Exploring TouchXPRT 2016

TouchXPRT 2016 is a benchmark for evaluating the capabilities of Windows 10 and Windows 10 Mobile devices.

March 28, 2016

TOUCHXPRT

TABLE OF CONTENTS

INTRODUCTION

TouchXPRT 2016 is a free tool for evaluating the performance of Windows 10 and Windows 10 Mobile devices while handling common media tasks. It is available for download in the Windows Store or on the BenchmarkXPRT website. The full test typically takes less than 15 minutes to run, but the runtime varies depending on the system. You can compare the devices you test or compare your device against scores posted on the benchmark results page. Technology manufacturers, developers, sellers, tech review websites, and consumers around the world use TouchXPRT.

TouchXPRT 2016 includes the same performance workloads as TouchXPRT 2014, but is built as a Universal Windows app for Windows 10 and includes updated content. Because of the updated content and changes in our results calculations for TouchXPRT 2016, TouchXPRT 2016 results are not comparable to those of TouchXPRT 2014.

This paper describes the development process for TouchXPRT 2016 and the other BenchmarkXPRT tools. It explains what's new in this release, how TouchXPRT 2016 calculates results, how to download and run the benchmark, and how to submit results. It also describes the BenchmarkXPRT Development Community and how you can become a member.

DEVELOPMENT PROCESS

We build the BenchmarkXPRT apps using a unique community-driven model. Instead of the closed, bottom-up approach used by many benchmarking efforts, we use an open, top-down approach that includes the BenchmarkXPRT Development Community (the community) throughout the design, development, and testing process.[1]

Our approach starts by taking input from the community and examining the most common use cases for the platforms we plan to benchmark. We then write a Request for Comment (RFC) proposing use cases to incorporate into the benchmark and publish it to the community. The community's input on the RFC guides the drafting of a design document. The design document drives the implementation of one or more community previews, which we release to the community for input. We make changes based on that input and finalize the code to create a general release.

We keep results stable between the community previews and the general release so that community members can publish results from the community preview without waiting for the general release.

TOUCHXPRT 2016: THE DETAILS

TouchXPRT runs five workloads that simulate the kinds of photo, video, and audio tasks you do on your device each day (Beautify Photos, Blend Photos, Convert Videos for Sharing, Create Music Podcast, and Create Slideshow from Photos). The full test typically takes less than 15 minutes to run, but the runtime varies depending on the system. It has an intuitive user interface that produces easy-to-understand results for each of the five test scenarios, plus an overall score. When you're done testing, you can share your results with others and see how your device compares on the TouchXPRT results page.[2]

---

[1] We discuss the community in the section titled "What is the BenchmarkXPRT Development Community?"
[2] principledtechnologies.com/benchmarkxprt/touchxprt/2016/results

What's new with TouchXPRT 2016

TouchXPRT 2016 runs the same five workloads as TouchXPRT 2014, but is built as a Universal Windows app that runs on Windows 10 and Windows 10 Mobile devices. We made changes related to that transition, including modifying icons to meet Universal Windows app requirements.

We also made some changes to make the benchmark look good on Windows 10 Mobile phones. The main screen now lists tests vertically on small displays, and the redesigned workload activity screen looks better on phones. Rather than displaying both the project summary and workload activity information throughout the test run, we show the project summary screen as an overlay for just a few seconds before and during the test. During the remainder of the test, the workload activity screen occupies the whole area.
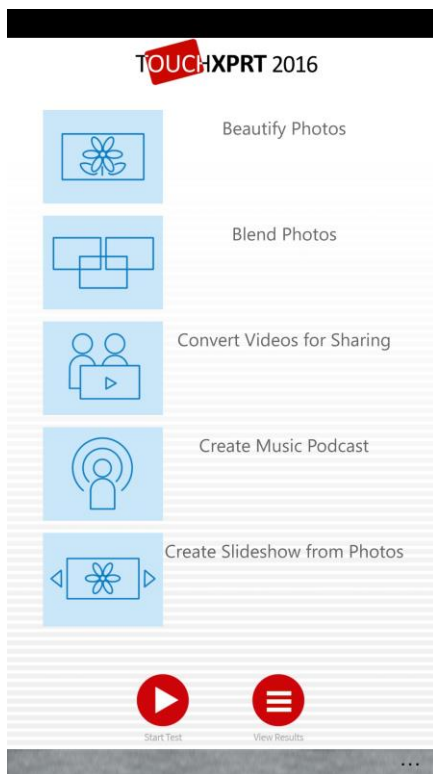

Figure 1: The UI in Windows 10 Mobile on a phone.

