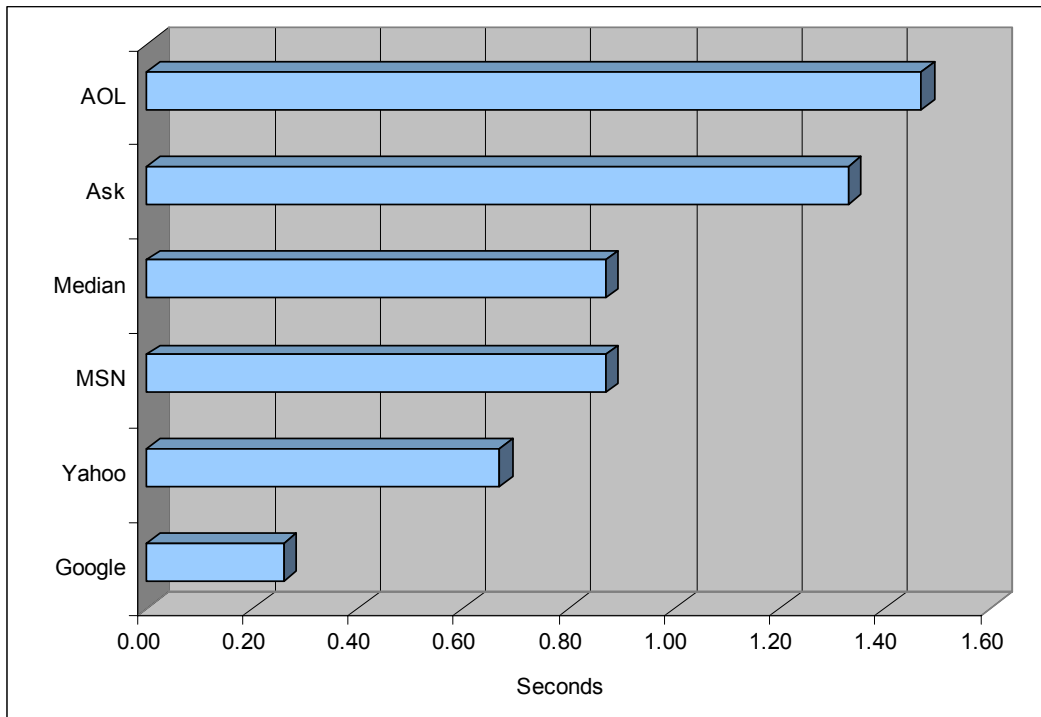
- Geometric mean of total transaction response times as seen from the US10 connected agents

- How long does it take to perform the entire transaction?
- AOL reported the longest transaction execution time at 2.9 seconds
- Google was the fastest site, executing the transaction in 0.52 second
- The industry average was 1.8 seconds
- Note: “industry average” = median



High Speed Response

■ Average Page Download Time



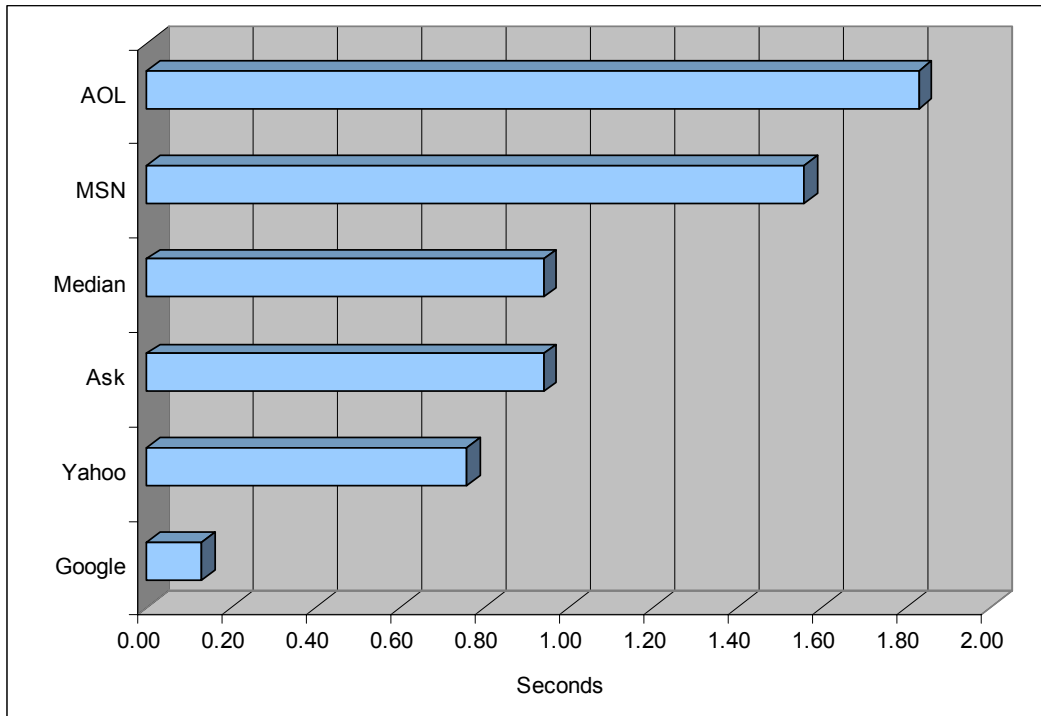
■ Geometric mean of page download time as seen from the US10 agents

- What is the average performance of the pages within the transaction?
- Each site has a different number of pages in the transaction. Rather than penalize a transaction with a longer transaction, the average page performance rank normalizes for transaction length
- The industry average was 0.87 second with Google reporting the fastest average page download time at 0.26 second



High Speed Response

■ Average Home Page Download Time



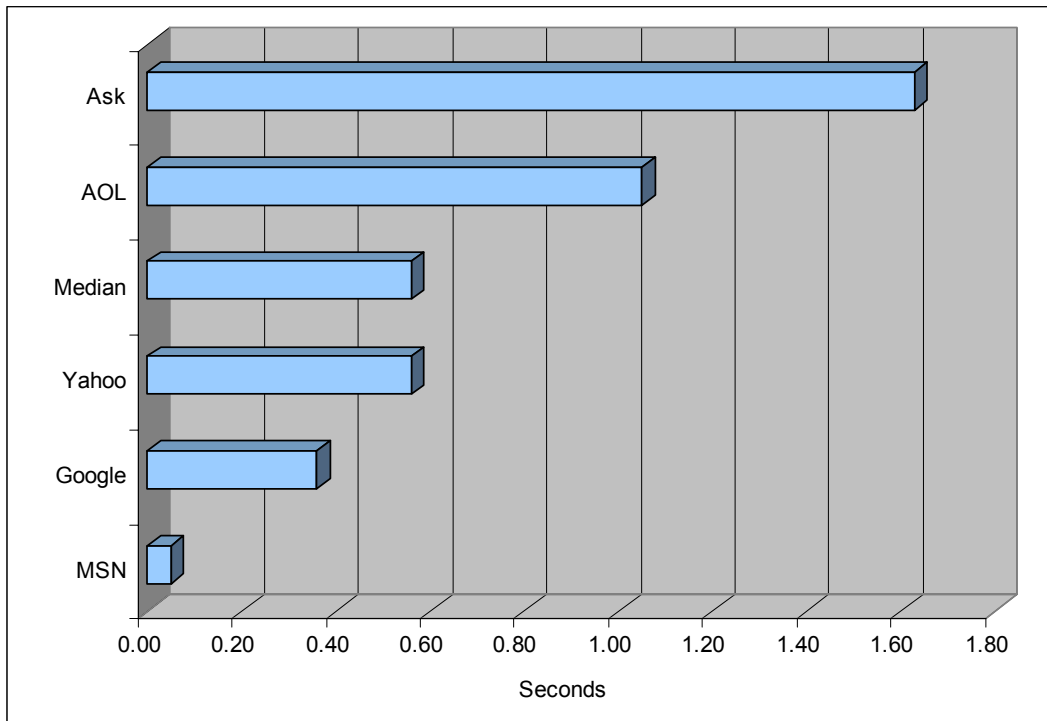
■ Geometric mean of page download time as seen from the US10 agents

