dynatrace

Dynatrace
Performance Benchmarks

Methodology

Contents

1. Overview: What are Dynatrace Benchmarks?

Competitive context is powerful. It can help us know what is possible and where we are falling short. Dynatrace provides over a hundred global benchmarks that include thousands of sites to provide our customers this competitive context. The benchmarks are in key industry verticals around the globe and cover both Home Pages and full transactions.

Using Dynatrace's Advanced Synthetic product (https://www.dynatrace.com/platform/offerings/synthetic-monitoring/) we collect technical quality (performance and availability) data on the top sites in each of these key verticals and that data is ranked with results published each week. While the data is captured using our synthetic product it does represent a typical customer experience. To this end, we use real browsers and measure from outside on the Internet. This data provides our customers an understanding of how fast is fast enough or how an industry is evolving over time.

1.1 How can you best use Benchmarks?

First, it is important to understand their limitations. If you are subscribed to a Home Page benchmark, know that while the Home Page is extremely important it is not the entire site. Further study should be done on other key pages. The same is true for even our transactional benchmarks, which will be measuring more pages but do not measure every page on the site. Any good performance management strategy should include additional synthetics measurements on the pages you know to be key.

Secondly, benchmarks do not capture every possible user configuration. We use a standard browser, operating system and connection. Your actual customers are not so consistent. If you want to capture the many variables of your customer experience, consider adding Real User Monitoring (RUM) (https://www.dynatrace.com/platform/offerings/dynatrace/) from Dynatrace. RUM captures all browsers, all users, all locations and this wider lens captures the entire experience of customers on your site.

Thirdly, the benchmarks provide objective ranking data focused on performance, availability and including high level details on front-end design. They are not meant to be a full analysis or deep dive of the competitive landscape. Said another way, benchmarks give you answers to the "what" questions not the "why" questions. Answering the "why" questions requires a deeper study. Dynatrace would recommend a deeper Competitive Assessment a few times a year to dive deeper into how your site can improve.

1.2 How do we choose who is included?

We align our benchmarks with our customer's industries and geographic regions. This is why you will see such a focus on financial services and retail, two of our top customer segments. Within each region and industry, we choose the top sites based on a combination of metrics that include traffic, sales or innovation. We are always open to feedback around new industries and/or additional sites to include in existing benchmarks. We'd love to hear from you at benchmarks@dynatrace.com.

1.3 How do we measure?

Synthetic measurement is perfectly suited for collecting competitive data. It doesn't require anything to be installed on the site under study or even for the site owner to be aware. Synthetic measurement is simply another user type accessing the site. We are very careful to keep our measurement sampling low and ensure that there is no adverse impact on the site.

We use the Chrome browser because it is the mostly widely used browser in the world. The goal of benchmarking is to capture the typical user experience not every possible user experience. Using the most popular browser allows us to do this. We have also found that for most modern browsers there are not many major differences between performance (if your site is slow in Chrome it is likely slow in IE and Firefox) though of course multi browser testing is very important for QA/functionality.

For our Mobile Home Page benchmarks, we use a mobile browser and emulate the connection speed (LTE/4G). Using real mobile connections is certainly possible with

dynatrace

Dynatrace but given the multiple variables of mobile we find emulation a bit fairer for a competitive comparison. This is especially true for mobile availability where we find that most errors are tied to the carriers themselves and not the site. Using an emulated LTE connection eliminates these errors and allows our benchmark customers to focus on what they have more direct control over.

For our Desktop benchmarks, we use backbone connections. These connections are faster than a typical user (but not that much faster) but eliminate the "noise" of last mile ISP providers.

For all our benchmarks we choose to measure multiple locations from each region. We always include at least two locations to eliminate the possibility of a single location skewing data. For most of our benchmarks we have at least five locations.

1.4 What do we measure?

We have two types of benchmarks, Home Page and Transactional. Home Page benchmarks measure a site's Home Page because of how important these pages are to the customer experience. Home Pages are still the most common entry point on most sites, set the expectation for performance and their performance is important for SEO rankings. You can tell a lot about a site by the performance of its "front door".