We updated some workloads with new content or output formats to either reflect current usages or to run successfully on Windows 10 Mobile phones.

Workload	TouchXPRT 2014	TouchXPRT 2016
Blend Photos	HDR photo with 3008x2000 resolution	Changed to HDR photo with 2497 x 1660 resolution due to memory issues on Windows 10 Mobile phones.
Convert Videos for Sharing	Input: 1080p video (1920x1080 ~ 84 sec)	Changed to 4K video (3840 x 2160 ~ 24 sec) to reflect the current 4K usages
	Output: SD video (854x480)	Changed to 1080p video (1920 x 1080)
Create Music Podcast	Merge Phase: output to .wma file (160 kbps)	Changed output to .m4a file (128 kbps) because Windows phones do not support .mp3 or .wma

Workload	TouchXPRT 2014	TouchXPRT 2016
		encode.
	Watermark phase: output to .mp3 file (160 kbps)	Changed output to .m4a file (128 kbps) because Windows phones do not support .mp3 or .wma encode.

Table 1: Workload content updates in TouchXPRT 2016.

TouchXPRT 2016 uses the same calculations for its overall and individual workload scores as TouchXPRT 2014. However, TouchXPRT 2016 uses a different calibration system to scale test results. Because of this difference and the changes in workload content, TouchXPRT 2016 scores are not directly comparable to TouchXPRT 2014 scores.

TouchXPRT 2016 Workloads

The five TouchXPRT 2016 workloads mirror common tasks related to creating and consuming media on Windows 10 and Windows 10 Mobile devices: Beautify Photos, Blend Photos, Convert Videos for Sharing, Create Music Podcast, and Create Slideshow from Photos.

To create a list of five tasks to test, we researched popular usages and typical media resolutions and file sizes. Obviously, this list does not reflect all the ways people use their touch devices, but we believe it accurately reflects the types of tasks that stress devices. When we measure such tasks accurately, they provide a solid basis for identifying performance differences between devices.

- **Beautify Photos.** This scenario represents a user applying one of three effects (Sharpen, Spotlight, and Emboss) to two sets each of four photos in a photo album (24 photos total, 16 megapixels each) and measures the elapsed time during each application. The scenario result (in seconds) is the sum of the elapsed times for applying each effect.
- **Blend Photos.** This scenario includes tasks for each of three different ways users blend photos. The first task, Photo Collage, blends three sets of eight photos from a photo album (24 photos total, 16 megapixels each) into three separate collages and adds a watermark to each collage. The second task, Smart Merge, removes moving objects from a burst of photos (5 photos total, 5 megapixels each). The third task, HDR, aligns photos and then performs an HDR operation (4 photos total, 4 megapixels each). The scenario result, reported in seconds, is the sum of the elapsed time for the five photos total output of the three blending tasks.
- **Convert Videos for Sharing.** This scenario represents a user converting two 4K videos in MP4 format (total length: approximately 24 seconds) to 1080p videos.[3] TouchXPRT measures the elapsed time in seconds for converting each video and sums those times for the scenario result.
- **Create Music Podcast.** This scenario represents a user creating an audio file for sharing. It mixes five audio files (each 8:16 minutes long, WMA, 160kbps), watermarks the mixed file, and then converts it to an AAC encoded audio file (8:15 minutes long, 131kbps) with an M4A extension. The scenario score is the sum of the elapsed time for applying the audio watermark and converting to AAC format. For TouchXPRT 2016, we