- How long does my Home Page take to download?
- The performance of a Home Page sets the expectation for every other page in the site. Keynote recommends that this page be the fastest of any page in the site
- Google had the fastest Home Page download at 0.13 second
- AOL had the slowest Home Page download, loading in 1.8 seconds



High Speed Response

■ Average Search Results Download Time



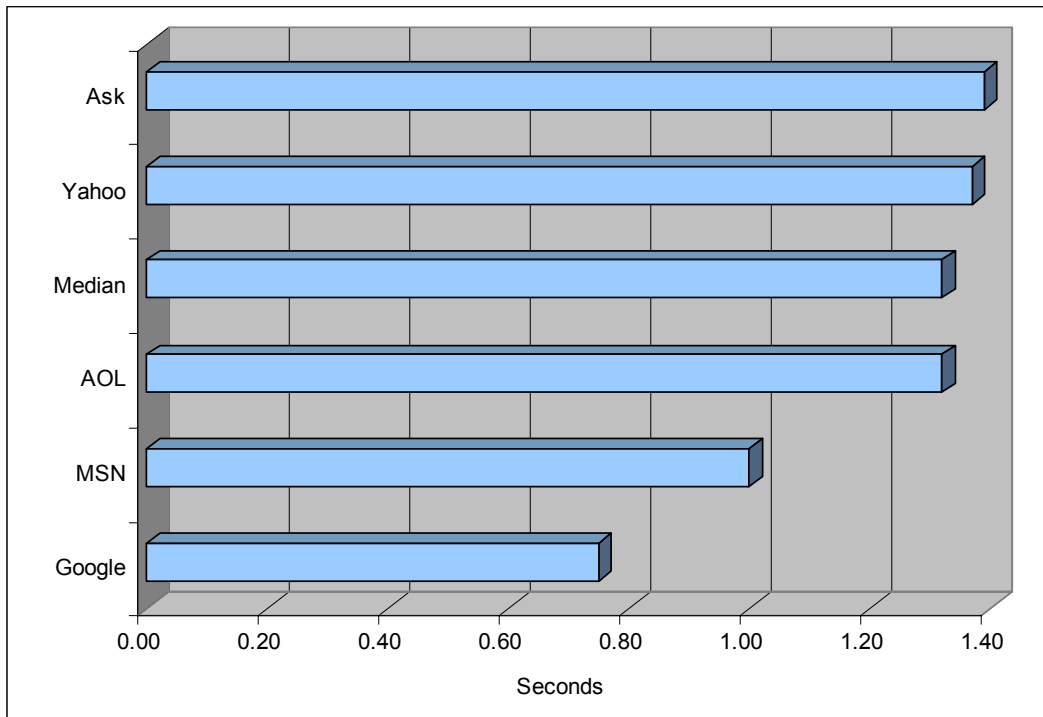
■ Geometric mean of page download time as seen from the US10 agents

- How long does the Search Results page take to download?
- The Search Results page measures how long it takes the search results to return and the page to be fully loaded.
- Ask had the longest Search Results pages, downloading in 1.6 seconds, compared to the industry average of 0.56 second
- MSN reported the fastest Search Results download time at 0.05 second



High Speed Response

■ Average Image Search Results Download Time



■ Geometric mean of page download time as seen from the US10 agents

- How long does the Image Search Results page take to download?
- The Image Search Results page measures how long it takes the image search results to return and the page to be fully loaded.
- Ask had the longest Image Search Results page, downloading in 1.4 seconds, compared to the industry average of 1.3 seconds
- Google reported the fastest Image Search Results download time at 0.75 seconds