Transactional benchmarks focus on key user journeys on a site. Almost all site visitors are doing more than just hitting a Home Page. They are accessing their bank accounts or buying products. Our transactional benchmarks try to capture these typical user paths. We can't capture every page and we certainly can't emulate the many paths users take through the site. So, we must choose a set of key pages that both represent a typical experience and can be done on EVERY site included in the benchmark. This doesn't mean the number of pages measured will always be equal, as each site has a unique UI design but the journey will be comparable.
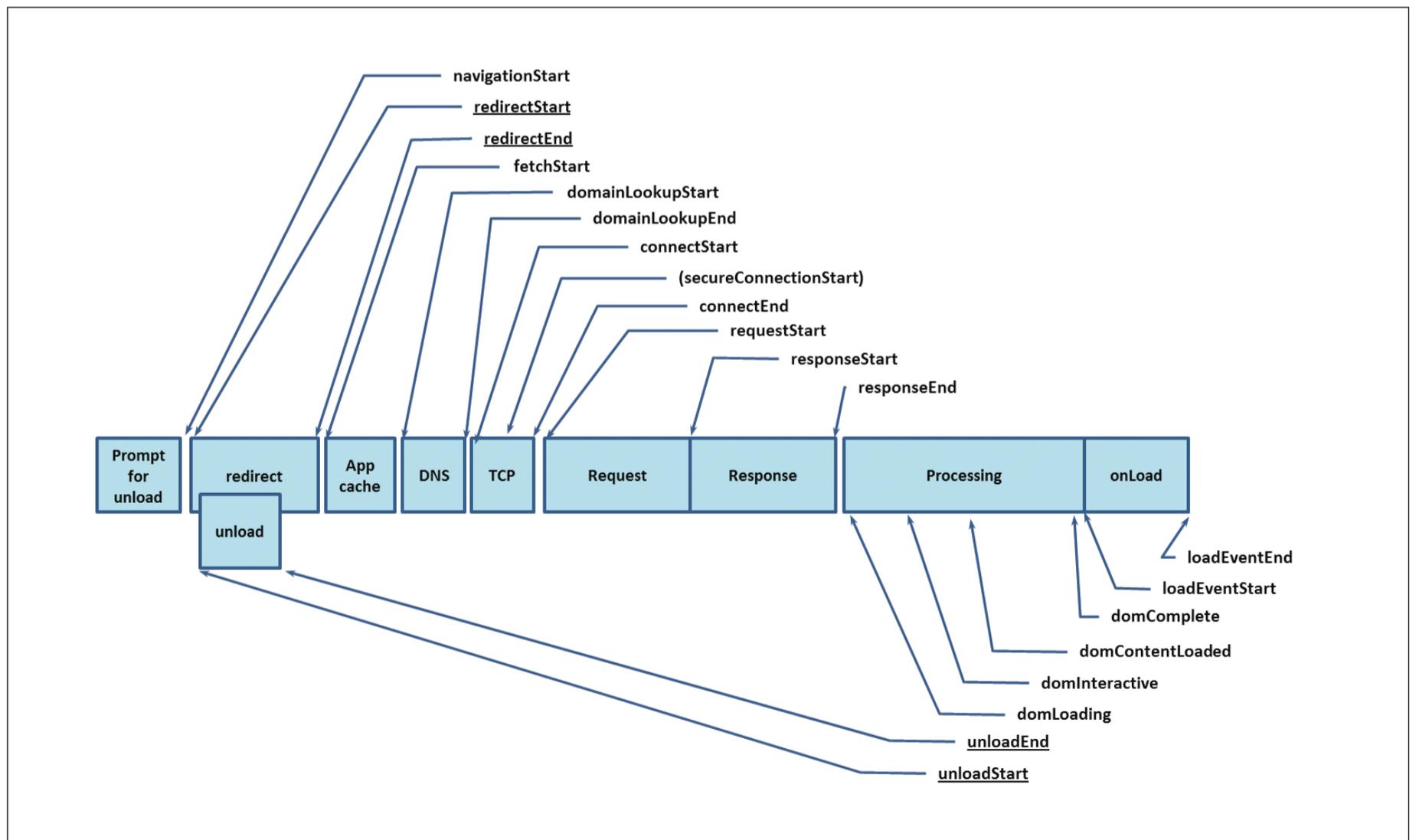
2. Ranking, Site Changes and Publication

2.1 How do we rank sites?

We rank sites based on two key metrics. The first is for performance and we are currently using loadEventEnd. When looking at performance no metric is perfect. The goal is to find a metric that in most cases represents when a page is usable and is generic enough to be used fairly on all sites. We realize that each page is unique and the case could be made for a unique metric for each page. For example, a login page might be more accurately captured using domInteractive while the Account Summary page might better use loadEventEnd. However, making this determination on hundreds of sites and thousands of pages is not possible. We have to choose the best fit for all types of pages.