---

[3] TouchXPRT v2016.6.51.0 end-of-test results screens incorrectly show an output video length of 84 seconds for the Convert Videos for Sharing workload. The correct video length is 24 seconds. Later builds will correct this"

changed the target audio file to AAC format from the MP3 format in TouchXPRT 2014 because Windows 10 Mobile phones don't support encoding in MP3 format.
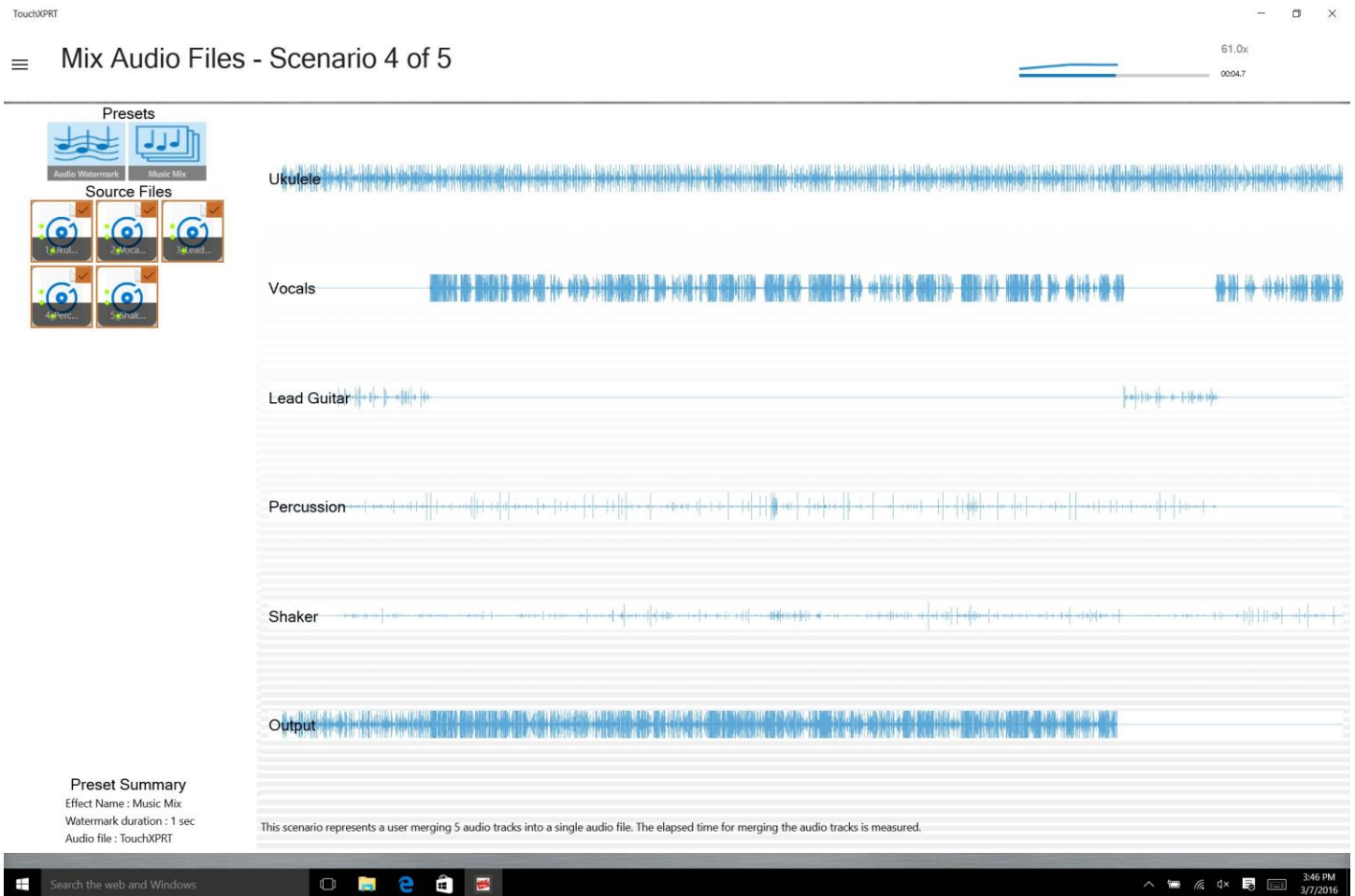

Figure 2: Mixing audio files during the Create Music Podcast test.