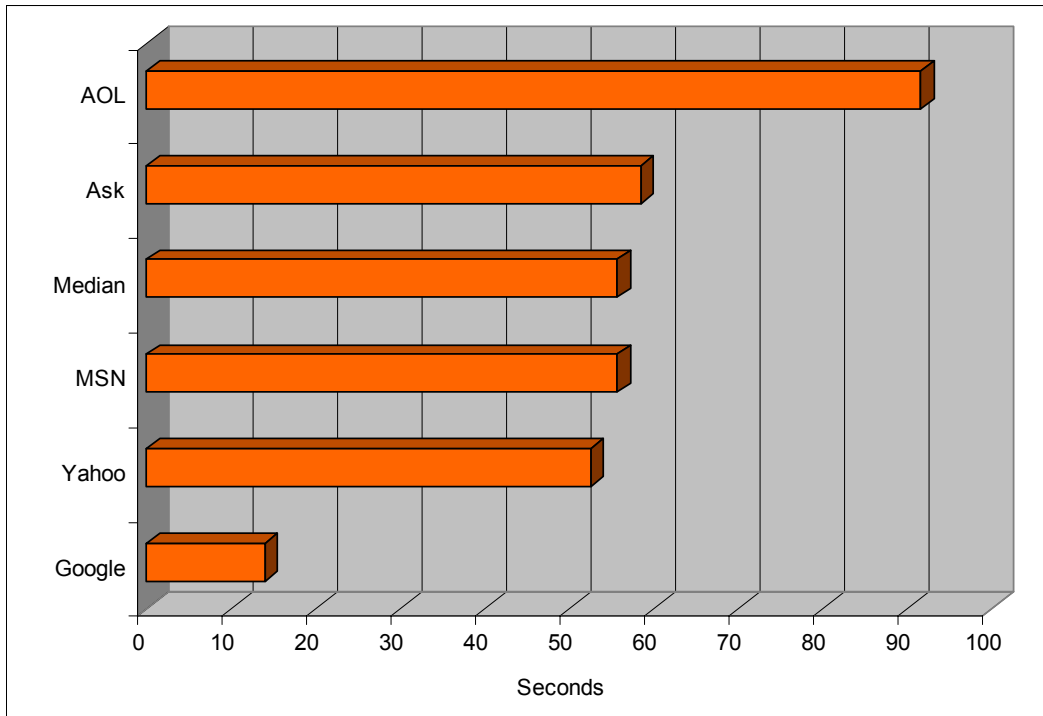
keynote™

The Mobile & Internet Performance Authority™

Dial-Up Response

Dial-Up Response

Overall Transaction Response Time



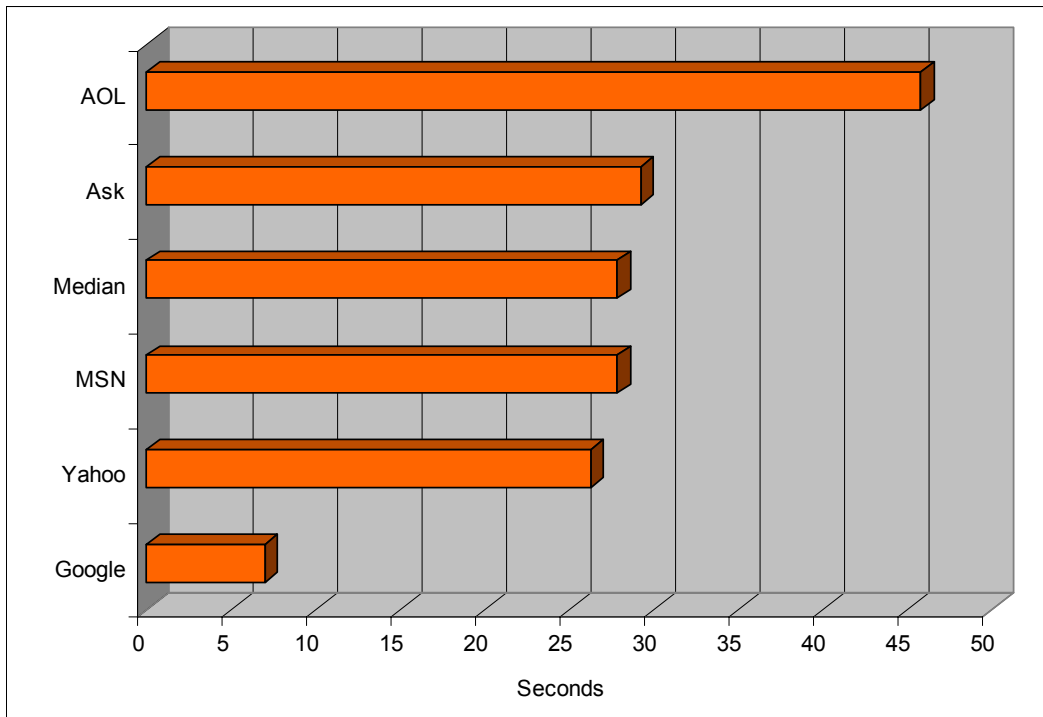
- Geometric Mean of total transaction response times as measured over dial-up.

- How long does it take to perform the transaction when using a dial-up connection?
- High speed performance is not always a good predictor of dial-up performance
- AOL was the slowest site over dial-up, executing the transaction in about 91 seconds
- The industry average for this metric was about 55 seconds
- Google had the fastest time, executing the entire transaction in about 14 seconds



Dial-Up Response

- Average Page Download Time



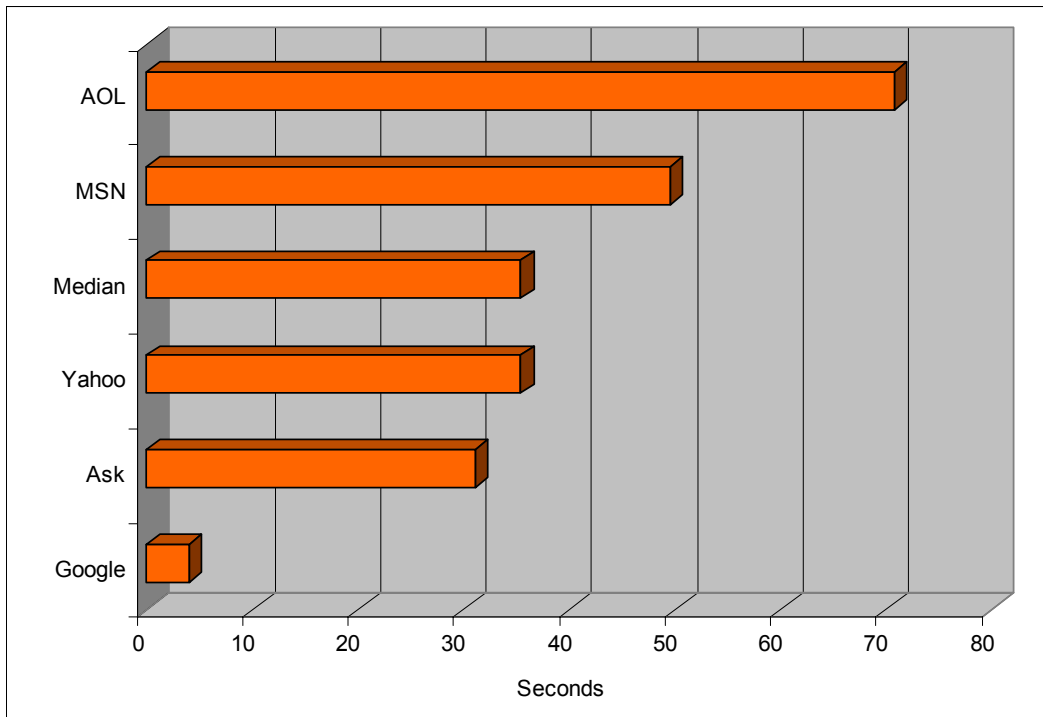
- Geometric Mean of page download times as measured over dial-up.

- How long does it take to download each of the pages in the transaction on average?
- Keynote finds that download times of less than 20 seconds in any industry should be considered excellent
- Google reported an average download time of about 7 seconds
- AOL had the longest average download of about 45 seconds



Dial-Up Response

- Average Home Page Download



- Geometric Mean of page download times as measured over dial-up.

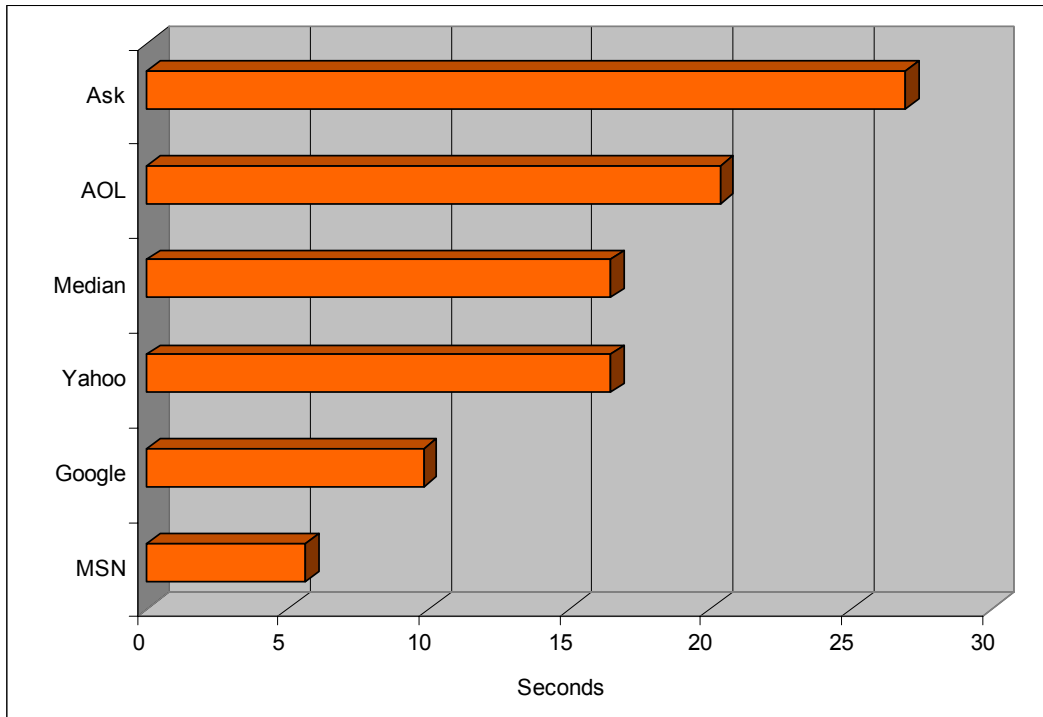
- How long does the Home Page take to download when using a dial-up connection?