We have chosen to use loadEventEnd because it does two things. It eliminates any lazy loaded content (1st party or 3rd party) that is firing after loadEventEnd. This would include lazy loaded images for upcoming pages, chat calls or 3rd party tags (if placed correctly) to name just a few. Secondly, it is far enough in the load cycle that it is quite likely that the page is ready to use. As additional metrics are available (like visually complete) we will be updating our strategy with the goal of a metric that's generic enough to be used on all sites and is as close to the user experience as possible.

The second key ranking is availability. Availability is a simpler calculation. We simply take all the attempted measurements and look at those that are successful. We only include fatal errors and not content errors. Fatal errors include DNS, network and server errors as well as errors wrapped in valid HTML pages (like a "I am sorry try again later page"). We do our best to exclude any script errors or Dynatrace node issues in this calculation.

2.2 How do we handle site changes or agent issues?

Sites are consistently changing and this requires us to update our synthetics measurements. This is especially a problem for transactional measurements. As sites are updating, we update our measurements. But this might take us a few days and leave inaccurate data. We have no SLA for when benchmark scripts will be updated and in some rare cases updates might take weeks.

We are committed to data accuracy and as part of our publication process we make every attempt to exclude any data that is not reflecting actual site performance. This "exclusion" process can remove data for periods when the script was failing or for any node issues.

We do not exclude Internet issues, third party issues, or maintenance periods outside of our defined time periods. While some issues are not under our customer's direct control they do impact actual site visitors. Given this publication is a service for all customers, we have the final word on what is excluded or what is not. To maintain the integrity of the benchmarks we cannot exclude data just because a customer asks us or it might make their site rank more poorly.

We apply exclusions in our reporting system and do not make changes to any raw data in your account.

2.3 How do we publish?

We publish the ranking results every week for the previous week. There are two versions of the publication, free and paid. The paid version is currently only available to customers who have purchased the detailed benchmark data and the free version is available to anyone.

For the free version, we will provide the industry snapshot of key averages and the detailed results only for the top three sites. The paid version includes the full rankings and a CSV of the output (to be sorted or ingested into your own dashboards).

The way we handle who has access to what, is based on the email domain. We know what customers have access to each benchmark and if anyone from that email domain subscribes they will get access to the full version.

Given customers expect digital properties to be available at all times of the day or night we measure and compare the sites in the benchmarking using all hours and all days. We do not exclude planned downtime.

dynatrace

3. Key Calculations / Terms

Geometric mean: For performance we use geometric mean versus an arithmetic average. All internet data is not evenly distributed (bell shape) but has a very long tail (slow outliers). Using an arithmetic mean is not ideal for this type of data as it will be skewed by the outliers. A geometric mean is not as affected by the outliers.

1st party and 3rd party: While not impacting the ranking we have informational data around 1st party JavaScript and 3rd party overhead. Dynatrace uses a third-party database to identify most third parties and will assume that any domain that includes the core customer domain is 1st party. For example, images.dynatrace.com is 1st party to dynatrace.com. While we know in today's hybrid hosting world this domain might be hosted by a Content Delivery Network or other vendor, it is still likely under the direct control of the site owner.

The reason we track 3rd party content so closely is that it has the largest impact on site performance. Every site will have some third-party content but poorly performing and poorly placed third party content can ruin a user experience. We generally find that sites with too much third-party content do not have good processes in place to audit and update third party tags or are not aware of their impact on site visitors.

Page Elements: Dynatrace captures network requests on each page. These are considered "page elements". They are all types of requests including JavaScript, CSS, images and all other requests. We only include requests that are on the network and therefore not cached objects.

Page Size: We also capture page size by adding up the size of each asset on the page (if compressed we use that size). Again, we are only including objects/requests that are on the network and not cached.

Caching: For our measurements we always begin with a clear cache, to capture a first-time site visitor experience. For transactional measurements, we handle caching as a user would as we move from page to page. This means that if the core site framework (JS and CSS) are loaded on the first page in a transaction and cache control is set correctly we will not redownload this content on the next page.

Pages without LoadEventEnd: Not all pages report this value. For those pages that are not reporting LoadEventEnd we fall back to geometric mean of end to end time.