- **Create Slideshow from Photos.** This scenario represents a user creating a slideshow video (720p, MP4) with photos from an album (24 photos, 16 megapixels each). The scenario result is the sum of the elapsed time for adding photos to the slideshow and creating a video.

What do all the test results mean?

By default, the benchmark runs all five tasks and reports individual workload results and an overall score. Optionally, you can run individual tests or subsets of tests, in which case the benchmark does not produce an overall score.

TouchXPRT presents a variety of data points about a device's performance during the test. The most important data point, the overall score, is in the upper-right corner of the results page at the end of the test. You can use the overall score to compare your device to others on TouchXPRT.com. However, TouchXPRT provides more than just the overall score; it also gives you information about each individual test.
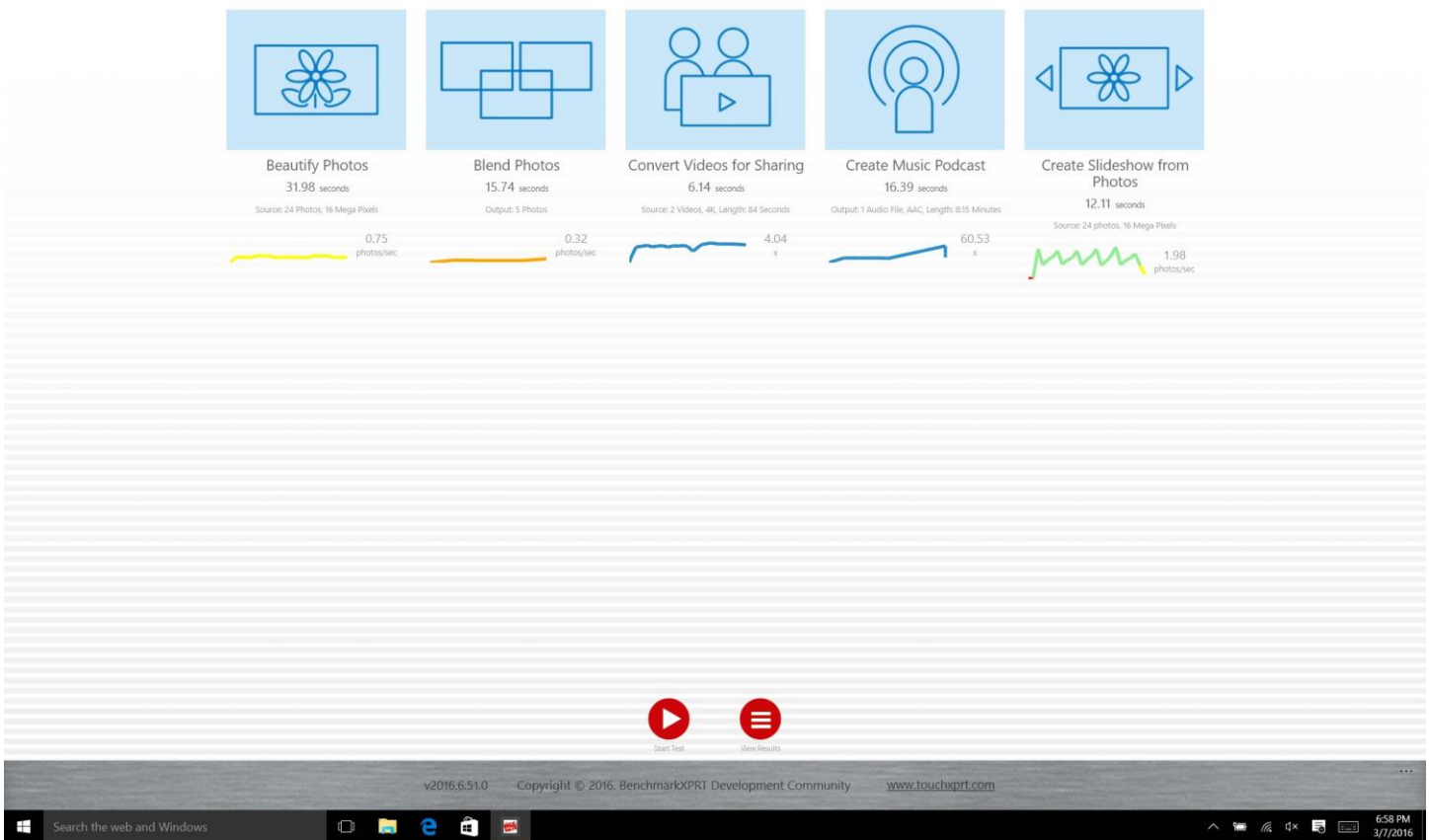
Figure 3: The TouchXPRT 2016 results page.