- AOL had the slowest Home Page over dial-up, downloading in about 70 seconds
- The industry average was about 35 seconds, with Google reporting the fastest Home Page download of about 4 seconds



Dial-Up Response

■ Average Search Results Download



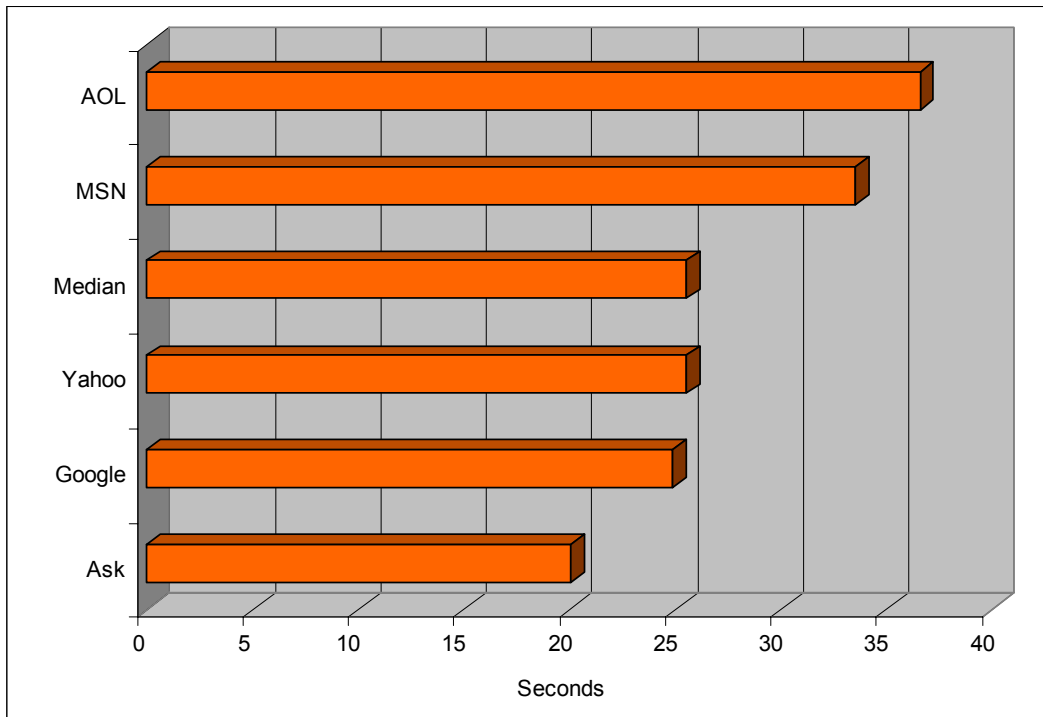
■ Geometric Mean page download time as measured over dial-up.

- How long does the Search Results page take to download when using a dial-up connection?
- Ask had the slowest Search Results page over dial-up, downloading in about 26 seconds
- MSN had the fastest Search Results page over dial-up, downloading in about 6 seconds



Dial-Up Response

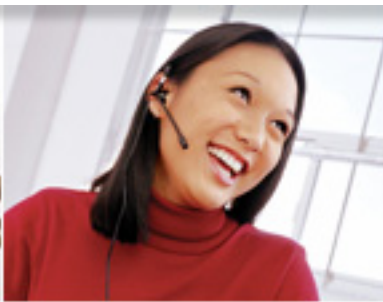
■ Average Image Search Results Download



■ Geometric Mean page download time as measured over dial-up.

- How long does the Image Search Results page take to download when using a dial-up connection?
- AOL had the slowest Image Search Results page over dial-up, downloading in about 36 seconds
- Ask had the fastest Image Search Results page over dial-up, downloading in about 20 seconds





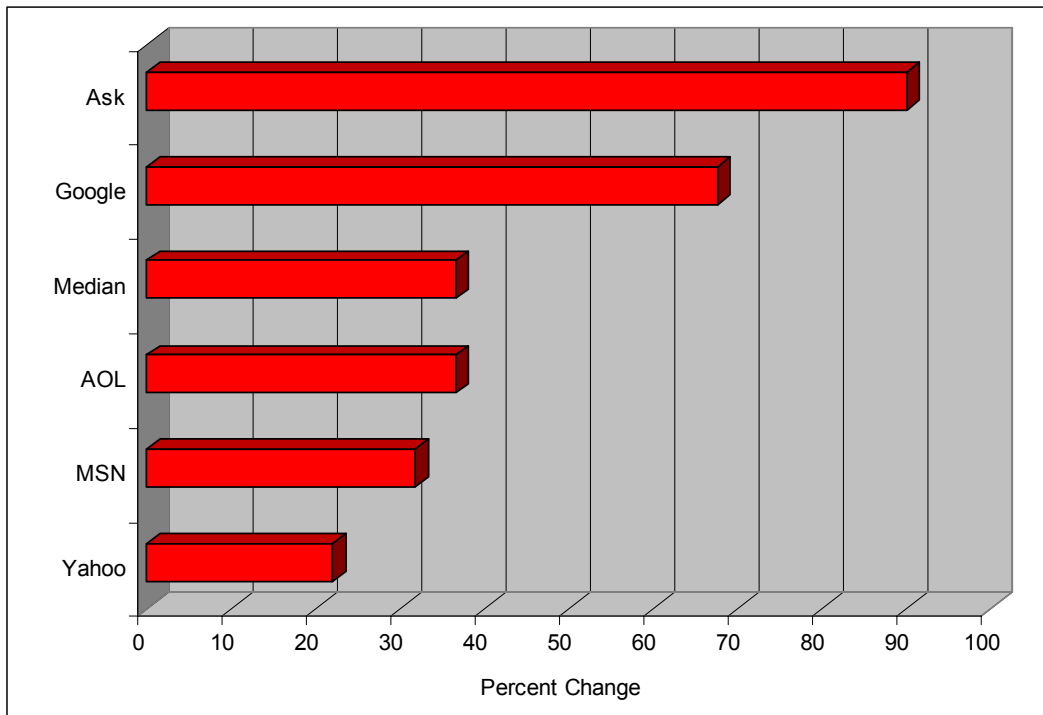
keynote™

The Mobile & Internet Performance Authority™

Response Time Consistency

Response Time Consistency

- Percentage Change Between 25th and 75th Percentiles



- Percentage change in overall transaction response between 25th and 75th percentiles (normalized by 50th percentile).
- Formula used: $100 \times (75^{\text{th}} \text{ percentile} - 25^{\text{th}} \text{ percentile}) / (50^{\text{th}} \text{ percentile})$