## Dynatrace Digital Performance Platform — it's digital business...transformed.

Successfully improve your user experiences, launch new initiatives with confidence, reduce operational complexity and go to market faster than your competition. With the world's most complete, powerful and flexible digital performance platform for today's digital enterprises, Dynatrace has you covered.

APPENDIX A

Annotated Sample Performance Benchmark Report
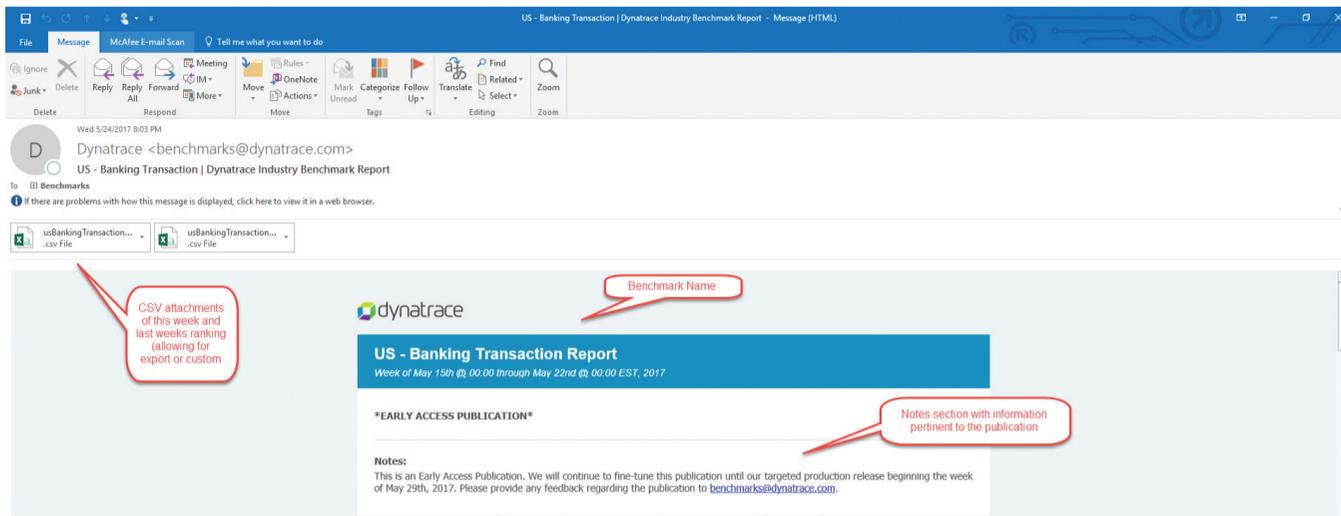
For subscribers, we include CSV files that include the underlying data for the current and previous week. This allows for sharing, sorting and long-term archiving of the data.
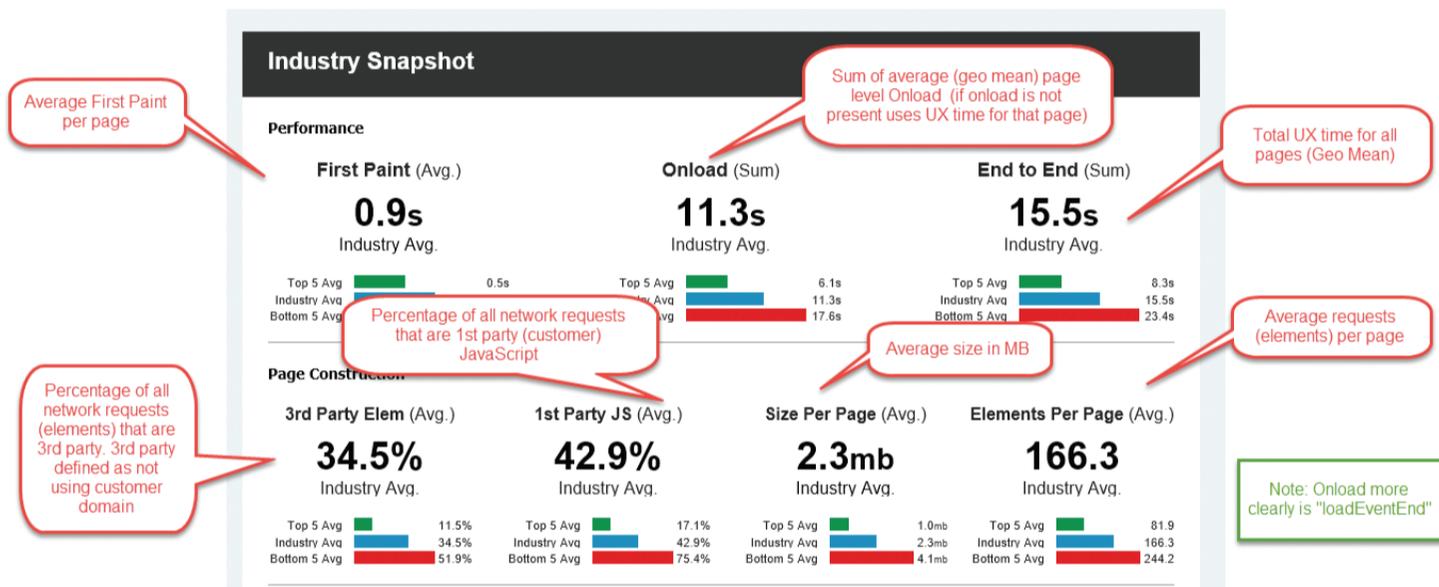
The Industry snapshot attempts to provide not only industry averages but also details on the top and bottom sites. Information about front-end design (3rd party, 1st party JS etc.) are included given that much of what drives performance is good or bad technical design. There is a very clear correlation between excessive elements or third parties and slower load times. Excessive 1st party JS can typically cause major delays within the browser engine and processing given the code overhead. These metrics are not used for ranking and are just provided for context.