Here is a description of the other data presented on the results page:

- **Colored lines.** These lines serve as a graphic indicator of how quickly the device is going at any given point in the test. The lines update multiple times per scenario; for example, once for each of the 24 photos in the Beautify Photos scenario and once for each of the five photo blends in the Blend Photos test.

  For the scenarios that show rates of photos per second, the colors of the lines indicate the following:
    - Elapsed time between 0 ms – 150 ms = green (instantaneous)
    - Elapsed time between 150 ms – 1 s = light green (quick)
    - Elapsed time between 1 s – 2 s = yellow
    - Elapsed time between 2 s – 5 s = orange
    - Elapsed time > 5 s = red.

  For the Convert Music for Sharing and Create Music Podcast scenarios, which do not consist of a series of discrete operations, the lines are a neutral blue.

- **Rates.** These will use one of the units:

  o **Photos/sec.** The photos/sec rate tells you how many photos the device processed per second during a task. The same description applies to the photo blends/sec rate. Higher is better for this kind of score.

  o **#x.** This is the ratio of the length of the song or video to the time the test took to run. This measure allows us to compare clips of any length. Higher is better for this type of score as well.

  o **Seconds.** This is the time TouchXPRT spent working on that test. Because finishing a task in less time is preferable, a lower number of seconds is better. If you watch the clock during the test, you will notice that the test result time is less than the lapsed time on the clock. This is because TouchXPRT realistically implements the way people work, including the way people pause during their work. However, the pauses are the same on all systems, so we omit them from the results.

  o **Source and output.** These TouchXPRT constants indicate the quantity and quality of the audio and visual media the test uses. They will be the same on all devices. If your device were a car, it would be like saying that the mileage test took place on a 40-mile track.

Example calculation of the overall score

In this section, we show the results from a Microsoft Surface Pro 3. For each test, TouchXPRT 2016 records the time each workload required to do its work. For the example run on the Surface Pro 3, the times were as follows:

Workload name	Time (seconds)
Beautify Photos	31.98
Blend Photos	15.74
Convert Videos for Sharing	6.14
Create Music Podcast	16.39
Create Slideshow from Photos	12.11

Table 2: Sample workloads times for the Microsoft Surface 3.

The timings include only the time spent working. Wait times are excluded, as are items that take a fixed amount of time. So, the timings reported by TouchXPRT will not typically match the elapsed wall time.

From the timings, TouchXPRT 2016 computes per-test rates and the overall score. While lower times are better, the per-test rates allow you to compare the tests with a "bigger is better" measure. The table below shows the rates for each test, along with a brief explanation of how they're calculated.

Test name	Rate	How the rate is computed
Beautify Photos	0.75 photos per second	Number of photos (24) divided by time.
Blend Photos	0.32 photo blends per second	Number of resulting photo blends (5) divided by time.
Convert Videos for Sharing	4.04x	Time to play the video divided by the time to convert the video. The result is how many times longer it would take to play the video than it took to do the conversion.
Create Music Podcast	60.53x	Time to mix and watermark the music divided by twice the time required to play the music (2 x 8:15 minutes).

Test name	Rate	How the rate is computed
Create Slideshow from Photos	1.98 photos per second	Number of photos (24) divided by time.

Table 3: Computing the rate of each workload.

To compute the Overall score, we scale the times for the test machine using times from a calibration device. For TouchXPRT 2016, the calibration device is a Microsoft Surface 3. The table below shows this calculation. This calculation also converts the lower-is-better time results to higher-is-better scaled results.

Test name	Test system results (seconds)	Calibration system results (seconds)	Scaled results (calibration result divided by test result)
Beautify Photos	31.98	68.2	2.13
Blend Photos	15.74	36.7	2.33
Convert Videos for Sharing	6.14	11.3	1.84
Create Music Podcast	16.39	31.0	1.89
Create Slideshow from Photos	12.11	29.4	2.43

Table 4: Scaling workload scores from the test machine using calibration system values.