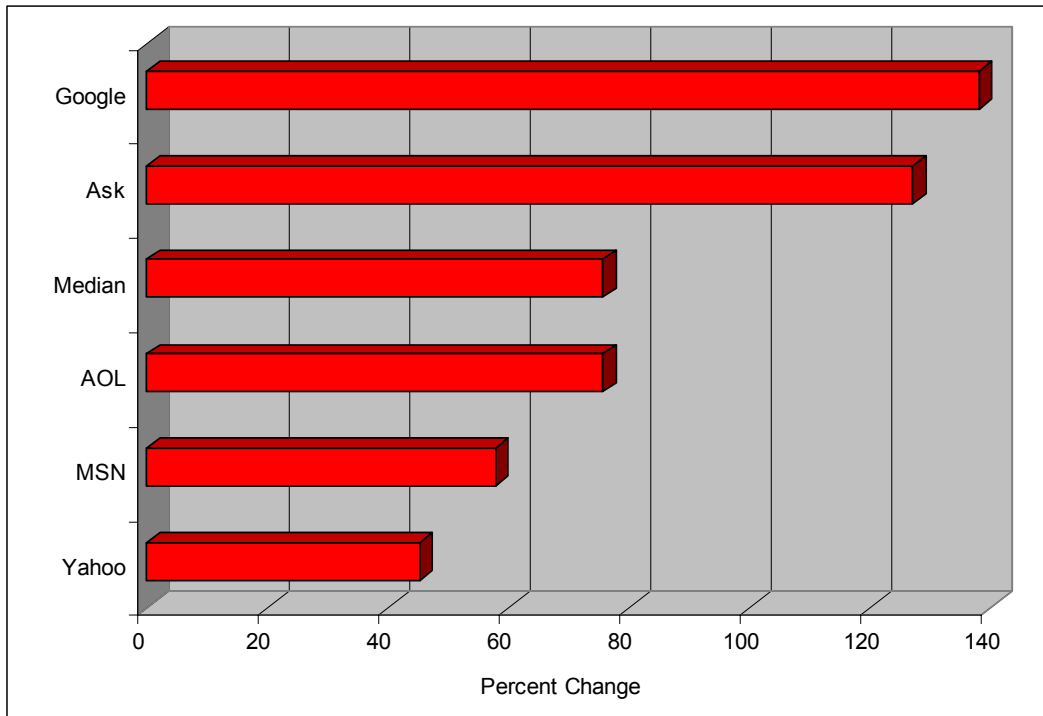
How consistently does each site perform? How much variation was there during the measurement period?

- Consistency is a good indicator of site health and stability
- Ask reported the greatest amount of instability, with 90% variation between the 25th and 75th percentiles
- The industry average was 37% with Yahoo reporting only 22% variation between the 25th and the 75th percentiles



Response Time Consistency

■ Percentage Change Between 10th and 90th Percentiles



- Percentage change in overall transaction response times between 10th and 90th percentiles (normalized by 50th percentile).
- Formula used: $100 \times (90^{\text{th}} \text{ percentile} - 10^{\text{th}} \text{ percentile}) / (50^{\text{th}} \text{ percentile})$

How consistently does each site perform? How much variation was there during the measurement period?

- All sites were more variable when looking at the extreme percentiles (which is expected)
- Google was the most variable, reporting a 138% variation between the 10th and 90th percentiles
- Yahoo continued to be the least variable site even at the most extreme percentiles





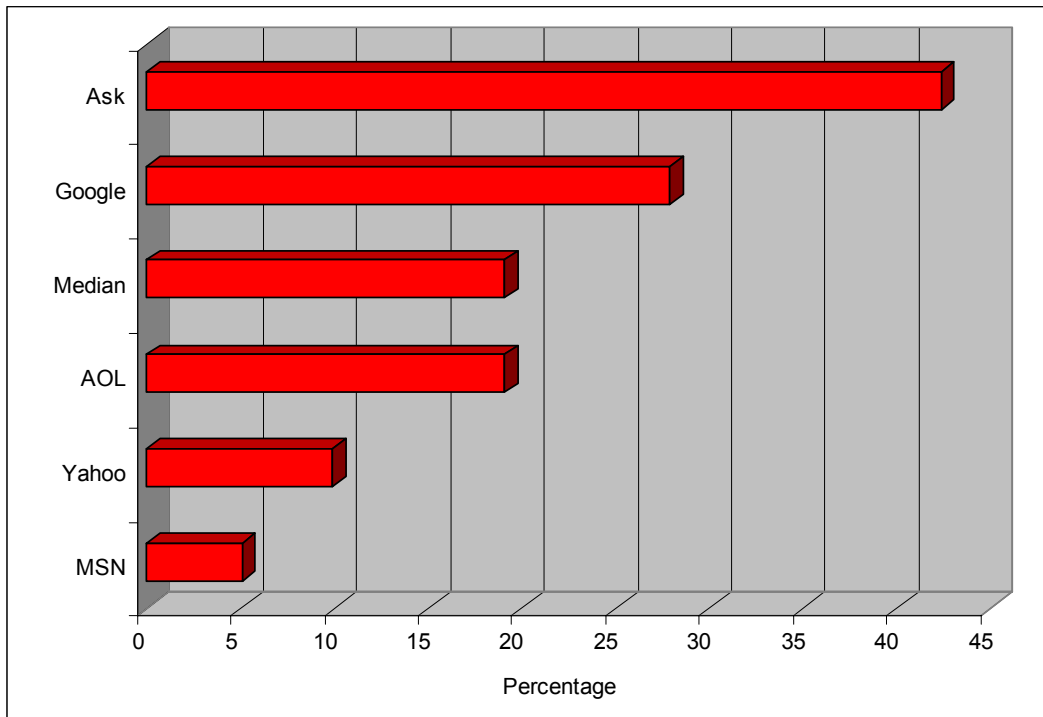
keynote™

The Mobile & Internet Performance Authority™

Geographic Uniformity

Geographic Uniformity

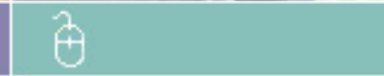
■ Variation Between Agents



- Standardized standard deviation percentage (normalized by average response time).
- Formula used: $100 \times \frac{\text{standard deviation of average response time of ten agents}}{\text{average response time}}$

- How uniform is the site's performance throughout the country and for users on different backbones?
- Ideally a site should provide a very consistent experience to users in all areas of the country and on all major backbones
- Some variation is normal given the nature of the Internet network and that most sites do not have a physical presence in each area of the country
- MSN showed the least amount of variation and Ask the most