Performance Benchmark Rankings - *Transaction*
Sorted by performance onload

Rank	Previous	Trending	Participant	Avg. First Paint	Onload	avg	End to End	avg
1	1	—	Charles Schwab	0.7s	**4.6s**	*0.9s*	**4.4s**	*0.9s*
2	2	—	Citizens Bank	0.7s	**5.1s**	*1.0s*	**15.1s**	*3.0s*
3	3	—	Wells Fargo	1.1s	**6.1s**	*1.5s*	**7.1s**	*1.8s*

Performance Benchmark Rankings - *Transaction*
Sorted by performance onload

Rank in previous week

Sorted by onload

Rank	Previous	Trending	Participant	Avg. First Paint	Onload	avg	End to End	avg
1	1	—	Charles Schwab	0.7s	4.6s	0.9s	4.4s	0.9s
2	2	—	Citizens Bank	0.7s	5.1s	1.0s	15.1s	3.0s
3	3	—	Wells Fargo	1.1s	6.1s	1.5s	7.1s	1.8s

Performance rankings are currently set to use onload (loadEventEnd) but will be updated as metrics evolve that more closely represent the customer experience.

The average values are shown given that some sites require fewer or more steps to complete the same business transaction. A high-ranking site with a high average likely indicates that they have fewer steps/pages in their transaction.

First Paint is included given the major impact render start has on the perception of load time. A page that takes longer to begins to render will feel slower than a similar page that beings to render much more quickly.

Page Construction Benchmark Rankings - *Transaction*
Sorted by performance onload

Sorted by Performance Rank Above

Rank	Previous	Trending	Participant	% 3rd Party	% 1st Party JS	Avg Size/Page	Avg # Elements/Page
1	1	—	Charles Schwab	0%	19.5%	0.2mb	25.7
2	2	—	Citizens Bank	10.2%	19.2%	1.5mb	130.6
3	3	—	Wells Fargo	5.9%	31.5%	0.9mb	50.7

The Page Construction section is just informational and therefore the ranking is tied to page performance. No metric is used for ranking in this section.

Page Construction Benchmark Rankings - *Transaction*
Sorted by performance onload

Sorted by Performance Rank Above

Rank	Previous	Trending	Participant	% 3rd Party	% 1st Party JS	Avg Size/Page	Avg # Elements/Page
1	1	—	Charles Schwab	0%	19.5%	0.2mb	25.7
2	2	—	Citizens Bank	10.2%	19.2%	1.5mb	130.6
3	3	—	Wells Fargo	5.9%	31.5%	0.9mb	50.7

Dynatrace works very diligently to exclude any script related errors and node related errors from each publication. The goal is to include errors that impact users.

Availability Benchmark Rankings - *Transactio...*

Transactional Availability during peak period. Exclusions applied for script or node issues

Rank	Previous	Trending	Participant	Peak Period Availability
1	8	+ 7	Huntington	100.0%
2	1	- 1	Branch Bank and Trust	99.9%
2	2	- 1	Wells Fargo	99.9%

dynatrace