To compute the Overall score, we take the geometric mean[4] of the resulting scaled values and multiply by 100. The geometric mean of the five scaled results rounds to 2.11. We get rid of the decimal by multiplying by 100. The result is 211, the Overall score reported by TouchXPRT 2016 for this run as shown earlier in Figure 2.

TOUCHXPRT INSTRUCTIONS

Where to get it

Simply download TouchXPRT5 from the Microsoft Store or from TouchXPRT.com6. The first time you launch TouchXPRT, it will prompt you to download the content for the tests. Go ahead and follow the instructions; you'll have to do this only once. This process should take about fifteen minutes.

How to run it

Before you start, do the following:

1. Go to your device's Settings menu.
2. Set your device to stay awake for at least 30 minutes.
3. Reboot your device.
4. Quit all the other apps that may have started automatically.

These steps will help you get the most accurate results. Now you're ready to test!

---

[4] en.wikipedia.org/wiki/Geometric_mean

[5] https://www.microsoft.com/en-us/store/apps/touchxprt-2016/9nblggh5g0bj

[6] www.touchxprt.com

When you click the Start Test button, it may seem as if something is taking over your device, but don't worry – this is how it's supposed to look! Running all the tests should take about 15 minutes. (This can vary widely by device, however, so don't worry if the tests take more or less time.)

Follow these three important rules to get accurate results:

1. Do not let the device fall asleep.
2. Do not use other apps while the test is running.
3. Do not try to interrupt or stop the tests before they finish.

You can also run single tests by tapping an individual tile, but you won't get an overall score that way.

Comparing results to the database

After you run the tests, TouchXPRT will show you its overall score, which you can compare to others on the TouchXPRT 2016 results page[7]. You'll find scores there from a wide range of Windows 10 and Windows 10 Mobile devices. To find detailed information on any set of scores, click the link under the Source column. Results for most recent test run are stored locally in Pictures\TouchXPRT2016\Docs\TouchXPRTResult.xml.

Submitting results

After trying out TouchXPRT2016, please submit your scores and send any comments to BenchmarkXPRTsupport@principledtechnologies.com. Click the View Results button to see the results of past runs. To submit a result, click the View/Submit button associated with that result. Enter your contact information, Windows version, the amount of RAM in the system, and any additional information you'd like to include, then click Submit.

---

7 principledtechnologies.com/benchmarkxprt/touchxprt/2016/results

Automation

TouchXPRT 2016 allows you to automate your testing procedure. You can run the benchmark from the command line or a script using the start command. The command line options include:

- Start running the test automatically on launch. You choose to run all scenarios and generate an overall score, or run any given scenario.
- Set the number of iterations you want the test to run.
- Change the delay before the test starts running after launch, by specifying the desired interval in milliseconds. By default, TouchXPRT 2016 waits for 5,000 ms before it starts running. This allows the app enough time to load and allows the system to stabilize.
- Change the output filename.

Running TouchXPRT from the command line or a script

There are two ways to launch the app with a command:

- From the command prompt or a batch file, launch the app using
  ```
  start touchxprt2016://
  ```
- From the Run window (Start+R keys), launch the app using
  ```
  touchxprt2016://
  ```

You can also use the options to automate one or more test runs. All switches are optional. However, a result of this behavior is that omitting the scenario option causes TouchXPRT 2016 to pause at its UI. As a general rule, we do not recommend mixing scripting and manual runs, so we encourage the use of the scenario option.

TouchXPRT 2016 ignores invalid switches or switches with invalid values and runs the test using TouchXPRT's typical defaults.

Selecting tests to run using the scenario option

To run all scenarios and generate an overall score:
```
touchxprt2016://options/scenario=all
```
To run any particular scenario, use the numbers 1-5, as in this example:
```
touchxprt2016://options/scenario=1
```

The scenarios are numbered as follows, in the order they show on the UI:

1. Beautify Photos
2. Blend Photos
3. Convert Videos for Sharing
4. Create Music Podcast
5. Create Slideshow from Photos

Note: You must specify the scenario option each time you launch TouchXPRT from the command line or a script unless you want it to stop at the UI.