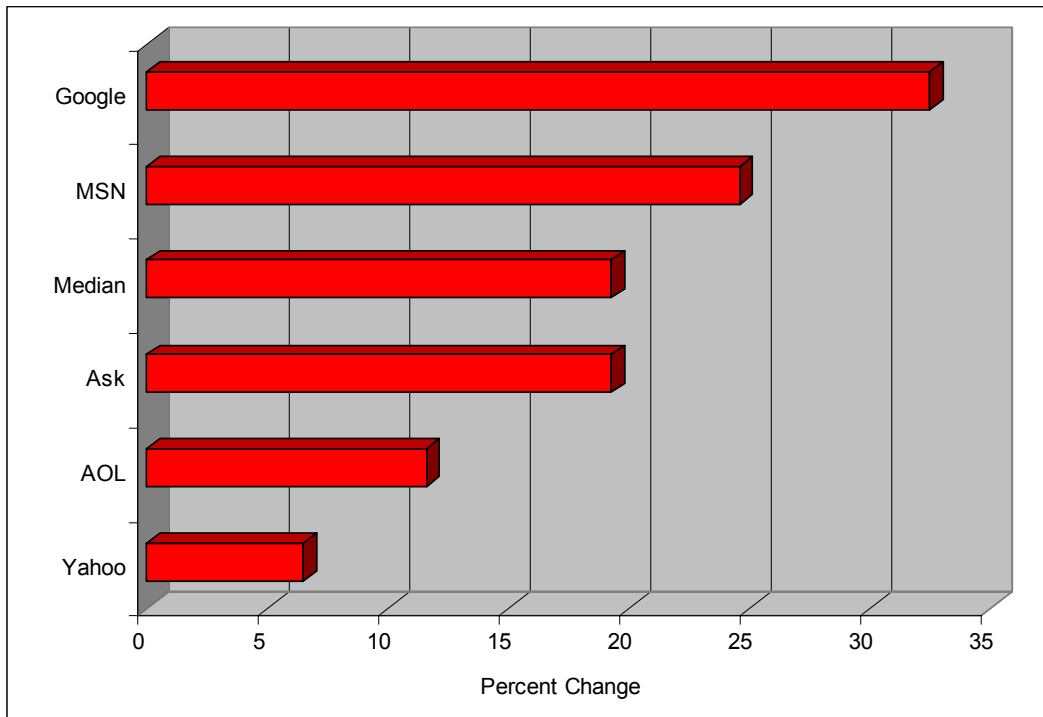
keynote™

The Mobile & Internet Performance Authority™

Load Handling

Load Handling

■ Peak Period Response Time Degradation



- Total percentage change in response time during the peak period. This indicates load sensitivity. Formula used: $100 \times (\text{Maximum response from just peak period} - \text{Minimum response from entire day}) / (\text{Minimum response from entire day})$
- Keynote had no knowledge of the load patterns or volumes on each of the sites. This rating simply captures the site's ability to service users consistently throughout each day, when user traffic on the Internet is in general higher.

- How does a site handle load?
- Typical Internet traffic peaks mid day, Monday through Friday. A site should be designed to handle daily increases in load without significant performance degradation
- A large peak period degradation indicates an underlying capacity issue or system bottleneck
- Most of the sites showed very little peak period slowdown. Google and MSN reported the largest slowdown of about 33% and 25% respectively





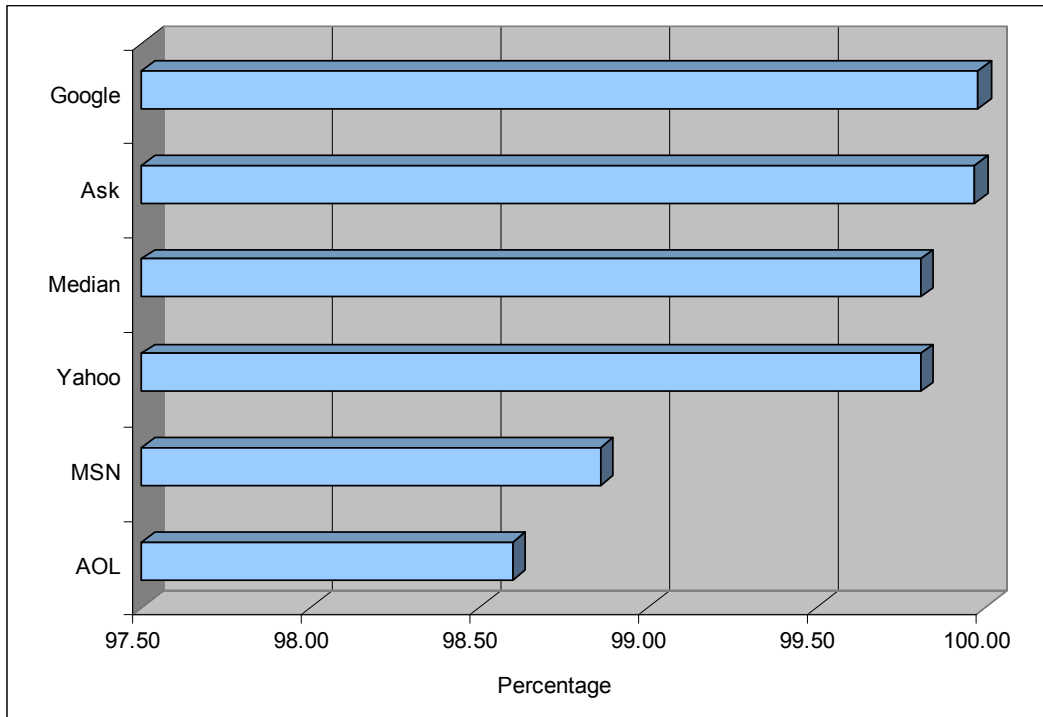
keynote™

The Mobile & Internet Performance Authority™

Availability

Availability

■ Total Transaction Availability (24X7)



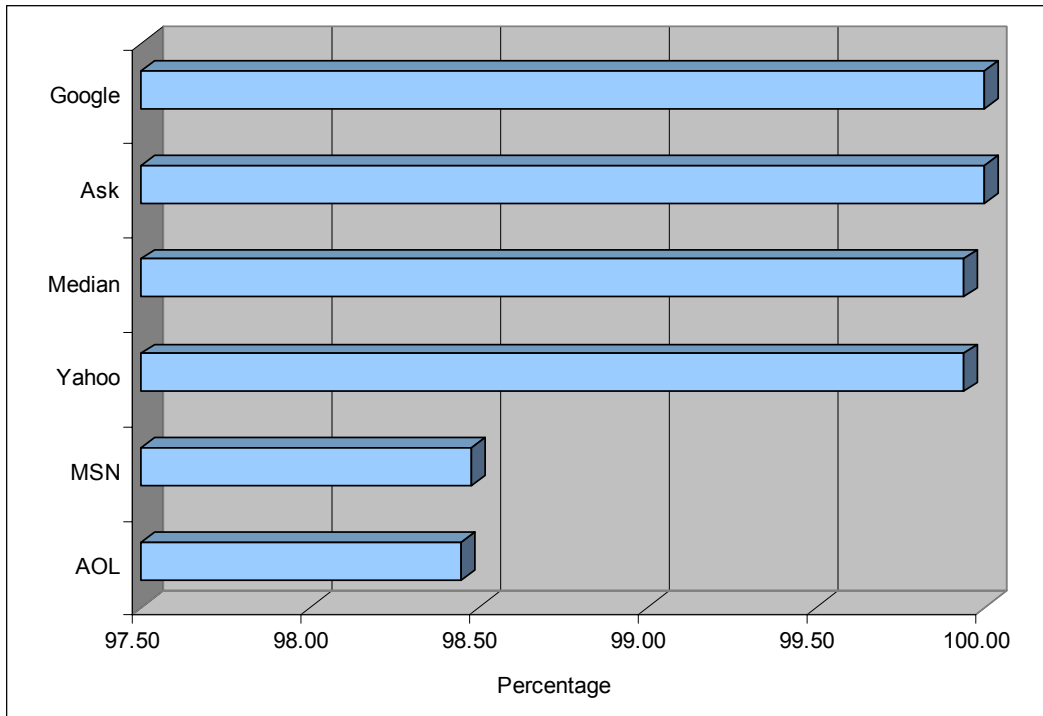
■ Availability for the entire measurement period. This does not include content errors.

- How reliable is the transaction at all times?
- Keynote marks a transaction as failed when any of the pages report a fatal (non-content) error.
- The industry average was 99.8% during all hours of the day and all days of the week
- Google reported the best 24X7 availability at 99.98%



Availability

- **Total Transaction Availability (Peak)**



- Availability for the entire measurement period. Taken during the peak period (Every Day 8am–Midnight EST). Does not include content errors.

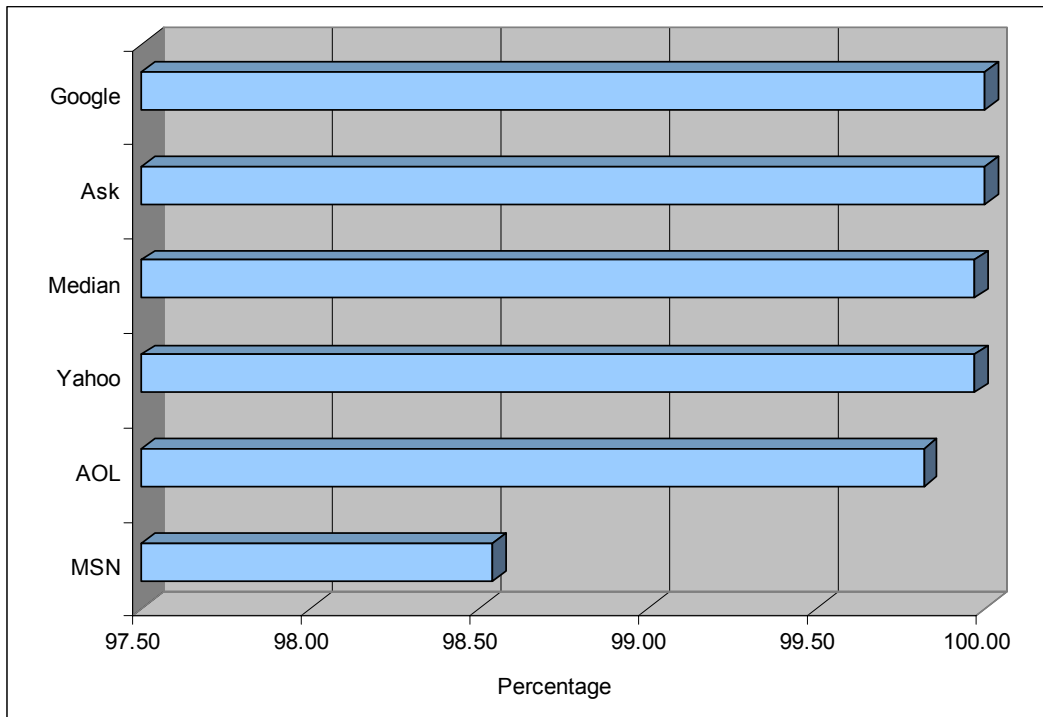
- How reliable is the transaction during the peak period?

- Keynote uses a peak period to exclude planned maintenance that a site might do very early in the morning
- The industry average during the peak period was 99.9%
- Ask and Google reported 100% average availability



Availability

- Home Page Availability (Peak)



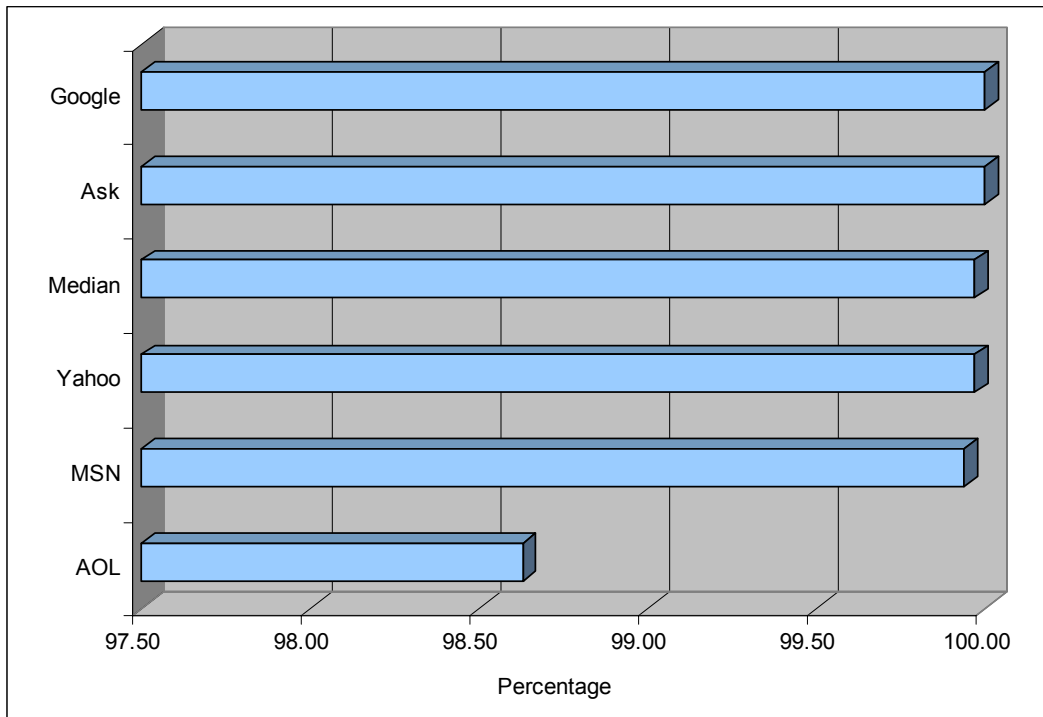
- Availability for the entire measurement period. Taken during the peak period (Monday–Sunday 8am–Midnight EST). Does not include content errors.

- How reliable is the Home Page?
- MSN had the least reliable Home Page, reporting 1.46% error during the peak period.
- The rest of the sites reported over 99.9% average availability



Availability

■ Search Results Page Availability (Peak)



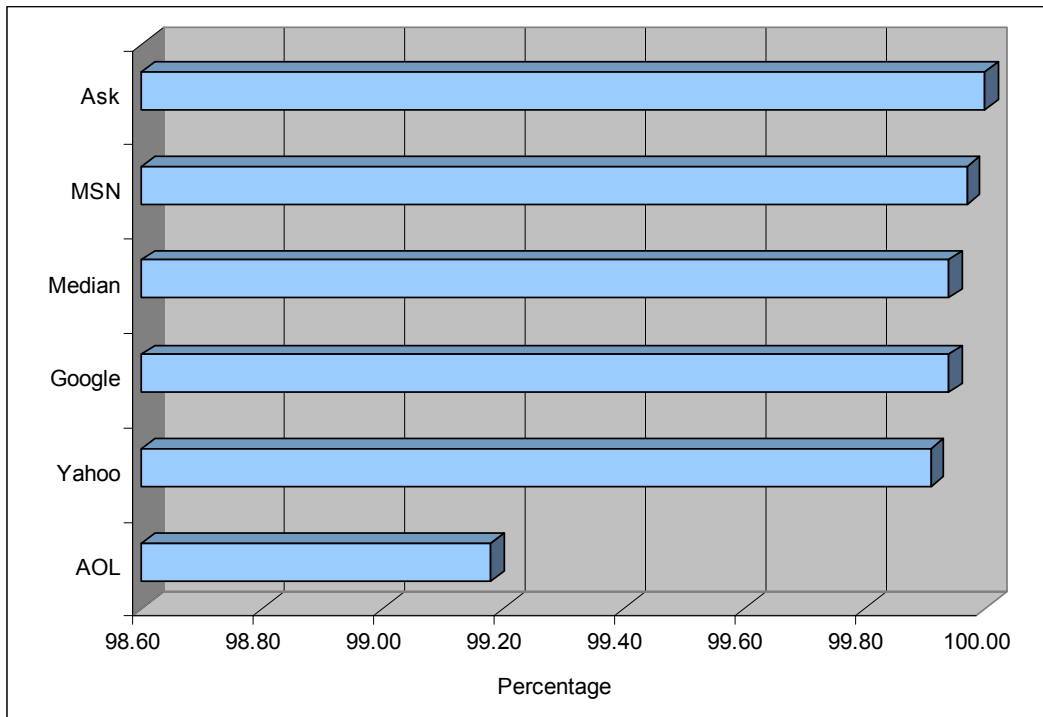
- Availability for the entire measurement period. Taken during the peak period (Monday–Sunday 8am–Midnight EST). Does not include content errors.