In the following examples, the scenario switch is set to all, but it can be set to the number of any given scenario.

Specifying the number of iterations

To specify the number of iterations:

```
touchxprt2016://options/scenario=all/iterations=2
```

Changing the default five-second delay

By default when the app is launched, it waits 5,000 ms before it starts running tests. This delay allows the app to load and the system to stabilize. To change the start delay, use the startdelay option to specify the delay in milliseconds. This example shows how to change the delay to 6,000 ms:

```
 touchxprt2016://options/scenario=all/startdelay=6000
```

Changing the output filename

All results save to Pictures\TouchXPRT2016\Docs. By default, the results are stored in a file named results.xml. However, you can change the name of the output file to identify the test according to your conventions.

For example, to change the filename to MyTestSystem.xml:

```
touchxprt2016://options/scenario=all/filename=MyTestSystem.xml
```

These commands can be combined as follows:

```
touchxprt2016://options/scenario=all/iterations=3/startdelay=6000/filename=MyTestSystem.xml
```

Batch file commands to check when test is complete

During the run, TouchXPRT 2016 stores its results in a temporary file. Scripts should test for the existence of the results file to determine that the test is complete. There are several ways to do this. We offer an example of how to test from a DOS batch file below. This example polls every 5 seconds for the MyTestSystem.xml file.

```
start touchxprt2016://options/scenario=all/filename=MyTestSystem.xml

:WAITRESULT

IF EXIST %homepath%\pictures\touchxprt2016\docs\MyTestSystem.xml GOTO DONE

    timeout /t 5 /nobreak > nul

goto WAITRESULT

:DONE
```

Proceed with the testing logic

ABOUT THE BENCHMARKXPRT FAMILY

The BenchmarkXPRT tools are a set of apps that help you test how well devices do the kinds of things you do every day. In addition TouchXPRT 2016 and TouchXPRT 2014, the BenchmarkXPRT suite currently includes the following tools:

- BatteryXPRT 2014 for Android measures the battery life of Android-based phones and tablets.
- CrXPRT 2015 tests the responsiveness and battery life of Chromebooks.
- HDXPRT 2014 uses commercial applications to test the capabilities and responsiveness of PCs.

- MobileXPRT 2015 tests the responsiveness of Android devices. MobileXPRT 2015 supports both 32-bit and 64-bit hardware running Android 4.4 or higher. To test systems running older versions of Android, or to test 32-bit performance on a 64-bit system, you can use MobileXPRT 2013. The results of the two versions are comparable.
- WebXPRT 2015 tests the performance and Web browsing capabilities of any device with Internet access.

We designed these apps to test a wide range of devices on a level playing field. When you look at results from XPRTs, you get unbiased, fair product comparison information. For a guide to which XPRTS to run on your devices, see our Infographic: Consult the right XPRTs.[8]

WHAT IS THE BENCHMARKXPRT DEVELOPMENT COMMUNITY?

We built BenchmarkXPRT around a unique community model. Community membership is open to anyone, and there are many different ways to participate. Members of the BenchmarkXPRT Development Community can be involved in every step of the process for TouchXPRT 2016 and the other XPRT benchmarks that are of interest to them. They can give input on the design of upcoming versions, contribute source code, and help test the benchmarks. Community members have access to previews of upcoming benchmarks and to the source code of released benchmarks.

The community helps us avoid the ivory tower syndrome. Diversity of input during the design process makes the tests more representative of real world activity. Giving community members access to the source code serves to increase confidence in the code and improves future versions of the benchmarks.

The community model differs from the open source model primarily by controlling derivative works. It is important that the BenchmarkXPRT benchmarks return consistent results. If the testing community calls different derivative works by the same name, the result would be that the test results would not be comparable. That would limit, if not destroy, the tools' effectiveness.

For more information, please watch the video introduction to the community [9] and consult the BenchmarkXPRT FAQ page.[10] To see just how far the reach of the BenchmarkXPRT community extends, check out our infographic.[11]

How to join

If you are not currently a community member, we encourage you to join. Registration is simple,[12] and does not obligate you to participate.

---

[8] principledtechnologies.com/benchmarkxprt/consult-the-right-xprts

[9] You will find a video introduction to the BenchmarkXPRT Development Community at www.youtube.com/embed/7Cb6zGHM4Xc.

[10] www.principledtechnologies.com/benchmarkxprt/faq.php

[11] www.principledtechnologies.com/benchmarkxprt/around-the-world-infographic

[12] You can register at www.principledtechnologies.com/hdxprt/forum/register.php

WHERE CAN I GET MORE INFORMATION?

Go to TouchXPRT.com or follow us on Twitter and Facebook. We announce breaking news on the BenchmarkXPRT blog (available to everyone) and the BenchmarkXPRT forums (available to members only). If you cannot find the answer to your question in this white paper, or you need help with TouchXPRT, send an email to our team at BenchmarkXPRTsupport@principledtechnologies.com.

CONCLUSION

We hope this paper has answered any questions you may have about TouchXPRT. If you have suggestions about ways to improve the benchmark, or if you have any other questions, please post them on the community forum or e-mail us at BenchmarkXPRTsupport@principledtechnologies.com. For more information, visit us at www.benchmarkxprt.com and www.touchxprt.com.

ABOUT PRINCIPLED TECHNOLOGIES

Principled Technologies, Inc.

1007 Slater Road, Suite 300

Durham, NC, 27703

www.principledtechnologies.com

We provide industry-leading technology assessment and fact-based marketing services. We bring to every assignment extensive experience with and expertise in all aspects of technology testing and analysis, from researching new technologies, to developing new methodologies, to testing with existing and new tools.

When the assessment is complete, we know how to present the results to a broad range of target audiences. We provide our clients with the materials they need, from market-focused data to use in their own collateral to custom sales aids, such as test reports, performance assessments, and white papers. Every document reflects the results of our trusted independent analysis.

We provide customized services that focus on our clients' individual requirements. Whether the technology involves hardware, software, websites, or services, we offer the experience, expertise, and tools to help our clients assess how it will fare against its competition, its performance, its market readiness, and its quality and reliability.

Our founders, Mark L. Van Name and Bill Catchings, have worked together in technology assessment for over 27 years. As journalists, they published over a thousand articles on a wide array of technology subjects. They created and led the Ziff-Davis Benchmark Operation, which developed such industry-standard benchmarks as Ziff Davis Media's Winstone and WebBench. They founded and led eTesting Labs, and after the acquisition of that company by Lionbridge Technologies were the head and CTO of VeriTest.

Disclaimer of Warranties; Limitation of Liability:

PRINCIPLED TECHNOLOGIES, INC. HAS MADE REASONABLE EFFORTS TO ENSURE THE ACCURACY AND VALIDITY OF ITS TESTING, HOWEVER, PRINCIPLED TECHNOLOGIES, INC. SPECIFICALLY DISCLAIMS ANY WARRANTY, EXPRESSED OR IMPLIED, RELATING TO THE TEST RESULTS AND ANALYSIS, THEIR ACCURACY, COMPLETENESS OR QUALITY, INCLUDING ANY IMPLIED WARRANTY OF FITNESS FOR ANY PARTICULAR PURPOSE. ALL PERSONS OR ENTITIES RELYING ON THE RESULTS OF ANY TESTING DO SO AT THEIR OWN RISK, AND AGREE THAT PRINCIPLED TECHNOLOGIES, INC., ITS EMPLOYEES AND ITS SUBCONTRACTORS SHALL HAVE NO LIABILITY WHATSOEVER FROM ANY CLAIM OF LOSS OR DAMAGE ON ACCOUNT OF ANY ALLEGED ERROR OR DEFECT IN ANY TESTING PROCEDURE OR RESULT.

IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC. BE LIABLE FOR INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH ITS TESTING, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC.'S LIABILITY, INCLUDING FOR DIRECT DAMAGES, EXCEED THE AMOUNTS PAID IN CONNECTION WITH PRINCIPLED TECHNOLOGIES, INC.'S TESTING. CUSTOMER'S SOLE AND EXCLUSIVE REMEDIES ARE AS SET FORTH HEREIN.