- How reliable is the Search Results page?
- Google and Ask had the most reliable Search Results Page reporting an average availability of 100%
- The rest of the sites, with the exception of AOL, reported an average availability of 99.9%



Availability

■ Image Search Results Page Availability (Peak)



- Availability for the entire measurement period. Taken during the peak period (Monday–Sunday 8am–Midnight EST). Does not include content errors.

- How reliable is the Image Search Results page?
- Ask had the most reliable Image Search Results Page reporting an average availability of 100%
- The rest of the sites, with the exception of AOL, reported an average availability of 99.9%



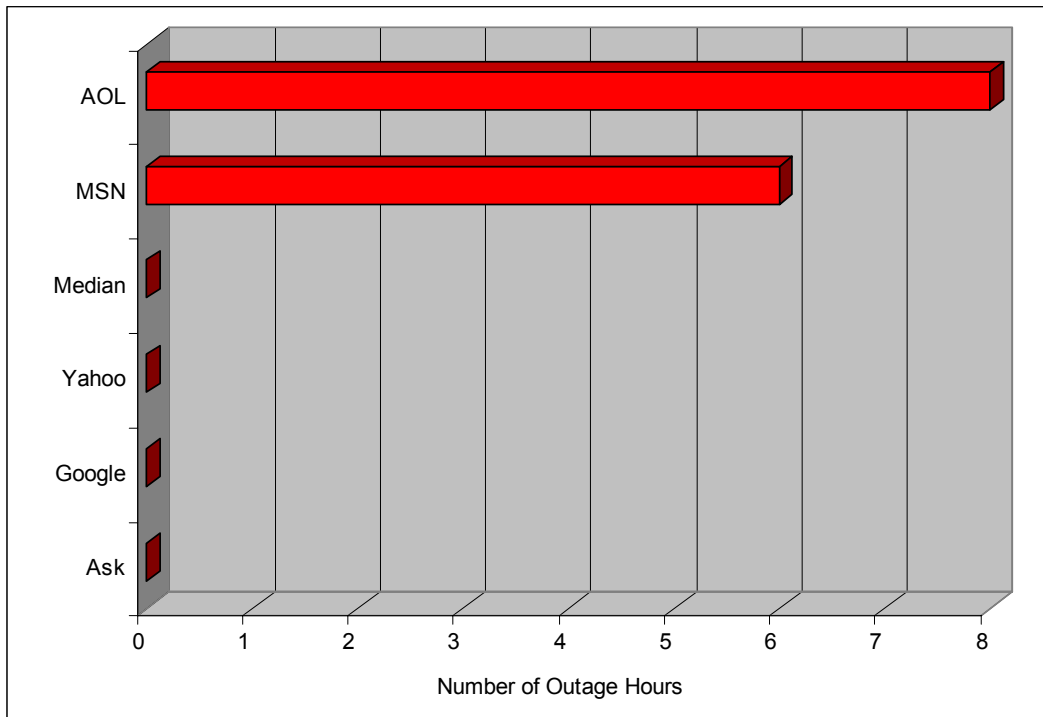


keynote™

The Mobile & Internet Performance Authority™

Outage Hours

Outage Hours

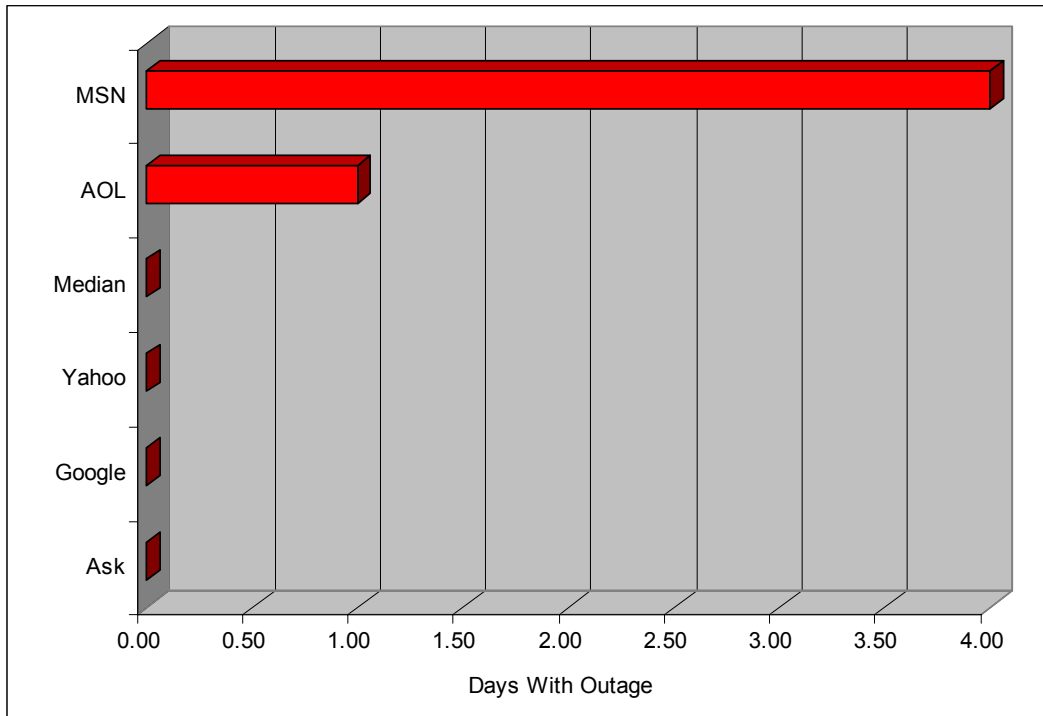


- An outage hour is defined as any hour in which the site reported greater than a 30% error rate. It was only calculated for the peak period (Every day 8am–Midnight EST).

- For how many hours did the sites have outages?
- Keynote defines an outage hour as any hour during the peak period where the transaction reported more than 30% failures.
- AOL and MSN reported 8 and 6 hours of outage respectively.



Days With Outages



- A day that has one or more than one outage hour is defined as an outage day. The outage day percentage is normalized by the total number of peak days (Monday through Sunday).

- How many days did the sites have outages?
- It is important to understand how many days during the study were impacted by outages. The Days with Outage metric provides this information.
- MSN reported 4 days with at least one outage hour.

